JUMS

Volume 7, Issue 4, September 2022

Advisory Editorial Board:

DOMINIK VAN AAKEN
FREDERIK AHLEMANN
JAN-PHILIPP AHRENS
BASTIAN AMBERG
THOMAS BAHLINGER
CHRISTOPH BODE
SULEIKA BORT
ROLF BRÜHL
KATRIN BURMEISTER-LAMP
JOACHIM BÜSCHKEN
CATHERINE CLEOPHAS
RALF ELSAS
KERSTIN FEHRE
MATTHIAS FINK
DAVID FLORYSIAK
GUNTHER FRIEDL
FRANZ FUERST
WOLFGANG GÜTTEL
CHRISTIAN HOFMANN
KATJA HUTTER
LUTZ JOHANNING
STEPHAN KAISER
NADINE KAMMERLANDER
ALFRED KIESER
NATALIA KLIEWER
DODO ZU KNYPHAUSEN-AUFSEß
SABINE T. KÖSZEGI
ARJAN KOZICA
CHRISTIAN KOZIOL
MARTIN KREEB
TOBIAS KRETSCHMER
HANS-ULRICH KÜPPER
MICHAEL MEYER
JÜRGEN MÜHLBACHER
GORDON MÜLLER-SEITZ
J. PETER MURMANN
ANDREAS OSTERMAIER
BURKHARD PEDELL
MARCEL PROKOPCZUK
TANJA RABL
SASCHA RAITHEL
NICOLE RATZINGER-SAKEL
ASTRID REICHEL
KATJA ROST
THOMAS RUSSACK
FLORIAN SAHLING
MARKO SARSTEDT
ANDREAS G. SCHERER
STEFAN SCHMID
UTE SCHMIEL
CHRISTIAN SCHMITZ
MARTIN SCHNEIDER
LARS SCHWEIZER
DAVID SEIDL
THORSTEN SELLHORN
ANDREAS SUCHANEK
TILL TALAULICAR
ORESTIS TERZIDIS
ANJA TUSCHKE
MATTHIAS UHL
CHRISTINE VALLASTER
CHRISTIAN VÖGTLIN
STEPHAN WAGNER
BARBARA E. WEIßENBERGER
ISABELL M. WELPE
HANNES WINNER
THOMAS WRONA
THOMAS ZWICK

JUNIOR MANAGEMENT SCIENCE

Xiaoyu Hu, Exploring How Macroeconomic Factors Affect REITs and Evaluating Its Downside Risk – Empirical Evidence From China and the US 874

Jonas Spengler, Moderation in Marketinganwendungen von Partial Least Squares Strukturgleichungsmodellierungen (PLS-SEM) 899

Sebestyén András Huszár, Die normalisierende Macht der Digitalisierung in der Arbeitswelt – Eine Analyse nach Michel Foucaults Machttheorie 932

Leon Kowalke, The Idiosyncratic Volatility Puzzle – Anomaly or Data Mining? 945

Marco Wall, One User – Two Viewpoints? An Examination of Information Privacy Concerns from the Employee and Consumer Perspective 986

Milan Schulze, Digitalisierung in der Finanzberichterstattung – Folgen und Konsequenzen für den Einsatz des neuen Mediums 1001

Jan Felix Csavajda, Die Organisation der Produktionsarbeit der Zukunft – Anforderungen und technische Lösungsansätze 1032

Marcel Schemmel, Digital Innovation in Corporations: Deriving a Practical Framework for the Measurement of Success of Digital Innovation Units 1098

Niclas Vogt, Die Rolle der Selbstaufwertung beim Einsatz von Statusanreizen im Vertrieb 1133

Malte Ludwigs, Bilanzierung von Spielerwerten im deutschen Profi-Fußball – Eine kritische Analyse der handelsrechtlichen Vorschriften 1147

Philipp S.R. Voss, Innovation Performance in Healthcare M&A: An Empirical Analysis 1164

Published by Junior Management Science e.V.

Junior Management Science 7(4) (2022) 874-898

Junior Management Science

journal homepage: www.jums.academy

Exploring How Macroeconomic Factors Affect REITs and Evaluating Its Downside Risk – Empirical Evidence From China and the US

Xiaoyu Hu

Technische Universität München

Abstract

Real Estate Investment Trust (REIT) is considered as a financial instrument operated and managed by professional management teams based on a range of income-producing real estate. The focus of this thesis is on publicly traded equity REITs. There are four research questions that this thesis attempts to answer. How did REITs develop in the United States (US)? What are the critical factors that incentivized the Chinese government to promote REITs, and what is the progress? Are REITs a good hedge against macroeconomic risk factors? How can the downside risk of REITs be evaluated? To begin, the first two questions have been answered using the literature review methodology. The VAR model is constructed to evaluate the relation between the REIT market and macroeconomic factors. Ultimately, downside risk of REIT market is assessed by the GARCH(1,1)-VaR model based on the student's t-distribution.

Keywords: Equity REITs; Macroeconomic risks; VAR; VaR; GARCH(1,1).

1. Introduction

This chapter describes the general research background and motivation. Research questions will also be mentioned and proposed. Related literature will be summarised. Afterwards, methodologies are briefly described. Finally, an overview of the thesis structure is provided.

1.1. Research Background and Motivation

Real Estate Investment Trust (REIT) is an alternative to direct physical real estate investment. REITs are typically publicly listed on exchanges, so they are easier to buy and sell. Block (2011) stated that there are two types of REITs. One is mortgage REITs, which is a debt investment based on real estate related collateral. On the other hand, equity REITs is much more common, enabling investors to own real estate shares and usually providing a source of income/cash flow for investors through dividends. Compared with mortgage REITs, equity REITs are less affected by the interest rate change and have better returns historically (Block, 2011). The focus of this thesis is on equity REITs. The REIT-related concepts in the thesis, if not specified, refer to publicly traded equity REITs.

The US was the first country to introduce REIT products to the market in 1960. Since then, approximately 40 countries/regions have adopted the US-based REIT approach according to the National Association of Real Estate Investment Trusts (NAREIT) (*Nareit Global real estate investment*, 2021). [1] In comparison, in China in 2008, an official report published by the State Council of China first discussed the concept of REIT. As part of the development of the financial system, complex products such as REITs are introduced slowly in China.

On 30th January 2021, the Shanghai and Shenzhen stock exchanges in China finalized the official rules for REITs listing in China. On 17th May 2021, 9 infrastructure REITs as a pilot program obtained the Initial Public Offering (IPO) permission in mainland China. Although REITs have begun a new phase in China, there is still not sufficient research on the risks of this asset class domestically.

This thesis seeks to review the framework and system of REITs in the US market in order to draw lessons for China. In addition, empirical methods would be applied to evaluate REITs performance in relation to various macroeconomic risks and measure REITs' downside risk.

How did REITs develop in the US? What are the critical factors that incentivized the Chinese government to promote REITs, and what is the progress? Are REITs a good hedge

[1] NAREIT is a trade association that represents US REITs and listed real estate companies. The association provides comprehensive industry data on the performance of the industries respectively.

DOI: https://doi.org/10.5282/jums/v7i4pp874-898

against macroeconomic risk factors? How can the downside risk of REITs be evaluated? This thesis attempts to answer these questions by literature review and building empirical models based on data from the US and China to evaluate the risks of REITs.

1.2. Literature Review

Several studies have been done to research further on REITs as an asset class. Goddard and Marcum (2012) systematically introduced a variety of real estate related financial instruments, such as the history, types and investment strategy of REITs (Goddard & Marcum, 2012). Yin (2019) focused on residential REITs and used development history in the US market as implications for analyzing the opportunities and obstacles for REITs in China (Yin, 2019).

There have been some studies that evaluated the risk-return performance of REITs in relation to macroeconomic factors. Early examples of research include Liu and Mei (1992) as well as Peterson and Hsieh (1997), both of which suggested that the expected excess returns of REITs investment are similar to stocks. The studies indicated that REITs are also dependent on business cycles as well as the broader macroeconomy (Liu & Mei, 1992; Peterson & Hsieh, 1997). Brooks and Tsolacos (1999) used the VAR model to evaluate the degree of influence that macroeconomic factors have on real estate. The factors considered were unemployment, inflation, dividend yield and nominal interest rates. Among these factors, the interest rate term structure and unexpected inflation were suggested to have contemporaneous effects (Brooks & Tsolacos, 1999). Park, Mullineaux, and Chew (1990) and Simpson, Ramchander, and Webb (2007) have made an extensive study on the relation between inflation and REIT returns. US data were used and the suggested result was that REITs could partially hedge inflation risk (Park et al., 1990; Simpson et al., 2007). However, Wong et al. (2017), who engaged in the use of the multi-factor asset pricing model using 20 years of data in Australia, indicated different results. Several macroeconomic variables have been used, which included stock market return, unexpected inflation, expected inflation, risk premium and term structure. In the research, REITs were found to have a negative relation to unexpected inflation (Wong et al., 2017). West and Worthington (2006) also performed similar studies using data from Australia based on the GARCH-M model, and suggested a significant relationship between commercial REITs return and the combined factors of interest rate, unexpected inflation and construction index (West & Worthington, 2006).

The factor interest rate has appeared in multiple studies with regards to its influence on REIT returns. Allen, Madura, and Springer (2000) explored the influence of stock-market returns, long-term and short-term interest rate on publicly traded REITs using seemingly unrelated regression. Afterwards, the study explored how REITs can alter their risk exposure by considering some characteristics, such as asset structure and category. The result suggested that those factors significantly impact REITs, although there are some slight sensitivity differences between equity and non-equity REITs (Allen et al., 2000). Such findings are consistent with the research of Swanson, Theis, and Casey (2002), where the result based on the Cobb-Douglas regression model suggested that interest rate impacts REITs returns and REITs are more sensitive to treaty bills than commercial bonds (Swanson et al., 2002).

The relationship between REITs and the macroeconomy has also been investigated for some Asian markets. Fang, Chang, Lee, and Chen (2016) had used the Autoregressive distributed lag model and Granger causality test, which found that a long-run equilibrium exists between the REIT index and the interest rate, inflation rate, and stock index in Japan, Singapore and China (Fang et al., 2016). The research of Loo, Anuar, and Ramakrishnan (2016) utilized data from seven Asian markets based on the Johansen cointegration test and Granger causality test and pointed out that the emerging markets are more sensitive to the change of macroeconomic environment (Loo et al., 2016).

Liow, Ibrahim, and Huang (2006) utilized principal components analysis, GARCH(1,1) model and Gaussian mixture model to investigate how macroeconomic risk variables impact excess returns of REIT using data from three Asian markets as well as the UK market. It was suggested that across different markets, the significance and impact of macroeconomic risk factors varies (Liow et al., 2006). The research of Kola (2016) used a similar three-step approach to evaluate the influences of the business cycle, price stability, exchange rates and interest rates on REIT pricing in the US, Bulgaria and South Africa. The idea was put forward that industrial production and inflation are significant factors in developed markets unlike the developing market (Kola, 2016).

1.3. Methodology Overview

Literature review method will be used to illustrate the history and development of REITs to answer the first two research questions.

To explore the relationship between macroeconomic factors and REIT returns, this study will use the Vector Autoregression (VAR) model, consistent with that from Brooks and Tsolacos (1999) (Brooks & Tsolacos, 1999). As an overview, the VAR model will attempt to capture the relationship between multiple time-series data with the end goal of highlighting the interaction between all variables. As the time series variables would need to be stationary, the Augmented Dickey Fuller (ADF) test would first be applied. Afterwards, the Granger causality test, impulse response analysis and variance decomposition analysis are applied as a supplement to unravel the relationship between variables (Lütkepohl, 2005). [2]

To evaluate the downside risk, Value at Risk (VaR), a measure of the risk of loss, has been used. In this thesis, the underlying research object is daily REIT logarithmic return. The

[2] See Lütkepohl (2005) pp. 13-66.

underlying assumption of directly calculating VaR requires normal distribution of returns. However, most financial time series data are usually characterized by the heavier tail and volatility clustering. Therefore, the Generalised Autoregressive Conditional Heteroskedasticity (GARCH) family model, which takes the special properties of financial series data into account, is widely used for an accurate VaR estimation. Miletic and Miletic (2015) had evaluated the performance of GARCH-VaR models based on data of five selected European capital markets. The result showed that GARCH models with the t-distribution in most analyzed cases provide better VaR estimation (Miletic & Miletic, 2015). This finding is consistent with Cerović Smolović, Lipovina-Božović, and Vujošević (2017), whose research indicated that GARCH(1,1) with the t distribution of residual is appropriate for capturing volatility clustering
(Cerović Smolović et al., 2017). Therefore, the classical GARCH(1,1)-VaR would be applied in this thesis to evaluate the downside risk of REITs.

1.4. Research Structure

This thesis consists of five chapters to provide a structured approach to answer the research questions. The detailed research flow chart is shown in Figure 1.

Chapter one describes the general research background and proposes the research questions. Literature review, including discussion of previous related studies and the methodologies, are also found in this chapter.

Chapter two focuses on the general concept and history of REITs and attempts to answer the first two research questions mentioned in section 1.1. Theories of REITs are also discussed in this chapter, with a focus on the US and China market. An initial summary will be made at the end of this chapter.

Chapter three discusses the relationship between various macroeconomic factors and REIT returns. Context to the data is first explained for both macroeconomic factors and REITs. The five macroeconomic factors used in this thesis are inflation, interest rate, growth in money supply, growth in industrial production, and the returns of the stock market. The VAR model would be constructed and results would be analyzed. A short summary at the end of the chapter will follow.

Chapter four examines the downside risk of REITs. First, the VaR, as well as the GARCH model will be discussed. Afterwards, training data and validation data will be specified. The Autoregressive Conditional Heteroskedasticity (ARCH) effect would be tested. Afterwards, the Kupiec test would be applied to validate the GARCH(1,1)-VaR model.

Chapter five draws the conclusion for this thesis and provides recommendations for further studies on this topic.

2. REIT Overview

This chapter introduces the basic concept and characteristics of REITs. The development of REITs in the US and the exploration of China will be specified.

2.1. Characteristics of REITs

REITs are financial instruments operated and managed by professional management teams based on a range of income-producing real estate. There are two main characteristics of REITs. First, its underlying assets are mainly real estate, and the primary source of income is rental income. Often, there is also the capital appreciation of the underlying real estate. Second, REITs typically distribute the majority of the annual net income to investors as dividends. REITs can be first traced back to the US by activities and discussion in Congress in 1960. The original idea of a REIT was to provide real estate investment opportunities to small private investors (or retail investors), which was previously only open to high-net-worth individuals or institutional investors. Meanwhile, REITs would broaden the traditional financing channel of the real estate industry and also improve capital allocation and efficiency. There are some fundamental legal requirements for a REIT in the US (Goddard & Marcum, 2012; Block, 2011): [3]

- Asset allocation requirement: At least 75% of its assets must be invested in real estate properties or equivalent
- Shareholder requirement: Must be 100 shareholders at a minimum and no more than 50% of the shares held by five or fewer individuals
- Income requirement: At least 95% of gross income must be received from real estate related activities
- Dividend requirement: Must distribute at least 90% of its annual taxable income to investors

In the years that follow, forty countries, including several emerging markets, have adopted the REIT framework learned from the experience of the US. Although REITs issued by different countries may differ slightly in terms of policy and management strategy, the definition of REITs in various countries worldwide is consistent. REITs can be thought of as a collection of funds through the issuance of shares, which are managed by a particular custodian. A professional institution is entrusted to carry out real estate related investment management, and the net profit of the investment is typically directly distributed to investors as dividends.

2.2. The Development of REITs in the US

The US has the longest history of publicly traded REITs and has one of the most mature REIT markets in the world. Figure 2 shows the historical development of US equity REITs market capitalization. Its development can be divided into the following four phases (Block, 2011; Yin, 2019).

> **Phase one: Emerging (1960s-1980s)** The US Congress passed the Real Estate Investment Trust Act in 1960, which paved the way and developed the REITs which

[3] See Block (2011) Section 3.1.

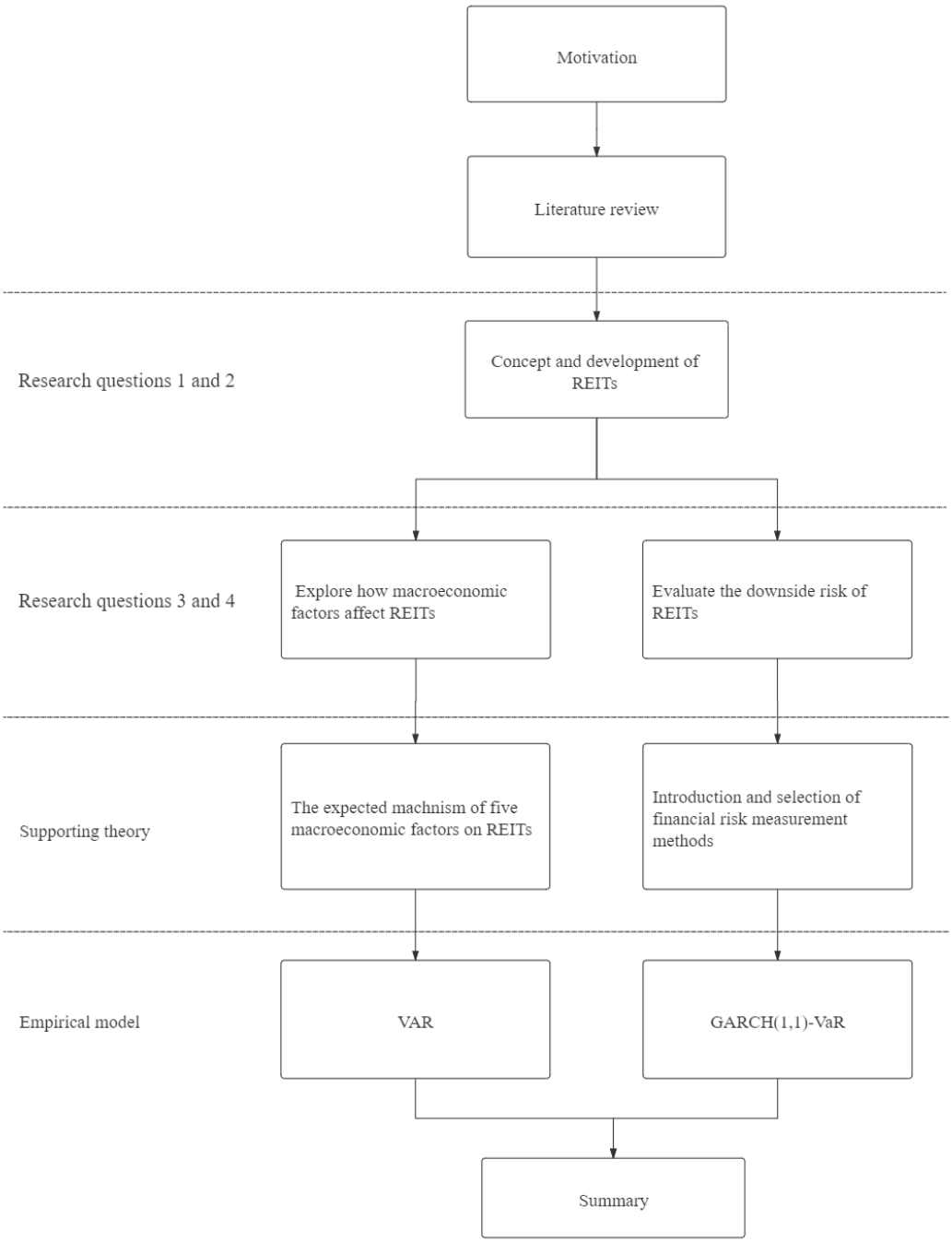


Figure 1: Research Flow Chart

Source: Author's representation

are known today. The main goal of the act was to authorize the mechanisms to provide for a real estate ownership structure and to have similar tax properties as that of a mutual fund. This way, the problem of double-taxation can be managed as the mechanisms of REITs were engineered such that they would distribute most of its earnings and capital gains. The problem of double-taxation was crucial as it would not be attractive to have an entity paying taxes at the corporate level in addition to investors (often private individuals) facing additional individual taxes on the dividends received. By mimicking elements of mutual funds, the REITs would not be subject to corporate taxes as most earnings (or almost all of its income) are passed to

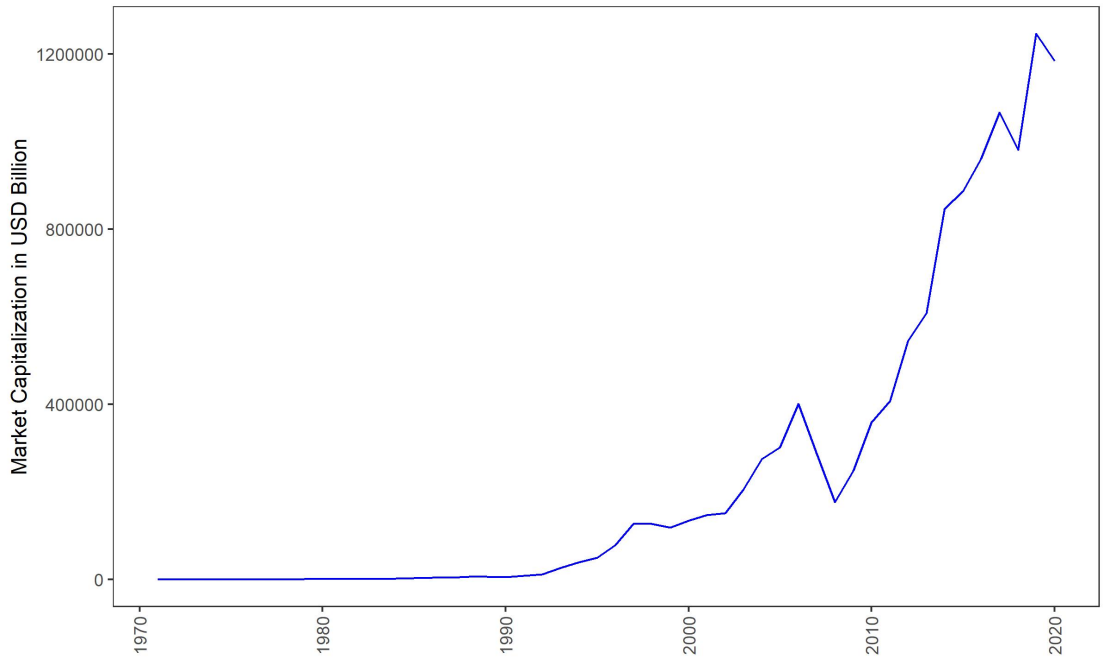


Figure 2: The US Equity REITs Market Capitalization From 1971 to 2020

Source: NAREIT (*Nareit US REIT industry equity market cap*, 2021)

investors as dividends, where they then pay individual taxes on these dividend income.

Although comparisons between REITs and mutual funds were made, there is a distinct difference that should be noted. Mutual funds typically purchase equities of companies, and they do not usually need to manage the operations of these companies. On the other hand, REITs hold properties that require active professional management, including operation (cleaning, repair work) and administration (tenant management, rental negotiation). As such, REITs would need to have a structure (or rather, a management team) in place to manage the properties. In the early years, REITs contributed little to the overall real estate investment market, and the need for professionally active and operational management could be one of the factors.

Phase two: Booming (1990s) Market awareness about REITs grew over the decades. The REITs market capitalization, number of REIT listings and investor interest skyrocketed in the 1990s. A number of reasons played a role on this topic, which will be discussed below.

First, there were significant changes to legislation. The Omnibus Budget Reconciliation Act of 1993 tweaked the ownership structure of REITs for pension funds. Due to this tweak, all investors in a pension fund could be counted as individuals for REIT investment purposes.[4] Ultimately, this helped to encourage and improve accessibility to REITs for the small private investors.

Second, the REIT Modernization Act of 1999 provided two significant points to support the REIT market. Mandatory income distribution of REITs was reduced from 95% to 90%.[5] Next, management of REITs were given more considerable independent flexibility on the topic of operational controls. REIT management could provide a variety of services to tenants without independent contractors and they also had a more extensive scope over investment decisions.

Next, financial innovation such as the umbrella partnership REITs also contributed to this skyrocketing growth of the REIT market. In such a structure, there is the standard REIT entity as well as an operating partnership. In this context, properties are owned by the REITs indirectly through the operating partnership. Owners of these operating partnerships can convert their units into shares of the REITs, thus deferring capital gain taxes until the actual transaction/conversion. This innovation increased the attractiveness of REITs as investors can better perform tax and financial planning.

[4] Specifically the shareholder requirements for REITs mentioned in Section 2.1

[5] Specifically the dividend requirements for REITs mentioned in Section 2.1

Phase three: Consolidating (1997-2008) Cycles happen in the market, and in this case, the REIT market was in a "correction" from 1997 to 1999, where the total market capitalization decreased from United States Dollar (USD) 127 billion to 118 billion (*Nareit US REIT industry equity market cap*, 2021). Various factors played a part, including overheating in the REIT market with too high a competition between lenders, which created unrealistic spreads and expected returns. Furthermore, there was the Asian financial crisis in 1997 as well as the Brazil financial crisis in 1999 - both of which had trickled impacts into the US market.

Afterwards, consolidation in the REIT market occurred. To remain competitive, companies were merging and the consolidation happened until 2008. Specialization of REITs also became a standard feature, as it became apparent that unique skill sets would be necessary to tailor for different types of REITs such as those involved in shopping centers, logistics, data centers or healthcare. This was an important discovery and development phase for the REIT market as a whole.

In 2007 the subprime mortgage crisis happened, and REITs had a significant crash in percentage terms. Kawaguchi, Sa-Aadu, and Shilling (2017) argued that the declining commercial mortgage 10 year Treasury yield spread during the Greenspan era (1994-2006) allowed the managers of REITs to be over leveraged with debt and had too much risk (Kawaguchi et al., 2017). Yin (2019) stated that such a crash did not result from overbuilding but rather due to excessive external debt to fund aggressive development activities (Yin, 2019).

Phase four: Mature (2008-present) Following the subprime mortgage crisis, REITs had the chance to acquire many undervalued properties to grow their portfolio. This was further supported by the easy monetary policies adopted by policymakers in response to the crisis. Significant capital inflow continued to flow into REITs as an asset class in the years that followed. According to Levy, Giano, and Jones (2015), there were two other reasons which contributed to the growth of the REIT market - the idea of renting and specialization (Levy et al., 2015).

An idea was put forward in the market that some industries can achieve higher efficiency by renting rather than owning the properties as part of their value chain. As a result, investors with capital can acquire properties and rent them to businesses who would prefer to rent it.

Next, there was to an extent wide belief that firms should specialize and have separate asset ownership. Thus, firms who previously owned real estate apart from their main operating business would have a tendency to dispose of or spin-off their real estate and turn to the market to fulfil this need. This contributed to the growth of REITs both from a supply and demand perspective.

2.3. The Exploration of REITs in China

The early development of REITs in China can be traced back to 2005 when the first REIT with underlying mainland Chinese properties called Yuexiu REIT was listed in the Hong Kong stock exchange. Subsequently, a few more REITs with underlying mainland China properties have been listed in markets overseas. Until May 2021, there are a total of 9 REITs with purely mainland China properties listed in the stock exchanges of Hong Kong and Singapore. Detailed information is specified in Section 3.2.2.

In the context of the REIT market in mainland China, the concept of REIT was first proposed by the State Council of China in 2008, which suggested that REITs could be an innovative financing mechanism to promote the healthy development of the real estate industry. During the time between 2008 to 2018, there was barely any substantial progress. However, over this period, China has launched various financial products that have relationships with real estates, such as commercial mortgage-backed securities, which is a type of fixed-income security collateralized by commercial real estate loans. The two most significant features of these products are privately traded and are a type of fixed income asset. Until May 2021, financial products underlying real estate have a total market value of approximately USD 62.8 billion (Deloitte China, 2021).

During the past three years, REITs have drawn much attention. According to Deloitte China, there are subsequently favorable policies for domestic REITs. These policies have provided strong political support from two aspects: by defining which type of REITs should be in the first trial and by setting up the specific IPO steps and regulations of REITs. In April 2018, the China Securities Regulatory Commission, together with the Ministry of Housing and Urban-Rural Development, advocated the secularization of the rental housing market and pushed the establishment of REITs. In April 2020, the China Securities Regulatory Commission confirmed infrastructure associated with technology as a pilot program of public REITs. This indicated the official launch of such a pilot program. Afterwards, there were various policy modifications as preparation for the development of the domestic REIT market (Deloitte China, 2021). In May 2021, the first batch of 9 publicly-traded REITs gained approvals. These 9 REITs were infrastructure REITs with underlying assets in highways, industrial parks, logistic warehouses and sewage treatment facilities, with a total capitalization of approximately USD 4.7 billion. The issuance of the pilot program of infrastructure REITs indicated that China is moving forward to develop REITs as an asset class (The Balance, 2021).

There are more listings of infrastructure REITs in the months that follow. The infrastructure REITs pilot targets five sectors. They are warehousing & logistics, transportation infrastructure, environmental protection, urban utilities and new infrastructure. Standard and Poor's (S&P) (2021) concluded the incentive of China to push the establishment of infrastructure REITs as "big infrastructure goals, limited financing options". The report suggested that REITs are part of an infrastructure financing framework with the aim to re-

lieve the debt burden for infrastructure developers in China, since REITs generally work with the context of equity instead of debt. This can be crucial to help send the message to developers to focus on profitability so as to align and be consistent with the demands of the REIT market. The issuance of pilot infrastructure REITs would bring new opportunities. Such REITs provide investors with great investment opportunities in the context of the fast development of the Chinese economy. Additionally, REITs can be traded easily in the secondary market. As a result, this increases the liquidity as compared to traditional real estate. It is a great way to improve the asset turnover rate, remove debt burden and speed up the reinvestment for infrastructure developers through selling off operating assets and accompanying project debt, while optimizing the resources and boosting economic growth (Standard and Poor's (S&P), 2021).

However, there are also some uncertainties. Deloitte China (2021) suggested that the REITs with infrastructure as the underlying assets should have the advantages of higher transparency and better liquidity. Although the first batch of REITs had such advantages, most infrastructure projects with financing needs in China may lack such characteristics. Some quality projects might not choose REIT as a financing channel. Therefore, it is questionable for the feasibility of financing through REIT in the long run (Deloitte China, 2021). Additionally, China REIT adopted a transaction structure of public fund plus Asset Backed Securities (ABS), which are relatively complex, as shown in Figure 3. There is an ABS in between which acts as a conduit for passing through the cash flows from the underlying assets to the public fund. There are several concerns regarding this framework. One of them is that such a structure may lead to duplicate or excessive taxation. The other concern is that under this multi-layer principal-agent relationship structure, the problem of information asymmetry and conflict of interest may be more serious. This may result in cash leakage and returns being diluted (Standard and Poor's (S&P), 2021).

The guideline of the pilot program is rigorous, with higher payouts and tighter caps on leverage than global norms. Such strict regulations indicate that China has a solid commitment to the development of the domestic REIT market. The pilot REITs program is an exploration that helps to test the market appetite and learn from challenges. It would provide great experiences and lessons for China to further scale up the REIT market in the future. According to the prediction of Standard and Poor's (S&P) (2021), infrastructure REITs in China could grow to a market size between USD 300 billion and 735 billion within a decade (Standard and Poor's (S&P), 2021).

3. Explore How Macroeconomic Factors Affect REITs

This chapter introduces the choice of data, with an emphasis on independent variables and dependent variables. The stationarity of the data will be tested before the construction of the VAR model. Afterwards, an additional three analysis will be applied to perform validity checks.

3.1. Choice of Macroeconomic Factors

According to the literature review discussed in Section 1.2, frequently considered macroeconomic variables are inflation risk, interest rate risk, money supply risk, industrial production risk and stock market risk. On this basis, this thesis applies these five factors as independent variables. Table 1 shows the macroeconomic variables and their selected proxies.

$$\gamma_t = \ln p_t - \ln p_{t-1} \tag{1}$$

Data for these five variables are obtained from Refinitiv with a time frame from February 2006 to March 2021. Consideration has been given to indicate the real economy and business cycles (182 months in total). It should also be noted that the data frequency applied is monthly. Among these variables, $GIND$, $GCPI$ and $GM2$ are MoM change data, which have been downloaded from Refinitiv (Refinitiv, 2021). The logarithmic return has been calculated for the stock market with the Formula 1, which is represented by R_S.

3.2. Choice of Target REITs

3.2.1. US REIT Data

In this thesis, the US REIT index created by Morgan Stanley Capital International (MSCI) is chosen to represent the US REIT market.[6] "The MSCI US REIT index is a free float-adjusted market capitalization weighted index. With 136 constituents, it represents about 99% of the US REIT universe and securities that are classified under the equity REIT industry (under the Real Estate Sector) according to the global industry classification standard" (MSCI, 2021a). The MSCI US REIT index is very comprehensive. Around 23.69% of components are specialized REITs, followed by residential REITs, industrial REITs, and retail REITs. The specific percentage of each category is shown in Figure 4 below.

The selected sample span is the same as that of macroeconomic factors. The monthly logarithmic return of MSCI US REIT index can be then calculated by the Formula 1. Symbol R_T represents the REIT variable in this thesis, and such variable of US is represented by US_R_T.

3.2.2. China REIT Data

As the REIT market in China is still relatively new, there is no existing REIT index (or equivalent) for mainland China, and one would have to self construct. Till date, there are 9 REITs (with underlying properties based in mainland China) listed in Hong Kong and Singapore. The detailed information of each of these REITs is shown in Table 2.

Compared with the MSCI US REIT index components, these 9 Chinese REITs are less diversified, with the main focus on commercial properties. Among them, the REITs with

[6] MSCI is an investment research firm that provides stock indexes, portfolio risk and performance analytic, and governance tools to institutional investors and hedge funds.

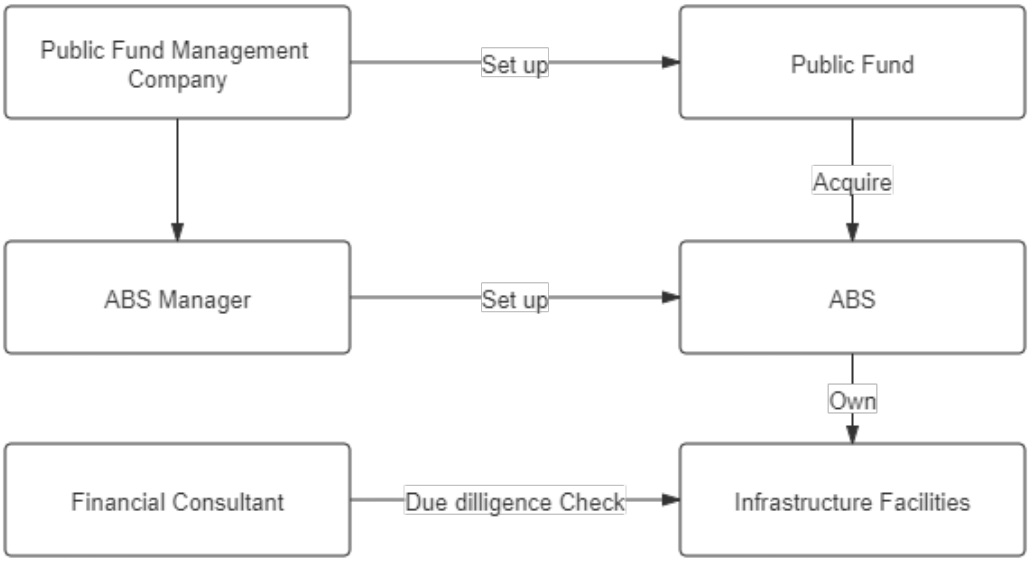


Figure 3: Public Fund Plus ABS Structure of China Infrastructure REITs

Source: Deloitte China (Deloitte China, 2021)

Table 1: Macroeconomic Variables and Proxies

Note: S&P 1500 Composite Index covers approximately 90% of US market capitalization (Standard and Poors (S&P), 2021); SSE Composite Index covers all stocks traded in SSE; MoM

Variables	Macroeconomic Factors	Proxy
R_0	Interest rate risk	10-year government bonds yields as interest free rate
$GIND$	Fluctuation of industrial production	Industrial production MoM change
$GCPI$	Fluctuation of inflation	CPI MoM change
$GM2$	Fluctuation of money supply	M2 money supply MoM change
R_S	Fluctuation of stock market	S&P 1500 Composite Index for US ; SSE Composite Index for China

a larger market capitalization are Yuexiu, Hui Xian and Capitaland. According to the financial report of New Century REIT, the tourism industry has suffered a significant financial loss due to the COVID-19 pandemic. Revenue of hotels has decreased 16% year over year due to travel restrictions and lock-downs. As a result, the New Century REIT was liquidated on the 14th of April 2021 (New Century, 2021). In this thesis, the index based on the 9 REITs mentioned before is used as a reference. The methodology to construct this REIT index is consistent with that of the MSCI US REIT index (Market free float-adjusted market capitalization weighted) (MSCI, 2021b). The formula is shown in 2.

$$p_t \& = \frac{\sum_{t=1}^{n} p_t x_t}{\sum_{t=1}^{n} x_t} \tag{2}$$

There are two key steps to be performed before the final index is calculated and constructed.

First, to maintain consistency with the MSCI US REIT index, dividends of REITs are to be considered and added (MSCI, 2021b). Often, after a stock goes ex-dividend, the share price drops by approximately the amount of the dividend paid. Therefore, the original close price downloaded from Refinitiv should be modified to add the equivalent cash dividend. Furthermore, since REITs have various IPO dates, the price and weights of unlisted REITs at time t would be set to 0. These 9 REITs are listed in Hong Kong and Singapore. As these two markets have different workdays, the workdays of Hong Kong are chosen for REIT index calculation. If the time t in Singapore is a bank holiday, the price and shares of that particular REIT would be adjusted to the previous work-

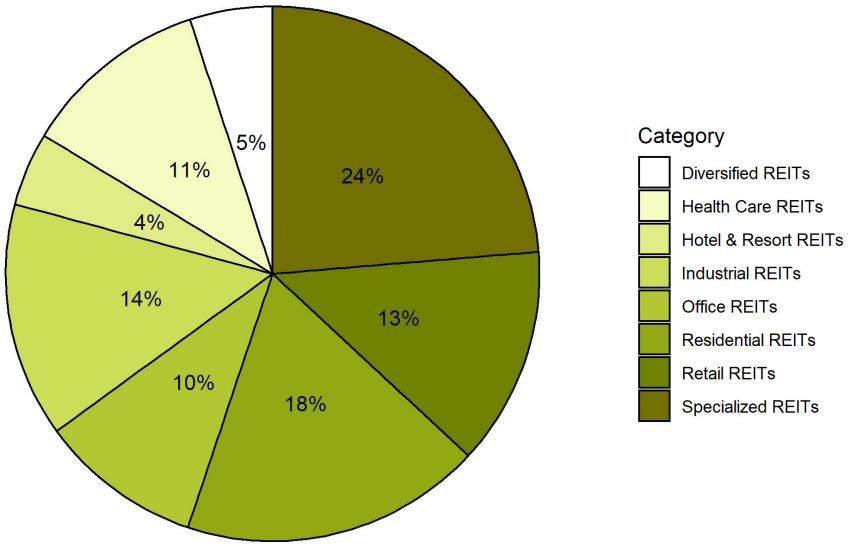


Figure 4: The Pie Chart of REITs Category of MSCI US REIT Index

Source: MSCI (MSCI, 2021a)

Note: Diversified REITs refer to REITs that own more than one category of properties; Specialized REITs refer to REITs that own properties which do not fit within the other category, such as casinos and farmland

ing day's value. Afterwards, the weighted average index can be calculated using Formula 2 and its logarithm return can be obtained by Formula 1.

As this index only has 9 components, the distribution of daily data should be checked to ensure that there is no outlier. The selected sample span is for a period from 1st February 2006 to 31st March 2021. There is a total of 3725 days of validated data and a box plot is constructed. As shown in Figure 5, there is a rate of return slightly above 30%. This data point is on the 08th December 2006, which is the IPO date of Capitaland China REIT. Before 08th December 2006, the index only consists of Yuexiu REIT. Thus, this resulted in an imbalanced calculation for the new addition to the index. To minimize this impact, this outlier is ignored.

To be consistent with the monthly macroeconomic factor data, a total of 182 months of data is used in this chapter. The REIT variable of China is represented by CN_R_T.

3.3. The Effect of Macroeconomic Factors on REITs

3.3.1. Descriptive Analysis

As mentioned in section 3.1, there are five independent variables, which are R_0, $GIND$, $GCPI$, $GM2$ and R_S. The dependent variable is R_T. The time series line plot of each variable and brief descriptions are provided.

Two distinctions are illustrated in Figure 6. First, yields on the 10-year bonds in both countries react in a similar fashion during times of crises, as they typically are expected to act as "safe-havens". For example, during the subprime mortgage crisis of 2008, yields on the 10-year bonds in both countries decreased sharply, indicating a "risk-off" event as prices on these bonds increased. From 2015 to 2016, the stock market bubble in China had burst, leading to a similar sharp decrease in yields on the 10-year bonds. In the most recent COVID-19 crisis, a similar sharp decrease in yields for both countries happened as well. Second, yields on the Chinese 10-year government bonds are within a band of approximately 0.04 and 0.019 throughout the period from 2006 until 2021. In comparison, yields on the US 10-year government bonds were in a downward trend for the same time period. A number of factors can help explain this divergence in trends, such as the reserve status of the USD and the growing trade deficit of the US with its trading partners.

Two interesting points can be noted in Figure 7, which illustrates the MoM change of industrial production. Using the same context of the two crises in the US from 2006 to 2021, the data highlights sharp volatility during the 2008 subprime crisis, where the financial system had a shock as well as a severe liquidity event that affected a large number of businesses. However, the impact looks minor compared to the COVID-19 crisis, where many businesses had to stop operations due to the uncertainty of COVID-19. Only in Q3 2020 were strategies employed to keep workers safe and to resume operations. There is, unfortunately, no publicly available information for the same MoM data set for China.

Table 2: Detailed Information for 9 REITs Which Have Purely Mainland China Properties

Note: HK refers to Hong Kong stock exchange and SG refers to Singapore stock exchange; The sector information of each REIT comes from the financial reports; The market capitalization data is downloaded from Refinitiv (Refinitiv, 2021), which is the data on 31.03.2021 shown in USD

Name	Listed	IPO Date	Category	Market Cap
Yuexiu	HK	21.12.2005	Office and Retail	$1,691,093,580
Hui Xian	HK	28.04.2011	Office, Retail,Hotels & Resort	$1,619,991,179
New Century	HK	09.07.2013	Hotels & Resort	$254,186,222
China Merchants	HK	10.12.2019	Office and Retail	$356,430,723
CapitaLand China	SG	08.12.2006	Office and Retail	$1,540,351,500
Dasin Retail	SG	20.01.2017	Retail	$422,312,494
BHG Retail	SG	11.12.2015	Retail	$209,315,134
EC World	SG	27.07.2016	Industrial	$436,201,412
Sasseur	SG	27.03.2018	Retail	$749,005,526

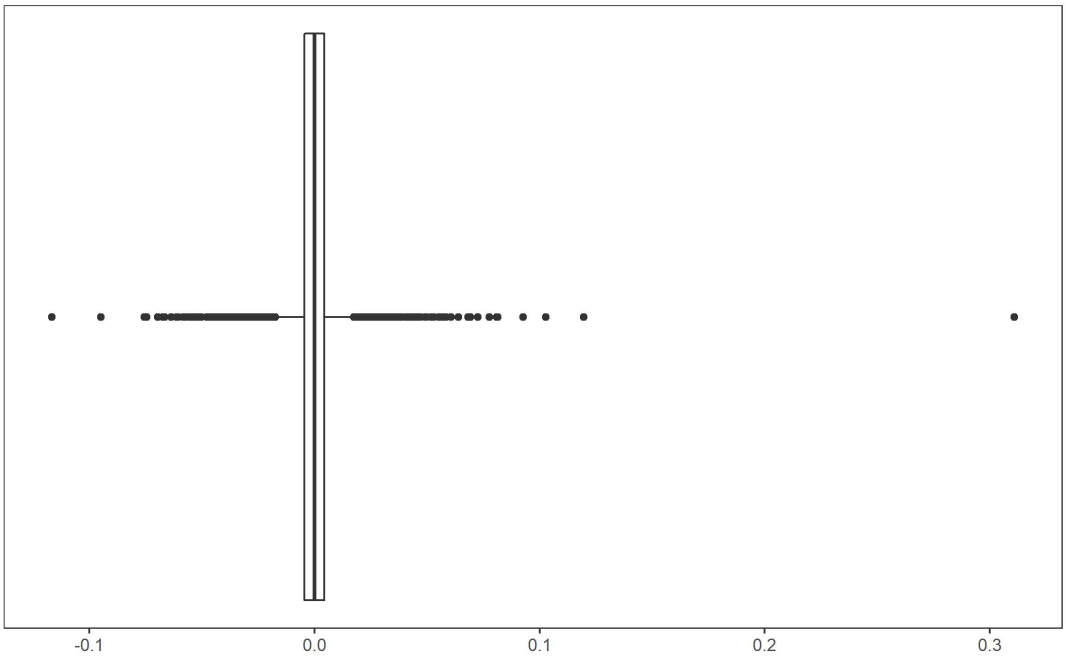


Figure 5: The Box Plot of Daily Logarithm Return of China REIT Index

Source: Author's representation

Thus, this independent variable will be ignored in the VAR model of China in later sections.

MoM change of CPI in China and the US is illustrated in Figure 8. The first takeaway is that CPI change in China is at a much larger magnitude compared to that of the US. Part of the reason is that as China is considered an emerging economy, a much more significant price change can happen due to the low base effect.[7] The second takeaway is that the MoM change in CPI in China did not appear to have any significant fall in the 2008 subprime mortgage crisis. On the other hand, it can be seen that the CPI MoM change in the US had a significant fall during the subprime mortgage crisis. One reason could be that China and its businesses had less exposure to the US economy during that time period, and China were still effectively increasing the number of the middle class, with rising wages which reduces any deflationary pressure and vice versa for the US.

Figure 9 illustrates the $M2$ money supply MoM change in China and the US. In China, the $M2$ money supply had generally been increasing MoM from 2006 until 2017 and had been in a rather consistent manner apart from a sharp spike in 2011 and 2015. From 2017 onwards, $M2$ money supply MoM change in China had higher volatility, which

[7]Low base effect refers to the tendency of a small absolute change from a low initial amount to be translated into a large percentage change

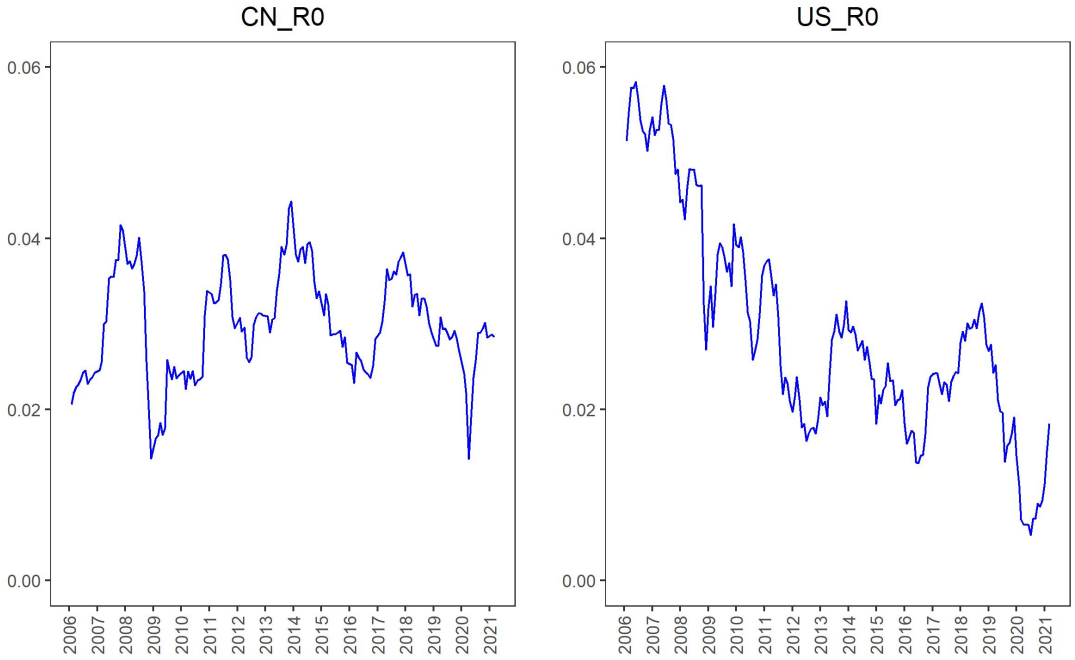


Figure 6: The Time Series Line Plots of CN_R_0 and US_R_0

Source: Refinitiv (Refinitiv, 2021)

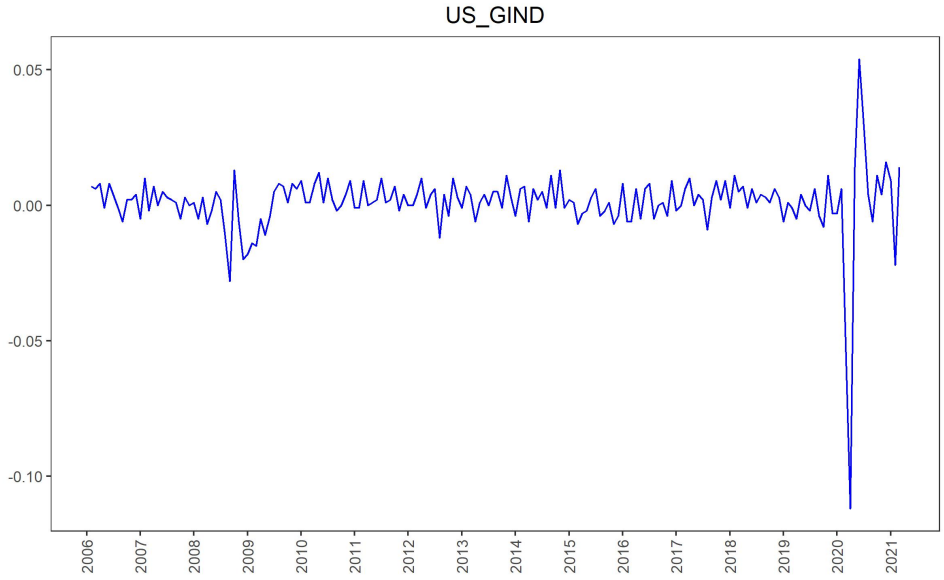


Figure 7: The Time Series Line Plot of US_GIND

Source: Refinitiv (Refinitiv, 2021)

could partly be attributed to a combined factor of a broadening definition of the $M2$ money supply as well as to react to the burst of the 2016 stock market bubble. In comparison, the $M2$ money supply in the US has been growing very consistently, except for the large liquidity injection in 2020 to combat the COVID-19 crisis. It is noteworthy that in the US in 2008, the money supply was mainly increased to recapitalize banks to combat the subprime mortgage crisis. However, in 2020, money was additionally handed out directly to individuals, which drastically increased the $M2$ money supply.

Figure 10 provides context to daily logarithmic returns of the SSE composite index and the S&P 1500 composite

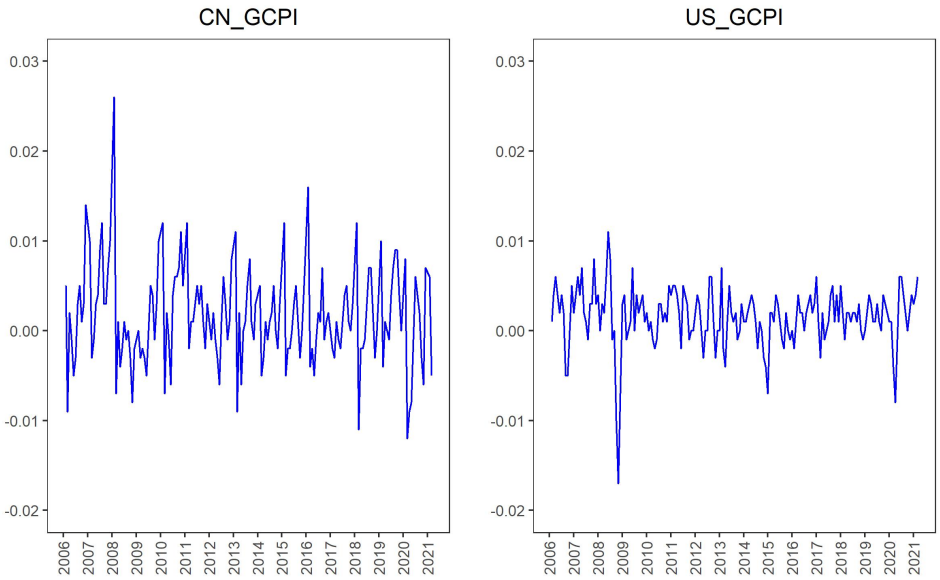


Figure 8: The Time Series Line Plots of CN_GCPI and US_GCPI

Source: Refinitiv (Refinitiv, 2021)

index in China and the US respectively. In China, three critical points of volatility can be identified. First, there was the subprime mortgage crisis in 2008-2009. Second, there was the burst of the stock market bubble in China in 2016. Third, there was the COVID-19 outbreak which affected businesses in China towards the end of 2019. In the US, two key points of volatility can be identified. First, and similarly, there was the subprime mortgage crisis in 2008. In 2020, the US reacted to the COVID-19 crisis where many businesses had to shut down, causing extreme volatility in the stock market. It can be seen that the magnitude of volatility during crises are consistent, which the use of tools such as circuit breakers in the stock market might have played a role. Another exciting finding is that the overall CN_R_S had higher volatility than US_R_S. This phenomenon can be explained by the characteristics of emerging markets, such as lower liquidity, which caused the market to be more volatile.

The daily logarithmic returns of the self-constructed China REIT index and MSCI US REIT index can be seen in Figure 11. Three key points can be taken away from both China and the US. First, R_T had much more volatility in the year 2006 leading to 2008. Volatility was due to the growing interest in real estate, as well as the speculative mindset on real estate which was evident throughout many countries in that period of time. Second, the period after the 2008 subprime mortgage crisis until 2019 has less volatile price action compared to the years before due to higher emphasis by policymakers on price stability, as well as deleveraging in many businesses. Third, like most other asset classes, REITs were also highly affected during the COVID-19 crisis, whereby an amalgamation of factors including low liquidity and uncertainty over the economic outlook affected the prices.

3.3.2. Detecting Stationary

A critical assumption of the primary, stationary finite order VAR model is that the variables should be Gaussian white noise. It should be a stationary and ergodic random process with zero mean (Lütkepohl, 2005). [8] A unit test could be used to test the stationary of time series data. In this thesis, Augmented Dickey Fuller (ADF) test would be applied, with the null hypothesis that a unit root is present or non-stationary.

As shown in Table 3 Panel A, only variable R_0 in both counties are non-stationary. Therefore, the first difference of R_0 is taken, represented by ΔUS_R_0 and ΔCN_R_0. ADF test has been used again for these two modified variables. Table 3 Panel B shows that both modified variables are significant at the 1 percent level, which indicates that they meet the stationary requirement. R_0 would be replaced by ΔR_0 to be used in the following sections.

3.3.3. VAR Model

The Vector Autoregression (VAR) model was first proposed by Christopher Sims in 1980. As Stock and Watson (2001) stated, "A VAR is an n-equation, variable linear model in which each variable is in turn explained by its own lagged values, plus current and past values of the remaining n-1 variables. This simple framework provides a systematic way to capture rich dynamics in multiple time series, and the statistical toolkit that came with VARs was easy to use and interpret" (Stock & Watson, 2001). The definition of VAR with p

[8] See Lütkepohl (2005), pp. 13-18.

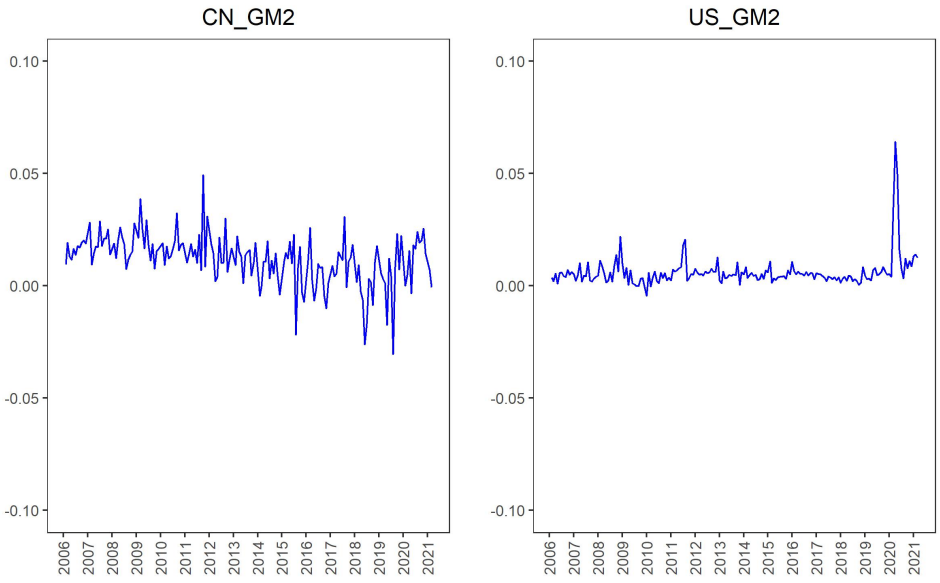


Figure 9: The Time Series Line Plots of CN_GM2 and US_GM2

Source: Refinitiv (Refinitiv, 2021)

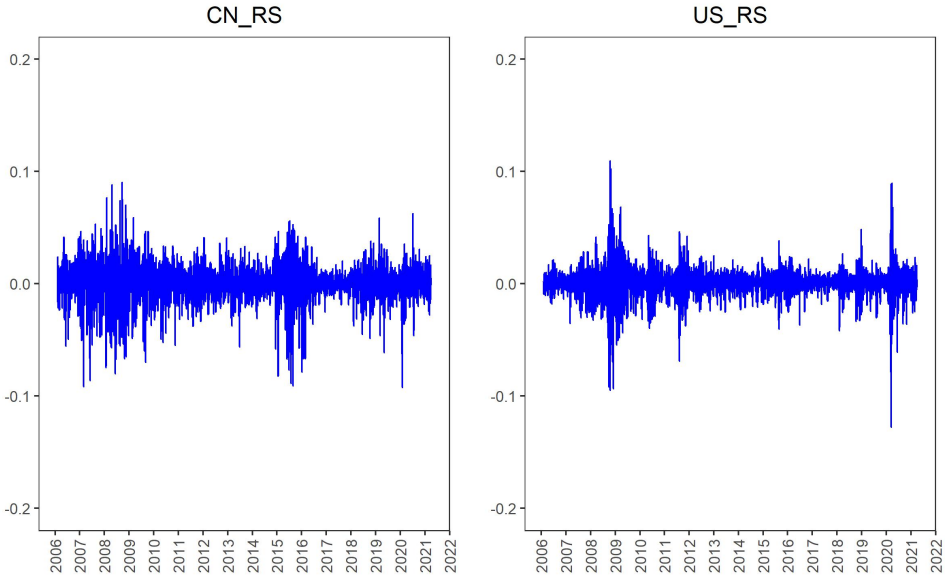


Figure 10: The Time Series Line Plots of Daily CN_R_S and US_R_S

Source: Refinitiv (Refinitiv, 2021)

lags can be represented in Formula 3, where v is a fixed (K x 1) vector of intercept terms, A_i is (K x K) coefficient matrices, y_t is (K x 1) dimensional random vectors, and u_t is K-dimensional white noise (Lütkepohl, 2005).[9]

$$y_t = v + A_1 y_{t-1} + ... + A_p y_{t-p} + u_t,\ t = 1, 2, ..., T \quad (3)$$

Choice of lags is crucial in terms of VAR model construction. A very often used criterion in selecting lags is the Akaike Information Criterion (AIC), which is based on approximate 1-step ahead forecast mean squared error (Lütkepohl, 2005).[10] The optimized lags according to AIC for the US and China are 3 and 2 respectively.

[9] See Lütkepohl (2005), pp. 13-18.; p represents the lags; T represents the total sample size; K represents the total variable amount

[10] See Lütkepohl (2005), p. 146.

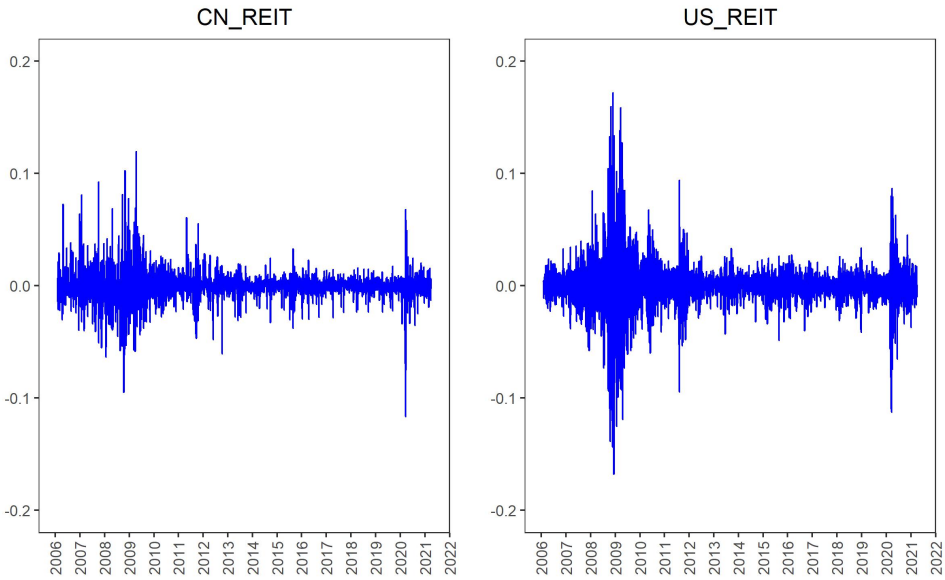


Figure 11: The Time Series Line Plots of Daily CN_R_T and US_R_T

Source: Refinitiv (Refinitiv, 2021)

The results are shown in the following Table 4 and 5. As the model involves too many variables, only the variables with statistically significant relationships would be analyzed in this thesis. The results of both VAR models are summarized as follows:

1) For REIT returns based on US model, US_R_T(-3), US_R_S(-1), US_R_S(-3), ΔUS_R_0(-3) and US_GIND(-2) are significant. Among these, the impacts of US_R_T(-3), US_R_S(-1) and US_R_S(-3) have stronger evidences. However, in the model of China, only CN_R_S(-2) is significant at the 5 percent level, which means only the stock market fluctuation of two months earlier can positively affect REIT. Other factors do not help to explain the changes in REIT return.
2) Looking at the fluctuation of 10-year government bonds yields, the US_R_T(-3), US_GIND(-1), US_GCPI(-1), US_R_S(-1) and US_R_S(-3) are significant. In the context of China, only CN_R_T(-1) and CN_R_T(-2) are significant at the 1 percent level. This indicates that ΔR_0 is significantly affected by factor of the the stock market, REIT market, industrial production and CPI in the US, but only observe the significant effect of REIT market in China.
3) In the analysis of MoM change in industrial production, US_R_T(-1), US_R_T(-3), ΔUS_R_0(-2), US_GIND(-1), US_GIND(-2), US_GIND(-3), US_GM2(-1) and US_GM2(-2) are significant. This indicates industrial production is significant correlated with $M2$ money supply, 10-year government bonds, REIT market and especially limited by its own growth level in the early period.
4) For the MoM change in CPI, US_R_T(-3), US_GM2(-2) and US_GCPI(-2) are significant at 10 percent level, US_R_S(-1) is significant at 5 percent level and only US_GCPI(-1) is significant at 1 percent level. The results of the US are consistent with the results of China, which shows the significance of CN_GM2(-1), CN_GM2(-2), CN_GCPI(-1) and CN_R_S(-2).
5) For the the MoM change in $M2$, US_R_T(-1), US_GIND(-1), US_GM2(-1) and US_GCPI(-3) are significant. However, in the context of China, only US_GM2(-1) and US_GM2(-2) have significant observations.
6) Looking at the fluctuation of S&P 1500 composite index returns, only US_R_T(-3) is significant at the 5 percent level. Similarly, for the fluctuation of SSE composite index returns, only two variables are significant. They are CN_R_S(-1) and CN_R_T(-2), which are statistically significant at the 5 percent level.

The formulas of both models are shown in Formula 5 and 4.

$$\begin{aligned} CN_R_T = {} & 0.051 * CN_R_T(-1) + 0.044 * \\ & CN_R_T(-2) - 0.725 * \Delta CN_R_0(-1) \\ & - 3.240 * \Delta CN_R_0(-2) - 1.222 * \\ & CN_GCPI(-1) + 0.265 * CN_GCPI(-2) \\ & + 0.534 * CN_GM2(-1) - 0.040 * \\ & CN_GM2(-2) + 0.064 * CN_R_S(-1) \\ & + 0.151 * CN_R_S(-2) - 0.006 \end{aligned} \quad (4)$$

Table 3: The Results of ADF Test of Time Series Variables

*Note: ***Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.*

Panel A: ADF Test Results of 11 Initial Variables

Variables	ADF Value	P Value	Result
US_R_0	-2.711	0.279	Non-Stationary
US_GIND	-8.044	0.010***	Stationary
US_GCPI	-7.723	0.010***	Stationary
US_GM2	-6.308	0.010***	Stationary
US_R_S	-7.247	0.010***	Stationary
US_R_T	-6.842	0.010***	Stationary
CN_R_0	-3.403	0.056*	Non-Stationary
CN_GCPI	-8.428	0.010***	Stationary
CN_GM2	-6.394	0.010***	Stationary
CN_R_S	-6.705	0.010***	Stationary
CN_R_T	-7.284	0.010***	Stationary

Panel B: ADF Test Results of ΔUS_R_0 and ΔCN_R_0

Variables	ADF Value	P Value	Result
ΔUS_R_0	-7.124	0.010***	Stationary
ΔCN_R_0	-6.909	0.010***	Stationary

$$
\begin{aligned}
US_R_T = & -0.105 * US_R_T(-1) - 0.141 * US_R_T(-2) \\
& + 0.305 * US_R_T(-3) - 2.346 * \Delta US_R_0(-1) \\
& + 3.395 * \Delta US_R_0(-2) + 4.425 * \Delta US_R_0(-3) \\
& + 0.582 * US_GIND(-1) + 1.023 US_GIND(-2) \\
& - 0.819 * US_GIND(-3) - 1.280 * US_GCPI(-1) \\
& + 1.522 * US_GCPI(-2) + 0.476 * US_GCPI(-3) \\
& + 0.391 * US_GM2(-1) + 1.509 * US_GM2(-2) \\
& - 1.670 * US_GM2(-3) + 0.499 * US_R_S(-1) \\
& + 0.001 * US_R_S(-2) - 0.422 * US_R_S(-3) \\
& - 0.001
\end{aligned}
\tag{5}
$$

3.3.4. Granger Causality Test

The Granger causality test, first proposed by Clive Granger in 1969, is a hypothesis test for determining whether one time series is useful in forecasting another. According to Lütkepohl (2005), the Granger Causality test is an essential part of structured analysis with the VAR model. "The idea is that a cause cannot come after the effect. Thus, if a variable x affects a variable z, the former should help to improve the predictions of the latter variable" (Lütkepohl, 2005). [11] The Null hypothesis of Granger causality test is that X does not Granger Cause (GC) Y.

The results of the Granger causality test can be seen in Table 6 Panel A and B.

1) For the results of the US model, ΔUS_R_0 and US_R_S are the significant Granger causality of US_R_T at 10 percent level. Meanwhile, ΔUS_R_0 and US_R_T are Granger causality of each other, same as US_R_T and US_R_S.
2) Similarly, the results of the China model shows the CN_R_T and CN_R_S are Granger causality of each other at 10 percent level. Other than that, there is no other significant Granger causality for CN_R_T.

3.3.5. Impulse Response Analysis

Impulse response analysis is an important tool to interpret the results of the VAR model. This analysis helps to quantify the response of one variable to the dynamic impulse in another variable. This tool investigates the inter-ration and causality of the model, and also provides information about the positive and negative directions of the response to exogenous shocks (Lütkepohl, 2005). [12]

Figure 12 and 13 show the responses of R_T to six variables based on US and China VAR models. The horizontal axis represents the number of retrospective periods in the impulse response, which is defined as 10 months. The vertical axis represents the response size of the dependent variable to the independent variable. It is noteworthy that the solid blue line refers to the actual value, and the red dot line represents the value at the 95% confidence interval. There are six findings as follows.

1) R_T is greatly influenced by itself. The evidence of the US model shows that when a positive impact is given

[11] See Lütkepohl (2005), pp. 41-51.

[12] See Lütkepohl (2005), pp. 51-63.

Table 4: The Results of VAR model Based on China Data

*Note: ***Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level. [] is t-value.*

	CN_R_T	ΔCN_R_0	CN_GCPI	CN_GM2	CN_R_S
CN_R_T(-1)	0.051	0.006***	0.003	-0.006	-0.142
	[0.593]	[2.698]	[0.521]	[-0.427]	[-1.461]
CN_R_T(-2)	0.044	0.007***	0.005	-0.013	0.230**
	[0.497]	[2.703]	[0.813]	[-0.928]	[2.308]
ΔCN_R_0(-1)	-0.725	0.094	0.127	0.425	-0.502
	[-0.279]	[1.256]	[0.668]	[1.034]	[-0.171]
ΔCN_R_0(-2)	-3.240	0.049	0.210	0.187	-1.342
	[-1.293]	[0.686]	[1.140]	[0.472]	[-0.475]
CN_GCPI(-1)	-1.222	-0.019	0.277***	-0.025	-0.471
	[-1.216]	[-0.639]	[3.759]	[-0.157]	[-0.416]
CN_GCPI(-2)	0.265	-0.023	-0.098	0.082	0.071
	[0.272]	[-0.817]	[-1.367]	[0.533]	[0.064]
CN_GM2(-1)	0.534	-0.008	0.113***	0.244***	0.740
	[1.086]	[-0.612]	[3.140]	[3.130]	[1.333]
CN_GM2(-2)	-0.040	0.007	-0.113***	0.207***	-0.614
	[-0.080]	[0.496]	[-3.078]	[2.624]	[-1.090]
CN_R_S(-1)	0.064	0.002	-0.008	-0.004	0.174**
	[0.856]	[1.129]	[-1.367]	[-0.357]	[2.064]
CN_R_S(-2)	0.151**	0.001	0.013**	0.001	0.045
	[2.009]	[0.356]	[2.365]	[0.078]	[0.527]
Constant	-0.006	1.450e-04	0.002***	0.007***	0.003
	[-0.672]	[0.546]	[2.582]	[4.498]	[0.321]

to US_R_T, the corresponding US_R_T will have at least eight phases of impact, of which the first phase is positive and the impact is more significant. The subsequent phase is negative and gradually followed by another two phases of positive impact. Afterwards, there are some minor fluctuations around 0. Similar evidence is shown in the China model, but with a short period of impact, which has only four phases of impact and afterwards, the influence gradually weakened to 0.

2) R_S also shows a great impact on R_T in both models. When R_S is given a positive shock, R_T would initially have a positive response, indicating that a booming stock market would benefit the REIT market. The US_R_S has a stronger impact than CN_R_S with the evidence that US_R_S has at least eight phases of impact, whereas CN_R_S shows three stages of effects.
3) The impact of CN_GM2 is relatively significant. When CN_GM2 is given a positive shock, CN_R_T would have a positive response, and such impact would last for four phases with a decreasing trend. However, when US_GM2 is given a positive shock, US_R_T would have an initial small negative response, followed by two stages of positive response and one phase of negative response.
4) US_GIND has lagged and lasting impact on US_R_T. The initial negative response is relatively small but becomes most vital in the fourth phase. Afterwards, the fluctuation hovers around 0 and ends until the eighth phase.
5) ΔR_0 also has lagged impact in both models. It can be seen that when a positive impact is given to ΔUS_R_0, it would cause an initial small negative impact response and gradually reach maximum positive repose at the fourth phase. The impact of ΔCN_R_0 is also very small initially and reaches a maximum negative response at the third phase. Afterwards, it comes to 0.
6) The impact of variable $GCPI$ is the weakest. In both models, the magnitude of the response is no more than 1%, which indicates that the impact on the REIT index is the smallest.

To sum up, both models produces consistent results. The R_T and R_S have the strongest impact. CN_GM2, US_GIND and US_R_0 are relatively significant. The impact of remaining variables can be ignored.

3.3.6. Variance Decomposition Analysis

The variance decomposition is an additional tool used for VAR model interpretation to simplify structures. In general, it indicates the amount of forecast error variance of the dependent variable that can be explained by shocks of other variables (Lütkepohl, 2005). [13]

[13] See Lütkepohl (2005), pp. 63-66.

Table 5: The Results of VAR Model Based on US Data

*Note: ***Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level. [] is t-value.*

	US_R_T	$\Delta\ US_R_0$	US_GIND	US_GCPI	US_GM2	US_R_S
US_R_T(-1)	-0.105	0.003	0.048**	0.002	-0.015*	0.033
	[-0.859]	[0.745]	[2.598]	[0.524]	[-1.747]	[0.397]
US_R_T(-2)	-0.141	-0.004	-0.011	0.003	-0.005	-0.012
	[-1.143]	[-0.951]	[-0.615]	[0.629]	[-0.549]	[-0.143]
US_R_T(-3)	0.305**	-0.012***	0.030*	-0.008*	0.016	0.018**
	[2.579]	[-2.966]	[0.092]	[-1.665]	[1.150]	[2.266]
ΔUS_R_0(-1)	-2.346	-0.072	0.571	0.034	-0.288	-0.432
	[-0.932]	[-0.809]	[1.497]	[0.328]	[-1.601]	[-0.250]
ΔUS_R_0(-2)	3.395	0.003	0.702*	0.092	-0.103	2.698
	[1.339]	[0.034]	[1.828]	[0.887]	[-0.568]	[1.553]
ΔUS_R_0(-3)	4.425*	0.005	0.593	-0.023	-0.195	1.596
	[1.777]	[0.054]	[1.573]	[-0.222]	[-1.097]	[0.936]
US_GIND(-1)	0.582	-0.046**	0.206**	-0.004	-0.172***	-0.076
	[0.943]	[-2.105]	[2.205]	[-0.149]	[-3.901]	[0.181]
US_GIND(-2)	1.023*	0.014	-0.260***	0.016	0.063	0.509
	[1.743]	[0.692]	[-2.938]	[0.676]	[1.498]	[1.266]
US_GIND(-3)	-0.819	0.029	0.170**	0.014	0.097	-0.023
	[-1.518]	[1.512]	[2.088]	[0.681]	[1.123]	[-0.627]
US_GCPI(-1)	-1.280	0.179**	-0.054	0.479***	-0.034	0.258
	[-0.638]	[2.523]	[-0.178]	[5.867]	[-0.234]	[0.187]
US_GCPI(-2)	1.522	-0.115	-0.353	-0.149*	[-0.181]	-0.079
	[0.699]	[-1.495]	[-1.071]	[-1.677]	[-1.159]	[-0.053]
US_GCPI(-3)	0.476	-0.026	0.276	-0.095	0.437***	-0.489
	[0.246]	[-0.376]	[0.942]	[-1.211]	[3.154]	[-0.369]
US_GM2(-1)	0.391	-0.061	-0.390**	-0.031	0.544***	0.725
	[0.300]	[-1.315]	[-1.977]	[-0.601]	[5.828]	[0.812]
US_GM2(-2)	1.509	0.071	0.919***	0.096*	-0.065	0.977
	[1.071]	[1.415]	[4.312]	[1.667]	[-0.645]	[1.013]
US_GM2(-3)	-1.670	-0.013	[-0.046]	-0.017	0.097	-0.324
	[-1.386]	[-0.316]	[-0.251]	[-0.346]	[1.123]	[-0.392]
US_R_S(-1)	0.499***	0.012*	0.020	0.018**	-0.009	0.100
	[2.611]	[1.772]	[0.687]	[2.377]	[-0.674]	[0.764]
US_R_S(-2)	0.001	0.010	0.022	-0.002	0.006	-0.106
	[0.004]	[1.418]	[0.717]	[0.629]	[0.441]	[-0.779]
US_R_S(-3)	-0.422**	0.015**	0.004	-0.008	0.016	-0.171
	[-2.127]	[2.236]	[0.120]	[-1.665]	[1.150]	[-1.257]
Constant	-0.001	-4.843e-05	-0.002	0.001*	0.002***	8.588e-05
	[-0.069]	[-1.393]	[-1.194]	[1.791]	[3.104]	[0.013]

The results of variance decomposition can be seen in Table 7 and 8.

1) In both models, R_T is affected most by itself but shows a decreasing trend. The CN_R_T has higher explanatory power than US_R_T for its own forecast error variance. In the last period, the value is standing at around 38.4% and 71.3% respectively.
2) R_S also has very strong explanatory power that is consistent with the results of impulse response analysis in the previous section. However, US_R_S explains around half the variance of the US model, which is higher than the portion explained by CN_R_S.
3) As R_T together with R_S can explain the variance at least 85.8% and 91.7% of US_R_T and CN_R_T respectively. The remaining macroeconomic variables can be ignored.

3.4. Chapter Summary

This chapter first introduces the choice of data and then uses the VAR methodology to conduct an econometric analysis of the macroeconomic risk of REIT. On the basis of the stationary test, the data was processed, and then the VAR model

Table 6: The Results of Granger Causality Test

*Note: ***Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.*

Panel A: The Results of Granger Causality Test Based on US Data			
Null Hypothesis	Sample Size	F Value	P Value
US_R_T does not GC ΔUS_R_0	174	8.095	4.501e-05***
ΔUS_R_0 does not GC US_R_T	174	2.195	0.091*
US_R_T does not GC US_GIND	175	11.428	7.169e-07***
US_GIND does not GC US_R_T	175	1.628	0.185
US_R_T does not GC US_GCPI	175	6.500	3.426e-04***
US_GCPI does not GC US_R_T	175	0.625	0.600
US_R_T does not GC US_GM2	175	6.048	6.136e-04***
US_GM2 does not GC US_R_T	175	0.610	0.610
US_R_T does not GC US_R_S	175	2.211	0.089*
US_R_S does not GC US_R_T	175	2.201	0.090*
Panel B: The Results of Granger Causality Test Based on China Data			
Null Hypothesis	Sample Size	F Value	P Value
CN_R_T does not GC ΔCN_R_0	176	12.889	6.028e-06***
ΔCN_R_0 does not GC CN_R_T	176	0.835	0.436
CN_R_T does not GC CN_GCPI	177	2.000	0.138
CN_GCPI does not GC CN_R_T	177	0.928	0.397
CN_R_T does not GC CN_GM2	177	0.430	0.651
CN_GM2 does not GC CN_R_T	177	0.472	0.624
CN_R_T does not GC CN_R_S	177	2.780	0.064*
CN_R_S does not GC CN_R_T	177	2.556	0.081*

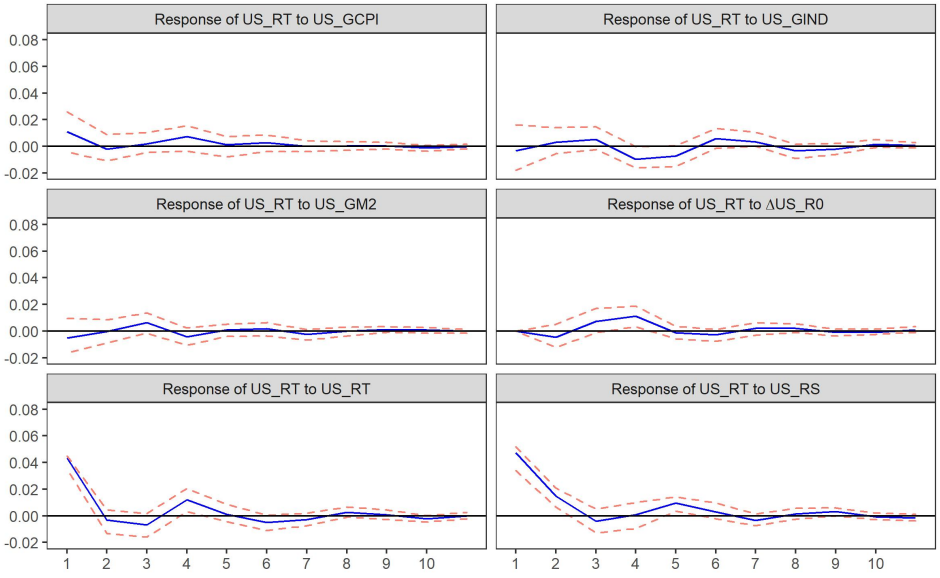


Figure 12: The Impulse Response Plots of US_R_T as Response

was constructed. To perform a structured VAR analysis, the Granger causality test was further carried out together with impulse response analysis and variance decomposition analysis.

The results of VAR based on US and China data lead to a consistent conclusion. In both models, the R_T and R_S are the most significant factors. This result is the same as the finding of Fang et al. (2016), who found the stock market has a significantly positive relationship with the REIT index based on evidence from Japan, Singapore, and China (Fang

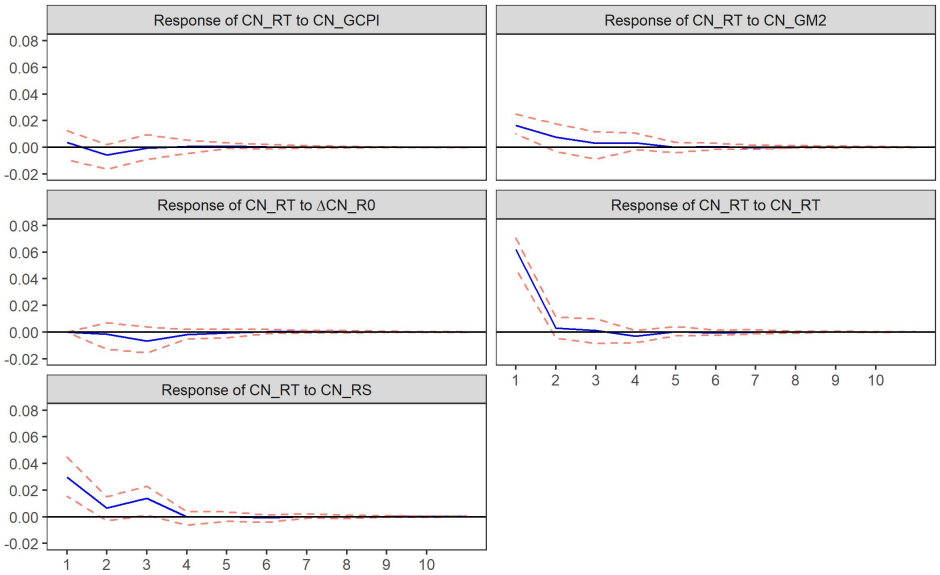


Figure 13: The Impulse Response Plots of CN_R_T as Response

Table 7: The Results of Variance Decomposition Analysis of US_R_T

Period	US_R_T	ΔUS_R_0	US_GIND	US_GCPI	US_GM2	US_R_S
1	0.436855	0.000000	0.002539	0.028611	0.006279	0.525713
2	0.413491	0.004638	0.004268	0.027903	0.005925	0.543773
3	0.406787	0.015966	0.009754	0.027541	0.014254	0.525696
4	0.399909	0.039144	0.027473	0.035942	0.016662	0.480866
5	0.389063	0.038329	0.036564	0.035278	0.016360	0.484404
6	0.387470	0.039154	0.042113	0.036037	0.016659	0.478564
7	0.385865	0.039743	0.044042	0.035752	0.017614	0.476980
8	0.385337	0.040432	0.046025	0.035593	0.017533	0.475078
9	0.384242	0.040441	0.046752	0.035502	0.017725	0.475335
10	0.384408	0.040492	0.047067	0.035698	0.017795	0.474538

Table 8: The Results of Variance Decomposition Analysis of CN_R_T

Period	CN_R_T	ΔCN_R_0	CN_GCPI	CN_GM2	CN_R_S
1	0.768411	0.000000	0.002509	0.053896	0.175183
2	0.748497	0.000422	0.008899	0.063823	0.178357
3	0.714481	0.008772	0.008549	0.062661	0.205536
4	0.712900	0.009304	0.008567	0.064630	0.204596
5	0.712670	0.009420	0.008756	0.064609	0.204542
6	0.712517	0.009435	0.008790	0.064651	0.204605
7	0.712509	0.009439	0.008792	0.064653	0.204604
8	0.712504	0.009444	0.008792	0.064654	0.204603
9	0.712503	0.009445	0.008793	0.064654	0.204603
10	0.712503	0.009445	0.008793	0.064654	0.204603

et al., 2016). This result can be explained by the research of Hoesli and Reka (2015). They found there is a demonstrated risk of contagion between REIT and stock in the US and their further study pointed out that such phenomenon is driven by behavioral and liquidity mechanisms, such as investor sentiment (Hoesli & Reka, 2015).

The second interesting finding is regarding the different explanatory power of R_S in the variance decomposition anal-

ysis. In the US model, US_R_S consists of 47.4%. In comparison, CN_R_S only consists of 20.4% in the China model. This is probably due to the different structure of the emerging market and developed market. Another possible explanation would be the proxies of REIT market. The underlying index of US_R_T is the MSCI US REIT index, and on the other hand, the CN_R_T is based on a self-constructed index. The MSCI US REIT has 136 constituents, whereas the latter only has 9. Therefore, it can be suggested that the self-constructed index might not be fully representative of the entire China REIT market.

Combining with the results of all analyses, the effects of other macroeconomic risk factors are relatively minor. Therefore, it is noteworthy to conclude that REITs are a good hedge against some macroeconomic variables, including the fluctuation of 10-year government bonds yields, industrial production, CPI, and $M2$ money supply.

4. Evaluate the Downside Risk of REITs

This chapter first introduces the concept of VaR and its combination with the GARCH model. Afterwards, the choice of data would be specified, and the characteristics of the data set would be mentioned. The ARCH effect is tested before the creation of the GARCH(1,1)-VaR model. The VaR at 95% and 99% would be evaluated, and the Kupiec test would be applied to test credibility.

4.1. Model Overview

Value at Risk (VaR) is defined as the maximum loss over a targeted horizon for a given level of confidence. It is used to estimate the tails of the empirical distribution of financial losses. The formula of VaR is shown in 6. Examining the variables, $Prob$ represents the probability of actual loss in a given period is greater than VaR at α significance level. The final VaR can be calculated using the conditional volatility of returns (σ) multiplied by the percentile of a given probability distribution at 1-α level (a) (Cerović Smolović et al., 2017).

$$\begin{aligned} Prob(\Delta P > VaR) &= \alpha \\ VaR &= \sigma a \end{aligned} \tag{6}$$

The aim here is to estimate VaR using an econometric approach, which is combined with the Generalised Autoregressive Conditional Heteroskedasticity (GARCH) model. This is because there are particular assumptions under the VaR model, such as normal distribution, which the financial time series data cannot fulfill in normal circumstances. Miletic and Miletic (2015) stated that "The GARCH-type model successfully captures several characteristics of the financial time series, such as thick-tailed returns and volatility clustering" (Miletic & Miletic, 2015). "In the classical GARCH models, the conditional variance is expressed as a linear function of the squared past values of the series." Tim Bollerslev proposed GARCH(p,q) in 1986, and the model can be specified as Formula 7 (Francq & Zakoian, 2019). [14] Only lower-order GARCH models are used in the most applications. Therefore, the classical GARCH(1,1) would be applied in this thesis — the formula is shown in 8 (Francq & Zakoian, 2019). [15] After getting the σ of the different periods based on the GARCH(1,1) model, the corresponding VaR can then be calculated using Formula 6.

$$\begin{aligned} \epsilon_t &= \sigma_t \eta_t,\ (\eta_t)\ i.i.d\ whitenoise(0,1) \\ {\sigma_t}^2 &= w + \sum_{i=1}^{q} \alpha_i {\epsilon^2}_{t-i} + \sum_{j=1}^{p} \beta_j {\sigma^2}_{t-j}, t \in Z \\ w > 0, \alpha_i >= 0, \beta_j >= 0, i = 1, ..., q, j = 1, ..., p \end{aligned} \tag{7}$$

$${\sigma_t}^2 = w + \alpha {\epsilon^2}_{t-1} + \beta {\sigma^2}_{t-1} \tag{8}$$

4.2. Descriptive Analysis

The data referenced in this chapter is the daily R_T, and specific methodologies of data preparation have been described in Section 3.2. The selected time span of data is from 1st February 2006 to 31st March 2021. The total sample size of the US and China are 3819 and 3725 respectively. To maintain data integrity and continuity of the model, the first thirteen years' data (from 1st February 2006 to 31st December 2018) are used to train the model. The remaining data would be used for model validation.

The main characteristics of US_R_T and CN_R_T have been measured. Meanwhile, the Jarque-Bera test has been applied for testing if sample data have the skewness and kurtosis matching normal distribution. The null hypothesis is that the data has a normal distribution. The histogram is also constructed to provide further evidence.

Table 9 indicates that the daily logarithm returns of both REIT markets are not normally distributed due to the fact that the skewness deviates from 0, and kurtosis is much higher than 3. These results are also supported by the Jarque-Bera tests, which are significant at the 1 percent level. The results confirm the presence of fat tails and are leptokurtic, which indicates that the R_T is not normally distributed and cannot apply methodologies with normal distribution assumption. These findings are consistent with histogram plots in Figure 14.

4.3. Test of ARCH Effect

The Autoregressive Conditional Heteroskedasticity (ARCH) effect refers to the relationship within the heteroskedasticity, often termed serial correlation of the heteroskedasticity. There are two available tests which are the Ljung–Box test and the the Lagrange multiplier test based on squared residuals of the mean. The null hypothesis of the Ljung–Box test is that there are no autocorrelations. In 1982, Engel suggested

[14] See Francq and Zakoian (2019) section 2.1.
[15] See Francq and Zakoian (2019) section 2.2.1.

Table 9: Descriptive Analysis of Daily US_R_T and CN_R_T

*Note: ***Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.*

Variable	Size	Max	Min	Mean	Jarque-Bera	Std. Dev.	Skewness	kurtosis
US_R_T	3252	0.172	-0.220	4.969e-05	35681***	0.021	-0.192	16.211
CN_R_T	3169	0.120	-0.094	-5.241e-05	16243***	0.013	0.432	11.048

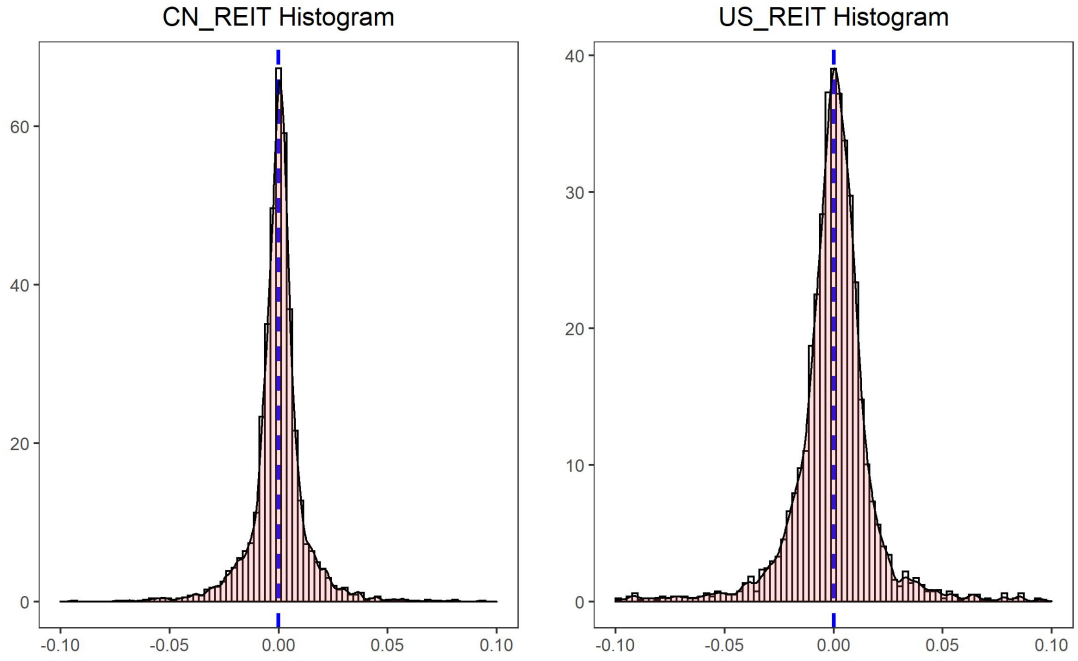


Figure 14: The Histogram Plots of US_R_T and CN_R_T

using Lagrange multiplier test to detect the ARCH effect and its null hypothesis is that there is no existing ARCH effect up to lag order q in the residuals (Tsay, 2005).[16]

The results of Table 10 indicate that the null hypothesis of both tests should be rejected. This suggests the evidence of autocorrelation and time varying volatility. This means that there are some patterns in the dataset that can be determined by other factors. In other words, the GARCH model would be appropriate to evaluate the σ and then VaR of US_R_T and CN_R_T.

4.4. GARCH(1,1)-VaR Model

As discussed earlier, the GARCH(1,1) model would be used. Evidence from previous research suggests that GARCH models based on the standardized residuals with t distribution can better capture the characteristics of financial data (Miletic & Miletic, 2015; Cerović Smolović et al., 2017).

Table 11 indicates the parameter estimates of the GARCH-(1,1) model. The AIC indicates that both models are good fits for the data. Additionally, The residuals obtained from the GARCH model show no evidence for the ARCH effect by Ljung–Box and Lagrange multiplier tests. Therefore, the models for US_R_T and CN_R_T do have appropriate statistical characteristics. The formula of GARCH(1,1) models of US_R_T and CN_R_T is shown in 9. Applying the predicted volatility and percentile of t distribution into formula 6, the VaR at a certain confidence level can be obtained.

$$\begin{aligned} US_R_T &: {\sigma_t}^2 = 0.100e-05 + 0.101 * {\epsilon^2}_{t-1} + 0.896 * {\sigma^2}_{t-1} \\ CN_R_T &: {\sigma_t}^2 = 0.100e-05 + 0.110 * {\epsilon^2}_{t-1} + 0.888 * {\sigma^2}_{t-1} \end{aligned} \quad (9)$$

Figure 16 and 15 shows the comparison between actual daily returns of REIT index and VaR. The blue line represents the actual R_T. The red and green lines are the expected maximum loss at 95% and 99% confidence levels respectively. Instinctively, in both models, there are only a few extreme cases in which the single-day loss of the REIT index exceeds the VaR value under the 95% confidence level. There is almost no single-day loss that exceeds the 99% VaR value. Therefore, it is reasonable to draw a preliminary conclusion that

[16] See Tsay (2005) section 3.3.1.

Table 10: The Results of Ljung–Box and Lagrange Multiplier Tests for US_R_T and CN_R_T

*Note: ***Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.*

Variable	LB[1]	LB[10]	LB[30]	LM[1]	LM[10]	LM[30]
US_R_T	131.590***	160.910***	314.530***	406.910***	1085.900***	1272.200***
CN_R_T	22.933***	73.508***	109.240***	98.637***	398.900***	568.810***

Table 11: Parameter Estimates of the GARCH(1,1) Model With t Distribution of the Standardized Residuals for US_R_T and CN_R_T

*Note: ***Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.*

Model	w	α	β	AIC	LB[1]	LM[7]
GARCH(1,1)-US_R_T	0.100e-05	0.101	0.896	-5.914	1.170	3.330
GARCH(1,1)-CN_R_T	0.100e-05	0.110	0.888	-6.642	1.960	3.552

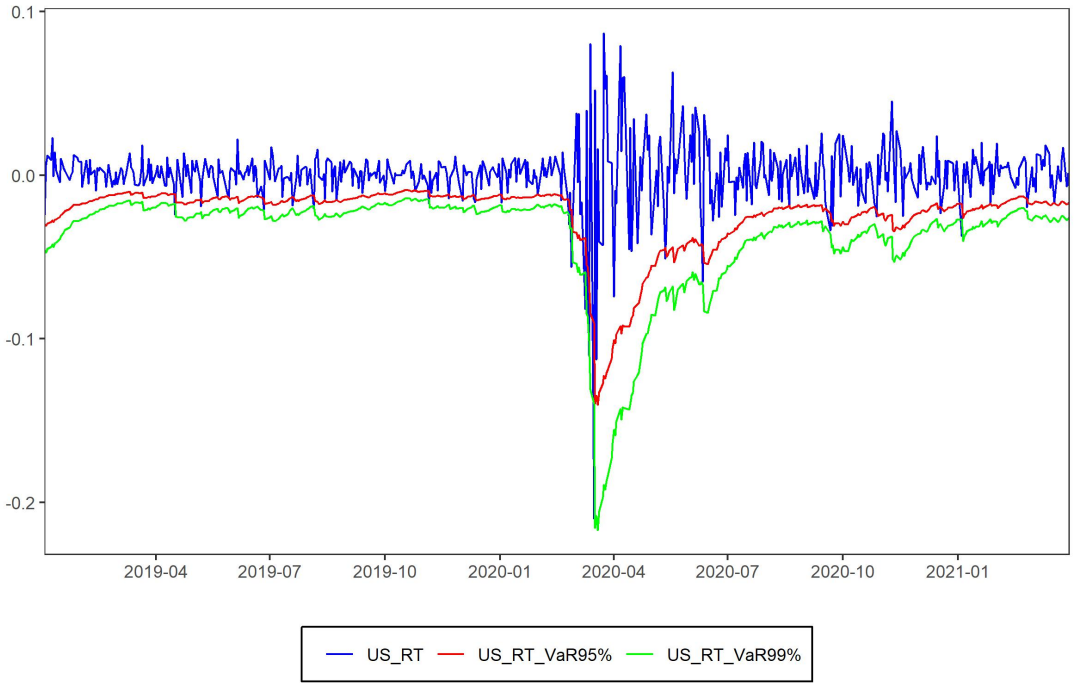


Figure 15: The Time Series Line Plot of Actual US_R_T, VaR95% and VaR99%

GARCH(1,1)-VaR model is quite effective in predicting the downside risk of R_T in both markets.

4.5. Backtesting

Backtesting refers to the statistical procedure used to estimate the reliability and accuracy of the VaR model. The main idea is to statistically examine if the frequency of exceptions is consistent with the chosen confidence level (Miletic & Miletic, 2015).

Kupiec test is most often used for backtesting. The Kupiec test has a chi-square distribution with one degree of freedom. The formula is shown in 10. It denotes the total sample size as T, and the number of exceptions as N. Thus, the failure ratio can be gotten by $\frac{N}{T}$. The null hypothesis is that the number of exemptions follows the binomial distribution, which means that the model is credible (Cerović Smolović et al., 2017).

$$LR = 2\ln[(1-\frac{N}{T})^{\text{T-N}}\frac{N}{T}^{\text{N}}] - 2\ln[(1-P)^{\text{T-N}}P^{\text{N}}] \quad (10)$$

Table 12 supports that VaR at 95% level has high reliability and accuracy for both US_R_T and CN_R_T. However, the test results of VaR at 99% level in both models reject the null hypothesis of Kupiec. This is because the Kupiec test would reject the null hypothesis for both higher and lower frequency of failures than the range at a certain significance

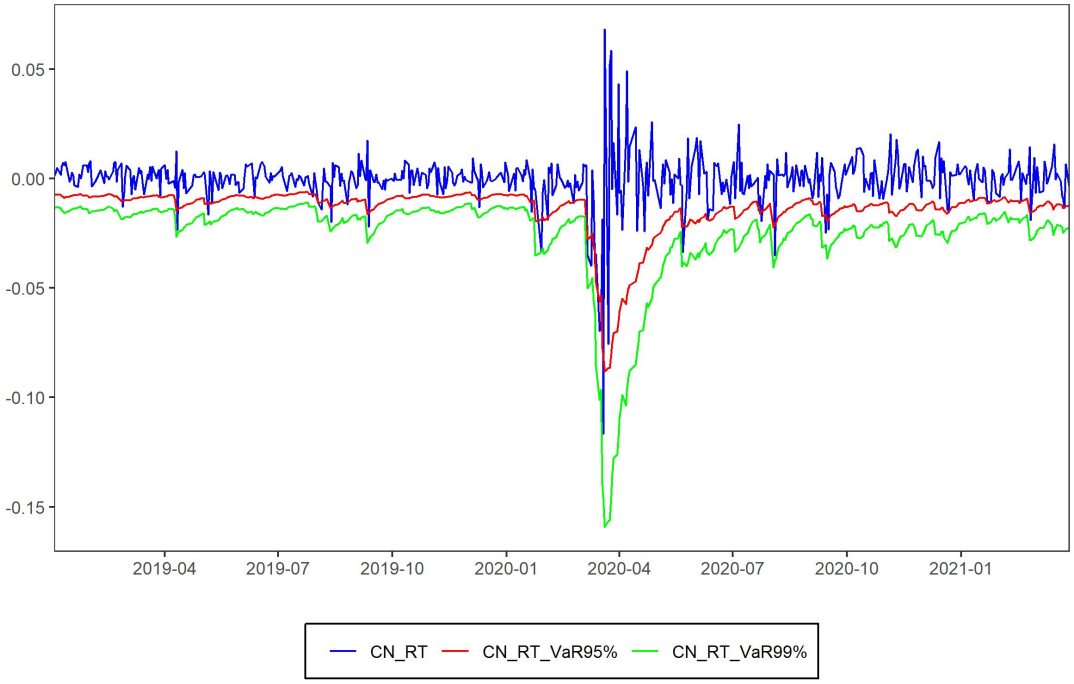


Figure 16: The Time Series Line Plot of Actual CN_R_T, VaR95% and VaR99%

Table 12: The Results of Kupiec Test

*Note: ***Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.*

Variable	Exceptions	Test statistics	P-Value
VAR95% of US_R_T	33	0.764	0.382
VAR99% of US_R_T	0	11.397	0.001***
VAR95% of CN_R_T	28	0.005	0.946
VAR99% of CN_R_T	1	5.677	0.017**

level. The number of exceptions for VaR at 99% has fewer exceptions than expected. It indicates that the predicted VaR at 99% is slightly higher than the actual, especially for the US REIT market.

4.6. Chapter Summary

There are a number of interesting findings from the training data set (from 1st February 2006 to 31st December 2018) and the GARCH(1,1)-VaR model.

From 2006 to 2018, the evidence of relatively low standard deviation indicates that the volatility of the US_R_T and CN_R_T were stable. In other words, the VaR concept can be used to predict the downside risk of REIT effectively.

The US_R_T and CN_R_T data from 2006 to 2018 have the features of fat tails and are leptokurtic, which are not matched with characteristics of normal distribution. Therefore, an econometric approach should be used to evaluate the VaR accurately.

The GARCH(1,1)-VaR is quite suitable for assessing the VaR at 95% level based on the evidence of non-significant Kupiec test results. However, it might be somewhat conservative in terms of estimating the maximum loss at the 99% level. It would be meaningful to perform further research on other econometric approaches of VaR measurement.

5. Conclusion and Outlook

This chapter provides a summary to answer the four research questions mentioned in Section 1.1. Afterwards, an outlook is provided regarding future potential research topics.

5.1. Conclusion

Research questions can be divided into three parts. To begin, the first two questions have been answered using the literature review methodology. The VAR model is constructed to evaluate the relation between the REIT market and macroeconomic factors. Ultimately, downside risk of REIT market is assessed by the GARCH(1,1)-VaR model based on the student's t-distribution.

1) REIT first originated in the US market in 1960, and its success can be attributed to four factors. One, tax benefits have played an essential role in the development of the REIT market. Next, favorable government legislation had also supported the growth of the REIT sector. Additionally, consolidation in the real estate sector helped to keep the sector healthy and growing in the long run. Lastly, REIT specialization helped managers develop specific expertise, which proved to be useful in value generation to promote growth (Parker, 2019).
2) REITs are relatively new in China. In May 2021, China officially launched a REITs pilot program, which focused on the infrastructure REITs. The motivation of the Chinese government was to relieve the debt burden and, at the same time, optimize resources to support economic growth. The issuance of the pilot infrastructure REITs was expected to uncover new opportunities. However, the feasibility of REIT is still questionable due to the transparency problems of traditional infrastructure projects and the complex REIT structure in China.
3) The VAR models using the US and Chinese data led to a consistent conclusion. The stock market is highly correlated with the REIT market. Such a relationship might be explained by investor sentiment and market liquidity (Hoesli & Reka, 2015). It is also noteworthy that the US stock market is found to have a more substantial impact on the REIT market. This could be explained by market maturity differences between the US and China and the possible flaws of the self-constructed China REIT index. The third finding is that REITs are a good hedge against some macroeconomic variables, including the fluctuation of 10-year government bonds yields, industrial production, CPI, and $M2$ money supply.
4) The logarithm returns of the REIT index between the year 2006 and 2018 in both countries have features of fat tails and are leptokurtic. As a result, an econometric approach is necessary to evaluate the accurate VaR. The results of the Kupiec test show that GARCH(1,1)-VaR assuming a t distribution is an excellent measure of the downside risk at the 95% confidence level. However, evidence indicated that it has the tendency to overestimate the maximum loss at the 99% confidence level.

5.2. Outlook

1) In this thesis, the China REIT index is self-constructed, which is based on 9 REITs with underlying mainland China properties listed in Hong Kong and Singapore. Quality of representation can be argued due to the small number of REITs used. It should be noted that with the launching of the infrastructure pilot REIT program in China, one can expect more listings of REITs in China in the future, which can lead to new REIT indexes for research purposes.
2) The target REIT data in this thesis is the equity REIT index, which is a mixture of different specializations. It might be helpful to perform similar research but categorize REITs into their different specializations, such as industrial and retail REITs. It can also be meaningful to conduct further research on mortgage REITs and hybrid REITs.[17]
3) The factors used in this thesis as representations of macroeconomic factors can be subjective. It might be useful to consider different types of factors to get new observations and findings. For example, in the context of the macroeconomic factor inflation risk, it can be helpful to consider factors such as unexpected inflation.
4) A very strong correlation between the REIT index and the stock index has been identified. Therefore, it can be meaningful to perform further research in this aspect to identify additional findings. One suggestion could be to include additional variables such as investor sentiment.
5) The results of the Kupiec test indicated that GARCH(1,1)-VaR assuming a t distribution tends to overestimate the VaR at 99% level. It can be helpful to investigate further on the VaR with different GARCH-family models or even using other econometric approaches.

[17] Hybrid REITs refer to those REITs, which are effectively a combination of equity and mortgage REITs

References

Allen, M. T., Madura, J., & Springer, T. M. (2000). REIT Characteristics and the Sensitivity of REIT Returns. *The Journal of Real Estate Finance and Economics*, *21*(2), 141–152. doi: 10.1023/A:1007839809578

Block, R. L. (2011). *Investing in REITs: real estate investment trusts* (Vol. 141).

Brooks, C., & Tsolacos, S. (1999). The impact of economic and financial factors on UK property performance. *Journal of Property Research*, *16*(2), 139–152.

Cerović Smolović, J., Lipovina-Božović, M., & Vujošević, S. (2017). GARCH models in value at risk estimation: empirical evidence from the Montenegrin stock exchange. *Economic Research-Ekonomska istraživanja*, *30*(1), 477–498.

Deloitte China. (2021). *Deloitte China china reits development report*. https://www2.deloitte.com/cn/zh/pages/real-estate/articles/china-reits-development-report.html. (Accessed: 2021-05-29)

Fang, H., Chang, T.-Y., Lee, Y.-H., & Chen, W.-J. (2016). The impact of macroeconomic factors on the real estate investment trust index return on japan, singapore and china. *Investment Management and Financial Innovations*, *13*(4), 242–253.

Francq, C., & Zakoian, J.-M. (2019). *GARCH models: structure, statistical inference and financial applications*.

Goddard, G. J., & Marcum, B. (2012). *Real estate investment: a value based approach*.

Hoesli, M., & Reka, K. (2015). Contagion channels between real estate and financial markets. *Real Estate Economics*, *43*(1), 101–138.

Kawaguchi, Y., Sa-Aadu, J., & Shilling, J. D. (2017). REIT stock price volatility and the effects of leverage. *Real Estate Economics*, *45*(2), 452–477.

Kola, K. V. (2016). *Macroeconomic risks and REITs: a comparative analysis* (Unpublished doctoral dissertation).

Levy, D., Giano, N., & Jones, K. (2015). Modern REITs and the Corporate Tax: Thoughts on the Scope of the Corporate Tax and Rationalizing Our System of Taxing Collective Investment Vehicles. *University of Chicago*.

Liow, K. H., Ibrahim, M. F., & Huang, Q. (2006). Macroeconomic risk influences on the property stock market. *Journal of Property Investment & Finance*.

Liu, C. H., & Mei, J. (1992). The predictability of returns on equity REITs and their co-movement with other assets. *The Journal of Real Estate Finance and Economics*, *5*(4), 401–418.

Loo, W. K., Anuar, M. A., & Ramakrishnan, S. (2016). Integration between the Asian REIT markets and macroeconomic variables. *Journal of Property Investment & Finance*.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*.

Miletic, M., & Miletic, S. (2015). Performance of Value at Risk models in the midst of the global financial crisis in selected CEE emerging capital markets. *Economic Research-Ekonomska istraživanja*, *28*(1), 132–166.

MSCI. (2021a). *MSCI MSCI US REIT index*. https://www.msci.com/documents/10199/08f87379-0d69-442a-b26d-46f749bb459b. (Accessed: 2021-05-29)

MSCI. (2021b). *MSCI MSCI US REIT index methodology*. https://www.msci.com/eqb/methodology/meth_docs/MSCI_Feb14_US_REIT_Methodology.pdf. (Accessed: 2021-05-29)

Nareit global real estate investment. (2021). https://www.reit.com/investing/global-real-estate-investment. (Accessed: 2021-05-29)

Nareit US REIT industry equity market cap. (2021). https://www.reit.com/data-research/reit-market-data/us-reit-industry-equity-market-cap. (Accessed: 2021-05-29)

New Century. (2021). *New Century REIT 2020 annual report*. https://www.ncreit.com/en/upload/202104/202104091732196931.pdf. (Accessed: 2021-05-29)

Park, J. Y., Mullineaux, D. J., & Chew, I.-K. (1990). Are REITs inflation hedges? *The Journal of Real Estate Finance and Economics*, *3*(1), 91–103.

Parker, D. (2019). The six success factors for REIT market development: lessons for China? *Journal of Property Investment & Finance*.

Peterson, J. D., & Hsieh, C.-H. (1997). Do common risk factors in the returns on stocks and bonds explain returns on REITs? *Real Estate Economics*, *25*(2), 321–345.

Refinitiv. (2021). *Refinitiv Data*. https://workspace.refinitiv.com/web. (Accessed: 2021-03-31)

Simpson, M. W., Ramchander, S., & Webb, J. R. (2007). The Asymmetric Response of Equity REIT Returns to Inflation. *The Journal of Real Estate Finance and Economics*, *34*(4), 513–529. doi: 10.1007/s11146-007-9023-0

Standard and Poor's (S&P). (2021). *S&P china's infrastructure reit market: From slow start to big bang?* https://www.spglobal.com/_assets/documents/ratings/research/100046937.pdf. (Accessed: 2021-05-29)

Standard and Poors (S&P). (2021). *S&P S&P composite 1500*. https://www.spglobal.com/spdji/en/indices/equity/sp-composite-1500/#overview. (Accessed: 2021-05-29)

Stock, J. H., & Watson, M. W. (2001). Vector autoregressions. *Journal of Economic Perspectives*, *15*(4), 101–115.

Swanson, Z., Theis, J., & Casey, K. M. (2002). REIT Risk Premium Sensitivity and Interest Rates. *The Journal of Real Estate Finance and Economics*, *24*(3), 319–330. doi: 10.1023/A:1015273532625

The Balance. (2021). *The Balance what are commercial mortgage-backed securities?* https://www.thebalance.com/what-are-commercial-mortgage-backed-securities-cmbs-416910#:~:text=CMBSs%20offer%20investors%20an%20alternative,are%20created%20from%20debt%20instrument. (Accessed: 2021-05-29)

Tsay, R. S. (2005). *Analysis of financial time series* (Vol. 543).

West, T., & Worthington, A. C. (2006). Macroeconomic risk factors in Australian commercial real estate, listed property trust and property sector stock returns. *Journal of Financial Management of Property and Construction*.

Wong, W. W., et al. (2017). The macroeconomic forces that drive REIT returns in Australia. In *24th Annual European Real Estate Society (ERES) Conference*.

Yin, Z. (2019). *A comparative analysis of apartment REITs in US and China: an examination of challenges and opportunities* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Junior Management Science 7(4) (2022) 899-931

Junior Management Science

journal homepage: www.jums.academy

Moderation in Marketing Applications of Partial Least Squares Structural Equation Modeling (PLS-SEM)

Moderation in Marketinganwendungen von Partial Least Squares Strukturgleichungsmodellierungen (PLS-SEM)

Jonas Spengler

Otto-von-Guericke-Universität Magdeburg

Abstract

This thesis investigates how current marketing research deals with simple moderation analysis in marketing applications of partial least squares structural equation modeling (PLS-SEM). In this context, different approaches to generate the interaction term as well as options of data treatment before analysis, among others, pose challenges to researchers. The goal of this thesis is therefore to highlight the most common sources of error through a review of 33 relevant empirical articles from the PLS-SEM context and to derive recommendations for research. Using a framework that shows how to preferably generate the interaction term, it is found that about 70% of the analyzed studies use a suboptimal approach. Furthermore, the product-indicator approach is the most frequently used but only in exceptional cases the optimal approach. Additionally, both the approach used as well as the effect sizes of moderation effects are generally rarely specified and finally uncertainties in the interpretation of effects in moderation analyses can be observed. This confirms the urgency of the need for educating researchers in terms of moderation analysis in the context of PLS-SEM to minimize common mistakes in the future that distort results of moderation analyses or make their meaning incorrect.

Zusammenfassung

Die vorliegende Arbeit untersucht den Umgang der aktuellen Marketing-Forschung mit einfachen Moderationsanalysen in Marketinganwendungen von Partial Least Squares Strukturgleichungsmodellierungen (PLS-SEM). Dabei stellen u.a. verschiedene Ansätze zur Generierung des Interaktionsterms und die Datenbehandlung vor der Analyse Forscher vor Herausforderungen. Das Ziel dieser Arbeit ist es deshalb, durch ein Review 33 einschlägiger empirischer Artikel aus dem PLS-SEM-Kontext die häufigsten Fehlerquellen zu beleuchten und Empfehlungen für die Forschung abzuleiten. Mithilfe eines Frameworks zur Generierung des Interaktionsterms konnte gezeigt werden, dass etwa 70% der analysierten Untersuchungen einen suboptimalen Ansatz verwenden. Weiterhin ist der Produkt-Indikator-Ansatz der am häufigsten genutzte, aber nur in Ausnahmefällen optimale Ansatz. Hinzu kommen die allgemein seltene Angabe des verwendeten Ansatzes, der Effektgrößen von Moderationseffekten und Unsicherheiten bei der Interpretation der Effekte bei Moderationsanalysen. Dies bestätigt die Dringlichkeit des Bedarfs nach einer Aufklärung von Forschern, damit häufige Fehler, welche die Ergebnisse von Moderationsanalysen verzerren oder inhaltlich falsch werden lassen, in Zukunft minimiert werden.

Keywords: Moderationsanalyse; Interaktionseffekt; Moderator; PLS-SEM; Partial Least Squares Strukturgleichungsmodellierung.

1. Einleitung

Das Testen komplexer Beziehungen zwischen einer Vielzahl von Variablen, die direkt (manifeste Variablen) oder nur unter Zuhilfenahme mehrerer messbarer Variablen beobachtbar sind (latente Variablen), erfolgt zunehmend in verschiedensten wissenschaftlichen Disziplinen mithilfe von Partial Least Squares Strukturgleichungsmodellierung (PLS-SEM) (Sarstedt, Ringle et al., 2020). Dabei nimmt besagte

DOI: https://doi.org/10.5282/jums/v7i4pp899-931

Komplexität der Beziehungen stetig zu, was u.a. darin resultiert, dass sich Forscher neben oftmals bereits untersuchten direkten Effekten zwischen Variablen vermehrt für Moderationseffekte (ME) interessieren (Henseler & Chin, 2010; Henseler, Fassott, Dijkstra & Wilson, 2012).

Moderation wird als der Einfluss auf die Stärke und/oder Richtung der Beziehung zweier Variablen beschrieben, der durch die Manipulation einer dritten Variable, der sog. Moderatorvariable (MV), verursacht wird (Baron & Kenny, 1986, S. 1174). Hierbei kann bspw. untersucht werden, unter welchen Umständen der positive Zusammenhang zwischen Kundenzufriedenheit und -loyalität wesentlich stärker oder schwächer ausgeprägt ist (Henseler & Fassott, 2010). Der ME wird dabei im aktuellen Kontext stets durch die Generierung eines sog. Interaktionsterms modelliert. Diese Modellierung von ME in Strukturmodellen ist jedoch nicht trivial und kommt mit einigen Herausforderungen einher, wie Memon et al. (2019) aufzeigen. Die Autoren weisen neben Schwierigkeiten bei der Hypothetisierung und Interpretation von ME insb. auf verschiedene Ansätze hin, zwischen denen Forscher wählen können, um den Interaktionsterm bei MA in PLS-SEM zu generieren.

So haben Forscher in der Regel die Wahl zwischen dem Produkt-Indikator-Ansatz, Zwei-Stufen-Ansatz und Orthogonalisierungsansatz (Becker, Ringle & Sarstedt, 2018). Eine Herausforderung besteht dabei darin, den passenden Ansatz zu wählen, da diese unterschiedliche Leistungen in Abhängigkeit von den Zielen von Forschern hervorbringen (Henseler & Chin, 2010). Vielmehr sorgen aber unterschiedliche Messmodelle, die latente Variablen in PLS-SEM messen können, für eine schwerwiegende potenzielle Fehlerquelle. Lediglich der Zwei-Stufen-Ansatz ist bspw. dazu in der Lage, eine inhaltlich korrekte MA durchführen zu können, wenn neben reflektiv gemessenen auch formativ gemessene latente Variablen innerhalb der MA vorkommen. Darüber hinaus kann die Behandlung der Daten vor der Analyse einen Einfluss auf die Schätzung des ME haben (Becker et al., 2018). Die Daten können hierbei standardisiert, mittelwertzentriert oder unstandardisiert in die Analyse eingehen. Diese Vielzahl von Entscheidungsmöglichkeiten kann für Verwirrung bei Forschern sorgen, wenn MA in PLS-SEM durchgeführt werden sollen (Hair, Sarstedt & Ringle, 2019).

Henseler und Fassott (2010) haben deshalb in ihrem Artikel einen Leitfaden gegeben, wie MA in PLS-SEM durchzuführen sind. Dieser beinhaltet jedoch nicht die Erkenntnisse der Simulationsstudie von Henseler und Chin (2010) zu den unterschiedlichen Leistungen der Ansätze und die Erkenntnisse von Becker et al. (2018) bezüglich der empfohlenen Datenbehandlung. Hierbei fehlt ebenfalls eine Berücksichtigung des Orthogonalisierungsansatzes. Aus diesem Grund soll in der vorliegenden Arbeit zunächst ein Framework entwickelt werden, das auf der aktuellen Forschung zur Durchführung von MA in PLS-SEM aufbaut. Wie der Leitfaden von Henseler und Fassott (2010) soll es dazu dienen, Forschern eine Vorgabe zu geben, MA in PLS-SEM bestmöglich umzusetzen.

Weiterhin kann u.a. eine bis heute andauernde Kontroverse bezüglich der Interpretation verschiedener Effekte, die bei MA auftreten, beobachtet werden. So vertreten bspw. Hair, Hult, Ringle, Sarstedt und Thiele (2017, S. 256-258) den Standpunkt, Forscher sollten ein Modell ohne den ME schätzen, um die Beziehung zweier Variablen unabhängig vom Moderator testen zu können. Hiergegen argumentieren Li, Sharp, Bergh und Vandenberg (2019), dass eine Interpretation dieser Beziehung nur innerhalb eines Modells möglich ist, das den ME ebenfalls beinhaltet, da sie nach den Autoren immer vom ME abhängig sein muss, insofern dieser signifikant ist.

Die potenzielle Verwirrung sowie die angesprochene Kontroverse von Forschern im Umgang mit MA in PLS-SEM führt zu der Frage, inwieweit den Empfehlungen der Literatur zur Analyse von ME mithilfe von Interaktionstermen in PLS-SEM gefolgt wird und wo häufige Fehler und Probleme auftreten. Zur Beantwortung dieser Frage werden nach der Aufstellung des Frameworks einschlägige Artikel durch eine systematische Literaturrecherche identifiziert und anschließend mithilfe eines Reviews untersucht. Dabei besteht das Ziel darin, aufzuzeigen, wie in aktuellen Forschungsartikeln mit den genannten Herausforderungen umgegangen wird und welche Empfehlungen dabei häufig vernachlässigt oder befolgt werden. Hieraus sollen schließlich Empfehlungen für die Forschung abgeleitet werden, um der Unsicherheit entgegenzuwirken und den angemessenen Umgang mit MA in PLS-SEM in der Zukunft zu motivieren.

Im Folgenden wird hierfür im zweiten Abschnitt der theoretische Hintergrund von PLS-SEM und MA dargelegt. Weiterhin werden bekannte Herausforderungen der aktuellen Forschung aufgezeigt und das Framework zur Vorgehensweise bei MA in PLS-SEM aufgestellt. Der dritte Abschnitt stellt die wissenschaftliche Methode hinter dem anschließenden Review dar. Darauf aufbauend umfasst der vierte Abschnitt das Review der identifizierten Artikel zur Bestimmung von Herausforderungen im Umgang mit MA in PLS-SEM. Abschließend werden Implikationen für die Forschung abgeleitet und Limitationen der vorliegenden Arbeit diskutiert, worauf die abschließend dargestellten Gelegenheiten für zukünftige Forschung basieren.

2. Moderationsanalysen in PLS-SEM

In dem nachfolgenden Abschnitt wird der theoretische Hintergrund der vorliegenden Arbeit erläutert, auf welchem das anschließende Review von Moderationen in Marketinganwendungen von PLS-SEM basiert. Hierbei wird zunächst die Funktionsweise und das Anwendungsgebiet von PLS-SEM dargelegt, wonach eine allgemeine Erläuterung zu Moderationen in Marketinganwendungen sowie eine Aufschlüsselung verschiedener Formen von MA folgt. Weiterhin werden die Vorteile von PLS-SEM bei der Analyse von ME aufgezeigt sowie näher auf die Berechnung der ME in PLS-SEM mithilfe der Generierung von Interaktionstermen eingegangen. Nachdem schließlich die Interpretation von ME und häufig auftretende Fehler behandelt werden, wird aus den aktuellen Erkenntnissen zu Analysen von ME in PLS-SEM mit Interaktionstermen ein Framework aufgestellt, welches als Anleitung

für die Durchführung solcher Analysen in zukünftiger Forschung dienen soll und das zum Teil die Grundlage für das Review im vierten Abschnitt dieser Arbeit bildet.

2.1. Partial Least Squares Strukturgleichungsmodellierung (PLS-SEM)

Liegt der Fokus von Forschern auf der Betrachtung komplexer Beziehungen zwischen zwei oder mehr Variablen, findet häufig eine Form der multivariaten Analyse Anwendung, die simultan die Beziehung mehrerer Variablen zueinander analysiert (Hair et al., 2017, S. 2). Basierend auf einem theoretischen Modell bilden die Variablen oftmals ein Netzwerk kausaler Beziehungen, welches durch statistische Methoden geschätzt wird, die unter Strukturgleichungsmodellierung (SEM) zusammengefasst werden (Esposito Vinzi, Trinchera & Amato, 2010). Insbesondere dient die auf der multivariaten Analyse basierende SEM dabei dazu, auch Netzwerke bestehend aus nicht direkt beobachtbaren Variablen, sog. latenten Variablen bzw. Konstrukten, welche indirekt durch mehrere beobachtbare Indikatorvariablen bzw. manifeste Variablen messbar sind, zu schätzen (Hair et al., 2017, S. 2-4). Konstrukte werden dazu genutzt, abstrakte Ideen und Konzepte zu messen, um menschliches Verhalten besser verstehen und voraussagen zu können (Sarstedt, Hair Jr, Nitzl, Ringle & Howard, 2020). Dabei stellen bspw. die Zufriedenheit mit einem Produkt oder Loyalität zu einem Unternehmen Konstrukte dar. SEM wird hierbei dazu verwendet, um die komplexen wechselseitigen Beziehungen zwischen latenten und manifesten Variablen simultan zu erfassen und dabei Messfehler zu berücksichtigen, die bei der Messung von Konstrukten auftreten (Sarstedt, Hair, Ringle, Thiele & Gudergan, 2016).

Netzwerke aus latenten Variablen werden nach Hair, Ringle und Sarstedt (2011) in der Regel mithilfe sog. Pfadmodelle visualisiert. Abbildung 1 zeigt eine beispielhafte Darstellung eines Pfadmodells mit insg. drei Konstrukten. Diese sind nach den Autoren von Y_1 bis Y_3 benannt, werden über einseitig gerichtete Pfeile (Pfade) miteinander verbunden und besitzen zum Teil eine verschiedene Anzahl von Indikatorvariablen, die sich von x_1 bis x_7 erstrecken. Weiterhin gehen die Autoren darauf ein, dass Y_1 und Y_2 bei der Betrachtung des im aktuellen Kontext sog. inneren Modells (Strukturmodells) exogene latente Variablen darstellen, die andere Konstrukte im Modell erklären, was bedeutet, dass lediglich Pfeile von diesen Konstrukten weggehen. Y_3 hingegen verkörpert nach den Autoren ein endogenes Konstrukt, das im Modell erklärt wird, d.h. es gehen sowohl Pfeile zu diesem Konstrukt hin, sie können jedoch darüber hinaus auch davon ausgehen. Auf Basis einer hinreichenden theoretischen Herleitung der Beziehung zwischen den Konstrukten können die Pfade auch als kausale Beziehungen zwischen den Konstrukten interpretiert werden (Hair et al., 2017, S. 11).

Das sog. äußere Modell (Messmodell) besteht in Abbildung 1 sowohl aus reflektiv als auch aus formativ gemessenen Konstrukten. Nach Hair et al. (2017, S. 13) werden Y_1 und Y_2 in der gezeigten Abbildung 1 reflektiv gemessen, was daran zu erkennen ist, dass die Pfeile ausschließlich von den Konstrukten zu den zugehörigen Indikatoren (von Y_1 zu x_1 und x_2; von Y_2 zu x_3, x_4 und x_5) verlaufen. Nach den Autoren wird Y_3 demnach formativ von den Indikatorvariablen x_6 und x_7 gemessen, da sich sämtliche Pfeile von den Indikatoren zu dem Konstrukt erstrecken.

Werden latente Variablen reflektiv gemessen, sorgen Änderungen im Wert des Konstruktes zu Veränderungen der respektiven Indikatorenwerte und vice versa bei formativ gemessenen Konstrukten (Hair et al., 2011). Reflektive Messungen von Konstrukten sind dabei dadurch gekennzeichnet, dass die Indikatoren die kausale Folge der latenten Variable darstellen und wegen desselben Ursprungs hoch miteinander korreliert sein sollten, weshalb sie bei nur geringen Auswirkungen auf das Konstrukt austauschbar und teilweise auch löschbar sind (Henseler & Fassott, 2010). Werden latente Variablen formativ gemessen, müssen die Indikatoren jedoch nicht wie bei reflektiven Messungen derselben Konstrukt-Domäne entspringen, sondern sie formen bzw. verursachen vielmehr die latente Variable, weshalb diese als die Konsequenz der zugrundeliegenden Indikatorvariablen betrachtet werden kann (Hair et al., 2017, S. 249, Henseler & Chin, 2010). Da jeder Indikator einen spezifischen Aspekt der Domäne des formativ gemessenen Konstruktes abbildet, sind die Indikatoren nicht austauschbar oder ohne weiteres löschbar, weil dies die Bedeutung des Konstruktes verändern würde (Hair et al., 2017, S. 47). Darüber hinaus können latente Variablen auch mithilfe nur eines einzelnen Indikators gemessen werden, wobei der Indikator in diesem Fall weder den Ursprung noch die Konsequenz der Variable darstellt und das Messmodell damit unerheblich ist (Hair et al., 2017, S. 108-109; Henseler & Fassott, 2010).

Hair et al. (2011) unterscheiden bei Methoden zur Schätzung von Strukturgleichungsmodellen zwischen Covariance-Based SEM (CB-SEM) und PLS-SEM. Die Autoren gehen darauf ein, dass PLS-SEM ein Ansatz zur kausalen Modellierung sei und das Ziel verfolge, die erklärte Varianz des betrachteten abhängigen Konstruktes zu maximieren. PLS-SEM wird deshalb vorrangig als explorative Methode betrachtet, bei der mit begrenztem Vorwissen zu den Beziehungen der zu inkludierenden Variablen nach möglichen Richtungen und Stärken der Beziehungen gesucht wird und bei der im Zuge dessen Theorien der Zusammenhänge zwischen den Variablen entwickelt werden (Hair et al., 2017, S. 3-4). Das Ziel von CB-SEM stellt hingegen die Überprüfung bestehender Theorien über strukturelle Beziehungen dar und die Verwendung erfordert die Erfüllung bestimmter Voraussetzungen, wie ein ausreichend großer Stichprobenumfang und eine Normalverteilung der Daten (Hair et al., 2011). Diese Voraussetzungen sind bei der Verwendung von PLS-SEM eher zu vernachlässigen. Deshalb sollten Forscher sich bei der Schätzung von Netzwerken aus latenten Variablen für die Verwendung von PLS-SEM entscheiden, wenn das Ziel die Voraussage von Beziehungen zwischen Konstrukten darstellt, ein kleiner Stichprobenumfang vorliegt, die erhobenen Daten nicht normalverteilt sind, formativ gemessene Konstrukte im Modell vorkommen und das Modell durch viele Konstrukte und Indikatoren sehr komplex ist

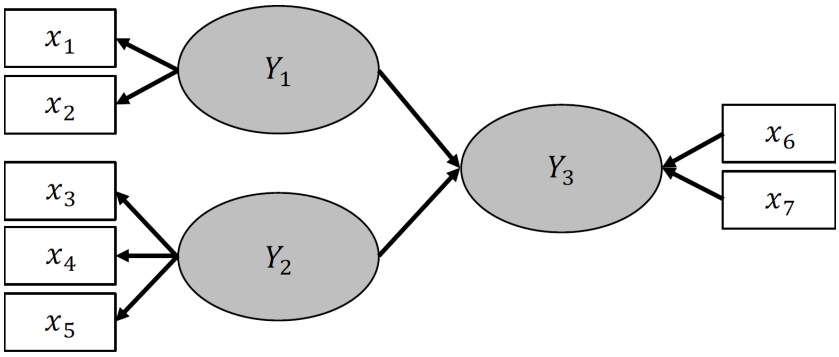


Abbildung 1: Beispielhafte Darstellung eines Pfadmodells (in Anlehnung an Hair et al., 2011)

(Hair et al., 2017, S. 23, Hair et al., 2011). In der vorliegenden Arbeit liegt der Fokus auf PLS-SEM, weshalb CB-SEM größtenteils von den nachfolgenden Betrachtungen ausgeschlossen wird.

Der varianzbasierte Algorithmus, welcher PLS-SEM zugrunde liegt, geht zurück auf Herman Wold (bspw. Wold, 1975) und basiert nach Hair et al. (2011) auf einem zweistufigen Ansatz mit dem Ziel der Maximierung der erklärten Varianz in den endogenen Konstrukten im Modell. Die erste Stufe umfasst nach den Autoren die Approximation der Werte der latenten Konstrukte (latent variable scores (LVS)) für jedes Konstrukt im Modell. Die zweite Stufe schätzt mithilfe der zuvor approximierten LVS, welche nun in je einer multiplen Ordinary Least Squares Regression pro endogenem Konstrukt als unabhängige und abhängige Variablen (je nach Position der betrachteten Konstrukte im Pfadmodell) dienen, die Pfadkoeffizienten des Strukturmodells sowie die finalen Gewichte und Ladungen der Indikatoren (Esposito Vinzi et al., 2010; Hair et al., 2011; Hair et al., 2017, S. 82-86). Gewichte und Ladungen umfassen nach Hair et al. (2011) dabei die Koeffizienten formativ bzw. reflektiv gemessener Beziehungen in PLS-SEM. Dabei wird nach den Autoren die erste Stufe so oft wiederholt, bis die Summe der Veränderungen der berechneten äußeren Gewichte bzw. Ladungen zwischen zwei Iterationen kleiner als ein zuvor festgelegtes Limit ist und die Gewichte bzw. Ladungen dementsprechend konvergieren. Die zweite Stufe bringt schließlich neben Pfadkoeffizienten und den finalen Gewichten und Ladungen auch die erklärte Varianz der endogenen latenten Variablen in Form des Bestimmtheitsmaßes R^2 hervor (Hair et al., 2017, S. 86).

Letztlich sollte nach Hair et al. (2017, S. 104-110) einer Interpretation der Ergebnisse von PLS-SEM zuvor die Evaluation der reflektiven und formativen Messmodelle sowie des gesamten Strukturmodells vorangehen. Dabei sollten nach den Autoren bei einem oder mehreren reflektiv gemessenen Konstrukten im Modell die interne Konsistenzreliabilität, konvergente Validität und diskriminante Validität evaluiert werden. Weiterhin weisen die Autoren darauf hin, dass formative Messmodelle im Modell ebenfalls im Hinblick auf konvergente Validität geprüft werden sollten, darüber hinaus aber auch auf Kollinearität zwischen den Indikatoren sowie auf die Signifikanz und Relevanz der äußeren Gewichte. Das gesamte Strukturmodell ist nach den Autoren anschließend auf das Bestimmtheitsmaß R^2, die prädiktive Relevanz Q^2, Größe und Signifikanz der Pfadkoeffizienten sowie auf die Effektgrößen f^2 und q^2 zu analysieren. Dabei wird bei der Interpretation der Ergebnisse die Signifikanz der Pfadkoeffizienten aufgrund der nichtparametrischen Eigenschaften von PLS-SEM, wegen der keine spezifische Verteilung von Daten voraussetzt wird, durch die Herleitung einer eigenen Verteilung der zugrundeliegenden Daten mithilfe des sog. Bootstrapping-Prozederes getestet (Hair et al., 2017, S. 87). Hierbei werden Standardfehler produziert, aus denen t-Werte und entsprechend auch p-Werte geschätzt werden, die wiederum zur Betrachtung der statistischen Signifikanz der Pfadkoeffizienten dienen (Hair et al., 2017, S. 185-186).

2.2. Moderation in Marketinganwendungen

2.2.1. Definition und Verwendung von Moderation

Moderation bzw. ME treten auf, wenn die Veränderung des Wertes einer dritten Variable (der MV) einen Einfluss auf die Stärke und/oder Richtung der Beziehung zweier anderer Variablen aufweist (Baron & Kenny, 1986, S. 1174). Die MV kann jedoch neben einem Einfluss auf eine spezifische Beziehung auch das gesamte Modell und demnach jede Beziehung in ihm beeinflussen, wobei die Zusammenhänge zwischen den Variablen hierbei nicht für die gesamte Stichprobe konstant sind, sondern in Abhängigkeit vom Moderator variieren (Hair et al., 2017, S. 243). In letzterem Fall kann ein ME auf alle Beziehungen im Modell mithilfe eines Vergleichs mehrerer Untergruppen der Stichprobe in Form einer Multigruppenanalyse getestet werden (Memon et al., 2019). Multigruppenanalysen werden aus den aktuellen Betrachtungen jedoch ausgeschlossen, da der Gegenstand dieser Arbeit insb. der Fall darstellt, in dem eine Moderatorvariable auf eine spezifische Beziehung im Modell wirkt. Diese sog. einfache MA wird in Abbildung 2 visualisiert, wobei Einkommen als MV auf den Zusammenhang zwischen der exogenen Variable Kundenzufriedenheit und der endogenen Variable Kundenloyalität wirkt (Hair et al., 2017, S. 243-244; Memon et al., 2019)[1].

[1] Im weiteren Verlauf der Arbeit meint eine MA immer eine einfache MA, bei welcher der Einfluss einer MV auf die Beziehung einer exogenen und einer endogenen Variable geschätzt werden soll.

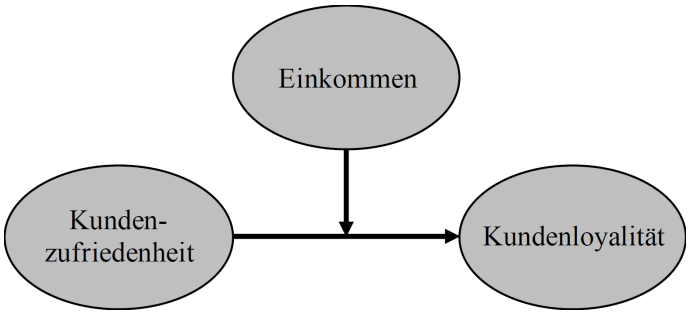


Abbildung 2: Beispielhafte Darstellung eines ME (in Anlehnung an Hair et al., 2017, S. 244)

Der hierbei auftretende ME wird auch als Interaktionseffekt bezeichnet, wobei Hall und Sammons (2014) Interaktionseffekte zunächst als den kombinierten Effekt zweier unabhängiger Variablen auf eine abhängige Variable beschreiben. Die Autoren weisen dabei darauf hin, dass ein Interaktionseffekt nur äquivalent zu einem ME sei, insofern die zugrundeliegende theoretische Grundlage dies gewährleistet. Ein signifikanter Interaktionseffekt kann demnach nur als ME interpretiert werden, insofern der ME a priori hypothetisiert wurde (Field, 2018, S. 486). Auch nach Hair et al. (2017, S. 243) sollte die moderierte Beziehung a priori hypothetisiert werden, wobei zu erkennen sein muss, ob die MV erwartungsgemäß einen Einfluss auf eine oder alle Beziehungen im Modell haben wird. Dabei sollte die Hypothese auf einer theoretischen Herleitung basieren, d.h. auf existierender Literatur und einer Begründung, wieso die Inklusion des Moderators zu einer besseren Erklärung einer Beobachtung führt und in welche Richtung der ME erwartungsgemäß wirken wird (Andersson, Cuervo-Cazurra & Nielsen, 2014; Frazier, Tix & Barron, 2004). Der theoretische Hintergrund kann nach Memon et al. (2019) dabei auf Grundlage von systematischen Literaturrecherchen oder Metaanalysen und den darin enthaltenen Vorschlägen zu zukünftiger Forschung und identifizierten potenziellen Moderatoren gebildet werden. Weiterhin dienen nach den Autoren zur Identifikation potenzieller ME auch Widersprüche in Erkenntnissen verschiedener Forscher, welche Beziehungen zwischen Variablen in unterschiedlichen Bedingungen getestet haben, oder Übertragungen von Erkenntnissen aus anderen wissenschaftlichen Kontexten in den aktuellen.

Unabhängig, ob MV, wie bspw. das Einkommen, Alter, Geschlecht oder die Nationalität, auf eine oder mehrere Beziehungen wirken, können sie dazu in der Lage sein, Heterogenität in den Daten zu erklären (Hair et al., 2017, S. 243). Forscher fragen sich demnach wann, unter welchen Umständen bzw. für wen eine oder mehrere zuvor definierte kausale Beziehungen zu beobachten sind, wenn sie Moderatorvariablen einbinden möchten (Hall & Sammons, 2014). Das übergeordnete Ziel besteht dabei meistens aus der Generierung neuer theoretischer Erkenntnisse und Theorien oder der Schließung von Forschungslücken (Andersson et al., 2014).

2.2.2. Arten von Moderatorvariablen

Die MV kann direkt oder nicht direkt beobachtbare Eigenschaften, wie bspw. die Einstellung zu einer Marke, umfassen, die entsprechend mit einem Indikator oder Skalen aus mehreren formativen oder reflektiven Indikatoren messbar sind (Hair et al., 2017, S. 244-245). Dabei wird nach Henseler und Fassott (2010) in erster Linie abhängig von der verwendeten Skalierung der MV zwischen kontinuierlichen und kategorischen MV unterschieden.

Kontinuierliche MV treten in Form von metrischen Variablen auf und werden deshalb durch Likert-Skalen, unter der Annahme gleichgroßer Abstände zwischen den Ausprägungen der Skala, und somit in Form von quasi-metrischen, sowie mit intervall- und verhältnisskalierten Variablen gemessen (Hair et al., 2017, S. 8-9; Henseler & Fassott, 2010). Der zuvor in Abbildung 2 gezeigte Moderator stellt bspw. eine kontinuierliche MV dar, wobei die Höhe der MV die Beziehung der exogenen und endogenen Variable in ihrer Stärke bzw. Richtung beeinflusst, insofern ein signifikanter ME vorliegt (Hair et al., 2017, S. 246). Die Beziehung wäre demnach konstant, wenn kein signifikanter ME wirkt.

Kategorische MV werden nominal- oder ordinalskaliert gemessen und beinhalten nur einen Indikator pro Konstrukt, welcher die Kategorien in Form von mehreren Antwortmöglichkeiten abbildet (Henseler & Fassott, 2010). Sie haben ebenfalls einen Einfluss auf die Richtung bzw. Stärke der Beziehung der exogenen und endogenen Variable zwischen verschiedenen Kategorien, können darüber hinaus aber auch zeigen, dass in einer Kategorie ein Effekt zwischen exogener und endogener Variable auftritt, während in anderen kein Effekt beobachtbar ist (Field, 2018, S.484). So kann bspw. das Geschlecht als MV die Beziehung zweier Variablen beeinflussen bzw. aufzeigen, dass ein spezifischer Effekt einer exogenen auf eine endogene Variable lediglich bei Frauen und nicht bei Männern beobachtbar ist. In der Regel werden die Kategorien dummy-kodiert, wobei die Referenzkategorie, mit welcher eine oder mehrere andere Kategorien verglichen werden sollen, mit „0“ kodiert wird (Hair et al., 2017, S. 245). Beinhaltet die MV mehr als zwei Kategorien, werden diese nach Henseler und Fassott (2010) mithilfe mehrerer Dummy-Variablen gleichzeitig in das Modell eingeschlossen.

Es besteht bei kontinuierlichen MV nach den Autoren ebenfalls die Möglichkeit mithilfe von Dichotomisierung, also

der Teilung der Stichprobe zur Schaffung mindestens zweier künstlicher Gruppen, Vergleiche von Kategorien durchzuführen. Dies ist jedoch mit dem Verlust von Informationen in Form von Varianz der MV und reduzierter statistischer Power[2] verbunden oder führt sogar dazu, dass Effekte gefunden werden, wo in der Population eigentlich keine auftreten (Frazier et al., 2004). Deshalb empfehlen Henseler und Fassott (2010) sich stets gegen eine Dichotomisierung kontinuierlicher MV zu entscheiden und sie nicht mithilfe eines Gruppenvergleichs zu analysieren.

2.3. Moderation in PLS-SEM

2.3.1. Modellierung des Moderationseffekts

Soll eine einfache MA in PLS-SEM, wie in Abbildung 3 gezeigt, mithilfe der Modellierung eines spezifischen ME (p_3) auf die Beziehung (p_1) einer exogenen (Y_1) und endogenen (Y_2) latenten Variable durchgeführt werden, erfordert dies die Inklusion einer MV (M) in das Pfadmodell, die ebenfalls einen direkten Einfluss (p_2) auf die endogene Variable hat (Hair et al., 2017, S. 247). Der Pfad p_2 verhindert dabei eine Überschätzung des ME (p_3) durch die Kontrolle des direkten Effektes der latenten MV auf das endogene Konstrukt, wobei die Steigung der exogenen latenten Variable nicht konstant sein, sondern linear von der Höhe des Moderators abhängen soll (Henseler & Fassott, 2010).

Der Einschluss des Moderatorkonstrukts erfolgt dabei zunächst in die folgende lineare Gleichung, welche die Main Effects (p_1 und p_2) von Y_1 und M auf Y_2 abbildet, wobei die Konstante p_0 den Schnittpunkt mit der Ordinate darstellt, d.h. wenn alle Steigungen (Regressionsparameter) der Variablen Y_1 und M (d.h. p_1 und p_2) Null sind (Henseler & Fassott, 2010):

$$Y_2 = p_0 + p_1 \cdot Y_1 + p_2 \cdot M + \varepsilon \tag{1}$$

Dabei bildet der Fehlerterm ε die nicht erklärte Varianz ab, d.h. den Messfehler, der bei der Messung der latenten Variablen auftritt (Becker et al., 2018; Henseler & Chin, 2010).

Der in Abbildung 3 dargestellte ME kann anschließend durch die Inklusion des Pfades p_3 nach (Hair et al., 2017, S. 247) sowie Henseler und Fassott (2010) mit der folgenden Gleichung ausgedrückt werden, in der die Stärke der exogenen Variable Y_1 vom Niveau von M abhängt ($(p_1 + p_3 \cdot M) \cdot Y_1$), weshalb p_1 nicht mehr als Main Effect sondern nun als Simple Effect bezeichnet wird:

$$Y_2 = p_0 + (p_1 + p_3 \cdot M) \cdot Y_1 + p_2 \cdot M + \varepsilon \tag{2}$$

Diese Gleichung kann umgestellt werden, um zu verdeutlichen, wie der ME in PLS-SEM integriert wird, wobei das Produkt $Y_1 \cdot M$ den sog. Interaktionsterm darstellt und der ME p_3 im Fall standardisierter LVS ausdrückt, inwieweit sich der Simple Effect (p_1) verändert, wenn M um eine Standardabweichung verringert oder erhöht wird (Becker et al., 2018; Hair et al., 2017, S. 247-248):

$$Y_2 = p_0 + p_1 \cdot Y_1 + p_2 \cdot M + p_3 \cdot (Y_1 \cdot M) + \varepsilon \tag{3}$$

Der als zusätzliche latente Variable in das PLS-SEM Pfadmodell eingebundene Interaktionsterm wird in Abbildung 4 visualisiert und demnach zusammen mit der exogenen latenten Variable Y_1 und dem Moderatorkonstrukt M in demselben Modell geschätzt (Field, 2018, S. 486; Hair et al., 2017, S. 248). Hierbei drückt der Simple Effect (p_1) die Stärke eines Effektes aus, wenn alle anderen Steigungen (p_2 und p_3) null sind, d.h. wenn die MV einen Wert von null hat, womit sich der Simple Effect in seiner Berechnung und Bedeutung vom Main Effect (Effekt, wenn nur $p_2 = 0$) unterscheidet und das Modell mit einem ME anders als das Modell ohne ME betrachtet werden muss (Hair et al., 2017, S. 257-258; Henseler & Fassott, 2010). Letztlich zeigt jedoch der Pfadkoeffizient des Interaktionsterms p_3 und damit der Interaktionseffekt (ME) zusammen mit einem hinreichenden zugrundeliegenden theoretischen Fundament auf, ob eine signifikante Moderation der Beziehung der exogenen und endogenen Variable vorliegt oder nicht (Field, 2018, S. 485-486).

Soll mehr als ein latenter Moderator auf dieselbe Beziehung wirken, was bspw. in Form einer „Cascaded MA“, d.h. einer moderierten Moderation, erfolgen kann, muss dieser nach Henseler und Fassott (2010) innerhalb einer Three-Way Interaction modelliert werden. So ist nach Hair et al. (2017, S. 248) bspw. der ME von Einkommen auf die Beziehung zweier Konstrukte nicht konstant, sondern schwankt ebenfalls durch die Veränderung der MV Alter. Dabei werden im Fall von zwei MV (M_1 und M_2) und jeweils einer exogenen (Y_1) und endogenen (Y_2) Variable nach den Autoren insgesamt acht latente Variablen modelliert, da neben der „Three-Way Interaktion“ $\left(p_7 \cdot (Y_1 \cdot M_1 \cdot M_2)\right)$ auch die Simple Effects ($p_(1$ bis p_3) und die zuvor diskutierten „Two-Way Interactions“ $\left(p_4 \cdot (Y_1 \cdot M_1) + p_5 \cdot (Y_1 \cdot M_2) + p_6 (M_1 \cdot M_2)\right)$ in der zugrundeliegenden linearen Gleichung beinhaltet sein müssen:

$$\begin{aligned} Y_2 = & p_0 + p_1 \cdot Y_1 + p_2 \cdot M_1 + p_3 \cdot M_2 + p_4 \cdot (Y_1 \cdot M_1) \\ & + p_5 \cdot (Y_1 \cdot M_2) + p_6 \cdot (M_1 \cdot M_2) + p_7 \cdot (Y_1 \cdot M_1 \cdot M_2) + \varepsilon \end{aligned} \tag{4}$$

2.3.2. Der Nutzen von PLS-SEM bei der Analyse von Moderationseffekten

Sollen ME innerhalb multivariater Analysen mit latenten Variablen analysiert werden, haben Forscher nach Sarstedt, Ringle et al. (2020) die Wahl zwischen CB-SEM und PLS-SEM, aber auch Regressionsanalysen mithilfe des PROCESS-Makros von Hayes (2013) finden hierbei Anwendung. Dabei unterliegen Regressionsanalysen nach den Autoren allerdings dem Nachteil, dass sie keine Messfehler bei der Berechnung von LVS in die Betrachtung mit einbeziehen. Hair et al. (2017) zeigen, dass eine Kombination aus Über- und Unterschätzung der Berechnung des gesamten Pfadmodells

[2] Nach Cohen (1988, S. 4) ist die statistische Power eines Tests die Wahrscheinlichkeit dafür, einen signifikanten Effekt zu finden, wenn dieser tatsächlich in der Population existiert.

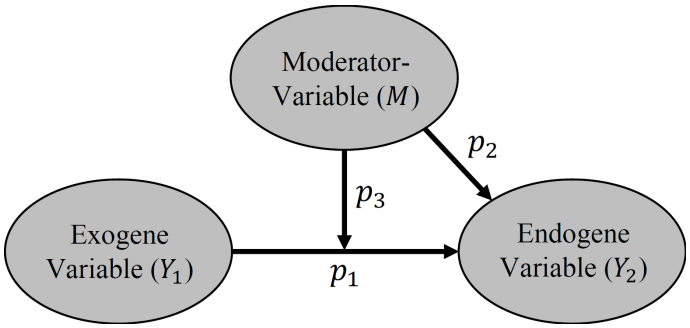


Abbildung 3: Moderationseffekt im Pfadmodell (in Anlehnung an Hair et al., 2017, S. 247)

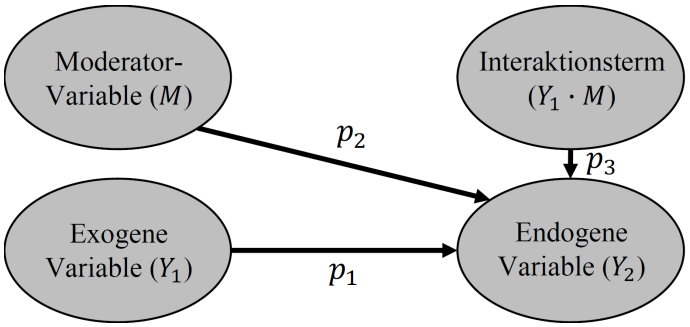


Abbildung 4: Einbindung eines Interaktionsterms (in Anlehnung an Hair et al., 2017, S. 248)

resultiert, wenn Messfehler nicht korrigiert werden. Die mit der Berechnung der LVS einhergehenden Messfehler können allerdings mithilfe von SEM berücksichtigt werden, was die Verzerrung des Pfadmodells verhindert (Sarstedt et al., 2016). Insbesondere bei der Analyse von MA ist dies von Bedeutung, da bei der Generierung des Interaktionsterms der verzerrende Effekt von Messfehlern durch die Multiplikation zweier fehlerbehafteter Konstrukte verstärkt wird (Li et al., 2019). Hinzukommt, dass ME zumeist klein ausfallen, was auch zu einer geringen statistischen Power beiträgt, weshalb bereits eine marginale negative Verzerrung des Effektes schwer ins Gewicht fallen und den Effekt nicht auffindbar machen kann (Aguinis, Beaty, Boik & Pierce, 2005; Sarstedt, Hair Jr et al., 2020). Somit ist eine Art der SEM einer Regressionsanalyse vorzuziehen, wenn ME in komplexen Modellen mit latenten Variablen geschätzt werden sollen.

Darüber hinaus ist nach Henseler und Fassott (2010) aufgrund des Ablaufes der zweiten Stufe des PLS-SEM-Algorithmus eine Schätzung von ME in PLS-SEM wie in multiplen Regressionsanalysen möglich. Nach den Autoren und nach Sarstedt, Ringle et al. (2020) ermöglicht die lineare Kombination der Indikatoren der exogenen Variable und der MV zu LVS eine analoge Vorgehensweise wie bei Regressionen, weshalb Forscher den Interaktionseffekt leicht in PLS-SEM abbilden und interpretieren können. CB-SEM hingegen liefert keine LVS, weshalb Interaktionen schwieriger abzubilden sind (Hair et al., 2017, S. 15). Hinzukommt, dass PLS-SEM mehrere Ansätze zur Modellierung des Interaktionsterms liefert, auf welche nachfolgend im Detail eingegangen wird und die es ermöglichen, sowohl reflektiv als auch formativ gemessene Konstrukte als MV in PLS-SEM einzubinden (Memon et al., 2019). Auch komplexe Modelle mit einer Vielzahl von Konstrukten und Indikatoren, sowie komplexe Beziehungen (d.h. Moderationen und Mediationen von Beziehungen) lassen sich ohne Probleme in PLS-SEM abbilden (Sarstedt, Hair Jr et al., 2020). Anhand der dargelegten Argumente kann PLS-SEM als der zu favorisierende SEM-Ansatz zur Durchführung von MA in Modellen mit latenten Variablen identifiziert werden.

2.3.3. Ansätze zur Modellierung des Interaktionsterms

Der aus latenten Variablen generierte Interaktionsterm zur Abbildung eines ME in PLS-SEM kann in MA mithilfe verschiedener Ansätze generiert werden: Produkt-Indikator-Ansatz, Orthogonalisierungsansatz, Zwei-Stufen-Ansatz und hybrider Ansatz (Henseler & Chin, 2010). Diese sind in Tabelle 1 zusammengefasst, die einen Vergleich der vier Ansätze anhand der Funktionsweisen, der Anwendbarkeit in Abhängigkeit von den Messmodellen der zugehörigen MV und der exogenen Variablen sowie der Vor- und Nachteile der Ansätze ermöglicht.

Der in Regressionsanalysen als Standardansatz zur Generierung des Interaktionsterms geltende Produkt-Indikator-Ansatz wurde erstmals von Chin, Marcolin und Newsted (2003) in den PLS-SEM-Kontext übertragen, um den zuvor diskutierten Nachteilen von Regressionsanalysen bei der Schätzung von ME im Zusammenhang mit latenten Variablen auszuweichen (Hair et al., 2017, S. 249; Henseler & Fassott, 2010). Henseler und Chin (2010) gehen darauf ein, dass die hierbei erfolgende Bildung des Interaktionsterms durch

die paarweise Multiplikation jedes Indikators der latenten exogenen Variable mit denen der MV relativ einfach in PLS-SEM möglich sei. Die so entstehenden Produkt-Indikatoren dienen nach den Autoren schließlich im Strukturmodell als die Indikatoren des Interaktionsterms. Abbildung 5 verdeutlicht die Funktionsweise des Produkt-Indikatoren Ansatzes im Fall von drei Indikatoren der exogenen Variable und zwei Indikatoren der MV (Hair et al., 2017, S. 250).

Es ist hierbei zu erkennen, dass sich aus den Produkten aller möglichen Kombinationen der Indikatoren der MV und exogenen Variable insg. sechs Produkt-Indikatoren ergeben, die den Interaktionsterm reflektieren.

Die multiplizierten Indikatoren müssen bei der Verwendung des Produkt-Indikator-Ansatzes von reflektiv gemessenen Konstrukten stammen, da nur so angenommen werden kann, dass diese auch tatsächlich jeweils zu genau einer entsprechenden inhaltlichen Domäne der MV bzw. der exogenen latenten Variable gehören und die Multiplikation dieser Indikatoren auch tatsächlich eine Interaktion der Variablen und ihrer Domänen wiederspiegelt (Hair et al., 2017, S. 249). Die Indikatoren müssen dafür austauschbar sein und recht hoch miteinander korrelieren (Henseler & Chin, 2010). Formativ messende Indikatoren können jedoch unabhängig voneinander sein und müssen nicht zwingend derselben Domäne eines Konstruktes entspringen, das sie formen, weshalb formativ gemessenen Konstrukten keine Domäne eindeutig zugeordnet werden kann (Hair et al., 2017, S. 249; Henseler & Fassott, 2010). Deshalb lassen Produkt-Indikatoren, die vollständig oder zum Teil aus einer Menge formativer Indikatoren bestehen, bei der Multiplikation der Indikatoren nicht zwingend auf denselben zugrunde liegenden Interaktionseffekt schließen (Chin et al., 2003). Dadurch wird die konzeptuelle Domäne des Interaktionsterms gestört und er verliert seine inhaltliche Bedeutung, was dazu führt, dass der Produkt-Indikator-Ansatz bei formativ gemessenen Konstrukten konzeptuell nicht abbildbar ist (Hair et al., 2017, S. 249).

Letztlich wird durch die Wiederverwendung der Indikatoren des exogenen Konstrukts und des Moderatorkonstrukts im Messmodell des Interaktionsterms Multikollinearität, also eine hohe Korrelation der Konstrukte, im Pfadmodell hervorgerufen (Henseler & Chin, 2010). Dies kann nach Hair et al. (2017, S. 249-251) zu verzerrten Schätzungen der Pfadkoeffizienten oder überschätzten Standardfehlern führen, da die Indikatoren des Interaktionsterms einen Teil der Varianz der Indikatoren der exogenen Variable und der MV teilen. Zur Reduktion der Multikollinearität und Erleichterung der Interpretation des ME werden deshalb nach den Autoren die Indikatoren der latenten MV vor der Generierung des Interaktionsterms häufig standardisiert (Umwandlung zu Variablen mit Mittelwert von null und Standardabweichung von eins).

Die Autoren weisen weiterhin darauf hin, dass die Standardisierung der Indikatoren vor der Generierung des Interaktionsterms jedoch die resultierende Multikollinearität nicht vollständig eliminiert und den Vergleich von Main Effect und Simple Effect bei der Interpretation des ME erschwert. Um die Korrelation des Interaktionsterms und der exogenen Variablen und der MV zu beseitigen, haben Little, Bovaird und Widaman (2006) nach den Autoren den Orthogonalisierungsansatz als Erweiterung des Produkt-Indikator-Ansatzes entwickelt, der ebenfalls die Bildung von Produkt-Indikatoren des Interaktionsterms auf dieselbe Art und Weise beinhaltet. Diese werden jedoch nicht direkt als Indikatoren des Interaktionsterms genutzt, sondern es wird eine Regression pro Produkt-Indikator durchgeführt, wobei diese als abhängige und die Indikatoren der latenten exogenen Variable und MV als unabhängige Variablen in der Regression dienen (Henseler & Chin, 2010). Im Fall des in Abbildung 6 gezeigten Modells würden aufgrund von drei Indikatoren des exogenen Konstrukts und zwei des Moderatorkonstrukts insgesamt sechs Regressionsmodelle folgen, da bei der paarweisen Multiplikation aller Indikatoren sechs Produkt-Indikatoren entstehen (Hair et al., 2017, S. 250-251):

$$x_1 \cdot m_1 = b_{1,11} \cdot x_1 + b_{2,11} \cdot x_2 + b_{3,11} \cdot x_3 + b_{4,11} \cdot m_1 + b_{5,11} \cdot m_2 + e_{11}$$
$$x_1 \cdot m_2 = b_{1,12} \cdot x_1 + b_{2,12} \cdot x_2 + b_{3,12} \cdot x_3 + b_{4,12} \cdot m_1 + b_{5,12} \cdot m_2 + e_{12}$$
$$x_2 \cdot m_1 = b_{1,21} \cdot x_1 + b_{2,21} \cdot x_2 + b_{3,21} \cdot x_3 + b_{4,21} \cdot m_1 + b_{5,21} \cdot m_2 + e_{21}$$
$$x_2 \cdot m_2 = b_{1,22} \cdot x_1 + b_{2,22} \cdot x_2 + b_{3,22} \cdot x_3 + b_{4,22} \cdot m_1 + b_{5,22} \cdot m_2 + e_{22}$$
$$x_3 \cdot m_1 = b_{1,31} \cdot x_1 + b_{2,31} \cdot x_2 + b_{3,31} \cdot x_3 + b_{4,31} \cdot m_1 + b_{5,31} \cdot m_2 + e_{31}$$
$$x_3 \cdot m_2 = b_{1,32} \cdot x_1 + b_{2,32} \cdot x_2 + b_{3,32} \cdot x_3 + b_{4,32} \cdot m_1 + b_{5,32} \cdot m_2 + e_{32}$$

Die sog. standardisierten Residuen (e_{11} bis e_{32}) werden anschließend als Indikatoren des Interaktionsterms genutzt, was auch als residuale Zentrierung bezeichnet wird und ebenfalls in Abbildung 6 dargestellt ist (Henseler & Chin, 2010). Dieses zweistufige Prozedere sorgt dafür, dass die Varianz des neuen orthogonalen Interaktionsterms von der Varianz der exogenen Variable und der MV unabhängig ist und nur den Interaktionsterm repräsentiert (Little et al., 2006). Dadurch ist der Interaktionsterm unkorreliert mit den zugrundeliegenden Konstrukten (Becker et al., 2018).

Hair et al. (2017, S. 251) verweisen hierbei darauf, dass die Parameterschätzungen des Main Effects und Simple Effects in diesem Fall sehr gleich und bei der Interpretation des ME äquivalent zu betrachten sind, was diese vereinfacht. Dieses Fehlen eines Simple Effects kann nach Echambadi, Arroniz, Reinartz und Lee (2006) und Becker et al. (2018) allerdings auch zu Problemen in der Interpretation von Effekten führen, da die Möglichkeit, den Simple Effect zu interpretieren ausbleibt und der Main Effect fälschlicherweise als Simple Effect interpretiert wird. Letztlich kann auch dieser Ansatz nur bei reflektiv gemessenem exogenen und moderierenden Konstrukt angewendet werden, da er auf der Bildung von Produktindikatoren beruht, für welche dieselben Voraussetzungen wie im Produkt-Indikator-Ansatz gelten (Hair et al., 2017, S. 251).

Damit auch formativ gemessene latente Variablen in MA untersucht werden können, haben Chin et al. (2003) den Zwei-Stufen-Ansatz vorgeschlagen, welcher durch die von PLS-SEM ermöglichte Schätzung der LVS von Konstrukten die MA in zwei Stufen durchführt (Henseler & Fassott, 2010). Wie in Abbildung 7 zu erkennen ist, wird in der ersten Stufe zunächst das Modell der Main Effects geschätzt, in welchem

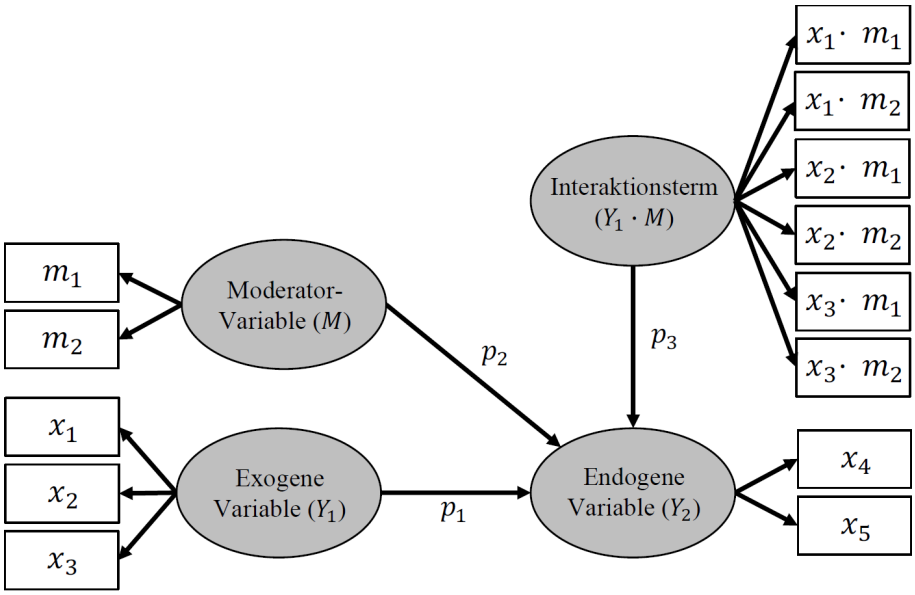


Abbildung 5: Produkt-Indikator-Ansatz (in Anlehnung an Hair et al., 2017, S. 250)

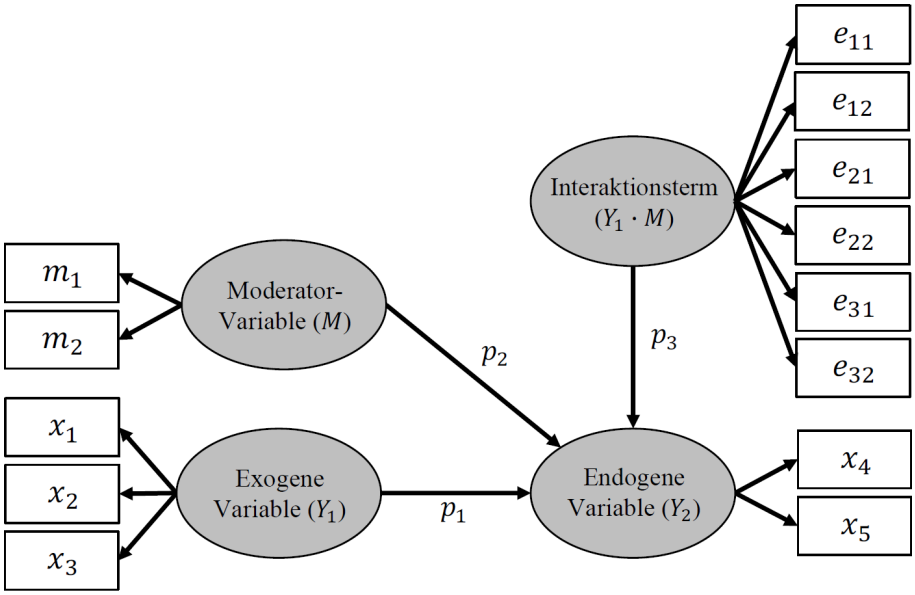


Abbildung 6: Orthogonalisierungsansatz (in Anlehnung an Hair et al., 2017, S. 252)

der Interaktionsterm noch nicht enthalten ist, und es entstehen die LVS der latenten exogenen Variable und der MV, die für die zweite Stufe gespeichert werden, in der die Multiplikation dieser LVS zu einem einzigen Indikator erfolgt, mit dem schließlich der Interaktionsterm gemessen wird (Hair et al., 2017, S. 251-252). Anschließend werden in der zweiten Stufe neben dem Interaktionsterm alle anderen latenten Variablen mit ihrem zuvor bestimmten LVS als unabhängige Variablen in einer multiplen linearen Regression auf die endogene Variable Y_2 genutzt (Henseler & Fassott, 2010).

Hierbei werden nach Henseler und Chin (2010) die LVS durch den PLS-SEM-Algorithmus stets standardisiert, der Interaktionsterm jedoch nicht, was ohnehin Schwierigkeiten in der Interpretation des Interaktionseffektes hervorrufen würde. Durch die Standardisierung werden auch die bei diesem Ansatz auftretenden Probleme reduziert, die durch Multikollinearität im Strukturmodell hervorgerufen werden (Becker et al., 2018; Fassott, Henseler & Coelho, 2016). Letztlich weisen Henseler und Chin (2010) darauf hin, dass der Zwei-Stufen-Ansatz auch bei ausschließlich reflektiv gemessenen MV sowie exogenen Variablen angewendet werden kann und nicht auf formative gemessene Konstrukte beschränkt ist.

Den finalen in der Literatur häufig besprochenen Ansatz stellt der hybride Ansatz dar, dessen Funktionsweise Wold (1982) geprägt hat und der von Henseler und Chin (2010) auf MA in PLS-SEM übertragen wurde.

Tabelle 1: Übersicht der gängigen Ansätze zur Modellierung des Interaktionsterms in PLS-SEM, deren Funktionsweise, die jeweils unterstützten Messmodelle und Vor- sowie Nachteile (in Anlehnung an Becker et al., 2018; Hair et al., 2017, S. 249-252; Henseler & Chin, 2010; Henseler & Fassott, 2010; Henseler et al., 2012)

Ansatz	Funktionsweise	Unterstützte Messmodelle	Vorteile	Nachteile
Produkt-Indikator-Ansatz (Chin et al., 2003)	Multiplikation der Indikatoren der exogenen Variable und MV zur Generierung von Produktindikatoren des Interaktionsterms	nur reflektiv gemessene exogene und Moderatorkonstrukte	einfache Anwendung in PLS-SEM	Multikollinearität durch Wiederverwendung der Indikatoren
Orthogonalisierungsansatz (Henseler & Chin, 2010)	Regression der Indikatoren der MV und des exogenen Konstrukts auf die Produktindikatoren des Interaktionsterms und Verwendung der Residuen als Indikatoren des Interaktionsterms	nur reflektiv gemessene exogene und Moderatorkonstrukte	keine Multikollinearität durch zu anderen Konstrukten orthogonalem Interaktionsterm	(Interpretation des Simple Effects nicht möglich)
Zwei-Stufen-Ansatz (Chin et al., 2003)	1. Stufe: Schätzung des Main Effects Models zur Generierung der LVS der MV und des exogenen Konstrukts 2. Stufe: Multiplikation beider LVS zur Generierung des einzigen Indikators des Interaktionsterms	reflektiv oder formativ gemessene exogene oder Moderatorkonstrukte	Anwendung bei formativ gemessener MV und/oder exogener Variable	Multikollinearität durch Wiederverwendung der Indikatoren in Form von LVS
Hybrider Ansatz (Henseler & Chin, 2010; Wold, 1982)	wie Zwei-Stufen-Ansatz, wobei Approximation des Interaktionsterms bei der Schätzung der LVS mit einbezogen wird	reflektiv oder formativ gemessene exogene und Moderatorkonstrukte	Anwendung bei formativ gemessener MV und/oder exogener Variable	in keiner gängigen Software implementiert

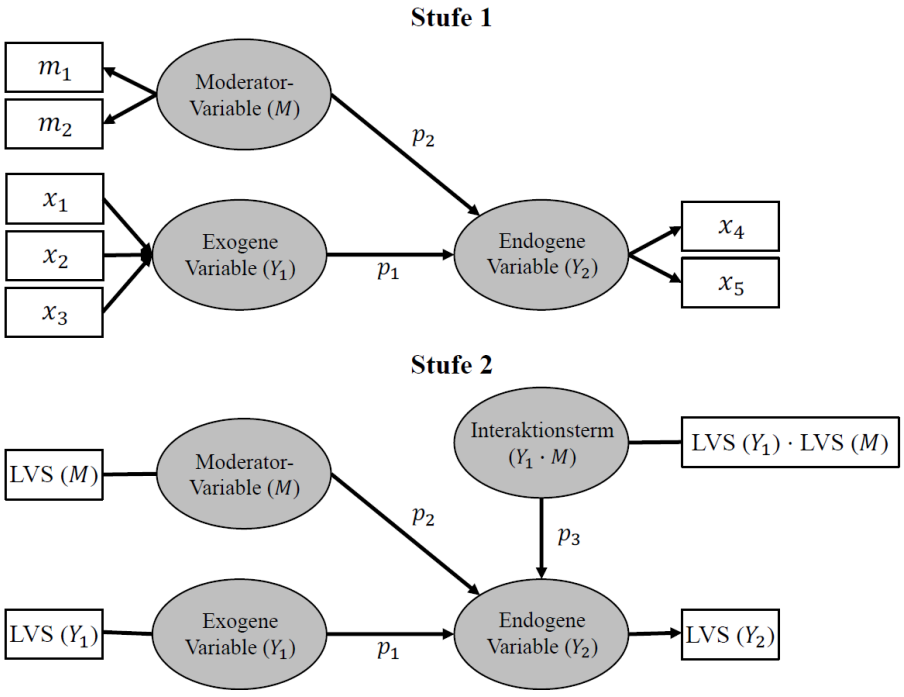


Abbildung 7: Zwei-Stufen-Ansatz (in Anlehnung an Hair et al., 2017, S. 253)

Nach Henseler und Chin (2010) kombiniert dieser Ansatz die Vorteile des Zwei-Stufen-Ansatzes und des Produkt-Indikatoren Ansatzes. Insbesondere gehen die Autoren darauf ein, dass der Interaktionsterm dabei wie bei dem Zwei-Stufen-Ansatz mithilfe des Produkts der LVS der MV und des exogenen Konstruktes gemessen wird. Darüber hinaus findet nach den Autoren wie bei dem Produkt-Indikator-Ansatz allerdings auch die Einbeziehung des Interaktionsterms selbst bei der Generierung der LVS statt, während der PLS-SEM-Algorithmus die LVS bestimmt. Aus dem Grund, dass hierbei LVS genutzt werden, können wie bei dem Zwei-Stufen-Ansatz auch formativ gemessene MV und exogene Konstrukte innerhalb der MA mit dem hybriden Ansatz genutzt werden (Henseler & Chin, 2010). Allerdings wird dieser Ansatz aus den weiteren Betrachtungen von MA in PLS-SEM ausgeschlossen, da er nach der Simulationsstudie von Henseler und Chin (2010) nicht besser als die anderen Ansätze bei der Schätzung von ME in ausschließlich reflektiv gemessenen Modellen funktioniert, nicht in gängiger Software zur Schätzung von PLS-SEM implementiert ist und aus diesem Grund bisher auch seltener von Forschern genutzt wurde (Becker et al., 2018).

Schließlich ist die Wahl des verwendeten Ansatzes unerheblich, wenn die exogene Variable und die MV mit nur einem Indikator gemessen werden, da das Messmodell hierbei keinen Einfluss auf die Messung der „latenten“ Variable hat (Henseler & Fassott, 2010).

2.3.4. Auswertung der Ergebnisse von Moderationsanalysen in PLS-SEM

Modellevaluation

Die zuvor besprochene Evaluation der Messmodelle der latenten Variablen und des gesamten Strukturmodells sollten nach Hair et al. (2017, S. 255) ebenfalls in Modellen mit MV durchgeführt werden. Auch hier muss nach den Autoren im Rahmen der Messmodellevaluation nach reflektiven und formativen Moderatorkonstrukten unterschieden werden. Sie weisen jedoch auch darauf hin, dass der Interaktionsterm nicht zu evaluieren ist. Werden bspw. der Orthogonalisierungsansatz oder der Produkt-Indikator-Ansatz bei der MA genutzt, liegen dem Interaktionsterm nach den Autoren durch die Multiplikation der Indikatoren beider Konstrukte anstatt einer spezifischen zwei konzeptionelle Domänen zugrunde. Vielmehr bringen die Autoren zum Ausdruck, dass das Messmodell des Interaktionsterm lediglich ein Hilfsmittel darstellt, um die wechselseitige Beziehung zwischen der latenten MV und dem exogenen Konstrukt im Pfadmodell zu integrieren. Insofern existiert der Interaktionsterm nur im Modell, drückt aber kein wirklich existentes Konstrukt aus, das tatsächlich gemessen werden kann, weshalb die Evaluation hinfällig ist.

Entgegen dieser auf den Prinzipien hinter der Messung latenter Variablen basierenden Sichtweise, bringen Aguinis, Edwards und Bradley (2017) zum Ausdruck, dass die Evaluation des Interaktionsterms durchaus eine große Bedeutung haben sollte. Sie beziehen sich wie auch Li et al. (2019) dabei vorrangig auf die Evaluation der Reliabilität des Interaktionsterms. Diese soll nach den Autoren insb. bei der Multiplikation von unkorrelierten Konstrukten aus dem Produkt der einzelnen Reliabilitätsmessungen bestehen und demnach wesentlich geringer ausfallen. Folgt man dieser Argumentation, spricht auch das Argument von (Hair et al., 2017, S. 255) für eine sorgfältige Evaluation des Interaktionsterms, da nach den Autoren die Wiederverwendung von Indikatoren der exogenen und moderierenden Variable bei der Bildung des Interaktionsterms für eine hohe Korrelation der Konstrukte sorgt und damit gegen die diskriminante Validität verstößt. Sie argumentieren dabei allerdings nicht für die Evaluation des Interaktionsterms, sondern dagegen, da die designbedingte Verletzung der diskriminanten Validität für die Autoren einen weiteren Grund darstellt, keine Evaluation durchzuführen.

Unabhängig von diesen unterschiedlichen Sichtweisen weisen die Autoren darauf hin, dass bei der Verwendung des Zwei-Stufen-Ansatzes keine Evaluation des Messmodells des Interaktionsterms erfolgen muss, da dieser lediglich mit einem Indikator aus dem Produkt der LVS der exogenen Variable und der MV gemessen wird. Im Anschluss an die Messmodellevaluation muss nach den Autoren das gesamte Strukturmodell ebenfalls im Hinblick auf die entsprechenden zuvor dargelegten Kriterien evaluiert werden, wobei die Effektgröße f^2 besondere Beachtung erhalten sollte.

Ergebnisinterpretation

Die Interpretation der Ergebnisse von MA umfasst insg. drei Aspekte: Betrachtung der Signifikanz des ME sowie des Simple Effects, Berechnung und Berichten der Effektgröße f^2, Erfassung der Stärke und Richtung des ME mithilfe eines Simple Slope Plots (Memon et al., 2019). Hierbei ist die inhaltliche Bedeutung dieser Aspekte in den Vordergrund zu stellen und mit dem zuvor dargelegten theoretischen Fundament zusammen zu interpretieren (Andersson et al., 2014). In der Regel sollte ein signifikanter ME und eine vom Forscher als substantiell eingeschätzte Effektgröße f^2 vorliegen, damit sichergestellt werden kann, dass der ME auch tatsächlich einen hinreichend großen Einfluss auf die Beziehung der Variablen ausübt und so wirklich zu den anschließend diskutierten Implikationen der MA für Forschung und Management führt (Henseler et al., 2012).

Die Untersuchung, ob die statistische Signifikanz des Interaktionsterms vorliegt, sollte demnach zunächst der Fokus der Interpretation der Ergebnisse einer MA sein (Hair et al., 2017, S. 256). Dies kann mithilfe des Effekts des Interaktionsterms auf das endogene Konstrukt betrachtet werden, wobei die Hypothese des ME unterstützt werden kann, wenn der entsprechende Pfadkoeffizient (p_3) signifikant ist (Henseler & Fassott, 2010). Durch das Bootstrapping-Prozedere wird dabei festgestellt, ob die MV einen signifikanten ME auf die Beziehung der exogenen und endogenen Variable ausübt, indem der Pfadkoeffizient p_3 darauf geprüft wird, ob sich der ME signifikant von null unterschiedet und damit tatsächlich in der Population existiert (Hair et al., 2017, S. 256; Henseler & Fassott, 2010).

Bei der Interpretation der ME ist nach

Hair et al. (2017, S. 256) ebenfalls eine Betrachtung des Simple Effects und ggf. des Main Effects zu erfolgen. Dabei sollte nach den Autoren der Wert null einen sinnvollen Bestandteil der Skala der MV darstellen, da der Simple Effect p_1 die Beziehung zwischen exogenem und endogenem Konstrukt darstellt, wenn die MV einen Wert von null annimmt. Deshalb wird nach Henseler und Fassott (2010) Mittelwertzentrierung der Indikatoren der MV und des exogenen Konstruktes vorgeschlagen, wobei der Referenzpunkt zum Mittelwert der Indikatoren verschoben wird. Hierfür müssen jedoch die Standardabweichungen der verschiedenen Variablen nach den Autoren in ihrer Interpretation untereinander vergleichbar sein. Hair et al. (2017, S. 257) empfehlen hierfür die Standardisierung der Indikatoren der MV, wodurch die Umwandlung zu Variablen mit dem Mittelwert bei null und der Standardabweichung von eins erfolgt. Somit werden sie in ihrer Bedeutung vergleichbar gemacht und die null stellt einen sinnvollen Bestandteil der Skalen der Variablen dar.

Weiterhin schlagen die Autoren die getrennte Betrachtung des Main Effects und Simple Effects vor, wenn der Produkt-Indikator-Ansatz oder der Zwei-Stufen-Ansatz verwendet werden, da sich der Simple Effect in seiner Bedeutung und dem geschätzten Pfadkoeffizienten vom Main Effect unterscheidet. Nach den Autoren sowie nach Chin et al. (2003) sollten Forscher, die sich für einen Test einer Hypothese interessieren, welche explizit den Main Effect beinhaltet und nicht den von der MV abhängigen Simple Effect, zunächst ein Modell ohne die MV heranziehen, um Verwechselungen der Effekte zu verhindern. Wird der Simple Effect für den Main Effect gehalten, d.h. aus dem Modell entnommen, das den ME enthält, kann dies nach Henseler und Fassott (2010) zu irreführenden oder inkorrekten Schlussfolgerungen führen. Hierbei sollte angemerkt werden, dass Aguinis et al. (2017) und Li et al. (2019) eine andere Sichtweise zum Ausdruck bringen. Die Autoren weisen darauf hin, dass eine Interpretation des Main Effects nicht sinnvoll sei, da das Vorliegen eines signifikanten ME bedeutet, dass kein konstanter Effekt der exogenen auf die endogenen Variable besteht, sondern dieser immer mit der Ausprägung der MV schwankt. Deshalb sollte nach den Autoren stets der Simple Effect auch im Rahmen von Hypothesen herangezogen werden, die explizit nur den Einfluss der exogenen auf die endogene Variable testen wollen.

Aus diesem Grund kann das von Echambadi et al. (2006) kritisierte Prinzip der residualen Zentrierung nach Becker et al. (2018) auch zu Problemen der Interpretation führen, wenn der Orthogonalisierungsansatz verwendet wird. Hierbei entsteht nach den Autoren durch die Orthogonalität des Interaktionsterms kein Simple Effect. Hair et al. (2017, S. 258) sind hierbei wie bereits erwähnt der Ansicht, dass dies zu keinen Problemen in der Interpretation führt, da der Fokus nach den Autoren nicht auf dem Simple Effect sondern auf dem Main Effect liegen sollte, wenn Forscher den konstanten Effekt der exogenen auf die endogene Variable betrachten möchten. Die Autoren bringen demnach zum Ausdruck, dass der ME und Main Effect aus demselben Modell entnommen werden sollen, während Aguinis et al. (2017) und Becker et al. (2018) hierbei das Problem sehen, dass die Betrachtung des Main Effects, als wäre er ein Simple Effect, zu inkorrekten Interpretationen führen kann.

Der zweite Aspekt der Ergebnisinterpretation umfasst die Berechnung und das Berichten der Effektgröße f^2, die aufzeigt, wie sehr die MV zur Erklärung des endogenen Konstruktes beiträgt (Hair et al., 2017, S. 256). Sie ist nach Carte und Russell (2003) einer Einschätzung der Größe des ME anhand des Pfadkoeffizienten p_3 vorzuziehen. Dabei berechnet sich f^2 nach Henseler und Fassott (2010) aus der Änderung der erklärten Varianz R^2 im endogenen Konstrukt, wenn der Interaktionsterm im Modell enthalten ist ($R^2_{included}$), entgegengesetzt dem Fall, dass er nicht enthalten ist ($R^2_{excluded}$). Diesbezüglich schlagen die Autoren folgende Formel zur Berechnung der Effektgröße f^2 vor:

$$f^2 = \frac{R^2_{included} - R^2_{excluded}}{1 - R^2_{included}} \quad (5)$$

Hierbei wird f^2 in der Regel nach Cohen (1988, S. 412-414) bei f^2=0,02 als schwach, bei f^2=0,15 als moderat und bei f^2=0,35 als hoch eingeschätzt (Henseler & Fassott, 2010). Diese Einschätzung von Effektgrößen ist nach Aguinis et al. (2005) jedoch kontextabhängig und sollte von jedem Forscher selbst eingestuft werden. Aus diesem Grund hat Kenny (2018) auf Basis eines Reviews von Effektgrößen und statistischer Power in MA kategorischer Variablen im Kontext von multipler Regression, das von Aguinis et al. (2005) stammt und in dem eine mittlere Effektgröße bei ME von nur 0,009 gefunden wurde, die folgende realistischere Einschätzung von f^2-Werten vorgeschlagen: f^2=0,005 sei schwach, f^2=0,01 sei moderat und f^2=0,025 sei hoch (Hair et al., 2017, S. 256). Dabei kann auch ein schwacher Effekt bspw. unter der Bedingung stark durch den ME beeinflusster Pfadkoeffizienten eine große Bedeutung haben und ist infolgedessen nicht direkt aus der Betrachtung auszuschließen (Chin et al., 2003; Aguinis et al., 2005). Aus diesem Grund sollte die Diskussion der Implikationen der MA für Manager und Forscher auch bereits auf Basis eines schwachen signifikanten Effektes nach Kenny (2018) erfolgen und nicht bei einer Unterschreitung eines schwachen Effekts nach Cohen (1988, S. 412-414) unterlassen werden, was bspw. von Henseler et al. (2012) empfohlen wird.

Um bei der Interpretation die Richtung und Stärke des ME korrekt einschätzen zu können, sollten die Resultate der MA in Form eines Simple Slope Plots dargestellt werden (Memon et al., 2019). Dabei werden die Steigungen der Kurven (Slopes) grafisch miteinander verglichen, welche die Beziehung der endogenen und exogenen Variable für verschiedene Niveaus oder Kategorien der MV darstellen (Field, 2018, S. 484-485). Der Pfadkoeffizient p_3, also die Beziehung zwischen Interaktionsterm und endogenem Konstrukt, gibt hier außerdem Auskunft über die Richtung des ME. Dabei zieht ein positiver Pfadkoeffizient einen positiven ME nach sich und vice versa.

Abbildung 8 zeigt ein Beispiel eines positiven ME. Hierbei

stellt man nach Hair et al. (2017, S. 258) die Beziehung des exogenen und endogenen Konstruktes für verschiedene Niveaus der MV dar, wobei der Moderator im Fall des hohen Niveaus meistens auf einer Standardabweichung über bzw. im Fall des niedrigen Niveaus unter seinem Mittelwert betrachtet wird. Ist der Moderator auf seinem Mittelwert, also auf einem mittleren Niveau, entspricht die Steigung der abgebildeten Kurve dem Simple Effect p_1, was in Abbildung 8 mithilfe der durchgezogenen Linie dargestellt ist. Es ist weiterhin zu erkennen, dass die gestrichelte Kurve, welche die Beziehung der exogenen und endogenen latenten Variable bei einem hohen Niveau der MV darstellt, steiler als die durchgezogene Kurve verläuft. Demnach ist die Beziehung auf dem hohen Niveau der MV stärker. Dies ist damit zu begründen, dass die Steigung der Kurve nun p_1+p_3 entspricht (Hair et al., 2017, S. 269). Gegenteiliges kann bei einem niedrigen Niveau (gepunktete Linie) der MV beobachtet werden, bei der die Beziehung zwischen der exogenen und endogenen Variable weniger stark ausgeprägt ist. In diesem Fall liegt wie bereits erwähnt ein positiver ME vor. Falls sich die Steigungen der Kurven bei verschiedenen Niveaus oder Kategorien der MV kaum bzw. gar nicht unterscheiden und die Kurven (eher) parallel verlaufen, deutet dies auf die Abwesenheit eines signifikanten ME hin (Field, 2018, S. 485). Letztlich ist bei einem negativen ME die Kurve steiler, welche die Beziehung der exogenen und endogenen Variable auf dem niedrigen Niveau der MV darstellt, als die Kurve auf dem Mittelwert der MV (Hair et al., 2017, S. 269). Wird der Moderator hierbei kleiner (größer), resultiert ein stärkerer (schwächerer) Zusammenhang der exogenen und endogenen Variable.

Wie eingangs erwähnt, sollte die Interpretation der Ergebnisse der MA und die Schlussfolgerung von Implikationen für die Forschung und/oder das Management stets auf die eingangs aufgestellte theoretische Grundlage zurückkommen, d.h. die Bedeutung der Ergebnisse im Kontext des theoretischen Hintergrunds in den Vordergrund stellen und analysieren, anstatt sich kontextlos auf die Signifikanz des ME zu fokussieren (Memon et al., 2019).

2.3.5. Herausforderungen bei Moderationsanalysen in PLS-SEM

Potenzielle Fehlerquellen

Bis hierhin sind bereits einige Fehlerquellen bei der Durchführung und Auswertung von MA herausgestellt worden, die nun zusammengefasst und um weitere Herausforderungen bei MA ergänzt werden sollen, was ebenfalls in Tabelle 2 dargestellt ist. Diese beinhaltet insg. neun potenzielle Quellen für Fehler, die kurz beschrieben werden und für die ggf. eine mögliche Lösung bzw. eine empfohlene Vorgehensweise zur Vermeidung der Fehler dargestellt wird.

Zunächst kann die Wahl des Ansatzes zur Generierung des Interaktionsterms dazu führen, dass inhaltliche Fehler entstehen, wie etwa durch die Nutzung des Produkt-Indikator-Ansatzes oder Orthogonalisierungsansatzes bei formativen Moderatorkonstrukten bzw. exogenen Konstrukten. Der resultierende ME und die Interpretation sowie die aus der MA geschlossenen Implikationen basieren nicht auf einer inhaltlich korrekten Bildung des Interaktionsterms, da den zugrundeliegenden formativ gemessenen Konstrukten keine eindeutige inhaltliche Domäne zugeordnet werden kann. (Hair et al., 2017, S. 249). Die Domäne des Interaktionsterms aus dem Produkt aus mindestens einem formativ gemessenem Konstrukt wäre folglich ebenfalls nicht bestimmbar und die inhaltliche Bedeutung des Interaktionsterms ist unklar. Deshalb darf bei einer formativ gemessenen latenten MV oder exogenen Variable lediglich der Zwei-Stufen-Ansatz gewählt werden.

Aber auch bei reflektiv gemessenen Konstrukten sind die Ansätze zur Bildung des Interaktionsterms nach einer Simulationsstudie von Henseler und Chin (2010) nicht gleichwertig zu behandeln, da sie abhängig vom Ziel der MA eine unterschiedliche Leistung erzielen. Wie eingangs erwähnt sorgt die Auswahlmöglichkeit aus den verschiedenen Ansätzen für Verwirrung bei Forschern (Hair et al., 2019). Welcher Ansatz unter welchen Bedingungen gewählt werden sollte, wird im nächsten Abschnitt der vorliegenden Arbeit näher beschrieben.

Der zu wählende Ansatz hängt in diesem Fall vordergründig von dem Ziel der Forscher ab, das sie bei der Durchführung von MA mit PLS-SEM verfolgen. Ein hierbei entstehendes Problem könnte die Unsicherheit von Forschern über ihr Ziel darstellen. In erster Linie werden im PLS-SEM-Kontext Effekte auf ihre Signifikanz getestet, was demnach das häufigste Ziel von Forschern darstellt und weshalb sich in den meisten Fällen für die Verwendung des Ansatzes mit der größten statistischen Power entschieden werden sollte. Oftmals wird jedoch die Nutzung von PLS-SEM damit begründet, dass das Hauptziel der Analyse die Vorhersage von Effekten sein soll. Daraus folgt, dass Forscher Probleme damit haben könnten, ihr Ziel korrekt einzuschätzen und infolgedessen einen nachteiligen Ansatz wählen.

Eine weitere Fehlerquelle stellt nach Becker et al. (2018) die Möglichkeit dar, bei der Anwendung des Orthogonalisierungsansatzes und Produkt-Indikator-Ansatzes über die Datenbehandlung der Indikatordaten vor der Generierung des Interaktionsterms zu entscheiden. Dabei besteht die Wahl zwischen Standardisierung, Mittelwertzentrierung oder keiner Behandlung der Daten (d.h. Verwendung unstandardisierter Daten). In ihrer Simulationsstudie haben die Autoren herausgefunden, dass alle Ansätze zur Generierung des Interaktionsterms bei der Genauigkeit der Schätzung der ME am besten mit der Verwendung von standardisierten Indikatorendaten funktionieren. Dies erleichtert wie bereits dargelegt auch die Interpretation des Simple Effects (Hair et al., 2017, S. 257). Aus diesem Grund sollte von Mittelwertzentrierung oder gar keiner Behandlung der Daten vor der MA abgesehen und die Daten stattdessen unabhängig vom verwendeten Ansatz standardisiert werden, um Verzerrungen der Schätzung des Interaktionseffektes zu vermeiden (Becker et al., 2018).

Nach der Schätzung des Pfadmodells in PLS-SEM folgt die Evaluation des Messmodells. Dabei ist wie bereits dargelegt die MV genauso wie die exogene und endogene Variable zu evaluieren, wogegen der Interaktionsterm nach Hair et al. (2017, S. 255) kein Gegenstand der Evaluation des Mess-

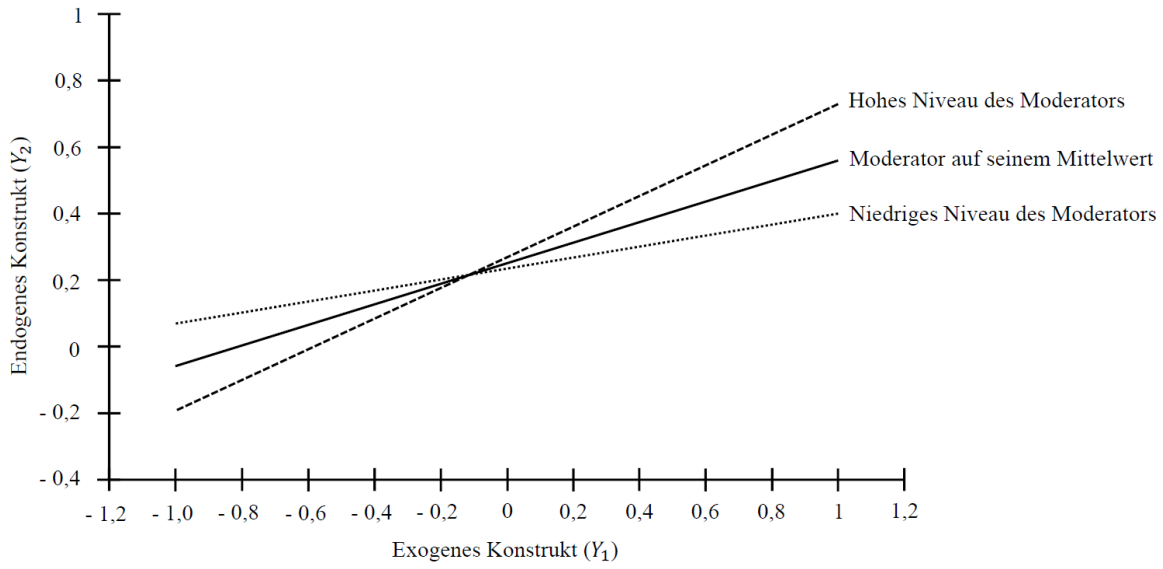


Abbildung 8: Beispielhafte Darstellung eines Simple Slope Plots (in Anlehnung an Hair et al., 2017, S. 259)

Tabelle 2: Potenzielle Fehlerquellen bei MA mit PLS-SEM mit Empfehlungen zur Problemvermeidung (in Anlehnung an Aguinis et al., 2017; Becker et al., 2018; Chin et al., 2003; Frazier et al., 2004; Hair et al., 2017, S. 249-258; Henseler & Chin, 2010; Henseler & Fassott, 2010; Li et al., 2019)

Fehlerquelle	Beschreibung der Fehlerquelle	Empfehlungen
Wahl des Ansatzes zur Generierung des Interaktionsterms (1)	nur Zwei-Stufen-Ansatz möglich bei formativ gemessenen Konstrukten	Verwendung des Zwei-Stufen-Ansatzes bei formativ gemessenen Konstrukten
Wahl des Ansatzes zur Generierung des Interaktionsterms (2)	Ansätze weisen verschiedene Leistung in Abhängigkeit vom Ziel der MA auf	siehe Framework in Abschnitt 2.4 für Umgang mit reflektiv gemessenen Konstrukten
Wahl des Ansatzes zur Generierung des Interaktionsterms (3)	Ziel der MA wird falsch von Forschern eingeschätzt	Ziel der MA ist im Zweifel das Herausstellen der statistischen Signifikanz
Datenbehandlung vor der Generierung des Interaktionsterms	Wahl zwischen Standardisierung, Mittelwertzentrierung und Verwendung unstandardisierter Daten	Standardisierung der Daten vor der Generierung des Interaktionsterms
Evaluation des Messmodells des Interaktionsterms	unterschiedliche Ansichten dazu, ob Evaluation (insb. der Reliabilität) erfolgen sollte	zumindest Angabe der individuellen Reliabilitäten des exogenen und moderierenden Konstrukts
Interpretation des Main Effects und Simple Effects	unterschiedliche Ansichten dazu, ob Main Effect und Simple Effect getrennt werden sollten	-
Signifikanztest des ME	Testen der MV auf Signifikanz	Testen des Interaktionsterms auf Signifikanz
Interpretation der Effektgröße f^2	keine Bestimmung von f^2 oder Interpretation nach Cohen (1988, S. 412-414)	Bestimmung von f^2 und Interpretation nach Kenny (2018)
Künstliche Dichotomisierung kontinuierlicher MV	künstliche Dichotomisierung sorgt für Verlust von Varianz in MV und geringerer Power	kontinuierliche MV nicht künstlich dichotomisieren

modells sein sollte. Insofern die Betrachtung von Messfehlern eine übergeordnete Rolle für Forscher spielt, sollten diese jedoch (zumindest) im Rahmen der Betrachtung der internen Konsistenzreliabilität nach Aguinis et al. (2017) auch für den Interaktionsterm berechnet werden. Forscher müssen sich deshalb entscheiden, ob sie dem inhaltlich korrekten Weg nach Hair et al. (2017, S. 255) folgen und Messfehler des Interaktionsterms ignorieren oder der empirisch begründeten Vorgehensweise nach Aguinis et al. (2017) nachgehen und Messfehler mit in ihre Betrachtungen einbeziehen. Nach

allen genannten Autoren sollten aber in jedem Fall die individuellen Messfehler der exogenen, endogenen und moderierenden Variable angegeben werden.

In ähnlicher Art und Weise können auch die verschiedenen Sichtweisen von Hair et al. (2017, S. 257-258), Chin et al. (2003) und Henseler und Fassott (2010) auf der einen Seite und Aguinis et al. (2017) und Li et al. (2019) auf der anderen Seite dazu führen, dass Unsicherheiten im Umgang mit den Main Effects und Simple Effects entstehen. Erstgenannte Autoren befürworten eine hierarchische Vorgehensweise bei der Analyse der Effekte, wobei jeweils ein Modell mit MV zur Berechnung des Main Effects und ein Modell ohne MV zur Berechnung des Simple Effects geschätzt werden soll. Somit werden die Effekte strikt voneinander getrennt und eine Interpretation des Main Effects bspw. zum Testen einer Hypothese des linearen Effektes der exogenen auf die endogene Variable ist unabhängig vom ME möglich. Dies ist nach den zuletzt genannten Autoren allerdings inhaltlich nicht sinnvoll, da ein signifikanter ME bedeutet, dass der Main Effect nicht existiert und nur der Simple Effect eine Bedeutung hat. Erneut könnten die verschiedenen Sichtweisen hierbei für Unsicherheit bei Forschern sorgen.

Darüber hinaus besteht die Möglichkeit, dass Forscher nicht den Interaktionseffekt als Grundlage für die Interpretation eines signifikanten ME heranziehen, sondern den Pfadkoeffizienten (p_2) von der MV zu der endogenen Variable im Modell auf Signifikanz prüfen. Dies würde dazu führen, dass der Effekt der MV auf das endogene Konstrukt getestet wird und die MV dadurch wie eine exogene Variable behandelt werden würde. Weiterhin besteht hierbei auch die Möglichkeit der Verwechslung von Moderation und Mediation, wobei fälschlicherweise der direkte Effekt der MV auf das exogene Konstrukt (oder vice versa) betrachtet und somit auch kein ME untersucht werden würde. Demzufolge wird hierbei nicht die Moderation der Beziehung einer exogenen und endogenen latenten Variable analysiert, die folglich nur durch den zusätzlichen Interaktionsterm im Modell und seinen Effekt auf die endogene Variable modelliert werden kann. Daher ist stets der Interaktionsterm auf Signifikanz zu prüfen, um auf einen signifikanten ME schließen zu können (Henseler & Fassott, 2010).

Wurde die Signifikanz eines ME festgestellt, sollte dieser auf seine Effektgröße f^2 geprüft werden. Eine erste potenzielle Fehlerquelle besteht hierbei darin, die Effektgröße nicht zu bestimmen. Dies stellt eine wahrscheinliche Quelle für Fehler dar, da die Berechnung bisher nur in wenigen Softwarepaketen implementiert und standardmäßig bei MA durchgeführt wird (Memon et al., 2019). Ein zweiter potenzieller Fehler wäre hierbei der verfrühte Ausschluss von zu schwachen signifikanten Effekten nach Cohen (1988, S. 412-414) aus der MA und der anschließenden Diskussion der Implikationen des ME. Henseler et al. (2012) empfehlen diese Vorgehensweise, sie sollte jedoch dem Vorschlag von Kenny (2018) angepasst werden. Demnach wird empfohlen, die Untergrenze für ME von $f^2 = 0{,}005$ heranzuziehen und nur bei einer Unterschreitung anzunehmen, dass der ME zu schwach ist, um Implikationen nach sich zu ziehen (Hair et al., 2017, S. 256).

Schließlich weisen Aguinis et al. (2017) neben den bisher besprochenen Fehlerquellen bei MA u.a. darauf hin, dass eine künstliche Dichotomisierung kontinuierlicher Variablen zu einer Reduktion der statistischen Power führt und die Größe des ME weiter verkleinert, da hierbei die Varianz der MV reduziert wird. Dies sollte demnach auch im PLS-SEM-Kontext bei MA vermieden werden und kontinuierliche Variablen sollten unverändert in die MA einbezogen werden (Henseler & Fassott, 2010).

Moderierte Mediation und Hierarchical Component Models

Im Kontext von MA in PLS-SEM stellt letztlich die Einbindung eines Moderators in eine Mediation eine besondere Herausforderung bei der Berechnung der Effekte und Interpretation dieser dar. Im Fall einer Mediation wirkt nach Hair et al. (2017, S. 247) eine dritte Variable in ähnlicher Art und Weise wie ein Moderator auf die Stärke der Beziehung zwischen zwei Variablen. Hierbei wird jedoch untersucht, inwiefern eine Änderung im exogenen Konstrukt zu einer Änderung im Mediatorkonstrukt führt, was nach den Autoren durch einen direkten Effekt zwischen diesen beiden Variablen modelliert wird. Moderierte Mediationsmodelle (Conditional Process Models) stellen die Kombination von Moderation und Mediation dar, wobei ein Moderator auf mindestens eine der direkten Beziehungen im Mediationsmodell wirkt und dieses damit abhängig von der Ausprägung des Moderators wird (Sarstedt, Hair Jr et al., 2020). Abbildung 9 stellt ein Beispiel eines Conditional Process Models dar, wobei die MV (M) einen Einfluss auf die Beziehung der exogenen Variable (Y_1) und der Mediatorvariable (Y_2) hat.

Die Berechnung und Interpretation des ME insb. im Zusammenhang mit der Mediation hängt hierbei allerdings von Faktoren, wie der Anzahl der Pfade, auf welche die MV wirkt, oder der Position des Pfades im Mediationsmodell, den die MV beeinflusst, ab (Hair et al., 2017, S. 261). Aus diesem Grund übersteigt die detaillierte Betrachtung von Conditional Process Models den Umfang der vorliegenden Arbeit.

Dies gilt ebenfalls für MA im Zusammenhang mit Hierarchical Component Models (HCM), welche Konstrukte umfassen, die nach Hair et al. (2017, S. 281-283) in mehreren Hierarchiestufen gemessen werden. Hierbei bilden nach den Autoren konkretere Bestandteile eines Konstruktes die Ebene erster Ordnung und das abstrakte Konstrukt die Ebene zweiter Ordnung. Weiterhin weisen die Autoren darauf hin, dass je nach Messmodell der Konstrukte erster oder zweiter Ordnung insg. vier verschiedene Arten von HCM entstehen können (d.h. reflektiv-reflektiv, reflektiv-formativ, formativ-reflektiv und formativ-formativ) und diese auf zwei verschiedene Weisen analysiert werden (d.h. mithilfe des Repeated Indicators Approach oder dem Zwei-Stufen-Ansatz). Becker, Klein und Wetzels (2012) gehen in diesem Zusammenhang auf die Schwierigkeiten bei der Analyse von ME im Zusammenhang mit HCM ein. Auch in diesem Fall wird von der detaillierten Betrachtung von HCM bei MA abgesehen, da eine Diskussion der zahlreichen Möglichkeiten bei der Analyse und Interpretation von HCM bei MA den Umfang dieser Ar-

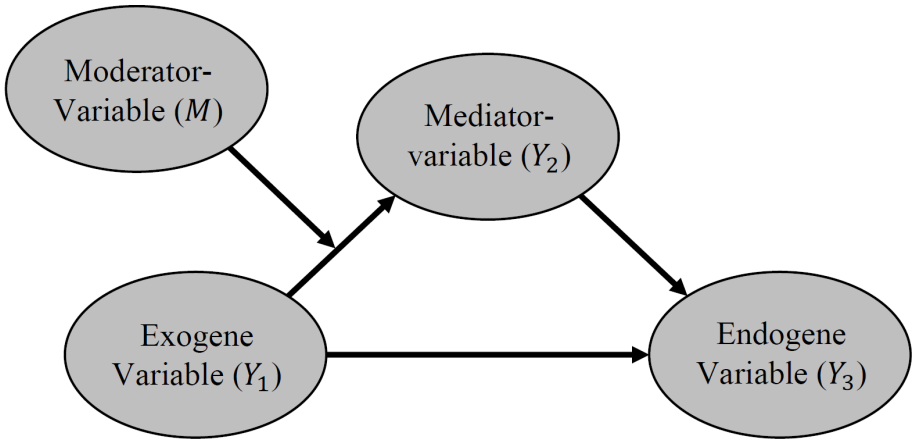


Abbildung 9: Beispielhafte Darstellung einer moderierten Mediation (in Anlehnung an Hair et al., 2017, S. 260)

beit übersteigt.

2.3.6. Framework zur Durchführung von Moderationsanalysen in PLS-SEM

Um die Unsicherheiten von Forschern bei der Berechnung und Interpretation von ME in PLS-SEM zu überwinden, wird nachfolgend ein Framework aufgestellt, das die Datenbehandlung vor der Generierung des Interaktionsterms, die Wahl des besten Ansatzes zur Generierung des Interaktionsterms und einen Teil der Vorgehensweise bei der Interpretation der ME und der Effektgröße f^2 zusammenfasst.

Damit das in Abbildung 10 dargestellte Framework zur Generierung des Interaktionsterms bei der Durchführung von MA verwenden werden kann, müssen nachfolgende Voraussetzungen erfüllt sein. Zunächst muss PLS-SEM zur Schätzung des Pfadmodells verwendet werden. Weiterhin sollte die Durchführung einer einfachen MA mit der Bildung eines Interaktionsterms zur Modellierung des ME erfolgen, wodurch Multigruppenanalysen zur Schätzung von ME ausgeschlossen werden. Außerdem kann die Verwendung des Frameworks, wie im vorigen Abschnitt dargelegt, nur mit Konstrukten erfolgen, die nicht Bestandteil eines HCM sind. Letztlich ist die Verwendung des Frameworks auf kontinuierliche latente bzw. manifeste MV (quasi-metrische, intervall- und verhältnisskalierte) limitiert, da die Variablen in den dem Framework zugrundeliegenden Simulationsstudien ebenfalls in jedem Fall kontinuierlich waren[3].

Wie in Abbildung 10 zu erkennen ist, sollten die Indikatoren der latenten MV (M) sowie des exogenen (Y_1) und endogenen Konstruktes (Y_2) vor der Generierung des Interaktionsterms standardisiert werden. Hierbei wird nach Henseler und Fassott (2010) sichergestellt, dass die Variablen untereinander in ihrer Bedeutung vergleichbar sind, was die anschließende inhaltliche Interpretation vereinfacht. Außerdem wird die bei dem Zwei-Stufen-Ansatz und Produkt-Indikator-Ansatz hervorgerufene Multikollinearität im Pfadmodell durch die Standardisierung verringert. Schließlich deuten auch die Erkenntnisse aus der Simulationsstudie von Becker et al. (2018), darauf hin, dass vor der Generierung des Interaktionsterms standardisierte Daten, im Gegensatz zu Daten, welche mittelwertzentriert oder unstandardisiert belassen wurden, die genausten Schätzungen der Pfadkoeffizienten des Moderatormodells hervorbringen. Die Autoren verweisen aber ebenfalls darauf, dass der Zwei-Stufen-Ansatz die LVS, aus denen der Interaktionsterm gebildet wird, standardmäßig standardisiert, weshalb bei der Verwendung dieses Ansatzes nicht zwingend auf eine Standardisierung vor der Analyse geachtet werden muss.

Um im Anschluss an die Standardisierung der Indikatorendaten den passenden Ansatz für das vorliegende Modell identifizieren zu können, sollte nach Henseler und Fassott (2010) zunächst zwischen den Messmodellen der exogenen bzw. moderierenden Variable unterschieden werden. Deshalb folgt auch im Framework im Anschluss an die Standardisierung der Indikatoren aller Konstrukte in der MA die Definition des Messmodells der exogenen Variable und der MV. Kann die Aussage, dass M und Y_1 beide reflektiv gemessen seien, bestätigt werden, sollte dem linken Ast gefolgt werden. Dies wird durch das Häkchen im Framework visualisiert. Henseler und Chin (2010) haben in ihrer Simulationsstudie mit ausschließlich reflektiv gemessenen Konstrukten alle gängigen Ansätze zur Generierung des Interaktionsterms untersucht, um den besten Ansatz für verschiedene Ziele von Forschern bei der Durchführung von MA herausstellen zu können [4]. Infolgedessen schließt sich an die Definition des Messmodells die Bestimmung des Ziels der MA an, wenn lediglich reflektiv gemessene exogene latente Variablen und MV identifiziert worden sind. Hierbei unterscheiden Henseler und Chin (2010) zwischen den folgenden Zielen, nach denen folglich auch im Framework differenziert wird: Enthüllung der statistischen Signifikanz (Maximierung der statistischen Power), Minimierung des Schätzungsfehlers des ME (Maximierung der Punktgenauigkeit) und Maximierung der

[3] Für eine Verwendung kategorischer Variablen sei auf die Ausführungen von Henseler und Fassott (2010) verwiesen.

[4] Wie Becker et al. (2018) haben auch Henseler und Chin (2010) in ihrer Studie standardisierte Indikatoren-Daten verwendet.

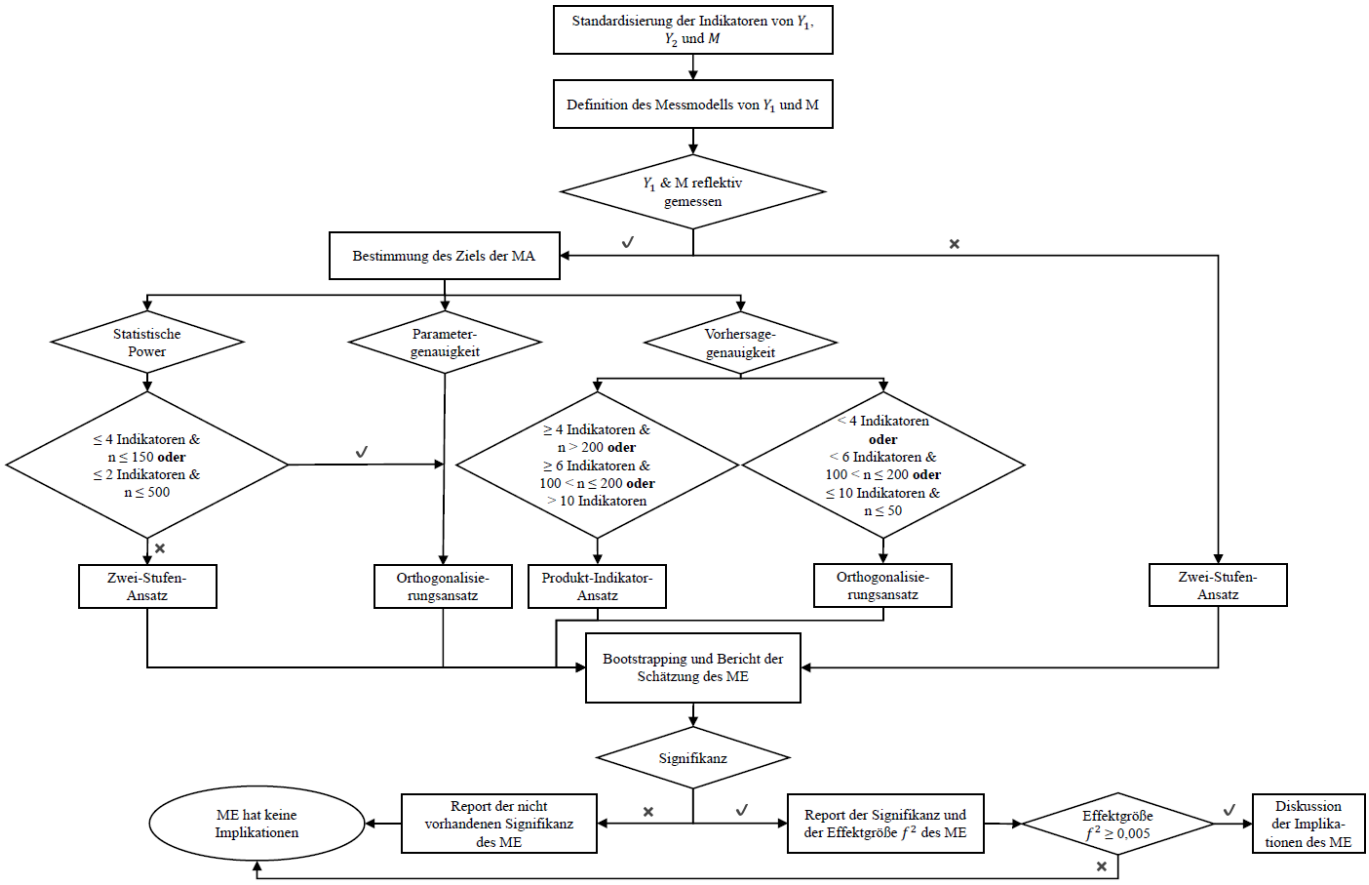


Abbildung 10: Framework zur Generierung des Interaktionsterms und der Interpretation von ME bei der Durchführung von MA in PLS-SEM (in Anlehnung an Becker et al. (2018); Henseler und Chin (2010); Henseler und Fassott (2010); Henseler et al. (2012); Kenny (2018))

Genauigkeit der Voraussage (Maximierung der Vorhersagegenauigkeit).

Bei der Betrachtung jedes Ziels wurde im Framework genauer als bspw. von Hair et al. (2017, S. 252-255) zwischen den Ansätzen differenziert, da die Ansätze innerhalb der Ziele teilweise unterschiedliche Leistungen in Abhängigkeit von Modellspezifikationen hervorbringen. Insbesondere weisen der Stichprobenumfang (n) und die jeweilige Anzahl der Indikatoren in den Konstrukten einen Einfluss auf die Leistung der Ansätze auf. So sollte nach Henseler und Chin (2010) zur Maximierung der statistischen Power bei einer gegebenen Effektgröße in der Population (= 0,136) und konstantem Signifikanzkriterium (= 0,05) bei sehr wenigen Indikatoren innerhalb der Konstrukte der Orthogonalisierungsansatz verwendet werden. Dabei dominiert dieser nach den Autoren bei sehr kleinen Stichprobenumfängen (n ≤ 150) und weniger als fünf Indikatoren in den Konstrukten die anderen Ansätze. Sollten sogar weniger als zwei Indikatoren in allen Konstrukten auftreten und der Stichprobenumfang übersteigt nicht 500, ist der Orthogonalisierungsansatz noch immer vorzuziehen. In den meisten Fällen wird nach den Autoren jedoch durch den Einsatz des Zwei-Stufen-Ansatzes die größte Power erzielt. Insofern die beschriebenen Grenzen nicht eingehalten werden und demnach eine mittlere bis große Anzahl von Indikatoren vorliegt oder der Stichprobenumfang groß genug ist, dominiert der Zwei-Stufen-Ansatz nach den Autoren demnach die anderen Ansätze.

Liegt der Fokus der Forscher jedoch auf der möglichst genauen Schätzung des ME und soll demnach die Punktgenauigkeit der Parameterschätzungen maximiert werden, bringt der Orthogonalisierungsansatz nach den Autoren stets die genausten Schätzungen hervor und sollte in diesem Fall genutzt werden.

Letztlich ist für eine möglichst präzise Vorhersage des ME nach den Autoren entweder der Produkt-Indikator-Ansatz oder der Orthogonalisierungsansatz zur Generierung des Interaktionsterms zu verwenden. Hierbei kann erneut in Abhängigkeit vom Stichprobenumfang und von den Indikatoren der Konstrukte unterschieden werden. Wie die Autoren zeigen, dominiert der Produkt-Indikator-Ansatz im Fall vieler Indikatoren in den Konstrukten (> 10) bzw. auch bei weniger Indikatoren und einem dafür größeren Stichprobenumfang (bspw. ≥ 4 Indikatoren und n > 200). Der Orthogonalisierungsansatz hingegen erzielt nach den Autoren genauere Vorhersagen des ME bei wenigen Indikatoren der Konstrukte (< 4) oder bei einem kleinen bis mittelgroßen Stichproben-

umfang und einer mittleren Anzahl von Indikatoren (bspw. ≤ 10 Indikatoren und $n \leq 50$).

Wird mindestens eins der Konstrukte formativ gemessen und kann die Aussage, dass M und Y_1 reflektiv gemessen seien, demnach nicht bestätigt werden, sollte dem rechten Pfad nach der Definition der Messmodelle im Framework gefolgt werden. Dies ist anhand des Kreuzes oberhalb des rechten Pfades in Abbildung 10 erkennbar. Hieraus kann geschlussfolgert werden, dass bei der formativen Messung der latenten MV oder der latenten exogenen Variable stets der Zwei-Stufen-Ansatz verwendet werden sollte. Nur mit diesem Ansatz ist die inhaltliche Korrektheit des Interaktionsterms bei der Multiplikation formativer Konstrukte, denen keine eindeutige inhaltliche Domäne zugrunde liegt, gewährleistet und der Interaktionsterm konzeptionell abbildbar (Henseler & Fassott, 2010). Darüber hinaus haben auch Henseler et al. (2012) in ihrer Simulationsstudie die Überlegenheit des Zwei-Stufen-Ansatzes über dem Produkt-Indikator-Ansatz und dem Orthogonalisierungsansatz bei MA mit formativ gemessenen Konstrukten im Hinblick auf die wie bereits von Henseler und Chin (2010) getesteten Aspekte festgestellt[5].

In jedem Fall sollte nach Henseler et al. (2012) die Prüfung auf Signifikanz des ME mithilfe des Bootstrapping-Prozederes erfolgen, wobei neben der Signifikanz auch die Schätzung der Höhe des Pfadkoeffizienten p_3 (d.h. Parameterschätzung des Interaktionseffektes) angegeben werden sollte. Im Anschluss an die Signifikanzprüfung sollte nach den Autoren die Größe des ME (f^2) bestimmt und berichtet werden, insofern der ME signifikant ist. Im Anschluss ist die Effektgröße nach den Autoren darauf zu kontrollieren, dass diese mindestens als klein angesehen werden kann. Hierfür wird im Framework die Untergrenze von Kenny (2018) herangezogen, ab welcher kleine Effekte, die nach Chin et al. (2003) ebenfalls bedeutsam sein können und Implikationen nach sich ziehen, mindestens 0,005 betragen müssen. Bei der Diskussion der Implikationen ist die theoretische Grundlage der Hypothese des ME erneut aufzugreifen (Andersson et al., 2014). Liegt kein signifikanter ME vor, sollte dieser dennoch berichtet werden (Henseler et al., 2012). Wie auch in dem Fall, dass der Effekt zwar signifikant, aber nicht mindestens als klein in seiner Größe einzustufen ist, zieht der ME dabei keine Implikationen nach sich.

3. Methode

Um die gestellte Forschungsfrage im Rahmen eines Reviews relevanter Untersuchungen in Forschungsartikeln beantworten zu können, thematisiert dieser Abschnitt die Durchführung der systematischen Literaturrecherche. Zunächst wird die Festlegung relevanter Inklusionskriterien beschrieben. Nachfolgend wird der Ablauf des Rechercheprozesses, dessen Ergebnisse und die Vorgehensweise bei der Kodierung der gefundenen Artikel dargelegt.

3.1. Inklusionskriterien

Vor der Literaturrecherche wurden Inklusionskriterien festgelegt, die jede Untersuchung erfüllen musste, um ein Bestandteil des Reviews werden zu können. Ein wichtiges Einschlusskriterium bei der Auswahl zu verwendender Untersuchungen war deren Analyse mindestens eines ME. Insbesondere muss dabei ein Interaktionsterm gebildet worden sein. Diese stellen den Fokus der vorliegenden Arbeit dar, weshalb MA mithilfe von Gruppenvergleichen (d.h. Multigruppenanalysen) ausgeschlossen wurden. Weiterhin musste zu erkennen sein, welcher Ansatz zur Berechnung von ME verwendet wurde, da anderenfalls keine klare Einschätzung der berechneten Modelle in den Untersuchungen mithilfe des Frameworks und der Auswirkungen der Wahl eines Ansatzes auf die Ergebnisse und Interpretation dieser möglich gewesen wäre. Aus diesem Grund wäre außerdem die Vergleichbarkeit mit den restlichen Artikeln eingeschränkt. Ist demnach in einer Publikation der Ansatz nicht klar zu erkennen, liefert sie keinen Mehrwert für das Review und kann nicht eingeschlossen werden.

Es wurden weiterhin nur empirische Untersuchungen eingeschlossen, die Ergebnisse durch die Verwendung des PLS-Algorithmus generiert haben, um sicherzustellen, dass die Bildung des Interaktionsterms den zuvor erläuterten Prinzipien folgt. Dies ermöglicht die Vergleichbarkeit der inkludierten Artikel, da alle auf demselben Algorithmus zur Berechnung des Pfadmodells basieren. Insbesondere müssen die Untersuchungen dabei aus Artikeln stammen, welche in einer vom Lehrstuhl für Marketing der Otto-von-Guericke-Universität Magdeburg zur Verfügung gestellten Datenbank enthalten sind. Diese setzt sich aus 239 empirischen Forschungsartikeln zusammen, die aus einem Suchprozess stammen, welcher die Vorgehensweise im Review von Hair, Sarstedt, Ringle und Mena (2012) für die Jahre 2011 bis 2020 repliziert. Hierbei wurden die 30 Top-Marketingjournale nach Hult, Reimann und Schilke (2009) auf die Schlagwörter „PLS“ und „Partial Least Squares“ in den Datenbanken EBSCO, Scopus und Web of Science sowie deren Websites in einer Volltextsuche durchsucht. Es folgten Filtermaßnahmen, die denen von Hair et al. (2012) nachempfunden wurden und aus denen 239 empirische Marketinganwendungen von PLS-SEM hervorgingen.

Gegenstand des vorliegenden Reviews konnten somit nur Artikel werden, welche empirische Marketinganwendungen von PLS-SEM enthalten und innerhalb der vergangenen zehn Jahre in Top-Marketingjournals veröffentlicht wurden. Dabei musste mindestens eine MA in einer Untersuchung enthalten sein, die mithilfe der Bildung eines Interaktionsterms durchgeführt wurde und klar erkennen lässt, mit welchem Ansatz die Berechnung des ME erfolgt ist.

[5] Die Autoren weisen darauf hin, dass der ebenfalls bei formativen Konstrukten anwendbare hybride Ansatz den Zwei-Stufen-Ansatz in allen Kategorien leistungsmäßig übertrifft, dieser aufgrund fehlender Software-Implementierung aber nur wenigen Forschern zur Verfügung steht. Sollte er jedoch verfügbar sein, empfehlen die Autoren, den hybriden Ansatz zur Generierung des Interaktionsterms bei formativen Konstrukten zu verwenden.

3.2. Systematische Literaturrecherche

Um die Marketinganwendungen von PLS-SEM identifizieren zu können, welche Moderationen mithilfe von Interaktionstermen analysieren, wurde die aus 239 Publikationen bestehende zur Verfügung gestellte Datenbank zunächst mithilfe der Schlagwörter „moderation“, „moderator“, „moderate“, „moderating“, „interaction“, „interacting“ und „interaction effect“ durchsucht. Aus der Volltextsuche ergaben sich 109 Artikel, die eine beliebige Form der MA enthielten. Diese wurden anschließend darauf geprüft, ob ein Interaktionsterm zur Analyse der ME gebildet wurde. Insbesondere mussten deshalb zunächst 27 Artikel ausgeschlossen werden, die ME mithilfe von Multigruppenanalysen untersuchen. Weiterhin wurden neun Artikel ausgeschlossen, welche MA mit dem PROCESS-Makro von Hayes (2013) im Rahmen von Regressionsanalysen oder mit einer ANOVA durchführen. Demnach resultierten nach der ersten Filterung der Artikel mit MA insgesamt 73 Artikel, die einen Interaktionsterm zur Analyse der Moderatoren gebildet haben.

Von diesen mussten jedoch weitere 40 Publikationen vom Review ausgeschlossen werden, da sie keine Informationen zum verwendeten Ansatz bei der MA angaben. Neben den bereits diskutierten Problemen hierdurch bei der Vergleichbarkeit mit anderen Studien, ermöglicht diese intransparente Vorgehensweise einerseits keine Einschätzung der inhaltlichen Korrektheit der Bildung des Interaktionsterms im Fall formativ gemessener Konstrukte, da nicht nachvollzogen werden kann, ob der einzige inhaltlich korrekte Zwei-Stufen-Ansatz hierbei verwendet wurde. Andererseits ist die Interpretation des Main Effects und Simple Effects nicht gänzlich nachvollziehbar, da unklar ist, ob die Effekte wie bei der Nutzung des Orthogonalisierungsansatzes als äquivalent anzusehen sind oder nach Hair et al. (2017, S. 257-258) einer getrennten Betrachtung bei der Interpretation der MA benötigen. Aus diesem Grund ist die Angabe des Ansatzes, welcher zur Bildung des Interaktionsterms herangezogen wurde, von großer Bedeutung. Anderenfalls sind die potenziell falsch analysierten und folglich fehlinterpretierten Ergebnisse der MA nicht mit den Ergebnissen anderer Artikel vergleichbar und darüber hinaus müssen deshalb die geschlussfolgerten Implikationen für die Forschung oder das Management angezweifelt werden. Die nach der letzten Filterung verbleibenden 33 Artikel, die das Ergebnis der systematischen Literaturrecherche darstellen, wurden anschließend nach den verwendeten Ansätzen zur Generierung des Interaktionsterms sortiert.

3.3. Kodierung der Moderationseffekte

In den ausgewählten 33 wissenschaftlichen Arbeiten, die in das Review zu Moderation in Marketinganwendungen von PLS-SEM aufgenommen wurden, konnten 35 Untersuchungen gefunden werden, die MA mithilfe von Interaktionstermen durchführen. Diese 35 Untersuchungen wurden im nächstens Schritt in einer Excel-Tabelle kodiert, wodurch eine Vielzahl an Datensätzen entstand, deren Analyse die Grundlage des Reviews und damit der Beantwortung der Forschungsfrage bildet (siehe den externen Anhang A1 für eine Auflistung der final inkludierten Artikel).

Das verwendete Kodierschema wurde im Hinblick auf das Ziel der vorliegenden Arbeit und den daraus folgenden Bestandteilen des anzufertigenden Reviews erstellt. Dabei fand zunächst eine Probekodierung von fünf Artikeln mit insgesamt sechs Untersuchungen statt, die Aufschluss darüber gaben, wie das Kodierschema strukturiert werden sollte und welche Eigenschaften der Artikel und der enthaltenen MA im Kodierschema bisher fehlten. Im Anschluss an die Festlegung der Struktur des Schemas und der Ergänzung fehlender Merkmale schloss sich die erneute Kodierung dieser fünf Artikel sowie die finale Kodierung der restlichen Untersuchungen an.

Die Struktur des Kodierschemas folgt dabei grob den Vorgaben zu Kodierschemata in Metaanalysen von M. (2014, S. 31-32). Der Autor verweist darauf, mehrere Hierarchiestufen im Kodierschema zu verwenden, um zwischen der übergeordneten Ebene der Artikel und untergeordneten Ebene der Untersuchungen unterscheiden zu können. Dies wurde in dem verwendeten Schema umgesetzt, indem ein Artikel mehrere Datensätze (gruppierte Zeilen) erhalten hat, insofern er mehrere Untersuchungen enthielt. Eine Untersuchung wiederrum konnte ebenfalls mehrere Zeilen umfassen, falls sie mehr als einen ME analysiert hat. Somit konnte zwischen Artikeln, Untersuchungen und ME differenziert werden, wobei letztere schließlich jeweils eine Zeile in einer Untersuchung erhalten haben.

Jeder Datensatz wurde schließlich im Hinblick auf drei wesentliche inhaltliche Abschnitte kodiert: Allgemeine Eigenschaften des Artikels, Merkmale des einzelnen ME und Hinweise zu den MA in der Untersuchung. Im finalen Kodierprozess wurden dementsprechend im ersten Abschnitt der Tabelle für jeden Artikel u.a. allgemeine Eigenschaften, wie die Autoren, Erscheinungsjahr und das Journal kodiert. Hinzukamen außerdem die Anzahl der kodierten ME und der Stichprobenumfang der betrachteten Untersuchung. Die wichtigsten Merkmale der einzelnen ME im zweiten Abschnitt wurden für jeden ME getrennt erfasst, sodass diese einzeln im Hinblick auf das Framework analysiert werden konnten. Hierfür wurden für jeden ME insbesondere die Skalierung, die Anzahl der Indikatoren sowie das Messmodell der Moderatorvariable und der exogenen und endogenen Variable erfasst. Hinzukamen der verwendete Ansatz zur Bildung des Interaktionsterms und die Datenbehandlung vor der Analyse dieses ME (Standardisierung vs. Mittelwertzentrierung vs. Verwendung nicht standardisierter Daten). Neben einer Einschätzung, ob die Analyse des ME auf Basis der vorliegenden Daten das Framework eingehalten oder gegen es in mindestens einem Aspekt verstoßen hat, wurden außerdem die Effektgröße f^2, das Ziel der MA und inwiefern das Messmodell der Moderatorvariable evaluiert wurde erfasst. Schließlich wurde im zweiten Abschnitt ebenfalls kodiert, ob der ME im Rahmen einer a priori aufgestellten Hypothese getestet wurde. Im letzten Abschnitt der Kodierung einer Untersuchung wurden Hinweise zu allen MA in einer Untersuchung aufgenommen, wobei festgehalten wur-

de, ob und in welcher Form nach Hayes (2013) eine MA im Zusammenhang mit einer Mediationsanalyse auftritt und inwieweit HCM Bestandteil einer MA sind. Schließlich wurde erfasst, ob der Main und Simple Effect bei der Interpretation der ME fälschlicherweise verwechselt wurden, wie bei der Analyse mehrerer MV vorgegangen wurde (gleichzeitige vs. separate Betrachtung der ME) und ob einer der Übersichtsartikel von Henseler und Chin (2010), Henseler und Fassott (2010), Hair et al. (2017) oder Becker et al. (2018) zitiert wurde.

Schließlich konnten jeder Untersuchung maximal 16 Merkmale im ersten, 53 Merkmale pro inkludiertem ME im zweiten und 24 Merkmale im dritten Abschnitt des Kodierschemas zugeordnet werden, wobei fehlende Werte zumeist mit „NA“ ausgefüllt wurden. Zusätzlich wurden Kommentare zu jedem ME verfasst, welche die Auswertung der insg. 183 kodierten ME erleichtern sollten. Um die Güte der Kodierung sicherstellen zu können, wurden die Ergebnisse mehrfach auf ihre Korrektheit geprüft.

4. Review von Moderationsanalysen in Marketinganwendungen von PLS-SEM

Die Frage, wie Forscher mit den Herausforderungen bei der Durchführung von einfachen MA in PLS-SEM umgehen, soll in diesem Abschnitt beantwortet werden. Hierfür wurden 33 relevante wissenschaftliche Publikationen auf den Umgang mit MA in PLS-SEM im Rahmen eines Reviews untersucht. Zunächst werden allgemeine Beobachtungen aus der Literaturrecherche dargestellt, nach denen der Umgang von Forschern mit den zuvor identifizierten potenziellen Fehlerquellen bei MA, in der u.a. auch Einhaltungen oder Verstöße gegen das zuvor aufgestellte Framework analysiert werden, folgt. Schließlich folgt die Diskussion sonstiger Beobachtungen aus dem Reviewprozess, bei denen weitere Probleme und Fehlerquellen identifiziert wurden. Summa summarum dient das Review zur Identifikation noch unbekannter Fehler- und Problemquellen sowie der Überprüfung bestehender Unsicherheiten der Forschung im Umgang mit MA für den PLS-SEM-Kontext.

4.1. Beobachtungen während der Literaturrecherche

Bereits bevor die ersten Beobachtungen im Rahmen des Reviews gemacht wurden, brachte die Literaturrecherche einige Erkenntnisse hervor , welche in diesem Abschnitt aufgegriffen werden. Die erste Beobachtung aus der Literaturrecherche bezieht sich auf das Verhältnis der 109 Artikel, die eine beliebige Form einer MA (sowohl einfache MA als auch MA mithilfe von Gruppenvergleichen) durchgeführt haben, zu den insg. 239 Artikeln aus der Datenbank, die keine MA enthielten. Fast die Hälfte (45,61%) der 239 Artikel in der Datenbank enthalten eine Art der Moderation, was deren Relevanz für die aktuelle Forschung hervorhebt. Bereits im Jahr 2011 haben Henseler et al. (2012) die Aussage getroffen, dass sich das Interesse von Forschern in Richtung nichtlinearer Effekte, wie Interaktionseffekte, verschoben hat, da lineare Effekte zwischen Konstrukten in vielen Fällen bereits identifiziert wurden. Diese Aussage wird von den vorliegenden Daten gestützt.

Darüber hinaus werden von diesen 109 Artikeln in insg. 73 Publikationen Interaktionsterme gebildet (66,97%), wobei demnach in 36 Fällen (33%) eine Multigruppenanalyse, Regressionsanalyse oder ANOVA durchgeführt wurde, um ME zu analysieren. Henseler und Chin (2010) weisen darauf hin, dass Multigruppenanalysen in PLS-SEM vorrangig bei kategorischen MV anzuwenden sind. Einfache MA hingegen können nach den Autoren sowohl mit kategorischen als auch mit kontinuierlichen MV bessere Ergebnisse als Multigruppenanalysen bei der Schätzung von ME in PLS-SEM erzielen. Aus diesen Gründen ist es naheliegend, dass auch in den meisten Fällen (66,97%) Interaktionsterme gebildet und damit die einfache MA vermehrt einer anderen Form der MA in PLS-SEM vorgezogen wurde.

In der Literaturrecherche wurden weiterhin insg. 40 Publikationen herausgefiltert, die zwar eine einfache MA mit der Bildung eines Interaktionsterms durchgeführt, aber nicht angegeben haben, mit welchem Ansatz der Interaktionsterm generiert wurde. Eine Begründung für diesen Ausschluss wurde bereits dargelegt. Zwischen 2011 und 2020 wurden somit 73 Artikel im PLS-SEM-Kontext veröffentlicht, von denen über die Hälfte (54,79%) keine Informationen enthalten, wie der Interaktionsterm gebildet wurde. Abbildung 11 zeigt die Verteilung der Artikel mit und ohne entsprechende Informationen.

Letztlich stellt das finale Ergebnis der Literaturrecherche eine weitere aufschlussreiche Beobachtung dar. Hierbei ist die Verteilung der Ansätze zur Generierung des Interaktionsterms so ausgeprägt, dass der Produkt-Indikator-Ansatz in 22 Artikeln vorkommt und dadurch mit 64,71% den überwiegenden Anteil der verwendeten Ansätze darstellt[6]. Dieser ist nach dem erstellten Framework der am seltensten optimale Ansatz, da er nur die beste Wahl für Forscher darstellen sollte, wenn diese die Vorhersagegenauigkeit des ME maximieren wollen und sich bspw. viele Indikatoren (> 10) in den verwendeten Konstrukten befinden. Es wird jedoch davon ausgegangen, dass das Finden und Testen signifikanter Effekte den eigentlichen Fokus der meisten PLS-SEM-Studien darstellt und die Vorhersage von Effekten ein recht spezielles Ziel ist, welches in der Praxis nur selten beobachtet wird. Dabei wird angenommen, dass das Vorhandensein von Hypothesen in einem Artikel darauf hindeutet, dass diese getestet werden sollen und dass Forscher hierfür den Ansatz mit der höchsten statistischen Power nutzen möchten.

In diesem Sinne ist das häufige Auftreten des Produkt-Indikator-Ansatzes kontraintuitiv und deutet bereits auf viele Verstöße gegen das aufgestellte Framework hin, die im kommenden Abschnitt genauer betrachtet werden. Dieser Ansatz

[6] Da in einem Artikel (Johnston, Khalil, Jain & Cheng, 2012) sowohl der Produkt-Indikator-Ansatz als auch der Zwei-Stufen-Ansatz verwendet wurden, wurde dieser beiden Kategorien zugeordnet und die Summe der Häufigkeiten der drei gefundenen Ansätze ergibt insg. 34, obwohl das Review nur 33 Artikel umfasst.

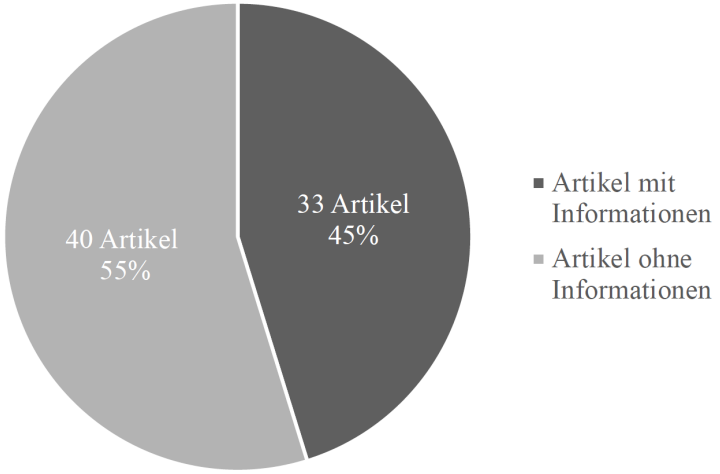


Abbildung 11: Artikel mit und ohne Informationen darüber, welcher Ansatz verwendet wurde

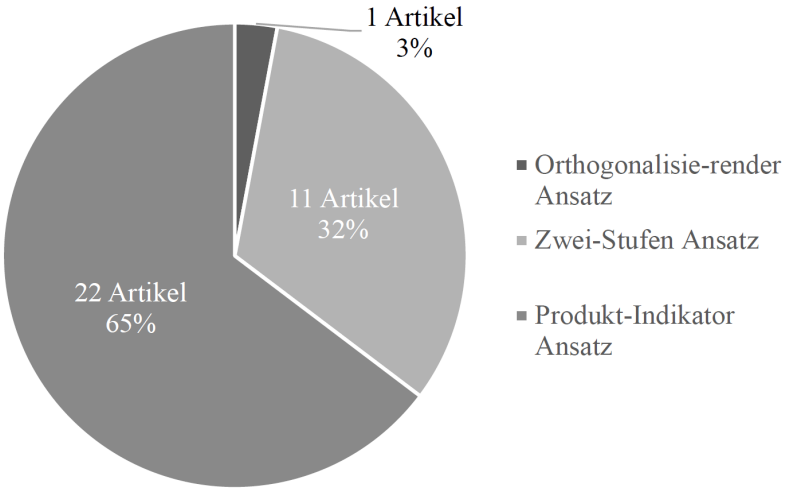


Abbildung 12: Verteilung der verwendeten Ansätze innerhalb der Artikel im Review

hätte erwartungsgemäß seltener gewählt werden müssen, wenn man davon ausgeht, dass Forscher nach den gängigen Vorgaben von Henseler und Chin (2010) den besten Ansatz für ihre MA auswählen würden. Wie im Framework gezeigt, bringen meist der Zwei-Stufen-Ansatz und ggf. der Orthogonalisierungsansatz in diesem Fall die größte statistische Power hervor, die jedoch mit elf (32,35%) bzw. einem Vorkommen (2,94%) innerhalb der Ergebnisse der Literaturrecherche die Minderheit der verwendeten Ansätze ausmachen. Die häufige Wahl des Produkt-Indikator-Ansatzes kann sich u.a. darauf zurückführen lassen, dass dieser den Prinzipien von MA im Regressionskontext entspringt, mit welchem multivariate Analysen insb. manifester Variablen i.d.R. durchgeführt werden (Hair et al., 2017, S. 2-3; Henseler & Fassott, 2010). Da sich die Vorgehensweise bei der Analyse von ME in Regressionsanalysen mithilfe des Produkt-Indikator-Ansatzes leicht auf PLS-SEM übertragen lassen, legt dies die Vermutung nahe, dass der Ansatz beliebt bei Forschern ist, die zuvor im Regressionskontext ME analysiert haben. Hierbei werden infolgedessen best-practice Vorgaben zur Wahl des besten Ansatzes unter den Gegebenheiten der Messmodelle, Indikatoren in den Konstrukten und des Stichprobenumfangs nicht berücksichtigt.

Diese Erklärung wird ebenfalls von den Daten gestützt, die während des Reviews aus den 35 inkludierten Untersuchungen erfasst wurden. Hierbei wird in lediglich zehn der 35 Untersuchungen (29,41%) darauf hingewiesen, dass sich an den Empfehlungen zur Generierung von Interaktionstermen in PLS-SEM aus den Übersichtsartikeln von Henseler und Chin (2010), Henseler und Fassott (2010), Hair et al. (2017) oder Becker et al. (2018) orientiert wurde. Somit weist die Mehrheit der Autoren der inkludierten Artikel (70,59%) keine Kenntnis der Vor- und Nachteile der einzelnen Ansätze zur Generierung des Interaktionsterms nach, was die unreflektierte Nutzung der Ansätze durch die Autoren erklären könnte.

4.2. Analyse der potenziellen Fehlerquellen bei Moderationsanalysen

Die in Abschnitt 2.3.5.1 herausgestellten potenziellen Fehlerquellen bei MA in PLS-SEM werden nachfolgend analysiert, um Rückschlüsse auf den aktuellen Stand der Forschung bezüglich des Umgangs mit diesen ziehen zu können. Dabei werden die Fehlerquellen inhaltlich gruppiert, sodass sich die Gebiete „Generierung des Interaktionsterms und Datenbehandlung vor der Bildung des Interaktionsterms“, „Interpretation des Moderationseffektes und Einschätzung der Effektgröße f^2“ und „Künstliche Dichotomisierung kontinuierlicher Moderatorvariablen“ differenzieren lassen. Letzteres Gebiet besteht dabei nur aus der namensgebenden potenziellen Fehlerquelle, die keinem anderen Gebiet zugeordnet werden konnte.

4.2.1. Generierung des Interaktionsterms und Datenbehandlung vor der Analyse

Die Bildung des Interaktionsterms geht mit insg. vier potenziellen Fehlerquellen einher, welche den Einsatz falscher Ansätze bei formativ gemessenen Konstrukten, den Einsatz eines abhängig vom Ziel der Forscher suboptimalen Ansatzes bei einer gegebenen Anzahl von Indikatoren in reflektiv gemessenen Konstrukten und dem Stichprobenumfang sowie die falsche Einschätzung des Ziels der Forscher bei der MA umfassen. Hinzukommt die Datenbehandlung vor der Generierung des Interaktionsterms. Dabei werden die genannten Problembereiche u.a. im Framework aufgegriffen, das Empfehlungen zum Umgang mit den Fehlerquellen beinhaltet und demnach dazu genutzt werden kann, eine Einschätzung bzgl. des Umgangs der Forschung mit der Generierung des Interaktionsterms und auch mit der vorherigen Datenbehandlung zu ermöglichen. Hierbei werden jedoch nur die Untersuchungen evaluiert, welche die Voraussetzungen zur Verwendung des Frameworks erfüllen. Aus diesem Grund werden in den Betrachtungen dieses Abschnitts lediglich 23 Publikationen mit insg. 25 Untersuchungen einbezogen, da sieben Artikel bei der Untersuchung aller ME ein HCM untersuchen und deshalb ausgeschlossen werden mussten. Dies bedeutet gleichzeitig, dass Artikel zumindest teilweise mithilfe des Frameworks analysiert werden können, bei denen nicht jede MA mit einem HCM zusammenhängt, was bspw. bei dem Artikel von Kumar, Singh, Pereira und Leonidou (2020) der Fall ist. Der Nichteinsatz von metrischen Variablen hat darüber hinaus zum Ausschluss drei weiterer Artikel geführt, die lediglich Effekte im Zusammenhang mit kategorischen exogenen oder moderierenden Variablen beinhalten[7].

Insgesamt werden in 20 Untersuchungen Modelle bestehend aus lediglich reflektiv gemessenen Konstrukten analysiert. Dabei musste jedoch berücksichtigt werden, dass in einer Untersuchung verschiedene Konstellationen von MA vorkommen können, die sowohl Modelle aus ausschließlich reflektiv gemessenen Konstrukten als auch Modelle mit mindestens einem formativ gemessenem exogenen oder moderierenden Konstrukt beinhalten. Demzufolge kann eine Untersuchung zur Menge von Untersuchungen mit ausschließlich reflektiv gemessenen Konstrukten, aber auch zur Menge mit mindestens einer entsprechenden formativen Messung gehören. Inklusive der Modelle mit zumindest teilweise formativ gemessenen Konstrukten (7 mal) ergeben sich durch die doppelte Betrachtung zweier Untersuchungen, die beide Modelle enthalten[8], somit 27 individuelle zu analysierende Fälle. Diese werden nachfolgend auf die ersten vier potenziellen Fehlerquellen überprüft, wobei die Herausforderungen bei der Generierung des Interaktionsterms, deren Vorkommen und passende Beispiele in Tabelle 3 zusammengefasst sind.

Zunächst wird mit der Verwendung eines falschen Ansatzes bei formativ gemessenen exogenen oder moderierenden Konstrukten der schwerwiegendste Fehler untersucht, der bei der Generierung des Interaktionsterms auftreten kann und demnach einen Verstoß gegen das Framework darstellt. Hierbei geht die inhaltliche Bedeutung des Interaktionsterms verloren, wenn dieser mithilfe des Orthogonalisierungsansatzes oder des Produkt-Indikator-Ansatzes gebildet wird, woraufhin die gesamte MA ihre Aussagekraft verliert (Henseler & Fassott, 2010). Insgesamt weisen Untersuchungen in Mahr et al. (2014) und Rollins et al. (2012) sowie eine MA in Matsuno et al. (2014) einen Verstoß gegen das Framework auf. Die erstgenannten Autoren nutzen formativ gemessene MV und die zweit- sowie drittgenannten Autoren verwenden eine formativ gemessene exogene Variable, um den Interaktionsterm fälschlicherweise mithilfe des Produkt-Indikator-Ansatzes zu generieren. Prozentual bildet nur ein kleiner Teil mit 11,11 Prozent aller 27 mit dem Framework analysierbaren Fälle den Interaktionsterm konzeptionell falsch. Insgesamt wird jedoch lediglich in sieben der 27 Fälle (25,93%) mindestens ein formatives Konstrukt im Zusammenhang mit einer MA genutzt. Demnach ist auffällig, dass mit drei von sieben Publikationen insg. 42,86% der Artikel den Interaktionsterm aus mindestens einem formativen Konstrukt nicht korrekt mit dem Zwei-Stufen-Ansatz bildet. Hierbei wird in keiner der gegen das Framework verstoßenden Untersuchungen auf einen der Übersichtsartikel von Henseler und Chin (2010), Henseler und Fassott (2010), Hair et al. (2017) oder Becker et al. (2018) hingewiesen. Dies legt die Vermutung nahe, dass die Verstöße mit einem allgemeinen Unwissen von Forschern über die Einschränkungen des Produkt-Indikator-Ansatzes und Orthogonalisierungsansatzes begründet werden könnten.

Darüber hinaus besteht auch bei der Wahl des Ansatzes in Modellen mit ausschließlich reflektiv gemessenen Konstrukten Unsicherheit in der aktuellen Forschung

[7] Die ME, die innerhalb einer moderierten Mediation auftraten, wurden nicht im Kontext der Mediation betrachtet, sondern als herkömmliche MA angesehen. Diese Vorgehensweise diente der Vereinfachung des Reviewprozesses und hat dennoch allgemeingültige Erkenntnisse des Umgangs von Forschern mit MA hervorgebracht.

[8] Die Artikel von Matsuno et al. (2014) und Johnston et al. (2012) weisen sowohl Modelle mit ausschließlich reflektiv gemessenen Konstrukten als auch Modelle mit mind. einem formativ gemessenem Konstrukt auf, wobei letzterer Artikel darüber hinaus zwei verschiedene Ansätze (Produkt-Indikator-Ansatz und Zwei-Stufen-Ansatz) in den MA verwendet.

Tabelle 3: Zusammenfassung der Beobachtungen zu bekannten Fehlerquellen bei der Generierung des Interaktionsterms, deren Vorkommen und Beispiele

Bekannte Fehlerquelle	Spezifizierung der Beobachtung	Vorkommen (absolut)	Vorkommen (relativ)	Beispiele
kein Zwei-Stufen-Ansatz (nur Untersuchungen mit mindestens einer formativen Messung)	-	3 von 27 Untersuchungen	11,11% der Untersuchungen	Mahr, Lievens und Blazevic (2014); Matsuno, Zhu und Rice (2014); Rollins, Bellenger und Johnston (2012)
	7 formative Messungen von Konstrukten in 27 Untersuchungen	3 von 7 Untersuchungen	42,86% der Untersuchungen	
Wahl eines suboptimalen Ansatzes nach dem Framework (nur reflektive Messungen)	Ziel der MA: 16 mal Herausstellen der Signifikanz und 4 mal Vorhersage von Effekten	14 von 20 Untersuchungen	70% der Untersuchungen	Chen, Peng und Hung (2015); Coelho und Henseler (2012) (beide Untersuchungen)
	Ziel der MA korrigiert: 20 mal Herausstellen der Signifikanz	16 von 20 Untersuchungen	80% der Untersuchungen	Kumar et al. (2020); Teller, Gittenberger und Schnedlitz (2013)
Zitierung eines Übersichtsartikels (nur reflektive Messungen)	gegen das Framework verstoßende Fälle; unveränderte Ziele	3 von 14 Untersuchungen	21,43% der Untersuchungen	Chen et al. (2015); Johnston et al. (2012); Wolter und Cronin (2016)
	gegen das Framework verstoßende Fälle; korrigierte Ziele	4 von 16 Untersuchungen	25% der Untersuchungen	Teller et al. (2013); Wolter und Cronin (2016)
Wahl eines suboptimalen Ansatzes nach dem Framework	Ziel der MA: 23 mal Herausstellen der Signifikanz und 4 mal Vorhersage von Effekten	17 von 27 Untersuchungen	62,96% der Untersuchungen	Matsuno et al. (2014); Mullins, Agnihotri und Hall (2020)
	Ziel der MA korrigiert: 27 mal Herausstellen der Signifikanz	19 von 27 Untersuchungen	70,37% der Untersuchungen	Mahr et al. (2014); Kumar et al. (2020)
Zitierung eines Übersichtsartikels	gegen das Framework verstoßende Fälle; unveränderte Ziele	3 von 17 Untersuchungen	17,65% der Untersuchungen	Chen et al. (2015); Johnston et al. (2012); Wolter und Cronin (2016)
	gegen das Framework verstoßende Fälle; korrigierte Ziele	4 von 19 Untersuchungen	21,05% der Untersuchungen	Johnston et al. (2012); Teller et al. (2013)

(Henseler & Chin, 2010). Es können hierbei prinzipiell alle gängigen Ansätze genutzt werden, um den Interaktionsterm zu bilden. In diesem Fall sollten Forscher jedoch stets den Ansatz wählen, der nach Henseler und Chin (2010) die beste Leistung im Hinblick auf das Ziel der MA bei der Generierung des Interaktionsterms hervorbringt. Das Ziel stellt dabei erfahrungsgemäß nur in Ausnahmefällen die Maximierung der Vorhersagegenauigkeit oder der Parametergenauigkeit dar. Aus diesem Grund wurde im Review bei der Identifikation des Ziels von Forschern bei MA konservativ vorgegangen und das Ziel im Zweifel stets als Maximierung der statistischen Power kodiert. Nur wenn eindeutig von Autoren herausgestellt wurde, dass der Fokus eines Artikels die Vorhersage oder Parametergenauigkeit umfasst, wurde nicht die Maximierung der Power als Ziel der MA anerkannt. Dabei zeigt sich bei der Analyse der 27 Fälle, die mit dem Framework evaluiert werden können, dass bei der Ausprägung der Ziele eindeutig das Testen von Effekten und damit die Maximierung der statistischen Power dominiert. So wird sich in insg. 23 Fällen (85,19%) auf das Testen der Signifikanz von Effekten und gleichzeitig in nur vier Fällen (14,81%) auf die Vorhersage von Effekten fokussiert.

Von diesen 27 Untersuchungen treten wie bereits erwähnt in 20 Untersuchungen Modelle mit ausschließlich reflektiv gemessenen Konstrukten auf. Betrachtet man nun die Wahl des Ziels der Forscher und zieht die Indikatoren der Konstrukte[9] innerhalb der MA sowie die Stichprobenumfänge heran, lässt sich erkennen, dass in 14 der 20 Untersuchungen (70%) mit ausschließlich reflektiv gemessenen Konstrukten der nach dem aufgestellten Framework suboptimale Ansatz verwendet wurde. Diese niedrige Anzahl der Verwendungen des besten Ansatzes hängt, wie bereits im

[9] In dem Fall, dass die Indikatoren in den Konstrukten unterschiedlich ausgeprägt sind und eine damit zusammenhängende im Framework abgebildete Grenze nicht eindeutig von allen Indikatoren unter- oder überschritten wird, wurde der Mittelwert aller Indikatoren gebildet und anhand dessen beurteilt, ob die Indikatorenanzahl in den Konstrukten die Verwendung eines Ansatzes begünstigt oder nicht.

Vorfeld antizipiert, vorrangig mit dem häufigen Einsatz des Produkt-Indikator-Ansatzes zusammen (15 mal), da das Ziel von Forschern meistens die Maximierung der statistischen Power darstellt (16 mal) und hierbei dieser Ansatz stets suboptimal ist. Insbesondere kann die Wahl des Produkt-Indikator-Ansatzes mit insg. 13 der 14 Verstöße in Verbindung gebracht werden (92,86%), während die Wahl des Orthogonalisierungsansatzes ein Mal gegen das Framework verstößt. Verglichen hiermit wird nur selten beobachtet, d.h. bei drei der 14 Verstöße (21,43%), dass die im Framework aufgestellten Grenzen bzgl. der Indikatoren pro Konstrukt oder des Stichprobenumfangs verletzt und damit ein anderer Ansatz als der im Artikel genutzte der beste Ansatz gewesen wäre. Demnach ist die hohe Anzahl der Verstöße gegen das Framework mit der häufigen Anzahl der Verwendungen des Produkt-Indikator-Ansatzes bei dem Ziel der Maximierung statistischer Power zu erklären. Darüber hinaus hält jeder der vier Einsätze des Zwei-Stufen-Ansatzes und auch der zwei des Produkt-Indikator-Ansatzes das Framework ein. Letzteres ist darauf zurückzuführen, dass in insg. vier Untersuchungen die Maximierung der Vorhersagegenauigkeit das Ziel der MA darstellt, wobei der Produkt-Indikator Ansatz in zwei Fällen den besten Ansatz darstellt und das Framework einhält.

Bei der Suche möglicher Gründe für die häufige Wahl eines suboptimalen Ansatzes fiel auf, dass nur in sechs der 20 Publikationen (30%) die genannten Übersichtsartikel zitiert wurden, wobei die Mehrheit der Publikationen (70%) folglich keinen Verweis der verschiedenen Vor- und Nachteile der Ansätze enthält. Hierbei wurde in drei der sechs Publikationen (50%), die einen der Übersichtsartikel zitieren, der Ansatz nicht korrekt gewählt. Dies ist bspw. bei Wolter und Cronin (2016) zu beobachten, die den Orthogonalisierungsansatz nach Henseler und Chin (2010) als den in jedem Fall besten Ansatz beschreiben, wobei der Zwei-Stufen-Ansatz aber die größere statistische Power hätte hervorbringen können. Da neben den genannten drei Fehldeutungen der Übersichtsartikel jedoch auch drei korrekte Deutungen auftreten, ist unklar, ob trotz der Zitierung eines Übersichtsartikels tendenziell der falsche Ansatz gewählt wird oder ob eine Zitierung eher mit der Wahl des besten Ansatzes in Verbindung gebracht werden kann. Demnach wird kein Zusammenhang zwischen Unsicherheit bei der Verwendung der Informationen aus den Übersichtsartikeln und einer häufigen Zahl gewählter suboptimaler Ansätze beobachtet, der von Hair et al. (2019) vermutet wurde.

Vielmehr könnte bei der Betrachtung der 14 Publikationen, in denen im Falle reflektiver latenter Variablen ein nach dem Framework suboptimaler Ansatz gewählt wurde, die geringe Anzahl an Zitationen einer der Übersichtsartikel erneut einen möglichen Grund für die hohe Anzahl der Verstöße liefern. Insgesamt erwähnen nur drei der 14 gegen das Framework verstoßende Artikel (21,43%) einen der Übersichtsartikel. Demnach könnte hier davon ausgegangen werden, dass nicht die falsche Anwendung der Vorgaben in den Übersichtsartikeln für die Unsicherheit der Forscher und häufige Wahl des Produkt-Indikator-Ansatzes sorgt, sondern Uninformiertheit über die Existenz der Übersichtsartikel und Unwissen darüber, dass die Ansätze eine unterschiedliche Leistung hervorbringen können. Forscher verwenden in diesem Fall unter Umständen den Ansatz, den sie aus dem Regressionskontext zur Durchführung von MA kennen, weshalb der Produkt-Indikator-Ansatz so häufig gewählt worden sein könnte.

Fasst man die 20 Modelle mit ausschließlich reflektiv gemessenen Konstrukten und die sieben mit zumindest teilweise formativ gemessenen Konstrukten zusammen, wird nun in insg. 17 Fällen gegen das Framework bei der Wahl des Ansatzes verstoßen (62,96%). Das bedeutet auch, dass in zehn Fällen der korrekte Ansatz bzgl. der Messmodelle gewählt wird (37,04%). Hierbei wird erneut der Zusammenhang zwischen der Wahl des falschen bzw. suboptimalen Ansatzes und dem nicht vorhandenen Verweis auf einen der besprochenen Übersichtsartikel deutlich, da in 14 der 17 Verstöße keine Zitation auftritt (82,35%). Demnach wird weiterhin vermutet, dass die Wahl des falschen Ansatzes mit dem Unwissen der Forscher über die Existenz der Empfehlungen in den Übersichtsartikeln zusammenhängt. Gleichzeitig kann die Einhaltung der Empfehlungen wie zuvor nicht eindeutig mit den erfolgten Zitierungen der Übersichtsartikel in Verbindung gebracht werden, da bei sechs der zehn korrekt gewählten Ansätze (60%) ein Übersichtsartikel zitiert wurde und bei vier Einhaltungen kein Verweis gefunden werden konnte.

Dabei sollte jedoch auch die dritte potenzielle Fehlerquelle bei MA in PLS-SEM einbezogen werden, um ein realistischeres Bild der Verstöße gegen die Vorgaben des Frameworks zur Bildung des Interaktionsterms erhalten zu können. Das Finden signifikanter Effekte stellt den primären Fokus einer Studie in PLS-SEM dar, insofern diese zu testende Hypothesen enthält. Dies macht die Differenzierung von Henseler und Chin (2010) nach den Zielen von MA obsolet, wobei höchstens in absoluten Ausnahmefällen die Maximierung des Bestimmtheitsmaßes R^2 und damit die Vorhersagegenauigkeit einen Fokus von Forschern darstellt. Dennoch wird in vier der 27 Fälle (14,81%) fälschlicherweise darauf verwiesen, dass der Fokus auf der Vorhersage von Effekten liegt. Korrigiert man diese vier Fälle dahingehend, dass das eigentliche Ziel in jedem Fall die Maximierung der statistischen Power ist, erhöht sich die Anzahl von Verstößen gegen das Framework von 17 auf 19 von 27 (70,37%) und die Anzahl der Einhaltungen verringern sich von zehn auf acht von 27 (29,63%). Dies kann damit begründet werden, dass zwei der vier korrigierten Fälle ohnehin bereits gegen das Framework verstoßen haben, da ein suboptimaler Ansatz gewählt wurde, um die Vorhersagegenauigkeit zu maximieren. Die verbleibenden zwei Fälle (Kumar et al. (2020); Teller et al. (2013)) wurden demnach korrigiert und verstoßen nun ebenfalls gegen das Framework, da beide den Produkt-Indikator-Ansatz nutzen, der nach dem Framework niemals der beste Ansatz zur Generierung des Interaktionsterms bei dem Ziel der Maximierung statistischer Power darstellt. Zusammengefasst verstößt der Einsatz des Produkt-Indikator-Ansatzes immer gegen das Framework, da die Wahl des besten Ansatzes bei dem Ziel der meisten Forscher vorrangig in dem Auffinden signifikanter Effekte besteht und hierbei in

den meisten Fällen der Zwei-Stufen-Ansatz und seltener der Orthogonalisierungsansatz bei der Maximierung der statistischen Power dominieren.

Werden nun die 19 fehlerhaften Wahlen der Ansätze näher betrachtet, kann erkannt werden, dass hierbei in 15 Fällen (78,95%) keiner der genannten Übersichtsartikel zitiert wurde. Darüber hinaus sind die acht Fälle mit korrekter Wahl des Ansatzes auch in diesem Fall fast gleichverteilt, da in fünf Fällen (60%) einer der Übersichtsartikel erwähnt wird und in drei nicht. Demnach kommt man auch nach der Korrektur mutmaßlich falsch vermuteter Ziele der MA von Forschern zu derselben Vermutung bezüglich der Begründung der häufigen Verstöße gegen das Framework wie zuvor. So kann das Unwissen über die verschiedenartige Leistungsfähigkeit der Ansätze als plausibler Grund für die Wahl des falschen Ansatzes herangezogen werden, anstatt des falschen Umgangs mit den bekannten Empfehlungen der Übersichtsartikel.

Letztlich stellt die Entscheidung über die Art der Behandlung der Indikatorendaten eine weitere potenzielle Fehlerquelle dar, die mit der Bildung des Interaktionsterms zusammenhängt. Wie bereits dargelegt, sollten die Daten der Indikatoren vor der Analyse standardisiert werden, was insbesondere bei dem Orthogonalisierungsansatz und dem Produkt-Indikator-Ansatz erfolgen sollte, da der Zwei-Stufen-Ansatz die Daten bei der Generierung des Interaktionsterms ohnehin standardisiert (Becker et al., 2018). Insgesamt weisen lediglich zwölf der 27 betrachteten Fälle (44,44%) eine Angabe zur Datenbehandlung auf, wobei die Daten hierbei in elf Fällen standardisiert und in einem Fall (Matsuno et al., 2014) mittelwertzentriert. Dementsprechend wird in 15 Fällen keine Information dazu gegeben, wie mit den Daten vor der Analyse verfahren wurde.

Dies stellt jedoch nicht in jedem Fall einen Verstoß gegen die Vorgaben des Frameworks dar, da die Standardisierung der Daten vor der Analyse bei der Verwendung des Zwei-Stufen-Ansatzes freiwillig erfolgen kann. Die Angabe keiner Informationen, kann jedoch bei der Verwendung des Produkt-Indikator-Ansatzes oder Orthogonalisierungsansatzes die Einschätzung von Problemen, wie einer Erschwerung der Interpretation der Ergebnisse oder einer potenziellen Verzerrung der Höhe der Pfadkoeffizienten, verhindern (Becker et al., 2018; Hair et al., 2017, S. 257). Schließt man die sechs Fälle bei dem Einsatz des Produkt-Indikator-Ansatzes und den einen Fall bei der Verwendung des Orthogonalisierungsansatzes, in denen keine Informationen zur Datenbehandlung gegeben werden, von der Evaluation mithilfe des Frameworks aus, wird deutlich, dass lediglich 20 der 27 Fälle für die Analyse verbleiben. Werden diese nun unter der Annahme, dass die Autoren stets das Ziel der Maximierung der statistischen Power verfolgen, ebenfalls auf die Korrektheit der verwendeten Ansätze geprüft, ergeben sich noch acht das Framework einhaltende Fälle (40%). Der relative Anteil verstoßender Fälle sinkt somit von 70,37 Prozent auf 60 Prozent, da sämtliche dem Framework folgende Artikel erhalten werden konnten, und der relative Anteil dem Framework folgender Fälle erhöht sich von 29,63 auf 40 Prozent. Wird jedoch davon ausgegangen, dass die sieben ausgeschlossenen Fälle auch weiterhin Verstöße darstellen, erhöht sich die Anzahl von Verstößen wieder auf 19, während die Anzahl der dem Framework folgenden Fälle bei acht verbleibt und das Verhältnis optimaler und suboptimaler bzw. falscher Generierungen des Interaktionsterms sich wie zuvor ausprägt (70,37% zu 29,63%). In dem vorliegenden Review hat die Datenbehandlung demnach keinen Einfluss auf die Einhaltung oder Verstöße gegen das aufgestellte Framework, da fast alle Artikel die Daten standardisieren und die keine Angaben zur Verfügung stellenden Untersuchungen sowie der Artikel mit der Mittelwertzentrierung bereits aus den zuvor diskutierten Gründen gegen das Framework verstoßen.

Zusammenfassend verstößt mit 19 Fällen die Mehrheit der mit dem Framework untersuchbaren 27 Fälle (70,37%) gegen die Empfehlungen zur Bildung des Interaktionsterms (und der vorangehenden Datenbehandlung). Dies könnte sich insb. auf die geringe Anzahl von Zitationen der Übersichtsartikel von Henseler und Chin (2010), Henseler und Fassott (2010), Hair et al. (2017) sowie Becker et al. (2018) und damit auf das Unwissen vieler Forscher zu den Stärken und Schwächen der Ansätze zurückführen lassen.

4.2.2. Evaluation der Messmodelle und Interpretation der Ergebnisse

Der zweite Bereich potenzieller Fehlerquellen umfasst neben der Interpretation der Ergebnisse einer MA zunächst die Evaluation der Messmodelle der Variablen, die innerhalb der MA geschätzt werden. Dabei sind die gemachten Beobachtungen mit der absoluten und relativen Anzahl der Vorkommnisse und passenden Beispielen sowohl für diesen Abschnitt als auch für den nächsten Abschnitt, d.h. für Beobachtungen zur künstlichen Dichotomisierung kontinuierlicher Variablen, in Tabelle 4 dargestellt.

Während die Evaluation der Messmodelle der latenten moderierenden, exogenen und endogenen Variable in herkömmlicher Art und Weise ablaufen sollte, weisen Forscher widersprüchliche Ansichten bezüglich der Evaluation des Interaktionsterms auf (Hair Jr, Hult, Ringle & Sarstedt, 2017, S. 255). Dabei konnte im Rahmen der vorliegenden Arbeit keine Empfehlung gegeben werden, wie mit der Evaluation insb. der Reliabilität des Interaktionsterms umzugehen ist.

Das Review aller 35 Untersuchungen[10] gibt jedoch Aufschluss über den Umgang der einschlägigen Literatur mit der Evaluation des Interaktionsterms bei einfachen MA in PLS-SEM. Es fällt auf, dass im Rahmen der Evaluation des Messmodells des Interaktionsterms lediglich die diskriminante Validität des Interaktionsterms geprüft wurde. Hierbei überprüfen dies nur Calantone und Rubera (2012) sowie Köhler et al. (2011), wobei in keinem der zwei Artikel ein Grund hierfür angegeben ist.

Insbesondere die Reliabilität des Interaktionsterms soll jedoch nach Aguinis et al. (2017) und Li et al. (2019) beachtet werden, um Verzerrungen der Ergebnisse durch Messfehler zu vermeiden. Die Autoren differenzieren dabei nicht

[10] Da eine Analyse mithilfe des Frameworks nun nicht mehr nötig ist, können wieder alle 35 Untersuchungen herangezogen werden.

Tabelle 4: Zusammenfassung der Beobachtungen bekannter Fehlerquellen bei der Evaluation der Messmodelle und bei der Interpretation der Ergebnisse

Bekannte Fehlerquelle	Vorkommen (absolut)	Vorkommen (relativ)	Beispiele
Evaluation des Interaktionsterms (diskriminante Validität)	2 von 35 Untersuchungen	5,71% der Untersuchungen	Calantone und Rubera (2012); Köhler, Rohm, De Ruyter und Wetzels (2011)
Evaluation des Interaktionsterms (interne Konsistenzreliabilität)	0 von 35 Untersuchungen	0% der Untersuchungen	-
Evaluation der reflektiv gemessenen exogenen und moderierenden Variablen (interne Konsistenzreliabilität)	27 von 27 Untersuchungen	100% der Untersuchungen	Heidenreich, Wittkowski, Handrich und Falk (2015); Pelser et al. (2015)
keine Trennung des Main Effects und Simple Effects	12 von 35 Untersuchungen	34,29% der Untersuchungen	Matsuno et al. (2014); Nakata, Zhu und Izberk-Bilgin (2011); Pelser et al. (2015)
Prüfung der MV auf Signifikanz	1 von 35 Untersuchungen	2,86% der Untersuchungen	Krishen, Leenders, Muthaly, Ziółkowska und LaTour (2018)
Angabe der Effektgröße f^2 (individuelle und kollektive ME)	10 von 35 Untersuchungen	28,57% der Untersuchungen	Johnston et al. (2012); Ahrholdt, Gudergan und Ringle (2019) (beide Untersuchungen)
davon Angabe der Effektgröße f^2 (individuelle ME)	7 von 10 Untersuchungen	70% der Untersuchungen	Johnston et al. (2012); Teller et al. (2013); Zobel (2017); Coelho und Henseler (2012) (beide Untersuchungen)
dabei Einschätzung der Höhe der Effektgröße f^2 nach Kenny (2018); (hiervon 3 mal ohne Angabe)	2 von 10 Untersuchungen	20% der Untersuchungen	Ahrholdt et al. (2019) (beide Untersuchungen)
künstliche Dichotomisierung kontinuierlicher Variablen	0 von 33 Untersuchungen	0% der Untersuchungen	-

zwischen den Messmodellen, wobei die Evaluation der internen Konsistenzreliabilität formativ gemessener Konstrukte im Allgemeinen nicht empfohlen wird (Hair et al., 2017, S. 137-138). Es wurden dennoch alle 35 Untersuchungen auf die Reliabilität des Interaktionsterms geprüft, da dieser nicht formativ gemessen wird. Jedoch gibt keine analysierte Untersuchung einen Wert für Cronbachs Alpha oder die Composite Reliability des Interaktionsterms an, welche die interne Konsistenzreliabilität von Konstrukten messen können (Hair et al., 2017, S. 106). Aguinis et al. (2017) empfehlen, dass zumindest die Angabe der Reliabilität aller exogenen Variablen und der MV, die Interaktionsterme bilden, damit sich im Fall nicht-signifikanter ME bei der Interpretation der Ergebnisse der MA hierauf bezogen werden kann. Im PLS-SEM-Kontext sollte eines der gängigen Reliabilitätsmaße demnach zumindest bei reflektiv gemessenen Konstrukten mit mehr als einem Indikator erfasst werden. Dem kommen sämtliche Untersuchungen im Review nach, da in jedem Fall reflektiver gemessener Konstrukte mindestens ein Reliabilitätsmaß angegeben wurde. In acht Fällen wurden keine Angaben zur Reliabilität gemacht, wobei hier mit formativ gemessenen Konstrukten oder latenten Variablen mit nur einem Indikator verfahren wurde und die Evaluation der Reliabilität nicht wie bei reflektiven Messungen möglich ist (Hair et al., 2017, S. 108-109, 137-138). Dies steht den Resultaten von Aguinis et al. (2017) gegenüber, die in ihrem Review herausfanden, dass in der Mehrheit der von ihnen untersuchten Artikel keine Angaben zur Reliabilität der Messungen machen. Demzufolge wird die Evaluation und insb. die Reliabilität der Konstrukte in Marketinganwendungen von PLS-SEM in allen Fällen berichtet, weshalb sie in diesem Kontext nicht als Problemquelle angesehen werden sollte.

Die Interpretation des Main Effects und Simple Effects stellt die nächste potenzielle Problemquelle dar, da hierbei ebenfalls gegensätzliche Meinungen in der Forschung vorherrschen. Insbesondere ist unklar, ob die Effekte nach Hair et al. (2017, S. 257-258), Chin et al. (2003) und Henseler und Fassott (2010) in getrennten Modellen, d.h. jeweils ohne und mit der latenten MV im Modell, oder nach Aguinis et al. (2017) und Li et al. (2019) in demselben Modell analysiert werden sollen. Letztere Autoren sprechen sich dabei dafür aus, den Main Effect bei Vorliegen eines signifikanten ME vollständig aus der Interpretation zu exkludieren und lediglich den Simple Effect zu interpretieren. Alles in allem wurden in den vorliegenden 35 Untersuchungen zwölf Mal der Main und Simple Effect nicht getrennt (34,29%), wobei in den meisten Fällen eine getrennte Berechnung des Main Effects- und Simple Effects-Modells erfolgt (65,71%). Bei-

spielsweise wurde im Artikel von Matsuno et al. (2014) nur ein Modell berechnet, in dem alle Effekte simultan geschätzt und berichtet wurden und der Main Effect demnach nicht vorkommt. Der Simple Effect wird in diesem Fall zum Testen einer Hypothese genutzt, was nach Hair et al. (2017, S. 257-258) nicht der inhaltlich korrekten Vorgehensweise bei der Interpretation der Ergebnisse entspricht, nach Aguinis et al. (2017) jedoch genau richtig ist.

Darüber hinaus beinhaltet die nächste potenzielle Fehlerquelle im Rahmen der Interpretation der Ergebnisse von MA die Prüfung der MV anstatt des Interaktionsterms auf Signifikanz. Nach Henseler und Fassott (2010) sollte stets der Interaktionsterm zur Überprüfung der Signifikanz eines ME herangezogen werden. In einer der 35 in das Review eingeschlossenen Untersuchungen (2,86%) konnte diesbezüglich festgestellt werden, dass der Pfadkoeffizient p_2 von einer MV auf eine endogene Variable auf Signifikanz überprüft wurde. So stellt der Artikel von Krishen et al. (2018) einen Fall dar, in dem der direkte Effekt der MV auf endogene Konstrukte getestet und interpretiert wurde, die mitunter als exogene latente Variablen in der MA hätten dienen müssen. Insbesondere wird hierbei keine MA durchgeführt, sondern eine Mediationsanalyse, da direkte Effekte der MV auf zwei endogene Konstrukte und direkte Effekte des einen endogenen Konstrukts auf das andere endogene Konstrukt hypothetisiert und schließlich geschätzt werden. Weiterhin finden sich in den Ergebnissen der Studie generell keine Interaktionsterme, wobei diese nach den Autoren dennoch gebildet wurden. Die Interpretation der MA, u.a. bestehend aus der Betrachtung des linearen Einflusses der „MV“ auf endogene Variablen, ist folglich inhaltlich nicht korrekt, da diese ebenfalls zu einer Mediationsanalyse passen würde. Die Ergebnisse der „MA“ sind demnach zwar im Rahmen einer Mediationsanalyse verwertbar, aber nicht für die Analyse von ME. Im Rahmen einer korrekt hypothetisierten und konzeptionell richtig modellierten MA wurde allerdings kein Fehler im Sinne der fälschlichen Prüfung der MV auf Signifikanz gefunden.

Schließlich können bei der Interpretation auch Fehler auftreten, die mit der Bildung und Einschätzung der Effektgröße f^2 zusammenhängen. Wie im Vorfeld vermutet wurde, weisen auch die 35 Untersuchungen im Review lediglich in zehn Untersuchungen (28,57%) die Angabe der Effektgröße des ME auf. Bei der Betrachtung der angegebenen zehn Effektgrößen konnte beobachtet werden, dass in den meisten Fällen (70%) herausgestellt wird, wie sehr ein individueller ME zur Erklärung der endogenen latenten Variable beiträgt. Demnach wurden in sieben Untersuchungen für jede MV einzeln die Größe des ME bestimmt. Drei Untersuchungen (30%) bestimmen die Effektgröße aller MV im Modell jedoch simultan, wobei nicht zu erkennen ist, wie sehr die einzelnen MV zur Erklärung der endogenen Variable beitragen. Dieses Vorgehen kann kritisch betrachtet werden, da bei der Interpretation der ME die Einschätzung der Höhe der Effektgröße in diesem Fall nicht für jeden ME einzeln möglich ist. So ist bspw. in dem Fall, dass eine MV auf mehrere Beziehungen wirkt, die alle mit derselben endogenen Variable zusammenhängen, unklar, wie hoch die Effektgröße der ME in jeder Beziehung tatsächlich ist. Unter anderem kann in beiden Untersuchungen im Artikel von Ahrholdt et al. (2019) nicht erkannt werden, wie hoch die Effektgrößen der fünf ME und damit die Größe der fünf Effekte ist, die auf die Beziehungen der fünf exogenen Variablen auf die endogene Variable wirken. Die Autoren Teller et al. (2013) hingegen prüfen bspw. jeden ME einzeln auf ihren Beitrag zur Erklärung der Varianz der endogenen Variable. Auch hier wirkt dieselbe MV auf mehrere Pfade im Modell.

Neben der Bildung der Effektgröße f^2 stellt auch die Einschätzung der Höhe der Effektgröße im Rahmen der Interpretation des ME eine Problemquelle dar. Diese kann nach Cohen (1988, S. 412-414) oder Kenny (2018) erfolgen, wobei in in der vorliegenden Arbeit die Einschätzung nach Kenny (2018) empfohlen wird, nach der ein kleiner Effekt zumindest 0,005 betragen sollte. Lediglich Ahrholdt et al. (2019) nutzen in ihren beiden Untersuchungen die Einteilung von f^2 nach Kenny (20%), wogegen in fünf Untersuchungen noch auf Cohen (1988, S. 412-414) zurückgegriffen wird (50%). Schließlich sind in drei Untersuchungen keine Informationen bezüglich der Einschätzung von Effektgrößen der ME angegeben (30%). Dies zeigt ein gewisses Unwissen der aktuellen Forschung darüber, dass ME i.d.R. sehr klein ausfallen. Schließlich könnte sich dieses Unwissen in Form von voreiligen Ausschlüssen einzelner ME aus der Analyse manifestieren, die keinen zumindest kleinen Effekt ausüben. Wird in einer Untersuchung mithilfe von Cohen (1988, S. 412-414) versucht f^2 einzuschätzen, ist denkbar, dass ein ME infolge eines zu kleinen Effekts von der Interpretation ausgeschlossen wird. Da Effektgrößen jedoch nach Aguinis et al. (2005) häufig unterhalb der unteren Grenze von Cohen (1988, S. 412-414) liegen, sollten diese nicht verfrüht bei einer Unterschreitung der unteren Grenze ausgeschlossen werden. Keine der inkludierten Untersuchungen weist bei den ME eine Unterschreitung der unteren Grenze nach Cohen (1988, S. 412-414) auf. Die Vermutung, dass es zu verfrühten Ausschlüssen von ME von der Interpretation gekommen ist, kann im Rahmen der vorliegenden Arbeit weder widerlegt noch bestärkt werden.

4.2.3. Künstliche Dichotomisierung kontinuierlicher Moderatorvariablen

Letztlich kann neben der Generierung des Interaktionsterms und der Evaluation der Ergebnisse auch der Umgang mit der Skalierung der MV bei einfachen MA eine potenzielle Problemquelle darstellen. Bei der Skalierung von MV unterscheiden Henseler und Fassott (2010) zwischen kategorischen und kontinuierlichen MV, wobei kontinuierliche Variablen durch künstliche Dichotomisierung in Kategorien eingeteilt und somit zu kategorischen MV umgewandelt werden können. Nach Aguinis et al. (2017) reduziert diese Vorgehensweise jedoch die Varianz der MV, wodurch die Analyse dieser beeinträchtigt wird. Im PLS-SEM-Kontext konnte bei der Betrachtung von 33 der 35 Untersuchungen (verbleibende zwei Untersuchungen beinhalten bereits ausschließlich kategorische MV) im Review keine Untersuchung identifiziert werden, die kontinuierliche Variablen künstlich di-

chotomisiert. Hierbei steht das vorliegende Ergebnis somit der Feststellung von Aguinis et al. (2017) gegenüber, die in 10,24 Prozent der von den Autoren untersuchten Artikel eine künstliche Dichotomisierung vorfanden. Da im aktuellen Review keine künstlichen Dichotomisierung gefunden wurde, folgen sämtliche inkludierte Artikel den Vorgaben von Henseler und Fassott (2010) und Aguinis et al. (2017), kontinuierliche MV nicht zu dichotomisieren sondern sie metrisch bzw. quasi-metrisch zu belassen.

4.3. Sonstige Beobachtungen

Der aktuelle Abschnitt umfasst weitere Beobachtungen, die erst während des Reviewprozesses gemacht werden konnten und deshalb im Vorfeld nicht theoretisch hergeleitet wurden. Aus diesem Grund beziehen sich die nachfolgenden Beobachtungen nicht auf bekannte Fehlerquellen, sondern repräsentieren neue Erkenntnisse bezüglich potenzieller und unerforschter Fehlerquellen bei Moderation in Marketinganwendungen von PLS-SEM. Diese Erkenntnisse sind ebenfalls in Tabelle 5 zusammengetragen und wurden darüber hinaus mit passenden Beispielen aus den in das Review eingeschlossenen Artikeln ergänzt.

Die erste Auffälligkeit in den kodierten Daten der 35 Untersuchungen im Review umfasst das häufige Vorkommen mehrerer ME, die innerhalb einer Untersuchung getestet wurden. Diese multiplen ME können dabei auf dieselbe oder verschiedene Beziehungen im Modell wirken. Alles in allem haben in 21 der 35 Untersuchungen (60%) mehrere MV einen Einfluss auf eine und/oder mehrere Beziehungen. PLS-SEM ist nach Sarstedt, Ringle et al. (2020) geeignet, um auch komplexe Modelle mit mehreren MV zu schätzen. Dennoch weisen Frazier et al. (2004) darauf hin, dass die simultane Analyse multipler MV in demselben Modell Probleme hervorrufen kann. Nach den Autoren kann demnach bei der Durchführung vieler MA in einem Modell die Fehlerquote vom Typ 1 erhöht werden. Die Betrachtung der insg. 21 Untersuchungen, in denen mehr als eine MV analysiert wird, gibt Aufschluss darüber, wie häufig die MV simultan bzw. separat, d.h. in demselben bzw. in unterschiedlichen Modellen, geschätzt wurden. Insgesamt wirken verschiedene MV in 16 Untersuchungen auf dieselbe Beziehung im Pfadmodell, wobei die ME dabei in 14 Fällen (87,5%) simultan berechnet wurden. Darüber hinaus haben in acht Untersuchungen multiple MV Einflüsse auf mehrere Beziehungen im Modell, wovon fünf (62,5%) auf einmal geschätzt wurden.

In diesem Fall erfolgte verhältnismäßig häufiger eine simultane Berechnung multipler ME, wenn die MV auf dieselbe Beziehung im Pfadmodell gewirkt haben. Hierbei wird allgemein deutlich, dass die Mehrheit der Untersuchungen mehrere MV analysiert und diese vorrangig simultan geschätzt werden. Infolgedessen bedarf es spezifischer Richtlinien für den Fall, dass mehrere MV in eine MA inkludiert werden. Hier ist bspw. fraglich, ob Forscher die Schätzung der ME simultan, durch sukzessives Hinzufügen der Interaktionsterme oder in gänzlich eigenen Modellen vollziehen sollten.

Weiterhin fiel auf, dass zehn der 35 Untersuchungen (28,57%) mindestens eine MV enthalten, die mit nur einem Indikator gemessen wird. Dies steht der Empfehlung von Hair et al. (2017, S. 246) gegenüber, MV mit mehr als einem Indikator zu messen. Dabei gehen die Autoren darauf ein, dass eine Messung mit nur einem Indikator weniger prädiktive Power (ausgedrückt in einem kleinen R^2) nach sich zieht als eine Messung mit mehreren Indikatoren und dies insb. das Auffinden kleiner, aber dennoch in der Population signifikanter, ME erschwert. Die Untersuchung von Teller et al. (2016) bspw. enthält die zwei MV „Attraktivität des Standorts des Ladens“ und „Kraft eines Ladens Laufkundschaft zu generieren“, die mit nur einem Indikator gemessen werden. Beide MV sind nicht als nicht signifikant herausgestellt worden, wobei ein Fehler vom Typ 2 aufgetreten sein könnte. Dies kann neben der verringerten prädiktiven Power ebenfalls u.a. mit einem zu geringen Stichprobenumfang oder zu kleinen Effektgrößen der ME zusammenhängen. Allerdings ist unklar, ob ein Fehler vom Typ 2 tatsächlich zu diesem Ergebnis geführt hat oder ob der ME tatsächlich nicht existiert. Werden hierbei auch exogene und endogene Variablen mit einbezogen, die mit nur einem Indikator gemessen werden, erhöht sich die Anzahl von Untersuchungen mit mindestens einer solchen Messung auf 15. Vor dem Hintergrund der potenziellen Nachteiligkeit solcher Messungen, entspricht dies mit 42,86 Prozent einem recht hohen Anteil an den inkludierten Untersuchungen im Review.

Bei der anschließenden Betrachtung der Anzahl signifikanter Ergebnisse aller inkludierter Untersuchungen fällt auf, dass nur zwei Studien keine signifikanten ME berichten (5,71%). Die demnach hohe Zahl signifikanter Ergebnisse ist dabei erwartungswidrig, da ME nach Aguinis et al. (2005) klein ausfallen und damit schwerer zu erkennen sind. Dies hätte zur Folge, dass ein größerer Anteil von nicht signifikanten Ergebnissen hätte gefunden werden müssen. Aus diesem Grund kann vermutet werden, dass die in das Review eingeschlossenen Artikel dem sog. Publication Bias unterliegen. Dieser stellt nach Rothstein, Sutton und Borenstein (2005, S. 1-3) die verzerrte Darstellung von Daten in wissenschaftlichen Publikationen dar. Dabei werden nach den Autoren u.a. selektiv vorrangig signifikante Ergebnisse veröffentlicht, was auch die geringe Anzahl nicht signifikanter und das häufige Vorkommen signifikanter Ergebnisse erklären könnte.

Darüber hinaus treten in nur zwei Fällen kategorische MV auf, wobei diese stets mit einem Indikator gemessen werden (Henseler & Fassott, 2010). Die Untersuchung im Artikel von Krishen et al. (2018) beinhaltet mit dem Geschlecht und dem Land zwei kategorische MV, während Rollins et al. (2012) mit der Kategorisierung nach verschieden hohen Anzahlen von Jahren, die ein Unternehmen Erfahrung mit Customer-Relationship-Management-Systemen hat, nur eine enthält. Somit kommen kategorische Variablen gegenüber kontinuierlichen Variablen jedoch wesentlich seltener (5,71%) in den 35 analysierten Marketinganwendung vor.

Neben den bisherigen sonstigen Beobachtungen sollten außerdem die neun Untersuchungen aufgegriffen werden, die mindestens einen ME innerhalb einer Mediation getestet haben (25,71%). Hierbei wird der Vorteil von PLS-SEM genutzt, auch sehr komplexe Modellbeziehungen wie moderier-

Tabelle 5: Sonstige Beobachtungen aus dem Review, deren Vorkommen und Beispiele

Beobachtung	Vorkommen (absolut)	Vorkommen (relativ)	Beispiele
multiple ME wirken auf dieselbe bzw. mehrere Beziehungen im Modell	21 von 35 Untersuchungen	60% der Untersuchungen	Ali et al. (2020); Mahr et al. (2014); Mullins et al. (2020)
Messung mind. eines Moderators mit nur einem Indikator	10 von 35 Untersuchungen	28,57% der Untersuchungen	Rollins et al. (2012); Teller, Alexander und Floh (2016); van der Borgh und Schepers (2014)
Messung mind. einer Variable mit nur einem Indikator	15 von 35 Untersuchungen	42,86% der Untersuchungen	Ashill und Jobber (2014); Teller et al. (2016); Wolter und Cronin (2016)
Anzahl nicht signifikanter Ergebnisse	2 von 35 Untersuchungen	5,71% der Untersuchungen	Jean, Sinkovics und Hiebaum (2014); Teller et al. (2016)
Anzahl kategorischer MV	2 von 35 Untersuchungen	5,71% der Untersuchungen	Krishen et al. (2018); Rollins et al. (2012)
mind. eine moderierte Mediation	9 von 35 Untersuchungen	25,71% der Untersuchungen	Kumar et al. (2020); Johnston et al. (2012)
mind. ein HCM im Zusammenhang mit MA	8 von 35 Untersuchungen	22,86% der Untersuchungen	Lam, Ahearne und Schillewaert (2012); Griffith, Lee, Yeo und Calantone (2014)
Hypothetisierung des ME	34 von 35 Untersuchungen	97,14% der Untersuchungen	Miao und Evans (2013); Sharma und Jha (2017); Siahtiri (2018)

te Mediationen ohne Probleme schätzen zu können (Sarstedt, Hair Jr et al., 2020). Diese wurden wie bereits angedeutet jedoch als einfache MA betrachtet und im vorliegenden Review so behandelt, als würden sie auf eine Beziehung zweier Konstrukte wirken, die nicht in einem Mediationsmodell beinhaltet sind. Dies zieht ohnehin vorrangig Konsequenzen für die Schätzung und Interpretation der mit der Mediation einhergehenden Effekte nach sich, da die Moderation auf die Mediation wirkt und nicht umgekehrt. Insbesondere die Struktur der Mediation und der Pfad bzw. die Pfade auf welchen eine oder mehrere MV wirken haben hierauf einen Einfluss (Hair et al., 2017, S. 261). Werden die Modelle von Hayes (2013) betrachtet, die diverse Arten von MA im Zusammenhang mit Mediationen darstellen, können insg. fünf der neun Mediationen einem Modell zugeordnet werden. So wirkt bspw. im Artikel von Kumar et al. (2020) die MV auf beide Pfade des indirekten Effektes, also den Pfaden zu der Mediatorvariable hin und von ihr weg, wobei Modell 58 von Hayes (2013) identifiziert werden konnte.

Mit acht Fällen konnten fast genauso häufig wie Mediationsmodelle auch HCM im Zusammenhang mit MA beobachtet werden. So ist ein HCM in acht Untersuchungen in mindestens einer MA zu finden (22,86%). Hierbei treten HCM jeweils zwei Mal an der Stelle der moderierenden, exogenen, endogenen sowie gleichzeitig der exogenen und endogenen Variable auf. Dabei haben Teller et al. (2016), Köhler et al. (2011), Lam et al. (2012) und Ali et al. (2020) HCM als endogene Variablen verwendet. Darüber hinaus wird nur in den Artikeln von Zobel (2017) und Ali et al. (2020) der Zwei-Stufen-Ansatz zur Repräsentation des HCM im Messmodell genutzt, während Heidenreich, Spieth und Petschnig (2017), Köhler et al. (2011), Kumar et al. (2020) und Lam et al. (2012) hierfür den Repeated Indicators Approach verwenden. In den Artikeln von Teller et al. (2016) und Griffith et al. (2014) wird keine Angabe zum verwendeten Ansatz gemacht.

Schließlich findet eine Hypothetisierung der ME in fast allen Fällen statt, da nur ein Mal im Fall in Wolter und Cronin (2016) beobachtet werden konnte, dass nicht versucht wurde, einen ME zu hypothetisieren. Somit wird in den meisten Fällen (97,14%) den Forderungen von Andersson et al. (2014) und (Field, 2018, S. 486) nach der Schaffung einer hinreichenden theoretischen Grundlage des vermuteten ME nachgekommen. Im Gegensatz dazu steht die Beobachtung, dass bei der Interpretation nach Memon et al. (2019) Simple Slope Plots verwendet werden sollen, um die Interpretation zu erleichtern und nachvollziehbarer zu machen. Es wurden jedoch nur in neun Fällen (25,71%) Simple Slope Plots abgebildet und für die Interpretation genutzt.

5. Diskussion

5.1. Zusammenfassung

Die vorliegende Arbeit hat den Umgang der aktuellen Forschung mit Moderationsanalysen in Marketinganwendungen von PLS-SEM untersucht. Die Wahlmöglichkeiten zwischen verschiedenen Ansätzen zur Generierung des Interaktionsterms, der zur Modellierung eines ME auf die spezifische Beziehung zweier Variablen benötigt wird, stellen Forscher dabei u.a. vor Herausforderungen. Das Ziel war es deshalb, anhand eines Reviews von 35 empirischen Untersuchungen aus 33 einschlägigen Artikeln herauszufinden, inwiefern bekannte potenzielle Fehlerquellen auftreten, wie mit diesen in der

Literatur umgegangen wird und wo weitere Probleme vorkommen. Hierfür wurde zunächst ein Framework aufgestellt, das Forscher bei der Generierung des Interaktionsterms unterstützt. Nach einer systematischen Literaturrecherche wurde dieses anschließend als eine Grundlage für das Review genutzt.

Die Recherche brachte zunächst die Erkenntnis hervor, dass in den meisten Fällen keine Angaben zum verwendeten Ansatz gemacht werden und der Produkt-Indikator-Ansatz der am häufigsten verwendete Ansatz ist. Im Review konnte die Unsicherheit der Forschung bei der Generierung des Interaktionsterms bestätigt werden, da insg. 19 von 27 mit dem Framework untersuchbare Fälle auftraten, die einen suboptimalen Ansatz verwenden. Darüber hinaus stellen die Behandlung der Indikatorendaten vor der Analyse sowie die Evaluation der Konsistenzreliabilität der exogenen und moderierenden Variable ein geringes Problem für Forscher dar. Demgegenüber spiegelt die Interpretation der Main Effects und Simple Effects eine größere Unsicherheit wider, da zwölf der 35 Untersuchungen diese nicht voneinander trennen. Weiterhin werden nur in zehn Untersuchungen Angaben zur Effektgröße f^2 des ME gemacht. Schließlich fielen außerdem eine hohe Anzahl mit nur einem Indikator gemessene Variablen, ein geringes Aufkommen ausschließlich nicht signifikanter Ergebnisse und einige HCM und moderierte Mediationen im Zusammenhang mit MA auf. Die Beobachtungen aus dem Review lassen den Schluss zu, dass trotz vorhandener Übersichtsartikel zur Durchführung von MA in PLS-SEM wie Henseler und Chin (2010) oder Henseler und Fassott (2010) die Empfehlungen häufig nicht in der aktuellen Forschung umgesetzt werden.

5.2. Implikationen

Die Ergebnisse des Reviews in dieser Arbeit tragen in erster Linie zum Verständnis von Herausforderungen der aktuellen Forschung zu PLS-SEM im Umgang mit MA bei. Dabei wird zunächst die Vermutung bestätigt, dass eine Vielzahl von Artikeln die Empfehlungen zur Generierung des Interaktionsterms nicht einhalten. Dies kann vorrangig darauf zurückgeführt werden, dass nur vier der 19 gegen die Vorgaben verstoßenden Fälle einen Übersichtsartikel zitieren. Hierbei kann eine gewisse Unkenntnis der aktuellen Forschung darüber erkannt werden, dass die verschiedenen Ansätze nicht bei allen Messmodellen gleichermaßen genutzt werden bzw. unterschiedliche Ergebnisse hervorbringen können. Einerseits kann dabei die Nutzung eines Ansatzes mit wenig statistischer Power dazu führen, dass Fehler vom Typ 2 auftreten. Andererseits kann im Fall formativ gemessener Konstrukte innerhalb der MA die Verwendung eines anderen anstelle des Zwei-Stufen-Ansatzes zu inhaltlich falschen Ergebnissen und infolgedessen inkorrekten Implikationen führen. Demnach sollten Forscher vermehrt auf die Vorteile und Probleme der Ansätze aufmerksam gemacht werden, was durch weitere Übersichtsartikel in der Art von Henseler und Fassott (2010) oder Zusammenfassungen in zugänglichen Werken wie Hair et al. (2017) erfolgen kann. Dies kann auch den seltenen Angaben des verwendeten Ansatzes, der bei der Interpretation eine Rolle spielt, der Effektgröße von ME, die zeigen, ob tatsächlich Implikationen aus den signifikanten ME folgen, und dem häufigen Einsatz von Konstrukten, welche mit nur einem Indikator gemessen werden und dadurch ebenfalls die statistische Power verringern können, entgegenwirken.

Im Gegensatz dazu stellt die Überprüfung der Konsistenzreliabilität im Rahmen der Evaluation reflektiv gemessener latenter Variablen nach den Ergebnissen des Reviews eine gängige Praxis in der aktuellen Forschung dar. In sämtlichen Artikeln wurden dementsprechend die Messmodelle der verwendeten Konstrukte evaluiert. Allerdings bestehen noch immer Unsicherheiten darüber, ob eine Evaluation der Reliabilität des Interaktionsterms erfolgen sollte, was sich zunächst an unterschiedlichen Meinungen in der Literatur erkennen ließ. Im Review konnte diese Unsicherheit nicht erkannt werden, da die Evaluation der Reliabilitäten in keinem Artikel durchgeführt wurde. Dies impliziert, dass in der Praxis i.d.R. den Vorgaben (u.a.) von Hair et al. (2017, S. 255) zur Evaluation von Messmodellen gefolgt wird und hier verglichen mit anderen Problemen bei MA in PLS-SEM kaum Aufklärung benötigt wird.

Mehr Unsicherheit in der Forschung besteht demgegenüber bei der Interpretation der Main Effects und Simple Effects. Die Ergebnisse des Reviews deuten hierbei darauf hin, dass Forscher teilweise unterschiedlich mit den Effekten umgehen und sie auf verschiedene Arten generieren. Hierbei wird deutlich, dass die Interpretation der verschiedenen Effekte im Zusammenhang mit MA weiter erforscht werden muss, um eindeutige Empfehlungen für die Auswertung empirischer Studien zur Verfügung stellen zu können.

Letztlich wurde ein relativ großer Anteil von Studien beobachtet, die komplexe MA durch die Inklusion von HCM oder in Form von moderierten Mediationen durchführen, deren Analyse und Interpretation ein hohes Maß an Wissen und Erfahrung fordert (Hair et al., 2017, S. 285). Deshalb sind die Ergebnisse einer entsprechend hohen Anzahl an Studien mit Vorsicht zu betrachten, was auch für die Durchführung eigener Untersuchungen mit komplexen MA gilt.

5.3. Limitationen und zukünftige Forschung

Die Durchführung des Reviews unterlag gewissen Limitationen, die nun diskutiert werden. Zunächst ist die Grundgesamtheit aller Artikel, die in das Review aufgenommen wurden, dadurch eingeschränkt, dass eine zur Verfügung gestellte Datenbank aus 239 empirischen Artikeln verwendet wurde. Es kann hierbei nicht ausgeschlossen werden, dass Artikel in der Datenbank fehlen, die in das vorliegende Review hätten mit aufgenommen werden können. Hinzukommt, dass die Datenbank lediglich Publikationen der Jahre 2011 bis 2020 enthält. Einerseits werden somit ältere Artikel ausgeschlossen, andererseits fehlen die Artikel aus dem Jahr 2021, welche ein aktuelleres Bild des Umgang mit MA in PLS-SEM hätten vermitteln können.

Darüber hinaus basiert das generierte Framework auf verschiedenen Simulationsstudien, die der Limitation unterliegen, dass Parameter, wie Effektgrößen oder Korrelatio-

nen zwischen Konstrukten, bei unterschiedlichen Einstellungen verschiedenartige Ergebnisse generieren. Dies wirkt sich demnach auf die nach dem Framework zu wählenden Ansätze aus und hat einen Einfluss auf die Beobachtungen im Review bzgl. der Befolgung oder des Verstoßes des Frameworks bei der Generierung des Interaktionsterms. Weiterhin konnten ausschließlich Untersuchungen ohne HCM und mit kontinuierlichen Variablen mithilfe des Frameworks ausgewertet werden, was den Umfang von einbezogenen Artikeln reduziert hat. Die Generierung des Interaktionsterms wurde bei Untersuchungen, die in jeder MA ein HCM bzw. kategorische Variablen beinhalten, nicht analysiert. Außerdem wurden ME, die innerhalb einer Mediation aufgetreten sind, nicht im Kontext der Mediation betrachtet sondern als einfache MA ausgewertet.

Insgesamt beleuchtet das vorliegende Review bereits eine Vielzahl von Herausforderungen der aktuellen Forschung, wobei eine Wiederholung des Reviews auf der Basis einer erneuten systematischen Literaturrecherche verschiedenster Datenbanken in Zukunft weitere bzw. andersartige Ergebnisse hervorbringen kann. Hierbei sollten ebenfalls Untersuchungen mit HCM und kategorischen Variablen auf die Vorgehensweise bei der Generierung des Interaktionsterms überprüft werden. Letztlich kann ein Review von Artikeln mit MA durch Multigruppenanalysen in PLS-SEM auf dieselbe Weise durchgeführt werden, um diesbezügliche Herausforderung aufzeigen und bessere Vergleiche beider Arten von MA zu ermöglichen.

Literatur

Aguinis, H., Beaty, J. C., Boik, R. J. & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: a 30-year review. *Journal of Applied Psychology*, *90* (1), 94–107.

Aguinis, H., Edwards, J. R. & Bradley, K. J. (2017). Improving our understanding of moderation and mediation in strategic management research. *Organizational Research Methods*, *20* (4), 665–685.

Ahrholdt, D. C., Gudergan, S. P. & Ringle, C. M. (2019). Enhancing loyalty: When improving consumer satisfaction and delight matters. *Journal of Business Research*, *94*, 18–27.

Ali, I., Ali, M., Salam, M. A., Bhatti, Z. A., Arain, G. A. & Burhan, M. (2020). How international SME's vicarious learning may improve their performance? The role of absorptive capacity, strength of ties with local SMEs, and their prior success experiences. *Industrial Marketing Management*, *88*, 87–100.

Andersson, U., Cuervo-Cazurra, A. & Nielsen, B. B. (2014). From the Editors: Explaining interaction effects within and across levels of analysis. *Journal of International Business Studies*, *25*, 1063–1071.

Ashill, N. J. & Jobber, D. (2014). The effects of the external environment on marketing decision-maker uncertainty. *Journal of Marketing Management*, *30* (3), 268–294.

Baron, R. M. & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51* (6), 1173–1182.

Becker, J.-M., Klein, K. & Wetzels, M. (2012). Hierarchical latent variable models in PLS-SEM: guidelines for using reflective-formative type models. *Long Range Planning*, *45*, 359–394.

Becker, J.-M., Ringle, C. M. & Sarstedt, M. (2018). Estimating moderating effects in PLS-SEM and PLSc-SEM: Interaction term generation* data treatment. *Journal of Applied Structural Equation Modeling*, *2* (2), 1–21.

Calantone, R. & Rubera, G. (2012). When should RD & E and marketing collaborate? The moderating role of exploration–exploitation and environmental uncertainty. *Journal of Product Innovation Management*, *29* (1), 144–157.

Carte, T. A. & Russell, C. J. (2003). In pursuit of moderation: Nine common errors and their solutions. *MIS Quarterly*, *27* (3), 479–501.

Chen, A., Peng, N. & Hung, K.-P. (2015). Strategic management of salespeople when promoting new products: Moderating effects of sales-related organizational psychological climate. *European Journal of Marketing*, *49* (9), 1616–1644.

Chin, W. W., Marcolin, B. L. & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a monte carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, *14* (2), 189–217.

Coelho, P. S. & Henseler, J. (2012). Creating customer loyalty through service customization. *European Journal of Marketing*, *46* (3/4), 331–356.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Publishers.

Echambadi, R., Arroniz, I., Reinartz, W. & Lee, J. (2006). Empirical generalizations from brand extension research: How sure are we? *International Journal of Research in Marketing*, *23* (3), 253–261.

Esposito Vinzi, V., Trinchera, L. & Amato, S. (2010). PLS path modeling: from foundations to recent developments and open issues for model assessment and improvement. In V. Esposito Vinzi, W. W. Chin, J. Henseler & H. Wang (Hrsg.), (S. 47–82). Berlin, Springer.

Fassott, G., Henseler, J. & Coelho, P. S. (2016). Testing moderating effects in PLS path models with composite variables. *Industrial Management & Data Systems*, *116* (9), 1887–1900.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5. Aufl.). Newcastle, Sage.

Frazier, P. A., Tix, A. P. & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology*, *51* (1), 115–134.

Griffith, D. A., Lee, H. S., Yeo, C. S. & Calantone, R. (2014). Marketing process adaptation: Antecedent factors and new product performance implications in export markets. *International Marketing Review*, *31* (3), 308–334.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M. & Thiele, K. O. (2017). Mirror, mirror on the wall: a comparative evaluation of composite-based structural equation modeling methods. *Journal of the Academy of Marketing Science*, *45*, 616–632.

Hair, J. F., Ringle, C. M. & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, *19* (2), 139–151.

Hair, J. F., Sarstedt, M. & Ringle, C. M. (2019). Rethinking some of the rethinking of partial least squares. *European Journal of Marketing*, *53* (4), 566–584.

Hair, J. F., Sarstedt, M., Ringle, C. M. & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the academy of marketing science*, *40* (3), 414–433.

Hair Jr, J. F., Hult, G. T. M., Ringle, C. M. & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Newcastle, Sage.

Hall, J. & Sammons, P. (2014). Mediation, Moderation & Interaction. In T. Teo (Hrsg.), *Handbook of quantitative methods for educational research* (S. 267–286). Rotterdam: SensePublishers. Zugriff auf https://doi.org/10.1007/978-94-6209-404-8_13 doi: 10.1007/978-94-6209-404-8_13

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, Guilford.

Heidenreich, S., Spieth, P. & Petschnig, M. (2017). Ready, steady, green: Examining the effectiveness of external policies to enhance the adoption of eco-friendly innovations. *Journal of Product Innovation Management*, *34* (3), 343–359.

Heidenreich, S., Wittkowski, K., Handrich, M. & Falk, T. (2015). The dark side of customer co-creation: exploring the consequences of failed co-created services. *Journal of the Academy of Marketing Science*, *43*, 279–296.

Henseler, J. & Chin, W. W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural equation modeling*, *17* (1), 82–109.

Henseler, J. & Fassott, G. (2010). Testing moderating effects in PLS Path Models: An illustration of available procedures. In V. Esposito Vinzi, W. W. Chin, J. Henseler & H. Wang (Hrsg.), *Handbook of partial least squares: Concepts, methods and applications* (S. 713–735). Berlin, Heidelberg: Springer Berlin Heidelberg. Zugriff auf https://doi.org/10.1007/978-3-540-32827-8_31 doi: 10.1007/978-3-540-32827-8_31

Henseler, J., Fassott, G., Dijkstra, T. K. & Wilson, B. (2012). Analysing quadratic effects of formative constructs by means of variance-based structural equation modelling. *European Journal of Information Systems*, *21* (1), 99–112.

Hult, G. T. M., Reimann, M. & Schilke, O. (2009). Worldwide faculty perceptions of marketing journals: rankings, trends, comparisons, and segmentations. *Trends, Comparisons, and Segmentations*, *3* (3), 1–23.

Jean, R.-J. B., Sinkovics, R. R. & Hiebaum, T. P. (2014). The effects of supplier involvement and knowledge protection on product innovation in customer–supplier relationships: A study of global automotive suppliers in china. *Journal of Product Innovation Management*, *31* (1), 98–113.

Johnston, W. J., Khalil, S., Jain, M. & Cheng, J. M.-S. (2012). Determinants of joint action in international channels of distribution: The moderating role of psychic distance. *Journal of International Marketing*, *20* (3), 34–49.

Kenny, D. A. (2018). *Moderation.* Zugriff auf http://davidakenny.net/cm/moderation.htm

Köhler, C. F., Rohm, A. J., De Ruyter, K. & Wetzels, M. (2011). Return on interactivity: The impact of online agents on newcomer adjustment. *Journal of Marketing*, *75* (2), 93–108.

Krishen, A. S., Leenders, M. A., Muthaly, S., Ziółkowska, M. & LaTour, M. S. (2018). Social networking from a social capital perspective: a cross-cultural analysis. *European Journal of Marketing*, *53* (6), 1234–1253.

Kumar, P., Singh, S. K., Pereira, V. & Leonidou, E. (2020). Cause-related marketing and service innovation in emerging country healthcare:

Role of service flexibility and service climate. *International Marketing Review*, *37* (5), 803–827.

Lam, S. K., Ahearne, M. & Schillewaert, N. (2012). A multinational examination of the symbolic–instrumental framework of consumer–brand identification. *Journal of International Business Studies*, *43* (3), 306–331.

Li, M., Sharp, B. M., Bergh, D. D. & Vandenberg, R. (2019). Statistical and methodological myths and urban legends in strategic management research: The case of moderation analysis. *European Management Review*, *16* (1), 209–220.

Little, T. D., Bovaird, J. A. & Widaman, K. F. (2006). On the merits of orthogonalizing powered and product terms: Implications for modeling interactions among latent variables. *Structural Equation Modeling*, *13* (4), 497–519.

M., E. (2014). *Metaanalyse* (1. Aufl.). München, Rainer Hampp Verlag.

Mahr, D., Lievens, A. & Blazevic, V. (2014). The value of customer cocreated knowledge during the innovation process. *Journal of Product Innovation Management*, *31* (3), 599–615.

Matsuno, K., Zhu, Z. & Rice, M. P. (2014). Innovation process and outcomes for large j apanese firms: Roles of entrepreneurial proclivity and customer equity. *Journal of Product Innovation Management*, *31* (5), 1106–1124.

Memon, M. A., Cheah, J.-H., Ramayah, T., Ting, H., Chuah, F. & Cham, T. H. (2019). Moderation analysis: issues and guidelines. *Journal of Applied Structural Equation Modeling*, *3* (1), 1–11.

Miao, C. F. & Evans, K. R. (2013). The interactive effects of sales control systems on salesperson performance: a job demands–resources perspective. *Journal of the Academy of Marketing Science*, *41*, 73–90.

Mullins, R., Agnihotri, R. & Hall, Z. (2020). The ambidextrous sales force: aligning salesperson polychronicity and selling contexts for sales-service behaviors and customer value. *Journal of Service Research*, *23* (1), 33–52.

Nakata, C., Zhu, Z. & Izberk-Bilgin, E. (2011). Integrating marketing and information services functions: a complementarity and competence perspective. *Journal of the Academy of Marketing Science*, *39* (5), 700–716.

Pelser, J., de Ruyter, K., Wetzels, M., Grewal, D., Cox, D. & van Beuningen, J. (2015). B2B channel partner programs: Disentangling indebtedness from gratitude. *Journal of Retailing*, *41* (9), 660–678.

Rollins, M., Bellenger, D. N. & Johnston, W. J. (2012). Does customer information usage improve a firm's performance in business-to-business markets? *Industrial Marketing Management*, *41* (6), 984–994.

Rothstein, H. R., Sutton, A. J. & Borenstein, M. (2005). *Publication bias in Meta-Analysis: Prevention, assessment and adjustments*. New Jersey, John Wiley & Sons Ltd.

Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O. & Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies! *Journal of Business Research*, *69*, 3998–4010.

Sarstedt, M., Hair Jr, J. F., Nitzl, C., Ringle, C. M. & Howard, M. C. (2020). Beyond a tandem analysis of SEM and PROCESS: use of PLS-SEM for mediation analyses! *International Journal of Market Research*, *62* (3), 288–299.

Sarstedt, M., Ringle, C. M., Cheah, J.-H., Ting, H., Moisescu, O. I. & Radomir, L. (2020). Structural model robustness checks in PLS-SEM. *Tourism Economics*, *26* (4), 531–554.

Sharma, R. & Jha, M. (2017). Values influencing sustainable consumption behaviour: Exploring the contextual relationship. *Journal of Business Research*, *76*, 77–88.

Siahtiri, V. (2018). Innovation at the service encounter in knowledge intensive business services: antecedents and boundary conditions. *Journal of Product Innovation Management*, *35* (5), 742–762.

Teller, C., Alexander, A. & Floh, A. (2016). The impact of competition and cooperation on the performance of a retail agglomeration and its stores. *Industrial Marketing Management*, *52*, 6–17.

Teller, C., Gittenberger, E. & Schnedlitz, P. (2013). Cognitive age and grocery-store patronage by elderly shoppers. *Journal of Marketing Management*, *29* (3-4), 317–337.

van der Borgh, M. & Schepers, J. J. (2014). Do retailers really profit from ambidextrous managers? The impact of frontline mechanisms on new and existing product selling performance. *Journal of Product Innovation Management*, *31* (4), 710–727.

Wold, H. (1975). Path models with latent variables: The NIPALS approach, InternationalPerspectives on Mathematical and Statistical Modeling. In *Quantitative sociology* (S. 307–357).

Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. G. Jöreskog & H. Wold (Hrsg.), *Systems under direct observations: Part II* (S. 1–54). Amsterdam, North-Holland.

Wolter, J. S. & Cronin, J. J. (2016). Re-conceptualizing cognitive and affective customer–company identification: the role of self-motives and different customer-based outcomes. *Journal of the Academy of Marketing Science*, *44* (3), 397–413.

Zobel, A.-K. (2017). Benefiting from open innovation: A multidimensional model of absorptive capacity. *Journal of Product Innovation Management*, *34* (3), 269–288.

Junior Management Science 7(4) (2022) 932-944

Junior Management Science

journal homepage: www.jums.academy

The normalizing power of digitalization in the world of work – An analysis based on Michel Foucault's theory of power

Die normalisierende Macht der Digitalisierung in der Arbeitswelt – Eine Analyse nach Michel Foucaults Machttheorie

Sebestyén András Huszár

Freie Universität Berlin

Abstract

Digitalization – a power that shapes the whole world and, in particular the world of work anew. Digitalization brings new rules, creates new conditions and establishes new standards. It has a strong normalizing effect. As Michel Foucault describes it in his work Discipline and Punish, normalizing is one of the predominant forms of power. Foucault's theory of power provides a perspective particularly suited to understand the processes of digitalization. The aim of this paper is to identify the most prominent effects of the normalizing power of digitalization in the world of work. In order to reach this goal, a profound qualitative content analysis is conducted, where articles from the German news magazines Die Zeit and Der Spiegel are analyzed using the method of concept coding. The result of the paper is a comprehensive category system, which describes the complex phenomenon, how digitalization normalizes the world of work. Digitalization converts the local nature of the world of work into global (Globaler Charakter), creates flexible working conditions (Flexibilität) and introduces a new level of connectedness and surveillance (Transparenz und Kontrolle). The paper provides a detailed discussion of the main components of these three categories. Results of the paper show, that digitalization changed the world of work fundamentally through normalization, and employees encounter novel challenges and opportunities.

Zusammenfassung

Digitalisierung – eine Kraft, die die ganze Welt und im Besonderen die Arbeitswelt umformt. Digitalisierung bringt neue Regeln, kreiert neue Bedingungen und schafft neue Standards. Sie hat eine starke normalisierende Wirkung. Normalisierung ist eine vorherrschende Form von Macht, wie Michel Foucault es in seiner Arbeit Überwachen und Strafen beschreibt. Foucaults Machttheorie bietet eine Perspektive, die besonders geeignet ist, die Prozesse von Digitalisierung zu verstehen. Ziel dieser Arbeit ist, die bedeutsamsten Auswirkungen der normalisierenden Macht der Digitalisierung in der Arbeitswelt zu thematisieren. Um dieses Ziel zu erreichen, wird eine tiefgreifende qualitative Inhaltsanalyse durchgeführt, indem Artikel von den Zeitschriften Die Zeit und Der Spiegel mit Hilfe der Methode Konzeptcodierung analysiert werden. Das Ergebnis der Arbeit ist ein umfassendes Kategoriensystem, welches das komplexe Phänomen beschreibt, inwiefern Digitalisierung die Arbeitswelt normalisiert. Digitalisierung wandelt die lokale Reichweite der Arbeitswelt ins Globale um (Globaler Charakter), flexibilisiert die Arbeitsbedingungen (Flexibilität) und leitet ein neues Niveau von Vernetzung und Überwachung ein (Transparenz und Kontrolle). Die Arbeit beschreibt die Hauptkomponenten dieser drei Kategorien im Detail. Die Ergebnisse der Arbeit zeigen, dass Digitalisierung die Arbeitswelt durch Normalisierung fundamental verändert und Arbeitnehmer neuartigen Herausforderungen und Angeboten begegnen.

Keywords: Digitalisierung; Foucault; normalisierende Macht; Arbeitnehmer.

1. Einleitung

„Zusammen mit der Überwachung wird am Ende des klassischen Zeitalters die Normalisierung zu einem der großen

DOI: https://doi.org/10.5282/jums/v7i4pp932-944

Machtinstrumente." schreibt Michel Foucault im Jahr 1977 in seinem revolutionären Buch *Überwachen und Strafen* (Foucault, 2019, S. 237)[1]. Heute, im digitalen Zeitalter, erleben wir, dass Digitalisierung in beiden Aspekten, Überwachung und Normalisierung, einen vorher unvorstellbar weiten Umfang verwirklicht. Die digitale Revolution wirkt sich auf alle Lebensbereiche aus. Sie verändert die Arbeitswelt und trägt zur Entwicklung von Trend bei, die die bisherigen Vorstellungen über die Arbeitswelt gewandelt haben (Rinne & Zimmermann, 2016).

Die Motivation hinter dieser Bachelorarbeit ist die Zusammensetzung dieser beiden Ideen – einerseits das Interesse an der Untersuchung der Digitalisierung der Arbeitswelt und andererseits den Abgleich mit der Theorie Michel Foucaults in Bezug auf die Digitalisierung. Die Wirkungen der Digitalisierung auf die Arbeitswelt werden im Zusammenhang mit Foucaults Machtbegriff der Disziplinarmacht, einer produktiven und normalisierenden Macht, betrachtet. Die Digitalisierung in Verbindung mit Foucaults Machtphilosophie wurde bereits mehrfach untersucht; jedoch fokussierten sich diese Arbeiten auf das Konzept der Bio-Macht statt der normalisierenden Macht (Šlesingerová, 2018) oder anstatt auf die Arbeitswelt auf die Armee und die Gesellschaft (Krasmann, 2017; Lifková, 2019; Maddalena & Packer, 2015).

Die vorliegende Arbeit beschäftigt sich mit der normalisierenden Macht der Digitalisierung, um umfangreiche Kenntnisse zu den Auswirkungen dieser Form der Macht auf die Arbeitswelt zu erlangen. Im Fokus der Arbeit steht der Arbeitnehmer, und deren Effekte der normalisierenden Macht der Digitalisierung, die auf den Arbeitnehmer wirken und wahrgenommen wurden.

Das Ziel der Forschung ist es herauszufinden, welche Erscheinungsformen die normalisierende Macht der Digitalisierung in der Arbeitswelt hat, und zu untersuchen, wie diese in der deutschen Presse präsentiert werden. Gesucht wird ein Überblick, wie die Digitalisierung die Arbeitswelt durch Normalisierung verändert hat.

In dieser Arbeit wird eine qualitative Inhaltsanalyse verwendet, um Artikel aus den Zeitschriften *Der Spiegel* und *Die Zeit* zu untersuchen. Durch Kodieren werden die Hauptmotive (Schlüsselthemen) der Artikel identifiziert und Ziel ist eine Ausgestaltung von Kategorien zu gewinnen, die die Hauptrichtungen des Wandels beleuchten und die Auswirkungen der normalisierenden Macht zusammenfassen. Es wurde eine qualitative Analyse gewählt, um zugrunde liegende Themen zu verstehen und erfassen, unter Beibehaltung der Möglichkeit einer interpretativen Auswertung.

Die vorliegende Arbeit ist folgendermaßen aufgebaut: Das erste Kapitel bildet den theoretischen Rahmen, stellt die Disziplinarmacht nach Foucault vor und gibt eine Übersicht über die aktuelle Literatur bezüglich Digitalisierung und Foucault. Anschließend wird die Datenerhebung und -analyse im nächsten Kapitel dargestellt (Kapitel 3), zusammen mit dem Rahmen der Inhaltsanalyse und der Codierungsmethode. Folglich wird der empirische Teil der Studie eingeleitet, in dem die Ergebnisse der inhaltsanalytischen Untersuchung präsentiert werden (Kapitel 4). Die Forschung identifiziert drei Hauptrichtungen (Kategorien), die zusammenfassen, wie die normalisierende Macht der Digitalisierung die Arbeitswelt durch neue Normen verändert hat. Diese drei Kategorien sind *Globaler Charakter*, *Flexibilität* sowie *Transparenz und Kontrolle;* im empirischen Kapitel erfolgt deren detaillierte Erklärung. Die Bachelorarbeit schließt mit den Schlussfolgerungen, in denen die Hauptergebnisse zusammengefasst werden und die Grenzen der Arbeit neben Empfehlungen für weiterführende Forschungen erwähnt werden.

2. Theoretischer Rahmen

2.1. Disziplinarmacht nach Foucault

Seit Mitte des zwanzigsten Jahrhunderts spielt der Machtbegriff von Michel Foucault eine zentrale Rolle in der modernen Politik, Soziologie und Philosophie. Foucault hat unter anderem die Relation zwischen Macht und Wissen sowie die Konstruktion und Disziplinierung von Subjekten untersucht. In dieser Arbeit beschäftige ich mich mit der Disziplinarmacht der Digitalisierung. Um die normalisierenden Effekte der Digitalisierung auf die Arbeitswelt zu betrachten, müssen wir zuerst Foucaults Machtbegriff betrachten.

Nach Foucaults Modell basiert Macht auf Beziehungen, es ist eher verbreitet als zentral (Detel, 1998). Foucault sieht Macht als etwas Dynamisches, was nicht von jemanden ausgeübt, sondern durch soziale Netzwerke verteilt wurde. Macht ist anonym und unmittelbar bevorstehend. Sie ist überall und wird durch Personen, Institutionen, Strukturen und Objekte reproduziert (Rothfusz, 2012).

Vor Foucault wurde Macht traditionell als repressiv verstanden. Laut ihm kann Macht, in diesem Sinne, nur negativ betrachtet werden, etwas, was mit Ausgrenzungsmechanismen operiert. Macht ist das, was 'nein' sagt (Allen, 2002). Dagegen argumentiert Foucault, dass Macht produktiv und integrierend ist: „In Wirklichkeit ist die Macht produktiv; und sie produziert Wirkliches. Sie produziert Gegenstandsbereiche und Wahrheitsrituale: das Individuum und seine Erkenntnis sind Ergebnisse dieser Produktion." (Foucault, 1977, S. 250) Die realen Wirkungen der Macht, die das Verhalten oder Gesten beeinflussen, nennt Foucault Mikroprozesse. Das Individuum ist Resultat dieser Mikroprozesse; es existiert nicht als Gegenüber der Macht, sondern es wird durch die Macht gestaltet (Thomas, 2009).

Disziplinarmacht passt menschliches Verhalten an soziale Regeln an. Sie formt die Individuen zu ‚gelehrigen' und produktiven Mitteln, damit sie die Macht in der Gesellschaft weiterzirkulieren (Messerschmidt, 2011). Diese „gelehrige[n] Körper" bezeichnet Foucault als Körper, die unterworfen, umgeformt und ausgenutzt werden können.

[1] Foucaults Werk *Überwachen und Strafen* erschien im Original 1977 (Surveiller et punir). Um den zeitlichen Kontext deutlich zu machen werde ich, deshalb im Folgenden nach dieser Jahreszahl zitieren. Es sei aber darauf hingewiesen, dass für die Arbeit die 21. Auflage des Suhrkamp Verlags verwendet wurde. Die Seitenangabe für direkte Zitate verweist der wissenschaftliche Genauigkeit halber weiterhin auf die verwendete 21. Auflage hin.

„Der menschliche Körper geht in eine Machtmaschinerie ein, die ihn durchdringt, zergliedert und wieder zusammengesetzt." (Foucault, 1977, S. 176) Die Disziplinierung steigert und schwächt die Kräfte des Körpers, damit sie dessen Nützlichkeit erhöht, aber ihn auch fügsam macht.

Folglich ist die Zielsetzung der Disziplinarmacht Normalisierung. Sie normalisiert das Leben der Menschen in dem Maß, dass sie alle Aspekte dessen innerlich durchdringt. Diese Techniken sind präzise und oft unauffällig, sie bestimmen eine politische und ausführliche Beschäftigung des Körpers, was Foucault ein neue „Mikrophysik der Macht" nennt (Foucault, 1977). Im Gegensatz zu der traditionellen Repressionsmacht funktioniert Disziplinarmacht präventiv. Die Integration bedeutet nicht nur die Standardisierung von Repräsentationen, sondern auch die Synthetisierung von bestimmten Anschauungen (Pongratz, 2013).

Foucault war interessiert an folgenden Fragen: Wie werden durch Konformität zu einem System von Regeln bestimmte Aktionen ermöglicht? Welche Regelsysteme sind diese tatsächlich? Welche historischen Brüche und Wandlungen sind in diesen Regelsystemen beobachtbar? Durch seine Untersuchung kommt er zu der Folgerung, dass diese Regeln produktiv sind, da sie neue Handlungsräume eröffnen und definieren (Detel, 2005).

2.2. Digitalisierung und Foucault

Digitale Technologien haben unsere Welt und insbesondere die Arbeitswelt verändert. Seit ein paar Jahren leben wir im digitalen Zeitalter. Konstante Überwachung, komplett computergesteuerte Fabriken und intelligente Roboter sind nicht mehr bloße Vision von Science-Fiction-Autoren, sondern eine Realität, die viele Arbeitnehmer täglich erleben. Es wird angenommen, dass es keinen Bereich des menschlichen Lebens gibt, der von der Digitalisierung unberührt bleibt.

Digitalisierung definieren Staab und Nachtwey (2016) als die Verbreitung von datenverarbeitenden Prozessen und den zugehörigen Anwendungen. Was es erschwert, den Begriff ‚Digitalisierung' umfassend zu verstehen, ist, dass diese Verbreitung (was Staab und Nachtwey in der Definition erwähnen) universell und umfassend ist. Digitalisierung ist überall präsent. Der einzige Paradigmenwechsel in der menschlichen Geschichte, der in Größe und allumfassender Eigenschaft mit Digitalisierung vergleichbar sein kann, war die Industrialisierung.

Digitalisierte Welt als Zustand weist ein neues Regelsystem auf und Digitalisierung als Prozess ist eine drastische Wandlung des bestehenden Regelsystems. Digitalisierung definiert und normalisiert neue Handlungsräume; deswegen ist es geeignet, die Effekte der Digitalisierung durch eine Foucaultsche Perspektive zu untersuchen. Machtbeziehungen verändern sich mit digitalen Technologien und ihre extensive Verteilung unterstützt Foucaults Konzeption von omnipräsenter Macht.

Hierbei muss erwähnt werden, dass digitale Technologien nicht nur passive Elemente unserer Umgebung sind; vielmehr nehmen sie häufig eine aktive Rolle ein und interagieren responsiv mit Menschen. Selbstlernende ADM-Systeme (Algorithmic Decision-Making Systems) sind geeignete Beispiele dafür. Die Entscheidungen von ADM-Systemen können sowohl einzelne Menschen als auch die ganze Gesellschaft betreffen (Zweig, 2019). Das Ausmaß des Einflusses kann sehr unterschiedlich sein, aber wie die Algorithmen der Google-Suchmaschine oder Facebook Newsfeed sowie die Algorithmen, die Terroristenidentifikation erzielen, sind ähnlich in dem Sinne, dass sie Macht produzieren.

Foucaults Machtbegriff ist auf mehreren Ebenen angemessen, um verschiedene Aspekte der Digitalisierung zu betrachten, und ich bin überzeugt, dass solche Forschungen ergiebig sein können. Obwohl diese Annäherung in der wissenschaftlichen Literatur zum Thema Digitalisierung wenig populär ist, existieren schon relevante Arbeiten. Zunächst möchte ich diese Arbeiten anführen und verdeutlichen, inwiefern meine Forschung von diesen abweicht.

Šlesingerová (2018) schreibt in ihrem Artikel ‚*Biopower imagined: Biotechnological art and life engineering*' über einen spezifischen biotechnologischen Bereich, Bio-Art. Kunstwerke, die mit Biotechnologie und lebendem Gewebe kreiert wurden, werden Bio-Art genannt. Šlesingerová verwendet Foucaults Konzept der biologischen Kontinuität und einen anderen Machtbegriff (der auch von Foucault definiert wurde) – Bio-Macht –, um die Effekte von kontemporären Biotechnologien und technologischer Rationalität auf Lebenskonzepte zu untersuchen. Als Bio-Macht bezeichnet Foucault Machttechniken, die „nicht auf den Einzelnen, sondern auf die gesamte Bevölkerung zielen." (Foucault, 2005, S. 230) Ein weiterer Schwerpunkt des Textes ist die Überwachung menschlichen Lebens in einer biopolitischen und genetischen Form. Weiterhin werden die Grenzen des Körpers, künstliches und synthetisches Leben sowie der Zusammenhang zwischen Bio-Art und Foucaults Bio-Macht behandelt.

Durch Šlesingerovás Artikel wird ersichtlich, wie relevant Foucaults Konzeption von Bio-Macht heutzutage ist. Jedoch werde ich mich in der vorliegenden Arbeit eher auf einen anderen Begriff, Disziplinarmacht, fokussieren. Bio-Macht orientiert sich an Kontrolle und Regulation der Population, wohingegen Disziplinarmacht auf der Ebene des Individuums wirkt. Sie produziert individualisierende Effekte und manipuliert den Körper, damit er nützlich und ‚gelehrig' wird (Foucault, 1977).

In ihrem Artikel ‚*The Digital Body: Telegraphy as Discourse Network*' verbinden Maddalena und Packer (2015) Foucault bereits mit Digitalisierung, obwohl durch einen ziemlich mittelbaren Weg. Ihre Arbeit betrachtet die Fahnentelegrafie, die beim ‚United States Army'-Signal-Corps während des Amerikanischen Bürgerkrieges als Vorreiter der digitalisierten Kommunikation verwendet wurde. Die Autoren argumentieren, dass die Körper der Soldaten durch Disziplinierung (im Foucaultschen Sinne) in ein Medium umgeformt werden. Mit anderen Worten ist der ‚Signal Corps'-Soldat als gelehriger Körper eine Form proto-technischen Mediums.

Hier wurde Disziplinarmacht, als produktive und normalisierende Macht, als ein Schlüsselfaktor auf dem Weg zur digitalisierten Kommunikation berücksichtigt. Dies ist erheb-

lich, da einerseits menschliche Normalisierung eine wesentliche Rolle in der Evolution von Digitalisierung gespielt hat, andererseits übt die derzeit universell verbreitete Digitalisierung normalisierende Effekte auf unser Leben aus. Die Begriffe Normalisierung und Digitalisierung sind historisch nicht separiert und eher verflochten; dementsprechend ist es zweckmäßig, Digitalisierung mit Normalisierung zusammen zu interpretieren.

In *‚Digital Power: Self-tracking Technologies through Michel Foucault Lens‘* betrachtet Lifková (2019) die Veränderungen der Machtbeziehungen aufgrund der Fortschritte in Bezug auf digitale Technologien. Digitalisierung erlaubt (und verursacht) eine Verbreitung von Macht in alle Richtungen und dieser Zustand entspricht exakt Foucaults Behauptungen. Der Text erklärt, wie selbstverfolgende Technologien Foucaults Panopticon[2] ähneln, da es ihnen gelingt, die Individuen an ihrer Selbstüberwachung teilnehmen zu lassen. Externe Überwachung wird verinnerlicht durch Selbstdisziplin, was Geräte wie eine Smartwatch oder ein Smartphone induzieren.

Lifková argumentiert, dass Foucaults Machttheorie deswegen geeignet ist, Digitalisierung zu interpretieren, da sie diese Technologien als einen digitalen panoptischen Machtapparat ansieht. Disziplinarmacht operiert auch durch digitale Werkzeuge, und sie überzeugt Individuen davon, sich Normen anzupassen. Macht zirkuliert durch Normalisierung, wo Standards ausgelegt sind, und folglich werden Individuen durch diese Normen kontrolliert.

Krasmann (2017) definiert das „digital subject" in ihrer wissenschaftlichen Arbeit *'Imagining Foucault. On the Digital Subject and ‚Visual Citizenship‘* als bezeichnenden Charakter unserer Zeit. Sie präsentiert eine neue Annäherung zu Foucaults *Überwachen und Strafen*, in deren Zentrum Visibilität steht. Wie wir in der digitalen Welt von Computern uns und die Welt sehen, was sichtbar ist und was ist unsichtbar, bleibt unbekannt. Sichtbar werden zu lassen und kennenzulernen ist sowohl eine Voraussetzung als auch ein Werkzeug, um Macht auszuüben.

Krasmann analysiert Systeme von Visibilität und Überwachung in der datengesteuerten Gesellschaft des 21. Jahrhunderts. Sie erklärt, dass das „digital subject" die Realität verdoppelt, da es „data doubles" produziert. Im Gegensatz zu der Disziplinargesellschaft, in der das Individuum primäres Interesse war, stehen in der digitalisierten Gesellschaft Muster von Verhalten und Bewegungen von Daten im Fokus. Diese werden gesammelt, um die Zukunft zu prognostizieren oder künftige Möglichkeiten zu gestalten. Die Macht hinter Computern und hinter diesen Prozessen der Datenverarbeitung ist sogar mehr anonym, sie operiert unbemerkt.

Aufgrund der Verbreitung von digitalen Technologien ist die Macht nicht in der Umgebung versteckt (wie es in Foucaults Gefängnismodell war). Eher ist die Umgebung selbst auch Ursprung von Macht, da, wie wir es anhand des Beispiels von ADM-Systemen gesehen haben, sie Daten verarbeitet und Informationen produziert. Somit wurde die normalisierende Macht der Digitalisierung auch in Wechselwirkung mit Technologien und Algorithmen verbreitet und verwirklicht.

Nach Foucault operiert Macht auf allen Ebenen der Gesellschaft und er hat Studien von Macht im Alltagsleben durchgeführt. Diese hat Foucault ‚Mikrophysik der Macht' genannt, in der er sich statt auf die Perspektive des Staates auf die Mikroebenen der Gesellschaft konzentriert. Er betrachtete dabei, wie diese Mikro-Mächte durch die Sprache, Wissen und Bewusstsein durchdringen und neue Wahrheiten schaffen.

In meiner Forschung wähle ich ein spezifisches Segment des Alltagslebens, die Arbeitswelt, und recherchiere die Effekte und Erscheinungsformen der normalisierenden Macht der Digitalisierung. Digitale Technologien haben die Arbeit neu definiert und derart auch das Alltagsleben von Menschen verändert sowie es mit neuen Konzepten und Normen verbunden. In dieser Arbeit beabsichtige ich, ein zusammenfassendes Bild über die Repräsentation solcher Normen in bestimmten deutschen Zeitschriften zu erlangen.

3. Methodik

3.1. Forschungsdesign

Um herauszufinden, welche Effekte die normalisierende Macht der Digitalisierung auf die Arbeitswelt und die Arbeitnehmer hat, wurde eine qualitative Untersuchung durchgeführt. Zu diesem Zweck wurden Zeitungsartikel untersucht, da Presseorgane eine wesentliche Rolle spielen in der Repräsentation und gleichermaßen bei der Gestaltung der öffentlichen Meinung (Lohmeyer, 2018). Durch diese Repräsentation beeinflussen Zeitschriften die Interpretation von Problemen sowie Phänomenen und normalisieren bestimmte Perspektiven (Hoffman, 1999). Dadurch produzieren Presseorgane Macht; insbesondere im Foucaultschen Sinne standardisieren sie Repräsentationen und synthetisieren Anschauungen.

Damit die gesammelten Daten einen repräsentierenden Charakter aufweisen, müssen sie aus Zeitungen oder Zeitschriften stammen, die einen möglichst großen Teil der Bevölkerung erreichen. Deswegen habe ich in meiner Forschung Artikel von zwei populären Akteuren der deutschen Mainstreampresse gewählt. Diese Akteure sind die liberale Wochenzeitung *Die Zeit* und die umsatzstärkste Publikumszeitschrift in Deutschland *Der Spiegel*. Beide Zeitungen gelten als

[2]Das disziplinarische Panopticon (das perfekte Gefängnis) ist ein Konzept, was ursprünglich vom englischen Philosophen Jeremy Bentham kreiert wurde, als eine Architekturform für bestimmte Institutionen wie Gefängnisse und Schulen. Im Panopticon sind die Zellen strahlenförmig angeordnet und in der Mitte steht ein Wachturm. Von dem Wachturm aus sind die Zellen zu jeder Zeit einsichtig, aber die Lichtverhältnisse im Turm sind speziell, damit die Beobachteten den Beobachter nicht sehen können. Das Hauptidee hinter diesem Konzept ist die Möglichkeit ständiger Überwachung, was eine Verhaltensverbesserung der Zelleninsassen erzielt (Altmeyer, 2011). Nach Foucault ist der Panoptismus ein Machtphänomen, da das Machtverhältnis aufgrund dieses Wissens internalisiert wird, und dementsprechend „übernimmt [der Überwachte] die Zwangsmittel der Macht, und spielt sie gegen sich selber aus". Der Überwachte wird das Prinzip seiner Unterwerfung (Foucault, 1977, S. 260).

deutschsprachige Leitmedien und beeinflussen als solche vorherrschend die öffentliche Meinungs- und Themenbildung. Sie etablieren Konsense und gestalten gesellschaftliche Kommunikation und Öffentlichkeit (Ziemann, 2011).

Alle ausgewerteten Artikel wurden im Zeitraum von 1995 bis 2018 publiziert. Für die Artikelauswahl (*Zeit Online*, *Spiegel Online*, *Die Zeit* und *Der Spiegel*) wurde die Datenbank WISO verwendet. Als Suchkriterien wurde nach Schlagworten in den Texten gesucht. Hauptsächlich genaue Schlagworte, wie ‚digitale Revolution' (132 Treffer), ‚Arbeitswelt' (292 Treffer) und ‚Arbeitnehmer' (237 Treffer) wurde angewandt, aber auch vagere, wie ‚Arbeit*' (3301 Treffer) und ‚Digitalisierung' (408 Treffer). Das Ziel war eine Sammlung von Artikeln zu erstellen, die die Auswirkungen der Digitalisierung für die Arbeitswelt in Deutschland thematisieren. Artikel mit weniger als 150 Wörtern und doppelt aufgeführte Artikel (manchmal erschien der gleiche Artikel sowohl analog als auch digital) wurden ausgeschlossen. Mit dieser Methode wurden aus den beiden Zeitschriften insgesamt 210 Artikel gesammelt (87 Artikel aus *Der Spiegel* und 123 Artikel aus *Die Zeit*).

Die Untersuchung orientiert sich an dem Arbeitnehmer, insbesondere an der Perspektive des Arbeitnehmers, und beobachtet die Konsequenzen der Digitalisierung (und die Erscheinungsformen deren Disziplinarmacht); deswegen wurden einige Artikel nicht in die Beobachtung eingeschlossen. Die beiden häufigsten Gründe für einen solchen Ausschluss waren ein fehlender (oder nicht hinreichender) Fokus auf die Arbeitswelt und die Arbeitnehmer sowie ein prognostizierender Charakter der Artikel. Die erste Gruppe beinhaltete Artikel, die zwar von der digitalen Revolution und Digitalisierung handelten, aber weniger von der Arbeitswelt. In ihrem Fokus stand meist Politik (auch spezifische Parteien), Geschichte oder Geschichte der Digitalisierung sowie Kinder und Schulen. Das andere entscheidende Ausschlusskriterium war, ob der Artikel eine Zukunftsprognose beschreibt. In meiner Beobachtung habe ich tatsächliche Phänomene gesucht, die wirklich oder mindestens existierend sind; deswegen konnte ich Ideen von bestimmten Autoren darüber, wie das Leben im Jahr 2040 oder 2050 sein wird, nicht einbeziehen. Durch die Anwendung der Ausschlusskriterien wurden 78 Artikel ausgegrenzt, während insgesamt 132 Artikel untersucht wurden.

Da ich sprachliches Material analysiere, das aus einer Form von Kommunikation stammt, habe ich als Art der Forschung die qualitative Inhaltsanalyse gewählt. Diese Vorgehensweise, um fixierte Kommunikation (in diesem Fall Zeitungsartikel) zu analysieren, ist besonders geeignet für diese Forschung, weil sie systematisch und regelgeleitet vorgeht. Eine weitere relevante Eigenschaft einer Inhaltsanalyse ist, dass sie theoriegeleitet operiert, somit der zu untersuchende Gegenstand an Erfahrungen anderer angeknüpft wird (Mayring, 2015). Foucaults Theorie von Disziplinarmacht und der normalisierende Charakter der Macht bilden den Theoriehintergrund für die Interpretation der Ergebnisse, da es Ziel der Arbeit ist, tatsächliche normalisierende Wirkungen der Digitalisierung (Mikroprozesse der Macht nach Foucault) in der Arbeitswelt zu betrachten.

Ziel der Analyse ist die Reduktion des Materials, um ein überschaubares Corpus zu erlangen; dementsprechend ähnelt sie einer zusammenfassenden qualitativen Inhaltsanalyse nach Mayring. Die Artikel wurden zuerst mithilfe der Datenanalysesoftware MAXQDA qualitativ kodiert. Letztere erleichtert die Datenanalyse, da sie unter anderem ermöglicht, die Texte in Gruppen zu ordnen und diese beliebig zu strukturieren, ein hierarchisches System von Codes zu erstellen und managen, zu der Code Memos oder Definitionen hinzufügen, wie auch das Codesystem visuell darzustellen.

Als Methode habe ich Konzeptcodierung gewählt, weil im Fokus der Untersuchung größere, symbolische Themen der Artikel standen. Konzeptcodierung ordnet Inhalte auf der Mesoebene oder den Makroebene des Textes zu Codes. Konzepte sind kurze Aussagen, die eine breitere Bedeutung als eine Aktion oder ein Gegenstand vorschlagen (Saldaña, 2016). Diese Vorgehensweise erlaubt größere interpretative Freiheit, was sich als hilfreich erwies, aufgrund der philosophischen Art des theoretischen Rahmens der Untersuchung.

3.2. Inhaltsanalyse

Im Fokus der Analyse waren solche Trends und Phänomene, die für den Arbeitnehmer die Arbeitswelt fundamental und konzeptuell beeinflussen und gestalten. Zuerst wurden alle Artikel aufmerksam durchgelesen und basierend auf den vorgelegten (und auch emergent identifizierten) Ausschlusskriterien solche aussortiert, die in Bezug auf ihre Thematik nicht unbedingt zu dem Untersuchungsfeld gehörten.

Im ersten Schritt wurden solche Normen, die aufgrund der normalisierenden Macht der Digitalisierung entstanden, in den Artikeln erkannt und mit unbestimmten Codes ausgestattet, ohne zu versuchen, strikte Kategorien zu bilden (Gioia, Corley & Hamilton, 2013). In dieser Phase wurden insgesamt 284 Codes erlangt. Für solches erste Rund Codierung waren Beispiele die Codierung des Satzes: „[...], Arbeitnehmer sind dank Smartphones per Mail, SMS oder Anruf rund um die Uhr erreichbar." (Artikel Nr. 45)[3] mit dem Code „erreichbar rund um die Uhr" oder die Codierung der Aussage „[...] muss jeder vierte Deutsche nach Feierabend für den Chef erreichbar sein. Viele gehen auch im Urlaub ran, wenn die Firma anruft [...]" (Artikel Nr. 155) mit dem Code „Erreichbarkeit auch während Urlaub und Feierabend".

Der zweite Schritt des Codierungsprozesses beinhaltet die Reduzierung der ursprünglich vielfältigen Codemasse. Wir suchen nach signifikanten Differenzen und Ähnlichkeiten, um die inhaltsgleichen Codes zu identifizieren und

[3] In dieser Arbeit benutze ich, wenn ich auf die untersuchten Artikel verweise oder die untersuchten Artikel zitiere, die Form ‚Artikel Nr. X' zur Notation. Um diese Notation zu interpretieren, befindet sich ein Excel-Dokument mit dem Namen Artikel (Die Zeit, Der Spiegel) im Anhang der Bachelorarbeit. Die erste Spalte dieser Excel-Tabelle wird ‚Artikel_ID' genannt; in dieser Spalte sind die Nummern der Artikel in aufsteigender Reihenfolge aufgelistet. Somit weist die in diesem Absatz adressierte Notation ‚Artikel Nr. 45' auf den Artikel „Roboter als Chefs und kaum noch Festanstellungen" hin, der am 6. Oktober 2014 in Die Zeit publiziert wurde.

zusammenzufügen. Die neuen und in ihrer Anzahl deutlich geringeren Codes bekommen neue Namen, die kurzer und exakter, aber trotzdem informativ sind (Gioia et al., 2013). Anhand des Beispiels der oben erwähnten Codierungen funktioniert es derart, dass die Codierungen „erreichbar rund um die Uhr" und „Erreichbarkeit auch während Urlaub und Feierabend" unter einem Code „ständige Erreichbarkeit" zusammengefasst wurden.

Diese Kategorien wurden mit Memos versehen, was MAXQDA ermöglicht. Laut Saldaña (2016) sind Memos persönliche, reflektive Texte oder Kommentare, und ihr Ziel ist es, kritisches Denken anzuregen, Annahmen zu bilden und später gegebenenfalls dieselben Annahmen anzufechten. Memos dienen dazu, Codes präziser zu bezeichnen und erleichtern dementsprechend Textstellen zu Codes anzuordnen.

In Tabelle 1 werden die Ergebnisse dieses zweiten Schrittes dargestellt. Am Ende der zweiten Phase wurde die Anzahl den Codes insgesamt auf zehn reduziert; neben allen Codes wird eine exemplarische Textstelle präsentiert.

Im dritten Schritt wurden, basierend auf den zehn Codierungen in Tabelle 1, Kategorien gebildet, die Tendenzen vertreten. Letztere zeigen, wie die Digitalisierung durch Normalisierung die Arbeitswelt verändert und umgeformt hat. Im Mittelpunkt steht weiterhin der Arbeitnehmer; alle diese Normen haben eine mittelbare oder unmittelbare Wirkung auf den Arbeitnehmer, eine Wirkung, was, gar unbewusst, Trends normalisiert, und demgemäß den Arbeitnehmer der Digitalisierung gestaltet.

Die drei festgelegten Kategorien sind die folgenden: *Globaler Charakter*, *Flexibilität* sowie *Transparenz und Kontrolle*. Unter *Globaler Charakter* gehören solche Codierungen, die deren Aspekt der Digitalisierung umfassen, dass die die Arbeitswelt globalisiert. Die Arbeitnehmer müssen ihre Mentalität und Denkweise dahingehend anpassen, dass die Arbeitswelt, die sie umgebt, nicht mehr nur eine lokale Reichweite hat, sondern eine globale. Als zugehörig zur nächsten Gruppe, *Flexibilität*, wurden solche Codierungen kategorisiert, die die Flexibilisierung der Arbeitsbedingungen darstellen, und sich darauf beziehen, welche Anforderungen die flexible Arbeitswelt an die Arbeitnehmer stellt. Unter der letzten Kategorie, *Transparenz und Kontrolle*, befinden sich solche Codierungen, die auf den mit digitalen Technologien ermöglichten schnellen und totalen Informationszugang hinweisen, und dessen Implikationen, wie die Vernetzung und Überwachung, betrachten. Die inhaltliche Beschreibung des Kategoriensystems erfolgt in Kapitel 4.

4. Normalisierende Macht der Digitalisierung

4.1. Darstellung des Kategoriensystems

Die Beschreibung des analysierten Datencorpus erfolgt sich mithilfe drei Kategorien und zehn Codes. Dieses Code- und Kategoriensystem, mit den Verbindungen zwischen Kategorien und Codes, ist graphisch in Abbildung 1 dargestellt. Als erstes Ergebnis erscheint, dass drei Hauptbereiche (Kategorien) festgestellt wurden, in denen die normalisierenden Wirkungen der Digitalisierung auf der Arbeitswelt zusammengefasst sind. Diese drei Hauptrichtungen für Normen, die das Leben der Arbeitnehmer im digitalen Zeitalter gestalten, sind: *Globaler Charakter*, *Flexibilität* sowie *Transparenz und Kontrolle*. In Abbildung 1 ist ersichtlich, dass zu diesen drei Bereichen verschiedene Phänomene gehören (die kurz auch in Tabelle 1 dargestellt werden).

In den folgenden drei Unterkapiteln werden diese drei Kategorien und die dazugehörigen Codes detailliert beschrieben, und, wo erforderlich, auch die Ähnlichkeiten und Unterschiede zwischen Codes geklärt.

4.2. Globaler Charakter

Diese Erscheinungsform der normalisierenden Macht der Digitalisierung, die durch die Kategorie *Globaler Charakter* gekennzeichnet ist, verweist darauf, wie fundamental sich die Denkweise der Arbeitnehmer mit der Verbreitung der Digitalisierung verändert hat. Die Arbeitnehmer müssen mit einer globalen Denkweise agieren, da aufgrund der Digitalisierung die geographischen Grenzen der Arbeit verschwunden sind, und wo vorher zahlreiche isolierte und zergliederte regionale Märkte waren, existiert jetzt ein globaler Markt. Die Arbeitswelt hat durch die digitalen Technologien ihren lokalen Charakter verloren. Diese globale Reichweite bietet Möglichkeiten, etabliert neue Methoden und birgt Gefahren, aber vor allem, beeinflusst sie Entscheidungen.

Boes et al. (2016) bezeichnen einen „global verfügbaren Informationsraum", eine neue Handlungsebene, deren Grundlage das Internet ist. Diese omnipräsente Handlungsebene (Mobile elektronische Mittel machen es fast überall verfügbar und die Signifikanz des Netzes durchdringt nahezu alle sozialen Sphären.) ermöglicht vielfältige Interaktionen zwischen Menschen weltweit, was auch die Arbeitswelt enorm verändert. Das Internet ist zentral in Bezug darauf, Arbeit zu suchen, anzubieten und zu erledigen: „Nie war es für Unternehmen so einfach, Unterstützung für Projekte zu finden – und nie war es für Selbstständige so einfach, ihr Können anzubieten, wie heute. Im globalen Netz spielen nicht nur Ländergrenzen keine Rolle mehr, hier ist für jeden etwas zu holen [...]." (Artikel Nr. 47)

Drei Codes gehören unter diese Kategorie, *Globaler Wettbewerb*, *Grenzübergreifende Arbeitsmodelle* und *Crowdworker/Crowdsourcing*; die Codes sind Resultate der Wirkungen der Macht der Digitalisierung.

Mit digitalen Technologien ist eine globale Infrastruktur entstanden, was die regionalen Grenzen des Arbeitsmarktes auflöst; der Code *Globaler Wettbewerb* erfasst dieses Phänomen. Das Internet ermöglicht es, ein Arbeitsverhältnis von überall auf der Welt einzugehen, in einem Arbeitsverhältnis zu stehen und es schließlich zu kündigen. Dementsprechend „konkurriert [im Netz] jeder Anbieter mit jedem anderen [...]." (Artikel Nr. 128) Einerseits ergibt sich die Möglichkeit für Arbeitnehmer, ihre potentiellen Kunden oder Arbeitgeber weltweit zu suchen; andererseits stehen sie jedoch unter steigendem Wettbewerbsdruck aufgrund der Globalisierung.

Tabelle 1: Codierleitfaden mit beispielhaften Zitaten. (Quelle: eigene Abbildung)

Code	Codebezeichnung	Beispielhafte Zitat
1	Globaler Wettbewerb	„An den Jobbörsen des Internets können Spezialisten ihre Fähigkeiten weltweit zu Spitzenpreisen anbieten, während digitale Tagelöhner ihre Arbeitskraft zum Dumpingpreis vermarkten müssen." (Artikel Nr. 149)
2	Grenzübergreifende Arbeitsmodelle	„In Datenwolken können Menschen über Kontinente verteilt in Echtzeit an einer Aufgabe arbeiten, als säßen sie in einem Büro." (Artikel Nr. 135)
3	Crowdworker/Crowdsourcing	„Zugleich sind selbstständige Crowdworker als digitale Tagelöhner von Weisungen ihrer Auftraggeber abhängig." (Artikel Nr. 193)
4	Flexible Arbeitszeiten	„Flexibilität: Dass Arbeitnehmer zeitlich flexibler sein sollen, gehört seit Langem zu den Hauptforderungen von Arbeitgeberverbänden." (Artikel Nr. 90)
5	Selbstbestimmte Arbeit	„Menschen wie [...] gehören zu jener Schicht, für die Globalisierung und Digitalisierung eine schöne neue Welt geschaffen haben. Sie können ihr Arbeitsleben weitgehend selbst gestalten [...]" (Artikel Nr. 206)
6	Mobiles Arbeiten	„Denn das Beantworten von E-Mails, Überarbeiten von Projektplänen oder die Organisation der Terminplanung der nächsten Woche ist mittlerweile in vielen Jobs auch in öffentlichen Verkehrsmitteln oder vom heimischen Sofa aus möglich und für viele Mitarbeiter mittlerweile Standard." (Artikel Nr. 97)
7	Totale Überwachung	„[...] das permanente Überwachen und Vergleichen von Mitarbeitern, das in der digitalen Wirtschaft möglich ist [...]" (Artikel Nr. 174)
8	Ständige Erreichbarkeit	„[...] Arbeitnehmer sind dank Smartphones per Mail, SMS oder Anruf rund um die Uhr erreichbar." (Artikel Nr. 45)
9	Vernetzte Unternehmen	„Das System funktioniert nach dem Prinzip der sozialen Netzwerke, es vernetzt die Mitarbeiter miteinander und berücksichtigt deren individuelle Umstände." (Artikel Nr. 149)
10	Alle Informationen zugänglich im Unternehmen	„[...] Ich kann jederzeit einsehen, woran Kollegen arbeiten, habe Zugriff auf ihren Kalender und ihre Projektplanung. Wenn alles für alle offen zugänglich ist, sich kein Machtwissen zurückhalten lässt, wird viel freier über neue Ideen oder Projekte diskutiert." (Artikel Nr. 157)

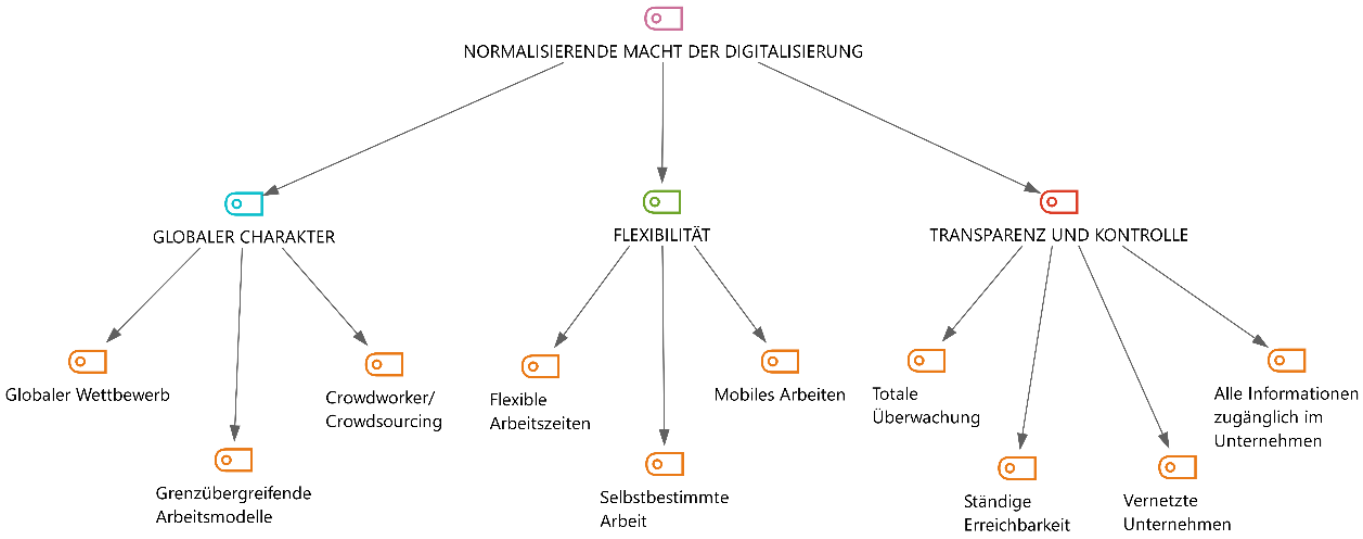


Abbildung 1: Code- und Kategoriensystem (Quelle: eigene Abbildung, gefertigt mit MAXQDA)

Im globalen Wettbewerb ist es zwingend erforderlich, dass Selbstständige wie auch Unternehmen Kundenkontakte weltweit halten, was nur durch die digitalen Informationskanäle, wie soziale Medien, möglich ist. Kunden zu gewinnen und zu verlieren, gestaltet sich in der digitalisierten Welt problemlos: „Was zählt, ist, dass in der Zeit von Smartphones und einer global vernetzten Menschheit die einst ehernen Grundregeln der Wirtschaftswelt nicht mehr gelten [...]."

(Artikel Nr. 147)

Weitere Implikation dieser globalen Konkurrenz ist, dass Arbeitnehmer nicht nur mit anderen ihres Heimatlandes oder der Nachbarländer konkurrieren, sondern auch mit Anbietern anderer Zeitzonen. Daraus resultiert in manchen Fällen eine Neugestaltung der Lebensbedingungen, da die Arbeitszeit auf den globalen Markt ausgerichtet werden soll. Die Nachfrage verändert sich schnell und Anbieter müssen darauf sofort reagieren, da Kunden leicht die Angebote miteinander vergleichen können.

Der Code *Grenzübergreifende Arbeitsmodelle* bezeichnet das Beschäftigungsmodell, dass Mitarbeiter, in einem Unternehmen oder Projekt, zusammenarbeiten (sogar in Echtzeit), obwohl ihre Arbeitsplätze weltweit verteilt sind. Es heißt, dass Kollegen „nicht mehr in einem Betrieb Tür an Tür [arbeiten], sondern weltweit verstreut, verbunden über das Netz." (Artikel Nr. 156) Arbeitsverträge werden nicht mehr auf regionalen Vereinbarungen aufgebaut und die Zusammenarbeit zwischen Arbeitnehmern in unterschiedlichen Zeitzonen wird immer häufiger.

Vorherrschende Form dieser grenzübergreifenden Arbeitsmodelle ist Cloudworking, was arbeiten in einer digitalen Datenwolke bedeutet. Cloudworking beinhaltet alle Formen gesellschaftlicher Arbeit, deren Objekte und Instrumente auf dem Informationsraum basieren und deren Kooperation als soziale Aktion über den Informationsraum vermittelt wird (Boes, Kämpf, Lühr, Langes & Ziegler, 2017). Via Datenwolke wird „digitale Arbeit [...] rund um die Welt verteilt; dorthin, wo sie am besten und oft genug am billigsten verrichtet wird [...]." (Artikel Nr. 149)

Der letzte Code dieser Kategorie ist *Crowdworker / Crowdsourcing* und obwohl Crowdsourcing nach Boes et al. (2016) als Subkategorie von Cloudworking angesehen werden soll, handelt es hier um einen separaten Code. Grund dafür ist, dass der Arbeitertypus ‚Crowdworker' ein geeignetes Beispiel bietet, wie die Disziplinarmacht der Digitalisierung die Arbeitnehmer umformt: „Der menschliche Körper geht in eine Machtmaschinerie ein, die ihn durchdringt, zergliedert und wieder zusammensetzt.", schreibt (Foucault, 1977, S. 176) in *Überwachen und Strafen*, und in diesem Sinne ist der Crowdworker ein Endprodukt des Machtprozesses der Digitalisierung.

Bei Crowdsourcing werden Aufgaben auf eine Menschenmenge (‚Crowd') ausgelagert, es handelt sich um eine Vermittlung von Arbeitskrafteinheiten mithilfe digitaler Plattformen (Staab & Nachtwey, 2016). Auftraggeber und Crowdworker werden durch die Plattform verbunden und das Arbeitsverhältnis dauert nur bis zur Fertigstellung des Auftrags, manchmal „kann man die Jobs in wenigen Minuten abarbeiten." (Artikel Nr. 111) Durch die Digitalisierung kann die Arbeit in beliebig kleine Aufgaben zerlegt werden und das Internet ermöglicht diese Form von Unverbindlichkeit. Mit diesem Modell sind Crowdworker die digitalen Tagelöhner des 21. Jahrhunderts und wie die Textstelle aus Artikel Nr. 77 zeigt, ist ihr Arbeitsverhältnis so unsicher wie das von Tagelöhnern der Industrialisierung: „Es war noch nie so einfach, einen Menschen für eine Arbeit von zehn Minuten einzustellen und dann wieder zu feuern."

Das Einkommen von Crowdworkern variiert stark; den Mindestlohn erreichen sie nicht unbedingt, da ihre Arbeit nicht nach Stunden, sondern nach Produkten oder Projekten bezahlt wird. Crowdworker müssen sich bei der Auftragsvergabe gegen die Konkurrenz durchsetzen; ihr Arbeitsverhältnis ist abhängig von den Weisungen der Auftraggeber und aufgrund der vereinfachten Austauschbarkeit prekär.

4.3. Flexibilität

Die Kategorie *Flexibilität* umfasst Auswirkungen der Digitalisierung und beschreibt, wie Digitalisierung die Arbeitsbedingungen flexibler neugestaltet. Die Arbeitnehmer müssen sich diesen veränderten Bedingungen der digitalen Arbeit anpassen, da Flexibilität nicht nur als Möglichkeit, sondern auch als Anforderung präsent in der Arbeitswelt ist. „Fest steht aber auch, dass die Digitalisierung die Anforderungen an die Mitarbeiter grundlegend verändert: Sie müssen mobiler, flexibler, unabhängiger arbeiten." (Artikel Nr. 204)

Seit den 1970er-Jahren entstanden Bestrebungen, das klassische Modell der Normalarbeit aufzuweichen. Die Versuche, neue Produktivitäts- und Flexibilitätspotentiale zu schaffen, erzielten eine Abweichung von den Rahmenbedingungen des fordistischen Arbeitsmodells, wie die Vollzeitbeschäftigung mit festen Arbeitszeiten oder die formalisierten Arbeitsbeziehungen (Gerber & Krzywdzinski, 2019). Durch die Verbreitung datenverarbeitender Prozesse und Techniken wurde diese Flexibilisierung der Arbeitswelt auf zeitliche, örtliche und entscheidungsmäßige Weise ermöglicht.

Drei Codes zählen zu dieser Kategorie (*Flexible Arbeitszeiten*, *Selbstbestimmte Arbeit* und *Mobiles Arbeiten*); diese repräsentieren die Schwerpunkte der Flexibilisierung der Arbeitsbedingungen.

Der Code *Flexible Arbeitszeiten* umfasst zwei Bedeutungen. Einerseits verweist er auf die Gelegenheit, dass Arbeitnehmer nicht innerhalb standardisierter, fester Arbeitszeiten arbeiten und zeigt andererseits, dass zeitliche Flexibilität oft als Anforderung erscheint. Diese Dualität wird in folgenden Beispielen dargestellt: „Die Arbeitgeber wollen, dass nicht mehr die tägliche Arbeitszeit gezählt wird, sondern die wöchentliche." (Artikel Nr. 54) Diese Aussage impliziert die Bemühungen der Arbeitgeber, die Arbeitstage von mehr als zehn Stunden zu normalisieren, da diese Verlängerung, in diesem Modell, durch kürzere Arbeitstage ausgeglichen werden kann; und „[...] früher nach Hause zu gehen, dort die Kinder ins Bett zu bringen und sich dann nach 20 Uhr wieder an den Rechner zu setzen." (Artikel Nr. 88) Die zweite Aussage stellt ein Beispiel dar, in dem die Flexibilisierung der Arbeitszeiten dem Interesse der Arbeitnehmer dient.

Flexibilisierung bringt eine Erosion der Grenzen zwischen Arbeit und privatem Leben, und die flexible Arbeitszeit ist ein Indikator dafür, dass durch die Digitalisierung eine beliebige Zergliederung möglich ist, während zuvor ein Tag in zwei klar abgegrenzte Segmente getrennt war. Diese Erosion, die durch die Digitalisierung verursacht wurde, erscheint auch in weiteren Codes, wie *Mobiles Arbeiten* oder *Ständige Erreichbarkeit*.

Selbstbestimmte Arbeit bezeichnet sich der wegen Digitalisierung erhöhte Grad von Selbststeuerung, bezüglich der Arbeit. Digitale Technologien erlauben einen Freiheitsfaktor hinsichtlich der individuellen Konfiguration der Arbeitsstruktur. Für den Arbeiter bedeutet dies, dass er selbst bestimmen kann, „[...] was ich arbeite, wann ich arbeite und wie. Das ist im Vergleich zu der herkömmlichen Situation eines Angestellten revolutionär." (Artikel Nr. 71) Im Vergleich zu den anderen Codes, die zur Kategorie *Flexibilität* zählen, liegt die Betonung nicht auf der Möglichkeit, dass die Arbeitnehmer an einem unbestimmten Ort oder zu einer unbestimmten Zeit arbeiten können, sondern auf der Möglichkeit der Entscheidung, dass sie ihre Arbeitsbedingungen selbst definieren können. Diese Tendenz erfordert einen höheren Bedarf an Autonomie von den Arbeitnehmern.

Diese Selbstbestimmung der Arbeit ist besonders charakteristisch bei selbstständigen Arbeitnehmern (Freelancern): „Sie erledigen Projekte, wann und wo immer sie wollen" (Artikel Nr. 204) und neben der freien Auswahl von Ort und Zeit können sie auch zwischen Aufträgen und Projekten selektieren (Das bringt eine gewisse Unsicherheit des Einkommens, allerdings schätzen Viele die Freiheit der Selbstbestimmung mehr als sicheres Einkommen.). Obwohl Selbstbestimmbarkeit typischer bei Freiberuflern ist, wird sie auch in Unternehmen immer häufiger zur Norm gehören. „In Teams dürfen sie selbst entscheiden, wie sie ihre Arbeitskraft und ihre Mittel einsetzen. Jeder hat eine Stimme." (Artikel Nr. 165) Mitarbeiter erhalten eine größere Teilhabe an der Entscheidung, wie sie arbeiten. Auf diese Weise normalisiert die Digitalisierung eine höhere Individualisierbarkeit der Arbeitswelt und erfordert gleichzeitig mehr Kreativität und Verantwortung.

Der Code *Mobiles Arbeiten* verweist auf die örtliche Flexibilisierung der Arbeit. Im digitalen Zeitalter ist die Arbeit nicht mehr ortsgebunden: „Zu Hause am Schreibtisch, auf der Couch oder im Café checken sie Mails, telefonieren mit Kunden oder arbeiten am neuen Projekt." (Artikel Nr. 190) Die Virtualisierung der Aufträge und Aufgaben sowie die Flexibilisierung der Arbeitsmittel erlaubt den Arbeitern, von überall aus zu arbeiten. Elektronische Endgeräte sind einfach transportierbar, somit ist es „seit Computer, Smartphones und Tablets unseren Alltag bestimmen [...] möglich, fast immer und überall zu arbeiten." (Artikel Nr. 90)

Bezüglich dieses Punktes muss ich erwähnen, dass dieser Code für eine lange Zeit während meiner Forschung in zwei separate Codes getrennt war. Der andere Code hieß *Der Arbeitsplatz ist kein ‚Ort'*; Grund für diese Trennung war, dass diese zwei Codes zwei Schlüsselmotive des Phänomens beschrieben. Der eine repräsentierte die Zerstörung des physischen Konzeptes des ‚Arbeitsortes' („Der Arbeitsplatz ist kein klar umgrenzter Ort mehr, stattdessen befindet er sich jeweils dort, wo ein Computer Zugang zum Netz hat." (Artikel Nr. 32)) und der andere stellte die Flexibilisierung der Arbeitsinhalte sowie Arbeitsmittel dar. Ich habe diese beiden Codes vereinigt, da die Auswirkungen der repräsentierten Prozesse dieselben waren, und weil diese beiden Codes zueinander in ihrer Bedeutung deutlich näher waren im Vergleich zu den beiden anderen Codes, die zu dieser Kategorie gehören (*Flexible Arbeitszeiten*, *Selbstbestimmte Arbeit*).

Da der Zugang zum Netz die Hauptvoraussetzung zum Arbeiten ist, erfolgt eine Dezentralisierung der Arbeit. Die Arbeit wird präsent in je mehr Bereiche das Leben der Arbeitnehmer, und das resultiert in der (beim Code *Flexible Arbeitszeiten* bereits betrachteten) Erosion der Grenzen zwischen Arbeit und privatem Leben. Es entsteht eine Gefahr der Verschiebung der Work-Life-Balance, die insbesondere bei einer der beliebtesten Form mobilen Arbeitens, dem Home-Office, als möglich erscheint. Für Arbeiter, die von zu Hause arbeiten, ist eine bewusste Selbstdisziplin erforderlich, um Überstunden und ungesunde Auslastung zu vermeiden. „Je mehr Möglichkeiten sich ergeben, von anderen Orten, zu anderen Zeiten oder mal kurz zwischendurch zu arbeiten, desto höher wird die Versuchung, dass der Arbeitnehmer dazu auch verpflichtet wird oder sich auch nur verpflichtet fühlt." Diese Textstelle von Artikel Nr. 101 beschreibt genau, wie in diesen Fällen die normalisierende Macht wirkt. Es ist nicht unbedingt ein System von Regeln, das die Wirkungen verursacht, vielmehr ist es ein „System von Normalitätsgraden, welche die Zugehörigkeit zu einem homogenen Gesellschaftskörper anzeigen" (Foucault, 1977, S. 237), und solche Normalitätsgrade sind in den Codes, unter der Kategorie *Flexibilität*, zusammenfasst.

Zu dem Code *Mobiles Arbeiten* gehören die flexiblen Büroformen, wie Desk-Sharing, und die Möglichkeit individuell konfigurierbarer Arbeitsplätze. Beide verwirklichen eine Form von Flexibilisierung der traditionellen Arbeitsbedingungen.

Die normalisierende Macht der Digitalisierung hat durch Flexibilisierung auch eine neue Form von Arbeitern entwickelt; dies sind die digitalen Nomaden, die mithilfe von digitalen Technologien unabhängig von bestimmten Orten arbeiten und ein multilokales Leben führen: „Ihr Büro ist der Ort, an dem sie sich gerade [aufhalten]: das Loft in Berlin, die Bar im kolumbianischen Medellín, der Strand von Koh Tao in Thailand." (Artikel Nr. 149)

4.4. Transparenz und Kontrolle

Die letzte Kategorie des Kategoriensystems, *Transparenz und Kontrolle*, beschreibt, wie digitale Technologien eine gänzlich neue Dimension totaler Transparenz und vollkommener Kontrolle im Unternehmen erbringen, und auf diese Weise die Arbeitsumgebung grundlegend verändern. Transparenz beinhaltet einerseits eine gewisse Durchsichtigkeit von Unternehmensprozessen und andererseits eine Nachvollziehbarkeit der Tätigkeiten der Arbeitenden. Diesen Zustand verstärken die Vernetzung von Unternehmen und die digitalen Kontrollsysteme, die lückenlose Überwachung ermöglichen.

Die Wirkungen der Macht können wir hier zweifach beobachten. Die Digitalisierung etabliert bezüglich der Gestaltung der Arbeit neue Normen im Unternehmen; daneben ermöglichen digitale Technologien eine detaillierte Beschreibung der Mitarbeiter und deren Arbeit. Transparenz erleichtert diese Beschreibung und resultiert in Kontrolle und Disziplinierung der Arbeitnehmer: „Die Disziplinarprozeduren [...] machen

aus der Beschreibung ein Mittel der Kontrolle und eine Methode der Beherrschung." (Foucault, 1977, S. 247)

Zur Kategorie *Transparenz und Kontrolle* zählen vier Codes (*Totale Überwachung*, *Ständige Erreichbarkeit*, *Vernetzte Unternehmen* und *Alle Informationen zugänglich im Unternehmen*), die die Wirkungen der normalisierenden Macht der Digitalisierung (auf die Arbeitnehmer) im Rahmen des Unternehmens betrachten.

Der Code *Totale Überwachung* bezeichnet das konstante Monitoring innerhalb eines Unternehmens, das permanente Überwachen und Vergleichen von Mitarbeitern. Mithilfe von digitalen Technologien können fast alle Arbeitsschritte überwacht werden und in dem Prozess werden zusätzlich Daten über die Mitarbeiter gesammelt. Die gesammelten Daten analysieren die Unternehmen, um Tätigkeitsprofile zu erstellen und diese zu vergleichen. Ziel solcher Analysen ist es, eine optimale Auslastung der Mitarbeiter zu verwirklichen und adäquate Aufgaben zu finden: „Wir messen, wie geeignet sie für bestimmte Positionen sind [...]." (Artikel Nr. 41)

Da zunehmend viele Aufträge im digitalen Raum erledigt werden sollen, ist eine minutiöse Überwachung der Arbeitenden problemlos. Digitalisierung ermöglicht das effektive Durchsetzen des Konzepts, was Foucault „zwingender Blick" nennt. Letzterer ist eine Anlage, „in der die Techniken des Sehens Machteffekte herbeiführen und in der umgekehrt die Zwangsmittel die Gezwungenen deutlich sichtbar machen." (Foucault, 1977, S. 221) Dieser Typ von Überwachung wird nicht ‚von außen' ausgeübt, sondern es geht hier um eine innere, intensive, stetige Kontrolle, die den gesamten Arbeitsprozess durchdringt und in deren Fokus die Tätigkeiten der Arbeitnehmer stehen.

Die Arbeitnehmer des digitalen Zeitalters arbeiten unter einem ständigen Druck, den der ‚zwingende Blick' hervorruft, und mittelbar werden sie beherrscht durch Algorithmen, die die gesammelten Daten auswerten. Mithilfe der Algorithmen weiß man, „was ein Mitarbeiter wann tut [...], wie man ihn am besten motiviert – oder zur Kündigung bewegt." (Artikel Nr. 175); in diesem Sinne setzt die normalisierende Macht der Digitalisierung die Arbeitnehmer in eine ‚verletzliche' Position.

Der Code *Ständige Erreichbarkeit* verweist auf den Zustand, dass Arbeitnehmer aufgrund elektronischer Endgeräte (Smartphones, Tablets, Computer etc.) per E-Mail, SMS oder Anruf Tag und Nacht erreichbar sind. Nicht nur an Werktagen, sondern am Feierabend oder während des Urlaubs: „in immer mehr Branchen sind Mitarbeiter [...] ganz selbstverständlich zu jeder Tageszeit erreichbar." (Artikel Nr. 54), dieses Phänomen führt zu der bei *Flexible Arbeitszeiten* und *Mobiles Arbeiten* erwähnten Auflösung der Grenzen zwischen Privatem und Arbeit. Wenn diese klare Trennung fehlt, bleiben Menschen konstant unter Spannung (mentaler, emotionaler sowie körperlicher Art).

Die Existenz von digitalen Technologien normalisiert ständige Erreichbarkeit (mobilisierte Arbeitsmittel in Zusammenhang mit digitalisierten Arbeitsaufgaben) und die ständige Erreichbarkeit der Arbeitenden produziert im Gegenzug Disziplinarmacht. Was die Präsenz der digitalen Technologien im Leben der Arbeitnehmer verwirklicht, ähnelt dem Panoptismus, den Foucault in *Überwachen und Strafen* definiert. Nach Foucault ist die Hauptwirkung des Panopticons „die Schaffung eines bewussten und permanenten Sichtbarkeitszustandes beim Gefangenen, der das automatische Funktionieren der Macht sicherstellt" (1977: S. 258). In dieser Nebeneinanderstellung ist der Gefangene der Arbeitnehmer und die Sichtbarkeit die Erreichbarkeit, und wohingegen bei Foucault es ein architektonischer Apparat ist, der die Schaffung und Aufrechterhaltung des Machtverhältnisses ermöglicht, funktionieren hier digitale Mittel an dessen Stelle.

Dieses Phänomen kann zur Überauslastung der Arbeitnehmer führen, wie es das Zitat von Artikel Nr. 97 darstellt: „Tatsächlich zeigen diverse Studien [...], dass die Beschäftigten [...] auch noch am späten Abend, im Feierabend und am Wochenende Arbeiten erledigen – oft obwohl der Arbeitgeber sie dazu gar nicht angewiesen hat." Hier ist die Wirkung der Disziplinarmacht beobachtbar; die Möglichkeit des Einsetzens kann fast so zwingend wirken, wie eine direkte Weisung.

Der netzwerkartige Charakter, und die, durch digitale Plattformen gesteuerte, Verbindung von Arbeitern, Arbeitsprozessen, Produktionsstellen und Betriebssystemen werden unter dem Code *Vernetzte Unternehmen* zusammengefasst. Mitarbeiter sind miteinander per Intranet oder soziale Medien vernetzt und derart entsteht konstante Kommunikation innerhalb des Unternehmens. Die Vernetzung umfasst neuartige Formen von Kooperation, wie die Zusammenarbeit mit intelligenten Robotern oder den Einsatz von virtuellen Teams von Mitarbeitern (verbunden über das Intranet oder das Internet). Eine weitere Implikation dieser Vernetzung ist das in manchen Unternehmen (wie Telekom und Amazon) bereits angewandte anonyme Feedback-System. Mit Letzterem wird „jeder Mitarbeiter von seiner Führungskraft, von seinen Kunden, von seinen Mitarbeitern und von seinen Kollegen beurteilt" (Artikel Nr. 63); das Feedback ist zu jederzeit möglich und der gesamte Prozess erfolgt anonym.

Essentielles Element in dem vernetzten Unternehmen ist, dass Dinge unmittelbar miteinander kommunizieren (Internet der Dinge), und somit ganze Arbeitsprozesse ohne direkte menschliche Intervention gesteuert werden (Däubler, 2018). Für Arbeitnehmer in digitalisierten Unternehmen erklären Apps Betriebsprozesse und wird gegebenenfalls auch ihre Arbeit bei Algorithmen gesteuert, die immer den optimalen Weg finden können. Beispielsweise ist die Lagerhaltung bei Amazon, Vorreiter der digitalen Wirtschaft, „organisiert durch ausgeklügelte Algorithmen, die den Amazon-Mitarbeitern per Smartphone die kürzesten Wege durch die Gänge weisen, damit sie [...] keine Zeit verlieren." (Artikel Nr. 120)

Der letzte Code unter der Kategorie *Transparenz und Kontrolle*, *Alle Informationen zugänglich im Unternehmen*, beinhaltet den schnellen und freien Zugang zu Informationen innerhalb des Unternehmens und die Teilung von Wissen unter Kollegen. Expertenwissen ist nutzbar sowie transparent in der Organisation und Mitarbeiter kennen auch die Unter-

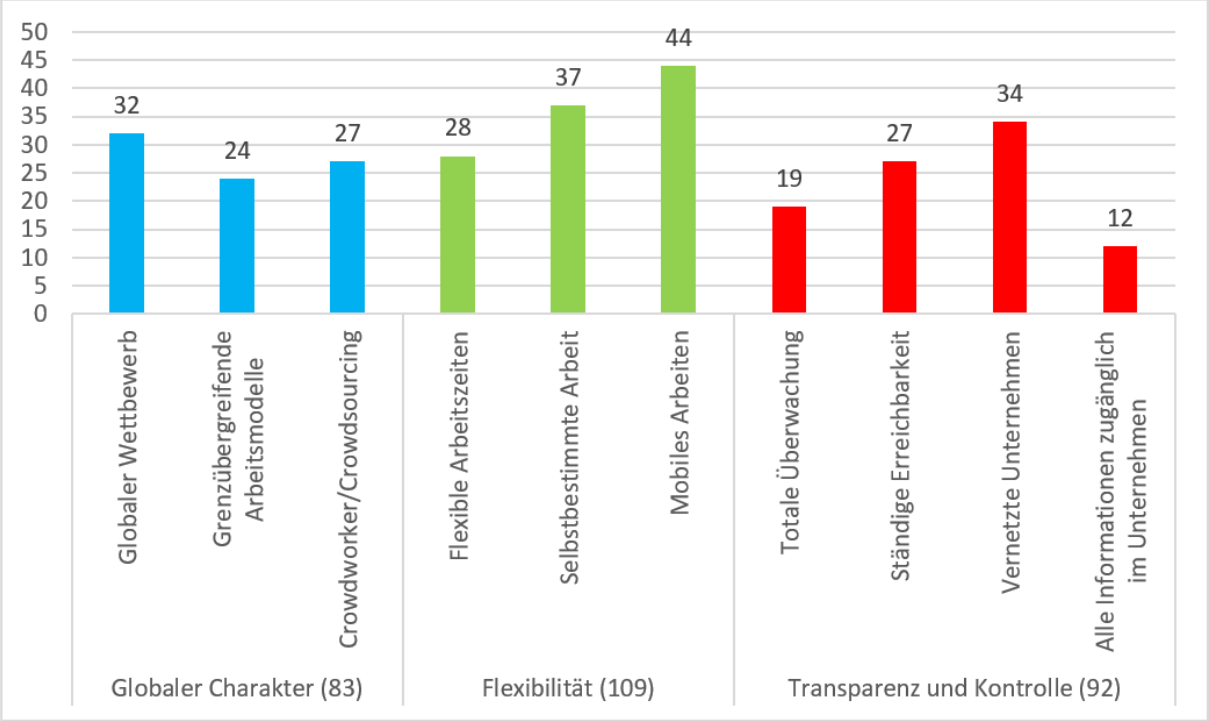


Abbildung 2: Häufigkeiten pro Code und Kategorie (von 1995 bis 2018). (Quelle: eigene Darstellung)

nehmenszahlen. Sie können jederzeit einsehen, woran ihre Kollegen arbeiten, sowie deren Projektplanung oder Termine. Diese Praktiken erleichtern eine offene Diskussion über Projekte sowie Produkte und zerstören Wissensbarrieren, da, wenn „alles für alle offen zugänglich ist, sich kein Machtwissen zurückhalten lässt, [...] viel freier über neue Ideen oder Projekte diskutiert [wird]." (Artikel Nr. 157)

4.5. Häufigkeit der Kategorien

Wenn wir nun die Verhältnisse zwischen den vorhin (Unterkapitel 4.2.-4.4.) beschriebenen Codes betrachten, ist es auffällig, dass diese unterschiedlich häufig erwähnt werden. In Abbildung 2 wird die Häufigkeit der verschiedenen Codes (Erscheinungsformen) und Kategorien (Hauptrichtungen) im analysierten Material in graphischer Form dargestellt.

Es zeigt sich, dass auch die Hauptrichtungen (in denen die Auswirkungen der normalisierenden Macht bestimmt wurden, demnach die Kategorien) unterschiedlich häufig vorkommen, aber ein signifikanter Kontrast in der Anzahl von Nennungen existiert nicht. Die Kategorie, die im untersuchten Zeitraum am meisten vorkommt, ist *Flexibilität* (insgesamt 109-mal), gefolgt von *Transparenz und Kontrolle* (92-mal) und *Globaler Charakter* (83-mal). Damit erscheint die Kategorie *Flexibilität* bedeutsamer als die anderen Kategorien. Glaubwürdiger Grund kann diesbezüglich sein, dass die Wirkungen der Macht, die unter *Flexibilität* zusammengefasst sind, unmittelbarer von Arbeitnehmern wahrgenommen werden als die unter den anderen beiden Kategorien.

Am häufigsten treten die Codes *Mobiles Arbeiten* (44-mal) und *Selbstbestimmte Arbeit* (37-mal) auf, somit ergeben sich die örtliche Flexibilisierung und die Selbstbestimmbarkeit der Arbeit als vorherrschende Erscheinungsformen der normalisierenden Wirkung der Digitalisierung. Deutlich seltener werden Codes wie *Alle Informationen zugänglich im Unternehmen* (12-mal) und *Totale Überwachung* (19-mal) erwähnt.

5. Schlussfolgerungen und Ausblick

Zu Beginn der Analyse wurden die Erwartungen gestellt, die Digitalisierung der Arbeitswelt mit Foucaults Machtphilosophie zu verbinden und die wesentlichen Erscheinungsformen der normalisierenden Macht innerhalb der Arbeitswelt zu kennzeichnen. Beide Erwartungen wurden durch die Analyse der Artikel erfüllt, jedoch wurde der Umfang der Analyse auf die Perspektive der Arbeitnehmer reduziert. Diese Reduktion des Umfangs erbrachte trotzdem befriedigende Resultate, da auch in Foucaults *Überwachen und Strafen* der Mensch im Fokus steht.

Um den Einfluss der normalisierenden Macht der Digitalisierung auf die Arbeitswelt zu untersuchen, wurde nach Mayring (2015) eine qualitative Inhaltsanalyse angewendet. Im Rahmen der Letzteren wurden insgesamt 132 Zeitungsartikel, zwischen 1995 und 2018, aus den Zeitschriften *Der Spiegel* und *Die Zeit* kodiert, die repräsentativ für die deutsche Presse stehen. Auf dieser Basis wurde die Reichweite der Forschung gewährleistet.

Nach der Anwendung der qualitativen Inhaltsanalyse anhand der Zeitungsartikel lässt sich die Forschungsfrage folgendermaßen beantworten: Die Ergebnisse meiner Untersu-

chung haben gezeigt, dass die normalisierende Macht der Digitalisierung mehrfach auf die Arbeitswelt wirkt. Durch Normalisierung verändert Digitalisierung die Arbeitswelt in drei Hauptrichtungen: sie stattet die Arbeit mit einem globalen Charakter aus; somit verlieren die geographischen Grenzen und dementsprechend auch die lokale Reichweite der Arbeitswelt an Bedeutung (1). Digitalisierung flexibilisiert die Arbeitsbedingungen, was sowohl als Möglichkeit als auch als Anforderung für die Arbeitnehmer erscheint. Sie erfüllt auf Seiten der Arbeitnehmer eine Anpassung an flexible Arbeit (2). Letztlich verwirklicht Digitalisierung mithilfe von Vernetzung und Überwachung eine völlig neue Dimension von Transparenz und Kontrolle in der Arbeitswelt (3).

Im Rahmen dieser Bachelorarbeit wollte ich zeigen, dass es förderlich sein kann, die Digitalisierung durch Foucaults Machtphilosophie zu betrachten. Ich bin der Ansicht, dass dieser Versuch erfolgreich war und die Forschung einen Beitrag zum Thema normalisierende Macht und Digitalisierung leistet.

Es muss jedoch berücksichtigt werden, dass sich diese Forschung, wie oben erwähnt, auf die Arbeitnehmer konzentriert, und demnach solche Wirkungen der Normalisierung betrachtet, die die Arbeitswelt unmittelbar für die Arbeitnehmer betreffen. Dementsprechend werden die Auswirkungen für Unternehmen nicht behandelt – es sei denn, dass diese auch die Arbeitnehmer primär betreffen (siehe der Code *Vernetzte Unternehmen*). Aus Praktikabilitätsgründen kann in dieser Arbeit keine umfassende Übersicht über den Arbeitsmarkt gegeben werden, obgleich bestimmte Aspekte teilweise angesprochen wurden (siehe die Kategorie *Globaler Charakter* und dessen Untercodes *Globaler Wettbewerb* und *Crowdworker/Crowdsourcing*). Aus diesem Grund kann keine generelle Aussage weder über die mit Digitalisierung verbundene Arbeitslosigkeit getroffen werden, noch über das Phänomen, wie digitale Technologien menschliche Arbeitskraft ersetzen.

Eine Empfehlung für weitere Forschung wäre, eine ähnliche Studie durchzuführen, aber anstatt Artikel von *Der Spiegel* oder *Die Zeit* (die die deutsche Presse repräsentieren) Artikel aus wirtschaftswissenschaftlichen Fachzeitschriften zu analysieren, beispielsweise aus der *Wirtschaftswoche* oder dem *Manager Magazin*; auf diese Weise wäre auch ein Vergleich der Ergebnisse möglich. Diese Bachelorarbeit stellt die Untersuchung der Disziplinarmacht in den Fokus; somit wäre es besonders interessant, als Ergänzung oder Fortführung dieser Arbeit, eine Studie zu betrachten, die gleichartig zu dieser vorgeht, aber die primär die Bio-Macht (die Foucault in seinem Werk ‚*Der Wille zum Wissen*' (1977) beschreibt) und deren Erscheinungsformen in der Arbeitswelt untersucht.

Literatur

Allen, A. (2002). Power, Subjectivity, and Agency: Between Arendt and Foucault. *International Journal of Philosophical Studies*, *10* (2), 131-149.

Altmeyer, S. (2011). Theologische Subjektorientierung im elektronischen Panoptikum. *International Journal of Practical Theology*, *14* (2), 238-258.

Boes, A., Kämpf, T., Langes, B., Lühr, T., Marrs, K. & Ziegler, A. (2016). Digitalisierung und "Wissensarbeit". *Aus Politik und Zeitgeschichte*, *66*, 32-39.

Boes, A., Kämpf, T., Lühr, T., Langes, B. & Ziegler, A. (2017). Cloud & Crowd: New Challenges for Labour in the Digital Society. *TripleC*, *15* (1), 132-147.

Detel, W. (1998). *Macht, Moral, Wissen: Foucault und die klassische Antike* (1. Aufl.). Frankfurt am Main: Suhrkamp. Zugriff auf https://primo.fu-berlin.de/FUB:FUB_ALMA_DS2183669024000288 3http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&local_base=BVB01&doc_number=007968298&sequence=000002&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA

Detel, W. (2005). *Foucault and Classical Antiquity: Power, Ethics, and Knowledge*. Cambridge, U.K.. Zugriff auf https://primo.fu-berlin.de/FUB:FUB_ALMA_DS511037551650002883

Däubler, W. (2018). Digitalisierung und Arbeitsrecht. *Soziales Recht* (Sonderausgabe Juli), 2-44.

Foucault, M. (1977). *Sexualität und Wahrheit 1, Der Wille zum Wissen* (1. Aufl.). Frankfurt am Main: Suhrkamp. Zugriff auf https://primo.fu-berlin.de/FUB:FUB_ALMA_DS21901603940002883

Foucault, M. (2005). *Analytik der Macht* (1. Aufl.). Frankfurt am Main: Suhrkamp. Zugriff auf https://primo.fu-berlin.de/FUB:FUB_ALMA_DS21924358760002883http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&local_base=BVB01&doc_number=013093801&sequence=000002&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA

Foucault, M. (2019). *Überwachen und Strafen: Die Geburt des Gefängnisses* (21. Aufl.). Frankfurt am Main: Suhrkamp. Zugriff auf https://primo.fu-berlin.de/FUB:FUB_ALMA_DS21784023350002883

Gerber, C. & Krzywdzinski, M. (2019). Entgrenzung in der digitalen Onlinearbeit am Beispiel von Crowdwork. In H. Hanau & W. Matiaske (Hrsg.), *Entgrenzung von Arbeitsverhältnissen: Arbeitsrechtliche und sozialwissenschaftliche Perspektiven* (1. Aufl., S. 25–48). Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG. Zugriff auf https://doi.org/10.5771/9783845296159-25

Gioia, D. A., Corley, K. G. & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational Research Methods*, *16* (1), 15-31.

Hoffman, A. J. (1999). Institutional Evolution and Change: Environmentalism and the U.S. Chemical Industry. *The Academy of Management Journal*, *42* (4), 351-371.

Krasmann, S. (2017). Imagining Foucault. On the Digital Subject and "Visual Citizenship". *Foucault studies*, 10-26.

Šlesingerová, E. (2018). Biopower imagined: Biotechnological art and life engineering. *Social Science Information*, *57* (1), 59-76.

Lifková, A. (2019). Digital Power: Self-tracking Technologies through Michel Foucault Lens. *Politické Vedy*, *22* (4), 81-101.

Lohmeyer, N. (2018). "Vocabularies of Motive": A Content Analytical Longitudinal Analysis of Motives for Corporate Social Responsibility in Manager Magazin Articles, 1971–2017. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, *70* (3), 277-308.

Maddalena, K. & Packer, J. (2015). The Digital Body: Telegraphy as Discourse Network. *Theory, Culture & Society*, *32* (1), 93-117.

Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12. Aufl.). Weinheim.

Messerschmidt, A. (2011). Michel Foucault (1926–1984). In B. Dollinger (Hrsg.), *Klassiker der Pädagogik* (3., durchgesehene Aufl., S. 289-310). Wiesbaden.

Pongratz, L. A. (2013). *Unterbrechung: Studien zur Kritischen Bildungstheorie*. Leverkusen.

Rinne, U. & Zimmermann, K. F. (2016). Die digitale Arbeitswelt von heute und morgen. *Aus Politik und Zeitgeschichte*, *66* (18/19), 3-9.

Rothfusz, J. (2012). Hanging around in suburbia. Understanding normalizing power in professional relationships with Dutch Caribbean migrants. *Journal of Social Intervention: Theory and Practice*, *21* (2), 63-82.

Saldaña, J. (2016). *The Coding Manual for Qualitative Researchers* (3. Aufl.). Los Angeles: SAGE. Zugriff auf https://primo.fu-berlin.de/FUB:FUB_ALMA_DS21775642560002883http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&local_base=BVB01&doc_number=028474356&sequence=000003&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA

Staab, P. & Nachtwey, O. (2016). Die Digitalisierung der Dienstleistungsarbeit. *Aus Politik und Zeitgeschichte*, *66*, 24-31.

Thomas, T. (2009). Michel Foucault: Diskurs, Macht und Subjekt. In A. Hepp, F. Krotz & T. Thomas (Hrsg.), *Schlüsselwerke der cultural studies* (S. 58–71). Wiesbaden: VS Verlag für Sozialwissenschaften. Zugriff auf https://doi.org/10.1007/978-3-531-91839-6_5

Ziemann, A. (2011). *Medienkultur und Gesellschaftsstruktur* (1. Aufl.). Wiesbaden.

Zweig, K. A. (2019). *Algorithmische Entscheidungen: Transparenz und Kontrolle*. Sankt Augustin; Berlin: Konrad Adenauer Stiftung.

Junior Management Science 7(4) (2022) 945-985

Junior Management Science

journal homepage: www.jums.academy

The Idiosyncratic Volatility Puzzle – Anomaly or Data Mining?

Leon Kowalke

Leibniz Universität Hannover

Abstract

In this study, I investigate the robustness of the idiosyncratic volatility puzzle to the configuration of the research design. Using the regression- as well as the portfolio-based concept, I start with the replication of the idiosyncratic volatility puzzle approving the findings of Ang, Hodrick, Xing, and Zhang (2006). However, when idiosyncratic volatility is estimated from monthly data and a time window spanning 1 or 5 years, the puzzle vanishes, regardless of the research method employed. Similar result hold if only stocks with a market capitalization above the cross-sectional median or those with a price higher than 10$ are used. Independent of the weighting scheme, the puzzle is also absent in the regression-based context when the risk premia are estimated by generalized least squares weighting returns by the inverse of their variance estimates. The same finding is derived in the portfolio-based context by extending the holding period to 12 months or controlling for the past month maximum daily return.

Keywords: Idiosyncratic Volatility; Cross-section of stock returns; Predictability; Risk Premium; Robustness.

1. Introduction

Common asset pricing models such as the CAPM argue that systematic market risk is the only type of risk that should be compensated with a return premium in the equilibrium. Idiosyncratic risk, on the other hand, should not be priced, as it is assumed to be perfectly diversifiable by holding a combination of the market portfolio and a risk-free investment. Nonetheless, some authors claim that perfect diversification might not be possible, if investors cannot attain the market portfolio due to market frictions such as transaction costs or incomplete information (e.g. Malkiel & Xu, 2006; Merton, 1987). Merton (1987) argues that in this case investors no longer only care for the market, but the total risk which includes the idiosyncratic volatility (IVOL hereafter) as special part. Therefore, idiosyncratic risk should be of importance to the investors decision making process and hence be compensated with a positive risk premium in the cross-section of average returns from a theoretical point of view. However, in an influential study, Ang et al. (2006) empirically discovered a negative relation between IVOL and subsequent stock returns instead. This finding is particularly puzzling when considering standard asset pricing theory and has started the debate about the importance of a pricing effect on idiosyncratic volatility. Many papers were published that verified the negative relation but claim to be able to explain it with new theories supported by empirical results, whereas other studies emerged that completely challenged the finding of a negative relationship. One example is the study of Fu (2009). His findings fit to the standard asset pricing theory, as he empirically confirmed a positive link between conditional IVOL and expected returns. This illustrates that not only the empirical findings of a consistent pricing effect induced by IVOL diverge from its theoretical predictions but also that the empirical evidence varies across studies. As there are various studies that appear to analyze the same theoretical concept, the question may arise why their results differ so fundamentally from each other. The research concepts related to the IVOL puzzle can be classified into two major groups which consist of the regression- and portfolio-based approach. Whereas the regression-based concept focuses on the estimation of the IVOL risk premium by means of monthly cross-sectional regressions, the portfolio-based method utilizes the idea of sorting stocks into portfolios based on their exposure to IVOL and searches for a systematic pattern in portfolio returns respectively. Even though both of these concepts are able to discover the IVOL puzzle as shown by Ang et al. (2006); Ang, Hodrick, Xing, and Zhang (2009), not all studies come to this same conclusion. Since Ang et al. (2006) for example use value-weighted portfolios for their analysis, Bali and Cakici (2008) find that using equal-weighted portfolio returns instead, lets the puzzle vanish. Also, in the regression-based

DOI: https://doi.org/10.5282/jums/v7i4pp945-985

concept, the IVOL puzzle is sensitive to the research design, as Bali, Cakici, and Whitelaw (2011) claim that the puzzle no longer exists when they integrate a control for the maximum daily return over the past month into their cross-sectional regressions. Hence, the choice of the research design is important for answering the question on the existence of an IVOL puzzle. On this background I aim to answer the question on how robust the IVOL puzzle is to a specific research design. To do so, I analyze in what aspects the studies on idiosyncratic volatility mainly differ and how these differences influence their conclusions. In addition, I incorporate further adjustments that I consider adequate and try to verify the robustness of the findings respectively. My analysis is structured along the regression- and the portfolio-based approach on the basis of which I investigate the differences in results that might emerge from adjustments in the research design. Here I start with implementation of adjustments that are applicable to both research concepts and afterwards I focus on those modifications that are related to the methodological peculiarities of the research concepts. By this methodology, I intend to settle the dispute on the existence and nature of the idiosyncratic volatility risk premium as well as to illustrate the influence a researcher has on the results obtained simply by choosing a specific research design. Using the example of idiosyncratic volatility, I try to highlight the importance to also consider the study's econometric design when interpreting its result.

My study is structured as follows. In section 2 I start with a brief overview on the literature related to the pricing effect of idiosyncratic risk as well as the corresponding research approaches and their findings. Afterwards section 3 sets out the methodology of my study. It begins with the description on how I measure idiosyncratic volatility, then gives the explanation on how the reference results based on the regression- and the portfolio-based research concept are derived and closes with an explanation on the kind of adjustments I consider as well how they are incorporated into the research concepts. Section 4 introduces the data used for this study. The empirical analysis starts in section 5 with some summary statistics on the sample. Afterwards I continue with the replication of the IVOL puzzle on the basis of the regression- and portfolio-based research concept that should later on act as a reference point for the further adjustment analysis. The empirical analysis of the general adjustments, consisting of idiosyncratic volatility and sample-related modifications, is located in section 6. In contrast, section 7 presents the evaluation of the method-specific adjustments. For the regression-based concept these adjustments involve changes in the risk-related control variables as well as the regression estimation procedure, while those adjustments for the portfolio-based method focus on modifications in the way portfolios are characterized and analyzed, as well as the interaction of the IVOL puzzle with other firm-specific effects when these are controlled for by means of bivariate dependent portfolio sorts. Lastly, section 8 concludes.

2. Literature Overview and Theoretical Motivation

2.1. The Pricing Effect of Idiosyncratic Risk

Markowitz (1952) was among the first to formulate the idea of a positive relationship between systematic risk and expected returns. In his study, he allows investors to only construct a portfolio-based on all stocks available in the market and measures risk as the portfolio return variance. His advice is then to select a portfolio laying on the efficient frontier, as these attain the optimal risk-return trade-off, which is why he terms them to be "mean-variance efficient" (see Markowitz, 1952). Mean-variance efficient portfolios contain the part of the variance that remains after diversification and thus determines the extent of risk that is equal for all assets in the portfolio. This type of risk is called "systematic risk" and should be compensated by a positive return premium to make a risky investment attractive for risk averse investors. Risk that can be diversified away is called "idiosyncratic risk" and should not yield any compensation, as no influence on the diversified portfolio's payoff is assumed. Based on the portfolio theory of Markowitz (1952) and his concept of mean-variance efficiency, Sharpe (1964), Lintner (1965) and Black (1972) formulate the well-known Capital Asset Pricing Model (CAPM hereafter). In their model, the investment strategy with the optimal risk-return trade-off in equilibrium involves investing in the mean-variance efficient market portfolio and a risk-free investment such as the risk-free interest rate. The market portfolio acts as a proxy for systematic risk and the asset's co-variation with this portfolio is captured by the market beta. This market beta allows for quantification of the asset's exposure to systematic risk and determines the part of the risk that requires compensation by a return premium. Hence, portfolios with a higher market beta are assumed to contain more systematic risk and should therefore deliver higher expected returns. In the CAPM, the linear relationship between securities' expected returns and their market betas implies that, in the equilibrium, idiosyncratic risk can be perfectly diversified and thus does not carry a risk premium.

However, authors have begun to question the CAPM assumptions including the irrelevance of idiosyncratic risk for pricing purposes and provided theoretical evidence that this kind of risk is indeed priced under plausible assumptions, even though the direction of the pricing effect has remained controversial. Merton (1987), for example, argues that when investors have incomplete information on the characteristics of securities, they are prone to hold assets whose risk-return profile they are familiar with. In consequence, they hold under-diversified portfolios and no longer care only for systematic market risk but instead demand a return premium for the total amount of risk including its idiosyncratic part. Hence, incomplete information leads to a situation in which diversifiable risk would be priced as investors require compensation for their imperfect diversification. Therefore, Merton (1987) predicts a positive relationship between idiosyncratic risk and cross-sectional stock returns. Conversely,

Miller (1977) sets the framework for an opposite theoretical prediction. His study is based on the assumption that investors diverge in their perception of expected returns for a specific firm and also constrains them in their short-selling activities. He argues that stocks which are characterized by dispersed opinions on their future payoffs are likely systematically overvalued, as their prices are mainly determined by market participants who are most optimistic about the firms expected returns and hence willing to pay the highest price. When this overvaluation is eventually corrected as relevant information becomes available to all investors over time, disagreements on future returns diminish and prices are corrected which leads to lower subsequent returns (see Miller, 1977). In addition Miller (1977) claims that this effect increases with the level of dispersion in opinions on future returns leading to even lower subsequent returns for stocks with a higher degree of disagreement. Diether, Malloy, and Scherbina (2002) measure the heterogeneity of opinions by the dispersion in analysts' forecasts and claim that it can proxy for idiosyncratic risk. They find that expected returns are negatively related to dispersion in analysts' forecasts and thus conclude also on a negative relationship to idiosyncratic risk which contradicts the notion of Merton (1987).

To support the above mentioned theoretical predictions, several studies have emerged that try to unveil the true pricing implications of idiosyncratic risk empirically. Among the first studies to do so is the one of Fama and MacBeth (1973). In a linear regression-based framework, they measure idiosyncratic risk as standard deviation of the least-squares regression residual that was generated by regressing the firm's stock excess return onto the CAPM. Based on their market beta, they sort all stocks into 20 portfolios and regress them each month onto the portfolio market beta, the squared portfolio market beta and the average idiosyncratic risk of all stocks in the portfolio to compute their regression coefficients that proxy for the variables' risk premia. They use monthly returns of all common stocks trading on the New York Stock Exchange (NYSE hereafter) in the period of January 1926 to June 1968 and find a statistically insignificant risk premium, wherefore they conclude that idiosyncratic risk is not priced in the cross-section of average returns (see Fama & MacBeth, 1973).

In a subsequent replication of the study from Fama and MacBeth (1973), Tinic and West (1986) use data from 14 additional years and conclude that idiosyncratic risk is indeed priced, carrying a positive risk premium which corresponds to the predictions of Merton (1987). Malkiel and Xu (1997) draw the same conclusion. Using portfolios sorted by idiosyncratic volatility they find that portfolios containing stocks with the highest idiosyncratic volatility also exhibit the highest returns. It is of note, however, that they do not report any statistics for the verification of statistical significance. In their follow up study, Malkiel and Xu (2006) utilize a regression-based method comparable to that of Tinic and West (1986) as well as Fama and MacBeth (1973) and show that the positive relationship between expected returns and IVOL holds up in statistical tests. Spiegel and Wang (2005) in addition to Fu (2009) argue that past realized idiosyncratic risk is not able to identify the true relationship between IVOL and expected returns, which is why they suggest to use conditional IVOL instead that is estimated using an Exponential Generalized Autoregressive Conditional Heteroscedastic (EGARCH) model. Both studies define idiosyncratic volatility relative to the Fama and French (1993) three-factor model (FF3 model hereafter) again, finding a positive return premium for IVOL (see Fu (2009); Spiegel and Wang (2005)). The FF3 model uses a market (MKT), size (SMB) and value (HML) factor for systematic risk-correction (see Fama & French, 1993).

Even though a positive risk premium for idiosyncratic risk violates the CAPM assumptions, it still retains the notion of Markowitz (1952) that risk averse investors require a return premium for taking on volatility risk. Ang et al. (2006) are among the first to disagree with previous studies as they propose a negative relationship between idiosyncratic risk and expected returns. This relationship is hardly explained by classic risk-return arguments, wherefore it is also known as the "Idiosyncratic Volatility Puzzle". Ang et al. (2006) sort all stocks traded on the American Stock Exchange (AMEX hereafter), National Association of Securities Dealers Automated Quotations (NASDAQ hereafter) and the NYSE into quintile portfolios based on their past month idiosyncratic volatility and point out that the portfolio in the highest IVOL quintile systematically underperform the one in the lower IVOL quintile by a statistically significant average total return difference of about -1.06%. This result also holds for risk-adjusted returns relative to the CAPM and FF3 model as well as when controlling for various cross-sectional effects in bivariately sorted portfolios (see Ang et al., 2006). Authors such as A. G. Huang (2009), Ang et al. (2009) and Duarte, Kamara, Siegel, and Sun (2014) are able to verify the finding of an IVOL puzzle. A. G. Huang (2009) uses both, a regression-based approach in fashion of Fama and MacBeth (1973) as well as a portfolio sorting approach comparable to that of Ang et al. (2006) and confirm that the underperformance of the highest IVOL portfolio remains statistically significant even if risk adjustment is induced by the Carhart (1997) four-factor (CH4 hereafter) model that augments the FF3 model by a momentum factor (MOM hereafter) (see Ang et al., 2006; Carhart, 1997). Ang et al. (2009) add international evidence on the IVOL puzzle by analyzing a data set consisting of 23 developed countries spanning a sample period from January 1980 to December 2003. In addition, they discover an international co-movement in portfolio returns that goes long the quintile portfolio containing firms with the highest IVOL and short those of the lowest IVOL quintile, hinting at systematic risk that is not captured by standard risk factors (see Ang et al., 2009). This idea is picked up by Duarte et al. (2014) who also confirm the negative relationship between idiosyncratic volatility and expected returns and construct a risk factor based on the return difference of a portfolio with high predicted IVOL as well as the one of low predicted IVOL. Adding this risk factor to the FF3 model, they improve in explaining the cross-section of returns, wherefore they argue that the IVOL puzzle arises due to an omitted systematic risk

factor (see Duarte et al., 2014).

2.2. Regression- vs. Portfolio-Based Method

To understand the contradictory findings on the relationship between idiosyncratic risk and the cross-section of expected returns as outlined in the previous section, it is necessary to get an overview on the research concepts and adjustments used in different studies. Two branches have emerged in the literature that investigate the pricing implications of idiosyncratic risk. The first branch follows the ideas of Fama and MacBeth (1973) and implements a regression-based methodology that is based on the risk premia estimation by means of cross-sectional linear regressions. In contrast, the second branch uses a portfolio-based methodology as proposed in Ang et al. (2006) where stocks are sorted into portfolios based on their idiosyncratic volatility and are then analyzed respectively.

The regression-based approach is based on the ideas of Fama and MacBeth (1973). They propose a two step procedure that starts with running a time-series regression of all stocks' monthly excess returns onto a risk-correction model which they choose to be the CAPM and compute the firm-specific risk exposure by the stock-related market beta. Afterwards 20 portfolios are formed based on the ranking of all stocks relative to their estimated market betas. These portfolios are constantly rebalanced and the stock-specific betas are regularly recomputed. Then, the simple average over all stock betas within a portfolio is used to compute the portfolio beta. The standard deviation of the regression residual from the time-series regression in the first step is used as a proxy for the stock-specific IVOL that is also averaged across all firms within a portfolio to generate a portfolio level IVOL measure. In the second step they compute cross-sectional regressions in each period and average the estimated coefficients over their sample time horizon to obtain an estimate on the expected risk premium for the corresponding regressor. Their study investigates the premium on the market beta as a proxy for the price of market risk, the premium on the squared market beta as a measure for non-linearity in market risk and the expected risk premium on their IVOL measure drawing the conclusions already mentioned in section 2.1 (see Fama & MacBeth, 1973). Ang et al. (2009) were among the first to verify the existence of the IVOL puzzle with such an approach but using single stocks instead of portfolios as dependent variable for the cross-sectional regressions.

In subsequent studies authors have begun to adjust the regression-based methodology to improve it by handling well-known econometric problems or simply updating the method based on recent empirical findings and new theoretical concepts such as more sophisticated risk-correction models. Malkiel and Xu (2006) try to handle the well-known "errors-in-variables" problem which arises in the IVOL estimation on the firm level. For this they use the ideas of Fama and French (1992) and apply the portfolio IVOL on all stocks within the portfolio. First they sorted all stocks into 200 portfolios based on its exposure to size and market beta. Thereafter they calculate the portfolio IVOL and assign this value to all stocks located in the specific portfolio. This IVOL is then used in the cross sectional regression with individual stocks (see Malkiel & Xu, 2006). Subsequently, Ang et al. (2009) measure IVOL relative to the FF3 model and pick up the idea of using single stock returns instead of portfolios in the cross-sectional regressions. They are also among the first to correct the t-statistics of the average cross-sectional risk premia estimates for serial correlation and heteroskedasticity using the adjustment from Newey and West (1987). Z. Chen and Petkova (2012) use two different techniques for the estimation of time-series risk factor loadings that focus on different degrees of time variation in these estimates. One method uses the complete sample for this time-series regression, whereas the other implements 60-months rolling estimation windows for that purpose which allows for better characterization of the dynamically changing firm-specific risk exposure. Furthermore, changes in the cross-sectional regression estimation technique have also been considered throughout the literature. Among others Ang et al. (2009) as well as Han and Lesmond (2011) suggest to incorporate value weighting into the regression approach of Fama and MacBeth (1973), as they weight all stock returns with the corresponding firm size at the beginning of the month to avoid small firm effects distorting their risk premia estimates. Moreover, some authors consider adjustments related to the control variables used in cross-sectional regressions. Various studies show that the choice of the risk-related control variables capturing systematic or firm-specific risk can have a notable impact on the IVOL risk premium. Boyer, Mitton, and Vorkink (2010) among others regress portfolio excess returns not only onto the firm's CH4 model risk factor loadings to account for systematic risk but also add those on a liquidity and a coskewness risk factor respectively. Additionally they include variables such as expected idiosyncratic skewness, size, the book-to-market ratio, momentum and turnover into their cross-sectional regressions to account for firm-specific risk which was not considered in the study of Fama and MacBeth (1973) (see Boyer et al., 2010). L. H. Chen, Jiang, Xu, and Yao (2012) propose the idea of a subsample analyses to dissect the finding of an IVOL puzzle from the effect of other firm characteristics. Hence they conduct part of their regression-based analysis in subsamples constructed based on attributes such as size, price and calendar time (see L. H. Chen et al. (2012)).

In contrast, the general idea of the portfolio-based research method is to sort stocks based on their idiosyncratic volatility into a pre-specified number of portfolios and rebalance these portfolios regularly. For all portfolios, the next period's return (total or in excess of the risk-free rate) as well as the Jensen's alphas relative to a select risk-correction model are computed each period. Based on the long-short portfolio, which goes long the portfolio of stocks with the highest IVOL and short the one containing those stocks with the lowest IVOL, the alpha and return difference between high and low IVOL portfolios is computed which should deliver evidence on the pricing implications from idiosyncratic risk (see e.g. Ang et al., 2006; Bali & Cakici, 2008; Rachwalski & Wen,

2016). Ang et al. (2006) are among the first to conduct such an analysis. They measure idiosyncratic volatility as standard deviation of the residual from a time-series regression of the stock-specific excess return onto the FF3 model using daily data over the past month. Based on this IVOL measure, they sort stocks into monthly rebalanced quintile portfolios for which they compute the value-weighted (VW hereafter) monthly total returns as well as the Jensen's alpha relative to the CAPM and the FF3 model that they then average over time for presentation purposes. Lastly, by investigating the return patterns from the quintile and the long-short portfolio, they discover the IVOL puzzle (see Ang et al., 2006).

The portfolio-based research concept has also been adjusted by various studies. Ang et al. (2006) argue that the trading strategy used for portfolio construction can influence the findings, which is why they implemented other strategies for robustness purposes but were not able to identify any influence on the IVOL puzzle in their study. In addition, also the number of univariate portfolios analyzed was subject to change. While Ang et al. (2006) and later Duarte et al. (2014) propose quintile portfolios for the analysis, Spiegel and Wang (2005) use decile portfolios and find a positive relationship between IVOL and expected returns instead. Bali and Cakici (2008) show that the choice of the return weighting scheme, the breakpoints necessary to subdivide stocks into portfolios and the data frequency used for IVOL estimation can lead to different conclusions on the existence of the IVOL puzzle. Their analysis of equal-weighted (EW hereafter), instead of VW portfolio returns, delivers no evidence on an IVOL puzzle. Neither do portfolios formed from IVOL breakpoints that should account for potential size effects and hence are calculated from stocks traded on the NYSE only, or are set such that stocks are subdivided into portfolios that each make up an equal market share. They draw the same conclusion when using monthly data over the past two or five years to estimate IVOL (see Bali & Cakici, 2008). Other studies also differ in the choice of the risk-correction model used for computation of the Jensen's alpha. Ang et al. (2006), Bali and Cakici (2008) and Fu (2009) report alphas that correct for the risk factors from the CAPM or FF3 model, whereas Spiegel and Wang (2005) as well as Rachwalski and Wen (2016) use the CH4 model for that purpose. Cao, Chordia, and Zhan (2021) adds the alpha relative to the Fama and French (2015) five-factor model (FF5 model hereafter) that augments the FF3 model by an investment (CMA) and a profitability (RMW) factor. Furthermore, Ang et al. (2006) and A. G. Huang (2009) use bivariate dependent sorted portfolios to investigate the robustness of the IVOL puzzle to firm characteristic effects. For that, they choose a control variable based on which they sort all stocks into portfolios first. Then all stocks within each of these portfolio are sorted again into portfolios based on their IVOL. Lastly, each of these IVOL portfolios is averaged across the control variable sorts resulting in average portfolios that capture the IVOL puzzle effect which is robust to the effects of the selected control (see Ang et al., 2006; A. G. Huang, 2009).

3. Methodology

This section introduces the methodology used throughout my study. Section 3.1 starts with the definition of my idiosyncratic volatility measure. In section 3.2 and 3.3 I present my benchmark specification for the regression- as well as the portfolio-based research approach, on the basis of which I replicate the IVOL puzzle finding and implement further adjustments. Lastly, section 3.4 sets out my adjustment analysis.

3.1. Estimation of Idiosyncratic Volatility

Usually idiosyncratic volatility is computed as the standard deviation of the residual from the regression of the firm's excess returns onto to a risk-correcting model like the FF3. Regardless of the model used, the following general time-series regression is run for each firm to obtain the corresponding idiosyncratic volatility estimate:

$$r_t^i = \alpha^i + \beta^i X_t + \epsilon_t^i, \tag{1}$$

where r_t^i is the return of stock i at time t in excess of the risk-free rate and α^i stands for the time-series intercept for stock i. $\beta^i X_t$ is a 1 x k vector of risk factor loadings for firm i that are estimated relative to the k x 1 vector of risk factors X_t that should proxy for systematic risk. ϵ_t^i presents the regression residual for firm i at time t, on the basis of which idiosyncratic volatility is generally computed as follows:

$$IVOL_i^X(t; t-s) = \sqrt{var_{t;t-s}(\epsilon_t^i)}. \tag{2}$$

Here $IVOL_i^X(t; t-s)$ presents the IVOL of firm i at time t measured relative to risk-correction model X over the time horizon from $t-s$ to t. The right-hand side of the equation states that IVOL is measured as the standard deviation of the time-series regression residuals over the time horizon from $t-s$ to t. Equation 1 allows for the integration of various risk-correcting models to compute IVOL which becomes important in the later adjustment analysis. If not otherwise stated, I use the FF3 model for risk-correction and compute IVOL from the following time-series regression for each stock separately while using daily data:

$$r_t^i = \alpha^i + \beta_{MKT}^i MKT_t + \beta_{SMB}^i SMB_t + \beta_{HML}^i HML_t + \epsilon_t^i, \tag{3}$$

β_{MKT}^i refers to the time-series regression factor loading on the MKT risk factor of firm i. β_{SMB}^i and β_{HML}^i stand for the factor loadings of firm i on the SMB and HML risk factor respectively. MKT_t, SMB_t and HML_t represent the level of the corresponding risk factor at time t. For estimation of the IVOL measure utilized throughout most of my analysis, I use the daily residuals from equation 3 and compute it as follows:

$$IVOL_{i,t}^{FF3} = \sqrt{var(\epsilon_t^i)}. \tag{4}$$

$IVOL_{i,t}^{FF3}$ is then defined as the idiosyncratic volatility relative to the FF3 model of firm i in month t that is measured

as the standard deviation of the daily residuals from regression 3 over the current month. For firms with less than 15 daily return observations, the IVOL measure is considered as missing. If not otherwise stated, this definition is used when referring to idiosyncratic volatility. Modifications to this definition are considered in the later adjustment analysis.

3.2. Regression-Based Approach

I use the study of Ang et al. (2009) as a reference point for implementation of the regression-based research approach, as it is among the first to verify the existence of the IVOL puzzle using a linear regression framework comparable to that of Fama and MacBeth (1973). This procedure is subdivided into two steps: a time-series regression and a cross-sectional regression step.

In the first step, a time-series regression equal to that of equation 1 is run for every firm. Here a pre-defined time window as well as risk-correction model are used to obtain the firm-specific risk exposure measured by the risk factor loadings. If not otherwise stated, I compute the risk factor loadings relative to the FF3 model based on equation 3. Similar to the IVOL estimation, monthly factor loadings of firm i are computed using rolling windows of one month that contain daily data. These factor loadings are then used in the second step of the procedure.

In the second step, I compute cross-sectional regressions at every month t across all firms for which data are available at that month. These have the following form:

$$r_i(t;t+1) = c+\gamma IVOL_i^X(t;t-s)+\lambda_b^{'} b_i(t;t+1)+\lambda_z^{'} z_i(t)+\epsilon_{t+1}^i, \tag{5}$$

where $r_i(t;t+1)$ is the stock's excess return from time t to $t+1$, c is the cross-sectional regression intercept, $b_i(t;t+1)$ is the m x 1 vector of time-series risk factor loadings for firm i computed over the window from time t to $t+1$, $z_i(t)$ is the k x 1 vector of firm characteristics for firm i observable at time t and ϵ_{t+1}^i is the cross-sectional regression residual of stock i at time $t+1$. γ, $\lambda_b^{'}$ and $\lambda_z^{'}$ define the 1 x 1, 1 x m and 1 x k vectors of cross-sectional regression coefficients on idiosyncratic volatility, the time-series risk factor loadings and the firm characteristics respectively. These coefficients should approximate the monthly risk premia for the corresponding regressors. Subsequently, I use the time series of monthly risk premia and the intercept to test whether the average premium on the chosen idiosyncratic volatility measure is statistically different from zero. For that I use a standard t-test where I correct the standard errors of the coefficient estimates for serial correlation by the methodology of Newey and West (1987). Lags are selected based on the maximum value proposed by the quadratic spectral or the Bartlett kernel that is then rounded downwards to the next integer value. In addition, I calculate the average monthly R^2 and number of stocks used in the corresponding regression.

Whereas the previous explanations are given generally to enhance the understanding of the later adjustment analysis, I will now concretize the regression-based research methodology that is used as reference point throughout my empirical analysis. I use monthly data to estimate the cross-sectional regression of equation 5 at each month t by means of the ordinary least squares (OLS hereafter). On the left-hand side of this equation, I use the next month excess return for stock i measured from the end of month t to $t+1$. As IVOL measure, I utilize the one defined in equation 4 that is computed from daily data over the current month. Risk-correction is implemented by usage of the time-series risk factor loadings on the FF3 model from equation 3 that are obtained from daily data over the next month $t+1$. For the calculation of the time-series risk factor loadings, I again require each stock to report at least 15 daily return observations during the month. As I rely on the concept of Ang et al. (2009) for the regression-based methodology, I apply almost the same firm characteristics and start further adjustment analyses from there on. Hence, my vector of firm characteristics at each t consists of the firms' size measured by the natural logarithm of market capitalization, the firm's book-to-market ratios of equity (BM hereafter) calculated based on Fama and French (1992) and the total return measured from the end of the previous month $t-1$ until the end of current month t to control for the one-month return reversal effect discovered by Lehmann (1990) and Jegadeesh (1990). In addition to these firm characteristics I augment the study by sequentially rotating among further firm-related control variables that might be interrelated with the IVOL puzzle effect. To reduce the arbitrariness in selecting these variables, I rely on the findings of L. H. Chen et al. (2012) as well as Hou and Loh (2016) who evaluated the pervasiveness of the IVOL puzzle across different explanations found in the literature. Barberis and Huang (2008) claim that investors have preferences for extreme events when evaluating stocks that they quantify by stock return skewness. Consequently I consider this effect by controlling for the skewness of stock returns over the current month, the expected idiosyncratic skewness measure of Boyer et al. (2010) and the coskewness measure over the recent past 5 years computed based on the procedure of Harvey and Siddique (2000). Lastly, I also account for the liquidity-related IVOL puzzle explanation and include the Amihud (2002) illiquidity measure as a control variable (see Han & Lesmond, 2011). At this step of the analysis I use 7 lags for the standard error correction based on Newey and West (1987). I term the above explained method as equal-weighted Fama-MacBeth regression procedure. However, similar to Ang et al. (2009) I also consider a value-weighted approach. For this I use a weighted least squares (WLS) method to estimate risk premia in the cross-section regression of equation 5, where all returns are weighted by the firm's current month market capitalization (MTCAP hereafter) in U.S. dollars. I call this procedure the value-weighted Fama-MacBeth regression.

3.3. Portfolio-Based Approach

For implementation of the portfolio-based research method I rely on the study of Ang et al. (2006). Here the idea

is that all stocks traded on the NYSE/AMEX/NASDAQ are sorted into a pre-defined number of portfolios based on their past idiosyncratic volatility. These portfolios are rebalanced regularly and analyzed, additionally to a long-short portfolio, to identify the relationship between IVOL and expected cross-sectional returns non-parametrically. The procedure starts by defining the trading strategy. Such a strategy consists of an estimation period E, a waiting period W and a holding period H. Using this notation, an $E/W/H$ trading strategy can be computed as follows. At month t the IVOL measure is computed for each firm over the E-month period from $t-E-W$ to month $t-W$ based on equation 2. Afterwards it is annualized by multiplication with $\sqrt{250}$. Based on the calculated IVOL, at each month t breakpoints are determined that subdivide stocks into a pre-defined number of portfolios. These portfolios are then held for the next H-months. As proposed by Ang et al. (2006), for portfolios with $E > 1$ and $H > 1$ I follow Jegadeesh and Titman (1993) and compute portfolios with overlapping holding periods. Hence, for a 6/1/6 strategy at month t the pre-defined number of portfolios are constructed from the IVOL estimate over the past 6 months ending 1 month prior to the formation date t and using daily return data. The same number of portfolios is constructed based on the IVOL estimate over the past 6 month but ending 2 months, 3 months up to 6 months prior to t. Then a simple average across all the sets of portfolios with different ending periods are calculated, such that only the set consisting of the number of pre-defined portfolios constructed from the averaging procedure remains. Consequently, 1/6th of the composition of these portfolios changes each month. Furthermore, each 1/6th of these portfolios consists of the portfolio from 1 month ago and 1/6th of the portfolio from 2 month ago up to the last 1/6th that consists of the portfolio from 6 month ago. In addition, I construct a long-short portfolio that goes long the portfolio of the highest IVOL firms and short the one with the lowest IVOL firms. This portfolio should directly reveal the effect of IVOL on the cross-section of expected returns. After having constructed the monthly time-series of IVOL portfolios introduced above, measures such as the portfolio's monthly average equal- and value-weighted total and excess returns as well as the Jensen's alphas are computed. Lastly, these averages are tested for significance by a standard t-test where standard errors are corrected by the concept of Newey and West (1987). The objective is to find out whether the average difference of the alphas from the long-short portfolio as well as its average equal- and value-weighted total and excess returns are statistically significant different from zero and have a negative sign which could be interpreted as finding of the IVOL puzzle in the data.

After having explained the general portfolio-based approach, I will now clarify the specification used as a reference point for the adjustment analysis. I mostly use a 1/0/1 trading strategy where IVOL is computed over the past month and equal- as well as value-weighted portfolio returns are computed over the subsequent month. Stocks are sorted based on their past month IVOL relative to the FF3 model into quintile portfolios. Then the long-short portfolio is constructed from the quintile portfolio containing stocks with the highest past month IVOL minus the one containing stocks with lowest IVOL. All these portfolios are rebalanced monthly. For all portfolios I compute the equal- and value-weighted total and excess returns as well as their time-series alphas relative to the FF3 and the Fama and French (2018) six-factor model (FF6 model hereafter) based on the equal- and value-weighted portfolio excess returns respectively. The FF6 model augments the FF5 model for better risk-correction with the momentum factor from Carhart (1997) (see Fama & French, 2018). Value-weighted monthly returns are calculated by averaging the returns of all stocks within the portfolio at month t while weighting them with their corresponding MTCAP. Alphas are calculated from equation 1 over the complete sample time span where the dependent variable is now the monthly equal- or value-weighted portfolio excess return. As further portfolio characteristics I calculate the time-series average of the monthly size and BM of stocks within the portfolio. The average monthly size and BM are computed as simple averages over the respective values for all stocks in the portfolio at the specific month. Furthermore, I compute the standard deviation of the monthly excess returns for every portfolio. For the t-tests of the average portfolio alphas and returns, I implement the Newey and West (1987) correction using 7 lags.

3.4. Adjustment Analysis

I proceed with the explanation of the adjustments to both reference research concepts introduced above with the aim of finding out how these affect the IVOL puzzle finding. I categorize them into the following three groups: general adjustments, regression-related adjustments and portfolio-related adjustments. The general adjustments are explained in section 3.4.1 and consist of all modifications that are applicable to both research methods. Sections 3.4.2 and 3.4.3 present the regression- and the portfolio-related adjustments, which contain alterations that are only employable in their respective research frameworks.

3.4.1. General Adjustments

In this section I introduce all adjustments which are applicable independent of the underlying research approach. These allow me not only to identify how research findings of a single approach are changing but also facilitates a cross-concept comparison of the effect from the same adjustment on both approaches respectively. I classify my general adjustments into the following two categories: idiosyncratic volatility-related adjustments and sample-related adjustments. They are applied sequentially and independently from each other to both reference concepts. Alongside these modifications, I investigate how findings on the IVOL puzzle change.

The idiosyncratic volatility-related general adjustments focus on the IVOL estimation procedure. Section 3.1 shows that several decisions have to be made by researchers when defining their IVOL measure. Among one of such decisions

is the choice of estimation window and the data frequency used for estimation of equation 2. Ang et al. (2009) use different estimation windows for their IVOL estimate showing that their evidence of an IVOL puzzle is robust to these adjustments in the regression-based research approach. In contrast, Bali and Cakici (2008) claim that when they change the data frequency used for the IVOL computation from daily to monthly data they are not able to find any significant relationship between IVOL and the cross-section of expected returns. Consequently I modify these two dimensions as well, while I keep using the FF3 model for risk-correction purposes. I begin by extending the estimation window using daily data over the past 12 months where I require a minimum of 200 days of stock return data for the estimation. Afterwards I estimate IVOL from monthly data using a rolling window of 1 and 5 years where I require at least 10 and 24 monthly stock returns to exist.

Another IVOL-related adjustment focuses on the choice of risk-correction model that is usually assumed to be the FF3 model. Malkiel and Xu (2006) use the FF3 model and the CAPM for IVOL estimation and find a positive relationship between IVOL and expected returns. So if the FF3 model insufficiently captures the relevant systematic risk, then some of it might remain in the risk that is assumed to be idiosyncratic and consequently could be the cause of the IVOL puzzle. Newer models have emerged that claim to explain returns better than the FF3 model does, which is why I decided to also use them for the IVOL estimation and analyze these estimates respectively (see e.g. Fama & French, 2018; Hou, Mo, Xue, & Zhang, 2020). Again I use daily data over the current month and compute for each stock i equation 1 using the respective risk-correction model to obtain the residuals, which are then used in equation 2 to compute the corresponding IVOL measure. As additional risk-correction models I use the FF6, the four-factor model proposed by Stambaugh and Yuan (2017) (SY4 model hereafter) and the five-factor model of Hou et al. (2020) (HOU5 model hereafter). The SY4 model consists of the MKT and the SMB factor from Fama and French (1993) as well as two mispricing-related factors, which are constructed from bivariate sorts of stocks relative to a total of 11 prominent asset pricing anomalies. One of these factors captures the exposure to anomalies related to quantities such as net stock issues that can be affected by the firms' management, whereas the other is directed at performance-related anomalies like momentum that cannot directly be influenced by the management (see Stambaugh & Yuan, 2017). With this model I try to account for the mispricing-related explanation of the IVOL puzzle put forward by Stambaugh, Yu, and Yuan (2015). They argue that only the overpriced stocks show an IVOL puzzle which is even more pronounced for overpriced stocks that are put under short selling restrictions (see Stambaugh et al., 2015). Hence I try to investigate whether researchers accounting for that fact can resolve the IVOL puzzle finding. On the other hand, the HOU5 model takes an investment-based asset pricing approach. The model consists of factors related to the market excess return, size, investment-to-assets, profitability as well as an expected-growth factor. The first four factors are constructed from triple sorting stocks according to a 2 x 3 x 3 strategy on size, investment-to-assets and return on equity, whereas the last factor is constructed by independent 2 x 3 double sorts related to market equity and the expected 1-year-ahead investment-to-assets changes (see Hou et al., 2020; Hou, Xue, & Zhang, 2015). I include the model here to incorporate the investment-based asset pricing idea and because Hou et al. (2020) argue that it outperforms most of the other asset pricing models including the FF6 model.

Besides the IVOL-related adjustments I investigate adjustments related to the sample used for the analysis. Here I sort all stocks into subsample based on stock-specific criteria and conduct my analysis within each of these subsamples separately. I form two kinds of subsamples based on the following criteria: price and size. Price-related subsamples should account for the fact that the IVOL puzzle might be specific to low price stocks and excluding these stocks from the analysis can diminish or even resolve the puzzle completely (see Bali & Cakici, 2008). To disentangle this relationship I subdivide stocks according to their price into three groups using the cut-offs from L. H. Chen et al. (2012). The first group contains so-called "penny stocks" with a price lower than 5$, whereas the second group contains low price stocks with a price higher than or equal to 5$ but lower than 10$. Lastly, high price stocks have a price higher than or equal to 10$ and hence form the last group. The prices used for subsample formation are adjusted for delisting based on the ideas of Shumway (1997). Bali and Cakici (2008) argue that size-related effects also affect the IVOL puzzle, wherefore I analyze size based subsamples as well. Size groups are constructed in reference to Fama and French (2008) who classify stocks as microcaps, small or big stocks based on the cross-sectional distribution of MTCAP at each month. Hence, each month the breakpoints used to categorize stocks are calculated as the 20th and the 50th percentile of the cross-sectional distribution of the market capitalization for all stocks traded on the NYSE. Consequently, stocks with a MTCAP lower than this 20th percentile are classified as microcaps during the corresponding month. In contrast a MTCAP within the 20th and the 50th percentile classifies stocks as small, whereas one larger than the 50th percentile characterizes them as big stocks.

3.4.2. Regression-Related Adjustments

This section presents all adjustments that are only applicable in the regression-based research concept as they address methodological peculiarities. I summarize them in the following two categories: risk-related control variables and estimation procedure.

The category named "risk-related control variables" refers to modifications in the variables that are used for risk-correction in equation 5. These variables can be grouped into two classes where the first consists of all variables that are related to the risk-correction model responsible to capture the systematic risk and the second class subsumes all the firm characteristic control variables that should characterize the left over firm-specific risk. Systematic risk-correction model-

related adjustments consist of modification to the time-series risk factor loadings used in equation 5 that should correct for risk when estimating the risk premia. In the reference concept explained above, the FF3 model was used for that purpose. Now I rotate independently among four further risk-correction models. Nevertheless, the computation of time-series factor loadings and their incorporation into the cross-sectional regression remains similar as explained in section 3.2. I employ the following models as an adjustment here: the FF6, the SY4 and the HOU5 model and a model that augments the FF3 model by a short and long-term reversal factor. These two short and long-term reversal factors are calculated as proposed by French (2021) from double sorting stocks relative to their past returns over the last month or the prior months from $t-60$ to $t-13$ and size respectively. I call the FF3 model augmented with the reversal factors the augmented Fama-French model (FFA hereafter). The second class of risk-related control variable adjustments focuses on the firm characteristic control variables from the vector of firm characteristics used in equation 5. Here I integrate 3 additional control variables. First I include the stock's monthly trading volume to account for possible trading volume-related effects that have also been considered in Ang et al. (2006). The effect of trading volume on expected returns was first discovered by Gervais, Kaniel, and Mingelgrin (2001). As a further control variable related to the liquidity effect, I include the monthly average over the daily bid-ask-spreads calculated from the difference between the ask and bid price, as Han and Lesmond (2011) argue that bid-ask bounces are among the drivers of the IVOL puzzle. If one or both of the bid and ask prices are not available for a specific day, I use the lowest and highest trading prices, or if these do also not exist, I choose the closing bid and ask prices for approximation of the bid-ask spread from that day instead. Lastly, I control for the maximum daily return during the current month computed based on the ideas of Bali et al. (2011) as an additional proxy for the investor's skewness preferences. I again rotate among all additional firm characteristic control variables sequentially such that they are integrated independently from each other to the benchmark setting and shifts in results can be investigated accordingly. In a last step, I integrate all additional control variables jointly and analyze the IVOL premium respectively.

The estimation procedure-related adjustment, on the other hand, refers to all modifications in the estimation of the first step time-series regressions and the second step cross-sectional regressions introduced in section 3.2. Related to the time-series regression step, I adjust the horizon over which the risk factor loadings are estimated. This idea is based on the arguments of Z. Chen and Petkova (2012) who claim that estimates over the full-sample would be most precise if the risk factor loadings are truly constant but as risk exposure of firms is likely not constant over time, the optimal time horizon maximizing the precision of the estimate remains an unsolved question. Therefore, I allow for varying degrees of time-variation in the risk factor loadings by using different window sizes for the estimation of equation 1 instead of the usual one month period. I choose window sizes covering the full-sample, the past 12 months and the past 60 months for which the existence of at least 800, 200 and 800 daily stock return observations are required. The cross-sectional regression step adjustments focus on the technique used for estimation of the risk premia by the cross-sectional regression of equation 5. This idea is motivated by the studies of Bali and Cakici (2008) as well as Hollstein and Prokopczuk (2020) who point out that findings and estimation precision of the risk premia change if returns are weighted differently. Hollstein and Prokopczuk (2020) claim that relying less on highly volatile stocks might improve the estimation precision of the risk premia, which is why they use a weighting matrix based on the empirical variances. I pick up this idea and apply a generalized least squares (GLS) technique where a diagonal weighting matrix consisting of the inverse of the estimated stock return variances is utilized. The estimation of the return variance is conducted over the past 1, 12 and 60 months as well as the complete sample using daily data where I require a stock to report at least 15, 200, 800 and 800 past daily return observations to allow for a valid estimate of the firm return variance at the specific month. These estimates are then annualized by multiplication with 250 and used in the weighting matrix respectively.

3.4.3. Portfolio-Related Adjustments

In the following I present all adjustments unique to the portfolio-based research concept that I consider in my study. I separate them into the following two groups: portfolio characterization and bivariate portfolio sorts.

In the category named "portfolio characterization" I subsume all modifications that are related to the way the univariate IVOL portfolios are computed and analyzed as explained in section 3.3. More precisely, I adjust the breakpoints used for portfolio sorting, the number of portfolios analyzed, the computation of risk-adjusted portfolio returns as well as the trading strategy used to construct portfolios.

Adjustments to the breakpoints used for sorting stocks into portfolios are primarily motivated by mitigation of the size effect that might distort the findings related to the IVOL puzzle as pointed out by Bali and Cakici (2008). These adjusted breakpoints again define the IVOL cutoffs on the basis of which I assign all stocks in my sample to portfolios subsequently. For the calculation of the breakpoints I first use only stocks traded on the NYSE and afterwards focus on breakpoints that sort stocks into equal market share portfolios. For the equal market share breakpoints I follow Bali and Cakici (2008) and rank all stocks in the sample according to their idiosyncratic volatility in the specific month. Based on this ranking, I adjust the breakpoints such that within all IVOL quintile portfolios the stocks in these portfolios together make up the same share of the total market capitalization. This means that, for example, the lowest IVOL quintile portfolio consists of all the stocks with the lowest IVOL that together account for approximately 20% of the total MTCAP.

In addition to the different breakpoints, the number of univariate portfolios is a further aspect that is adjusted here.

The idea is to vary the degree of dispersion in the data and hence also the measures of comparison including returns and alphas by increasing the number of portfolios analyzed. Additionally to the standard quintile portfolios, I sort all stocks into 10 and 15 portfolios respectively and implement the reference research concept accordingly. Ten portfolios have also been selected in the study of Spiegel and Wang (2005) who found a positive IVOL risk premium.

The adjustment in risk-adjusted portfolio returns addresses the choice of the risk-correction model used to calculate the portfolio alphas. These models are responsible for capturing the systematic risk in stock returns and should be able to explain the stock return differences by their exposure to the model specific systematic risk. As various studies use different models for this purpose, the choice might influence findings consequently (see e.g. Bali & Cakici, 2008; Bali, Del Viva, Lambertides, & Trigeorgis, 2020; Boehme, Danielsen, Kumar, & Sorescu, 2009). In addition to the models used in the reference concept, I consider the SY4, the HOU5 and the model proposed by Daniel, Hirshleifer, and Sun (2020). The Daniel et al. (2020) (DANIEL3 hereafter) model consists of three factors where one is the market factor also used in the Fama-French models and the other two are theory-based behavioral factors that should capture short- and long-term mispricing dynamics which are induced by investor behavior. The short-term mispricing factor is related to the post-earnings announcement drift phenomenon, whereas the long-term mispricing factor is derived from investors issuance and repurchase activities as well as related misperceptions (see Daniel et al. (2020)).

Lastly, the trading strategies are also modified and analyzed as done by Ang et al. (2006). In addition to the usual 1/0/1 strategy, I compute the results related to the following trading strategies: 1/1/1, 1/1/12, 1/0/12, and 12/0/12. These are implemented based on the procedure of Jegadeesh (1990) as explained in section 3.3. I include the strategies that include a one month waiting period to find out whether this waiting period would allow researchers to mitigate the IVOL puzzle effect and hence deliver evidence on a short-term reversal effect based explanation as argued by W. Huang, Liu, Rhee, and Zhang (2010). The longer holding period of 12 months should account for the explanation that the IVOL puzzle is driven by a short-term overreaction and can even reverse over a period longer than one month, as found by Rachwalski and Wen (2016). Similar to the study of Ang et al. (2006), the 12 months formation period is included again to mitigate the effect of potential short-term events.

As previously only univariate IVOL portfolios were considered, effects from further firm characteristic-related variables that might influence the IVOL puzzle were not taken into account. Hence, I now construct average portfolios out of bivariate dependent portfolio sorts, as also proposed by Ang et al. (2006) and Boyer et al. (2010), that should account for further firm-specific effects. I control for the following variables: size, BM, the stock's current month return, the Amihud (2002) illiquidity measure calculated over the current month, return skewness over the current month, the expected idiosyncratic skewness measure of Boyer et al. (2010), the Harvey and Siddique (2000) coskewness measure computed over the past 5 years, the stock's monthly trading volume, the current month average daily bid-ask-spreads and the maximum daily return of the current month as proposed by Bali et al. (2011). Based on each of these control variables I form the bivariate portfolios using the usual 1/0/1 strategy and again compute EW and VW portfolio returns. First I sort all stocks each month respective to the control variable currently considered into quintile portfolios. Within these quintile portfolios, the stocks are again sorted into quintile portfolios relative to their IVOL on the basis of which also the corresponding long-short portfolios are formed. To disentangle the IVOL effect from those of the control variables, I calculate the average IVOL quintile portfolios where each has the same dispersion in the control variable. So every month I average each IVOL quintile portfolio, as well as the long-short portfolio, over all the quintiles of the control variable to obtain the monthly time-series of these six IVOL-related average portfolios. Finally the time-series average across the monthly average IVOL portfolios is calculated. With these average portfolios I can investigate the return pattern of the quintiles IVOL portfolios that remains after controlling for the effect of the control variable.

4. Data

My study is primarily based on US-based individual stock data that is derived from the Center for Research in Security Prices database database (CRSP hereafter). From CRSP I obtain the closing stock returns, their closing prices, trading volume and end-of-period market capitalization on a daily and monthly basis. I use all stocks that are traded on the NYSE, AMEX and NASDAQ regardless of their security type captured by the share code. Further information on the stocks including information on the security type, the exchange on which the corresponding stocks are listed, the Standard Industrial Classification (SIC) code that classifies the companies into industry groups as well as the shares outstanding are also derived from the CRSP database. To calculate the bid-ask spreads I derive the daily and monthly bid and ask price as well as the corresponding lowest bid and the highest ask price again from CRSP. The lowest bid (highest ask) price either consist of the lowest (highest) trading price or the closing bid (ask) price during the time interval respectively. Throughout the analysis I generally use delisting adjusted returns and prices following the methods of Shumway (1997) for which I obtain delisting data such as the delisting code, the delisting date, the delisting price and the delisting return from CRSP as well. The accounting data including stockholders equity, deferred taxes, investment tax credit, preferred stock redemption value, preferred stock liquidating value and preferred stock par value are obtained from the CRSP/Compustat Merged (CCM) database. In general my sample period covers the horizon from July 1963 until December 2020. This starting point is chosen, as at that point in time most of the accounting-related and factor data are

available for the majority of the models. Data on the risk factors from the FF3 and FF5 as well as the MOM, long-term and short-term reversal factors are obtained on a daily and monthly basis from the Kenneth French's Data Library (see French (2021)). From here also the data on the risk-free rate, that is the one-month Treasury Bill rate, is acquired. Daily and monthly factor data for the SY4 and HOU5 model come from Robert Stambaugh's and the global-q webpage respectively (see Hou, Xue, & Zhang, 2021; Stambaugh, 2021). However, it is of note that the SY4 model data is only available until December 2016 and hence a shorter sample is used when this model is applied. The factor data related to the DANIEL3 model are only available at a monthly frequency up to December 2018 and are obtained from Kent Daniel's webpage (see Daniel, 2021). Lastly, the data related to the expected idiosyncratic skewness measure of Boyer et al. (2010) is obtained from Brian Boyer's webpage until December 2016 and is afterwards calculated manually based on their proposed concept until December 2020 (see Boyer, 2021).

5. Reference Results

In this section I replicate the findings of the IVOL puzzle using the reference research concepts introduced in section 3.2 and 3.3. Prior to that I provide cross-sectional summary statistics on the data used for replication in section 5.1.

5.1. Summary Statistics

The summary statistics for all relevant variables from the reference concepts can be found in table 1. These allow for a preliminary explorative analysis that can reveal first insights on interactions between the IVOL puzzle and related firm characteristic effects. Panel A provides the time-series average of all parameters that characterize the monthly cross-sectional distribution of the corresponding variable, whereas Panel B shows the time-series average of their monthly cross-sectional Pearson product-moment and Spearman rank pairwise correlations.

The mean column in Panel A shows the cross-sectional means of the variables in the average month that can be interpreted as the level of the respective variable for the average stock in the sample. Hence, the average stock has risk factor loadings relative to the MKT, SMB and HML of about 0.806, 0.643 and 0.202, an annualized IVOL of 40.83%, a size measured as the natural logarithm of the MTCAP of 4.601, a book-to-market ratio of equity of 3.214, a monthly total return of 0.966%, a one-month ahead excess return of 0.596%, a monthly return skewness of 0.209, an Amihud (2002) illiquidity measure over the current month of 5.396, an expected idiosyncratic skewness of 1.157 and a coskewness of -1.085. Comparing mean and median it can be seen that the median values are mostly smaller than the cross-sectional means on average except for $CoSkew$ and that they lay rather close together. The only exceptions are BM, R_t, r_{t+1} and $Illiq_{1M}$ that have a median in the average month of 0.711, 0.019%, -0.352% and 0.310 respectively. However, all these values remain within a bandwidth of one standard deviation from their corresponding mean. About half of the variables including IVOL, BM, R_t, r_{t+1}, $Illiq_{1M}$ and $CoSkew$ are positively skewed with an average monthly cross-sectional skewness ranging from 2.085 for $CoSkew$ up to 32.125 for BM, whereas the remaining variables are rather symmetric around the mean. Additionally, almost all variables except for $Size$ are leptokurtic with an average cross-sectional excess kurtosis ranging from 2.634 for $Skew_{1M}$ to 1337.492 for BM when excluding $Size$. The high level of kurtosis and the positive skewness of IVOL, BM, R_t, r_{t+1}, $Illiq_{1M}$ and $CoSkew$ shows that the cross-sectional distribution of these variables is characterized by a small number of extremely positive observations which also cause most of their variability. This can also be seen by the percentiles of their distribution as the right tails are longer than their left tails because the average distance between the 95th percentile and the maximum value is larger than the one between the 5th percentiles and the minimum in all these cases. The extreme observation might also cause the differences between the mean and median values of the BM, R_t, r_{t+1} and $Illiq_{1M}$. Nevertheless, it is of note that the distribution of IVOL, BM, R_t, r_{t+1} and $Illiq_{1M}$ has heavy right tails partly due to the mechanical reason, as these variables have a lower bound and are unbounded from above because IVOL, BM and $Illiq_{1M}$ are not defined for values below zero and R_t as well as r_{t+1} cannot become smaller than -100%. All other variables except for $Size$ are more concentrated around their means as they are leptokurtic but barely skewed. However, they have a heavy tailed distribution as the average distance between the minimum and 5th percentile as well as 95th percentile and maximum exceeds, except for $Skew_{1M}$, a distance of three standard deviations. The last column of Panel A depicts the average number of stocks for which the corresponding variable is available. Excluding BM these values range from 4646 for $EIdioSkew$ to 5730 for r_{t+1}. BM has on average 3727 valid observations.

Panel B of table 1 presents the average monthly cross-sectional pairwise correlations for all variables, where the upper-diagonal matrix present the Pearson product–moment and the lower-diagonal one the average Spearman rank correlations. The aim is to understand the cross-sectional relations between the variables with a focus on the relationship to IVOL. The Pearson product-moment correlations reveal linear relationships, whereas the Spearman rank correlations detect monotonic but not necessarily linear dependencies. I begin with the analysis of the pairwise correlations involving IVOL to derive insights on potential drivers of the IVOL puzzle. Then I briefly repeat the analysis for r_{t+1} to investigate the variables return predictability properties. Finally I comment on some of the remaining correlations.

The Pearson product-moment (Spearman rank) correlation between IVOL and the MKT, SMB and HML risk factor loadings are low with values of 0.064 (0.138), 0.096 (0.148) and 0.000 (0.010) indicating that the information content captured by these variables and IVOL is likely different. Nevertheless, stocks with high IVOL are also more likely to have a high loading on the MKT and SMB. The correlations to the

Table 1: Cross-Sectional Summary Statistics

This table presents summary statistics for all variables of my sample that are related to the replication of the reference research concepts. The sample covers the period from July 1963 until December 2020 and includes all stocks traded on the NYSE/AMEX/NASDAQ. $\beta(MKT)$, $\beta(SMB)$ and $\beta(HML)$ refer to the time-series loadings onto the MKT, SMB and the HML risk factors respectively that were calculated for each stock based on daily data over the current month using equation 3. IVOL is the annualized idiosyncratic volatility (in percentage points) calculated via equation 4 by benchmarking the daily excess returns to the FF3 model over the current month. *Size* refers to the natural logarithm of the firm's market capitalization. BM gives the firm's book-to-market ratio of equity for the specific month. R_t is the stock return at the current month. r_{t+1} is the excess stock return over the next month. $Skew_{1M}$ reports the total return skewness calculated over the current month. $Illiq_{1M}$ is the Amihud (2002) illiquidity measure computed over the current month. $EIdioSkew$ reports the expected idiosyncratic skewness of a stock calculated by the method of Boyer et al. (2010). $CoSkew$ stands for the firm's coskewness measure computed as proposed by Harvey and Siddique (2000). Panel A presents the time-series means of the monthly mean (*Mean*), standard deviation (*SD*), skewness (*Skew*), excess kurtosis (*Kurt*), minimum (*Min*), fifth percentile (*5%*), 25th percentile (*25%*), median (*Median*), 75th percentile (*75%*), 95th percentile (*95%*), and maximum (*Max*) values of the cross-sectional distribution for each variable. The column labeled *n* indicates the average number of stocks for which the corresponding variable is available. Panel B reports the time-series averages of the monthly cross-sectional Pearson product-moment and Spearman rank pairwise correlations between each of the variables. Here the above-diagonal entries present the average Pearson product–moment correlations, whereas the below-diagonal entries present the average Spearman rank correlations.

Panel A: Cross-Sectional Distribution

	Mean	*SD*	*Skew*	*Kurt*	*Min*	*5%*	*25%*	*Median*	*75%*	*95%*	*Max*	*n*
$\beta(MKT)$	0.806	1.996	-0.071	67.008	-26.265	-1.804	-0.060	0.695	1.625	3.736	26.523	5716
$\beta(SMB)$	0.643	2.906	0.379	56.398	-37.052	-3.115	-0.560	0.432	1.748	4.966	38.887	5716
$\beta(HML)$	0.202	3.437	-0.017	77.171	-47.098	-4.561	-1.095	0.156	1.511	4.989	46.163	5716
$IVOL$	40.833	37.413	5.567	110.012	0.785	9.259	19.455	31.354	50.661	101.802	820.888	5715
$Size$	4.601	1.935	0.297	-0.015	-2.294	1.644	3.229	4.473	5.858	8.002	11.771	5734
BM	3.214	41.670	32.125	1337.492	0.005	0.146	0.405	0.711	1.167	3.168	1893.320	3727
R_t	0.966	15.165	3.555	84.860	-83.581	-18.509	-5.848	0.019	6.249	22.654	292.276	5725
r_{t+1}	0.596	15.164	3.527	83.627	-83.822	-18.888	-6.222	-0.352	5.883	22.297	291.173	5730
$Skew_{1M}$	0.209	0.990	0.061	2.634	-4.000	-1.354	-0.276	0.182	0.688	1.851	4.155	5662
$Illiq_{1M}$	5.396	40.307	22.698	902.555	0.000	0.007	0.056	0.310	1.809	19.288	1883.481	4885
$EIdioSkew$	1.157	0.664	0.616	16.003	-2.659	0.258	0.684	1.089	1.590	2.219	7.409	4646
$CoSkew$	-1.085	9.442	2.085	86.961	-77.130	-14.323	-4.909	-0.952	2.609	11.543	151.906	4744

Panel B: Correlations

	$\beta(MKT)$	$\beta(SMB)$	$\beta(HML)$	$IVOL$	$Size$	BM	R_t	r_{t+1}	$Skew_{1M}$	$Illiq_{1M}$	$EIdioSkew$	$CoSkew$
$\beta(MKT)$		0.213	0.296	0.064	0.108	-0.007	0.020	-0.011	0.030	-0.039	-0.100	0.011
$\beta(SMB)$	0.237		0.211	0.096	-0.035	-0.012	0.021	-0.010	0.028	0.001	0.011	-0.003
$\beta(HML)$	0.280	0.197		0.000	-0.010	0.004	0.005	-0.003	0.010	0.009	-0.001	0.005
$IVOL$	0.138	0.148	0.010		-0.382	-0.004	0.150	-0.026	0.195	0.366	0.386	-0.039
$Size$	0.168	-0.018	-0.002	-0.412		-0.091	0.058	-0.005	-0.032	-0.284	-0.674	0.077
BM	-0.069	-0.029	0.078	-0.014	-0.252		0.001	0.001	-0.005	0.084	0.128	0.007
R_t	0.023	0.011	0.010	-0.011	0.109	0.020		-0.037	0.350	-0.009	-0.019	0.005
r_{t+1}	-0.007	-0.014	-0.002	-0.078	0.051	0.021	-0.030		-0.010	0.009	-0.007	-0.002
$Skew_{1M}$	0.038	0.030	0.011	0.179	-0.042	0.010	0.351	-0.024		0.025	0.076	-0.005
$Illiq_{1M}$	-0.155	0.047	0.021	0.478	-0.915	0.269	-0.067	-0.044	0.059		0.201	-0.017
$EIdioSkew$	-0.132	0.021	-0.001	0.441	-0.742	0.121	-0.081	-0.057	0.077	0.696		-0.069
$CoSkew$	0.020	-0.007	0.005	-0.069	0.117	0.022	0.010	0.005	-0.013	-0.116	-0.122	

BM and $CoSkew$ amount to -0.004 (-0.014) and -0.039 (-0.069), showing that these variables likely capture different information than IVOL does, wherefore they are improbable to explain the IVOL puzzle. Conversely, $Skew_{1M}$, $Illiq_{1M}$ and $EIdioSkew$ show correlations to IVOL of 0.195 (0.179), 0.366 (0.478) and 0.386 (0.441) respectively. Hence, stocks with higher IVOL likely exhibit high average return skewness, expected idiosyncratic skewness and are highly illiquid. To phrase it differently, it might be stocks with high skewness, high expected idiosyncratic skewness or those which are highly illiquid that drive the IVOL puzzle. Following the arguments of Barberis and Huang (2008) as well as Boyer et al. (2010), the low returns from high IVOL stocks might indeed be explained by their strong exposure to $Skew_{1M}$ and $EIdioSkew$, as stocks with high levels of these characteristics are assumed to have low returns respectively. The liquidity-related explanation of Han and Lesmond (2011) could also explain the IVOL puzzle, as high IVOL stocks also appear to be illiquid and for these stocks the bias in the IVOL estimate from microstructure effects is often high. An illiquidity premium as found by Amihud (2002), however, is unlikely explaining the puzzle, as then the high IVOL stocks would have to be liquid in order to posses low returns. $Size$ is the only variable with a non-negligible negative correlation to IVOL of -0.382 (-0.412), indicating that mainly small stocks exhibit high IVOL and consequently drive the IVOL puzzle. Ultimately, IVOL is positively correlated to the current month's return and negatively to the one-month ahead excess return with correlations of 0.150 (-0.011) and -0.026 (-0.078). Despite that the correlations to the next month excess returns are low, they still hint at the existence of the IVOL puzzle. It is of note that, even though the Pearson product-moment and the Spearman rank correlations to R_t have a different signs, they are low in magnitude wherefore an impact on the results is unlikely. For all variables the Pearson product–moment and Spearman rank measures are similar enough such that no non-linearity concerns in the relations arise that might distort the results.

Considering the Pearson product-moment correlations between r_{t+1} and all other variables as evidence on the direction of return predictability, I find that $\beta(MKT)$, $\beta(SMB)$, $\beta(HML)$, IVOL, $Size$, R_t, $Skew_{1M}$, $EIdioSkew$ and $CoSkew$ negatively predict the cross-section of next-month returns, whereas BM and $Illiq_{1M}$ do so in a positive direction. These relations are all in line with their theoretical predictions from the literature (see e.g. Amihud, 2002; Ang et al., 2006; Barberis & Huang, 2008; Boyer et al., 2010; Fama & French, 1993; Harvey & Siddique, 2000; W. Huang et al., 2010). It has to be noted that all correlations involving r_{t+1} are low ranging from -0.057 in case of the Spearman rank correlation with $EIdioSkew$ up to 0.009 for the Pearson product-moment correlation with $Illiq_{1M}$ and hence can only provide a rough evidence on the corresponding return predictability, meaning that there might be economically relevant relationships that are not found here.

The negative correlations of $Size$ to BM, $Illiq_{1M}$ and $EIdioSkew$ show that small stocks often have a high BM, high $EIdioSkew$ and are less liquid than big stocks. Especially the relationship to $Illiq_{1M}$ and $EIdioSkew$ with correlations of -0.284 (-0.915) and -0.674 (-0.742) has to be highlighted. The Spearman rank correlation between $Size$ and $Illiq_{1M}$ exceeds the Pearson product-moment correlation significantly and therefore indicates a non-linear relationship between the variables. BM is positively related to $Illiq_{1M}$ and $EIdioSkew$, wherefore value stocks are likely less liquid and have a high expected idiosyncratic skewness. R_t is only non-negligibly positively related to $Skew_{1M}$ implying that stocks with high returns in the current month also have on average positively skewed returns during that month. Finally illiquid stocks seem to exhibit on average a high expected idiosyncratic skewness and low coskewness.

5.2. Regression-Based Analysis

In this paragraph I empirically implement the regression-based reference research concept introduced in section 3.2 with the objective to replicate the IVOL puzzle that implies a negative cross-sectional relationship between stocks' IVOL and their expected returns. The corresponding results can be found in table 2. Here I display the average coefficients from the monthly cross-sectional regressions of equation 5, where I regress the next-month stock excess returns onto the stock-specific risk factor loadings relative to the FF3 model as well as further firm characteristics. In Panel A I use an OLS regression-based approach in the cross-sectional regression step to estimate the risk premia where all stock returns are weighted equally. Conversely, in Panel B I repeat the cross-sectional regression step with a WLS method where the stock excess returns are weighted by their contemporaneous MT-CAP.

Panel A of table 2 shows that the average coefficient on IVOL is negative and significant in all regression settings except for the simple regression in column (1) that includes IVOL as unique independent variable. No matter which control variable is added to the benchmark regression in column (2), the IVOL puzzle from Ang et al. (2006) can successfully be replicated. I discover that the effect is strongest, with a coefficient of -1.248 and a Newey-West robust t-statistic of -4.843 indicating statistical significance at the 1% level, when the Amihud (2002) illiquidity measure is added as control in column (6). In contrast, excluding the simple model of column (1), the effect is least significant in column (5) where I control for the coskewness measure of Harvey and Siddique (2000). There the average coefficient is -0.456 and has a robust t-statistic of -1.907 making it statistically significant at the 10% significance level. Considering all IVOL coefficients in Panel A, their magnitude ranges from -0.389 in the simple regression of column (1) to -1.248 when controlling for illiquidity in column (6). To interpret the economic magnitude of the IVOL puzzle, I compute the excess return effect related to a one standard deviation change in IVOL while all other characteristics are held constant. For that I multiply the average IVOL coefficients with the average cross-sectional standard deviation of IVOL computed in Panel A of table 1. Hence, a one standard deviation increase in IVOL is associated with

Table 2: Regression-Based Results - Reference Results

This table presents the average coefficients from monthly Fama and MacBeth (1973) cross-sectional regressions for individual stocks where each column refers to a different cross-sectional regression specification. Using equation 5, each month I regress the next-month excess firm returns (in percentage points) on a constant, the annualized idiosyncratic volatility measure (in decimals) calculated from equation 4 using daily data over the current month, the stock-specific risk factor loadings relative to the FF3 model over the next month, the end of month firm size defined as the natural logarithm of the market capitalization, the end of month firm book-to-market ratio of equity, the monthly return (in percentage points) computed from the end of the previous month to the current month, as well as additional control variables related to further firm characteristics that are incorporated into the regression equation successively. Among these firm characteristic variables are the following: $Skew_{1M}$ which is the monthly return skewness measured over daily return data from the current month, $EIdioSkew$ that is the expected idiosyncratic skewness measure from Boyer et al. (2010), $CoSkew$ which is the coskewness measure from Harvey and Siddique (2000) and $Illiq_{1M}$ being the Amihud (2002) illiquidity measure computed over the current month. In Panel A I use simple stock excess returns to run the OLS regression wherefore this procedure is named to be the equal-weighted Fama-MacBeth regression as all stocks are weighted equally. On the other hand, in Panel B I estimate the Fama-MacBeth regression with a weighted-least squares approach, where all individual monthly stock returns are weighted by their current month market capitalization, which is why this method is named value-weighted Fama-MacBeth regression. I report the t-statistics testing the null hypothesis that the average coefficient is equal to zero in parenthesis below each coefficient where corresponding standard errors are corrected for autocorrelation and heteroskedasticity by implementation of the Newey and West (1987) method using 7 lags. *, ** , and *** indicate significance at the 10%, 5%, and 1% level, respectively. The row "Adjusted R^2" reports the time-series average of the cross-sectional adjusted R^2's. The last row reports the average number of stocks used for the monthly cross-sectional regressions. The sample period covers July 1963 to December 2020.

Panel A: Equal-Weighted Fama-MacBeth Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Constant$	0.706***	1.328***	1.326***	1.779***	1.318***	1.389***	1.845***
	(3.151)	(3.594)	(3.579)	(4.140)	(3.546)	(3.828)	(4.579)
$IVOL$	- 0.389	-0.616**	-0.627**	-0.431**	-0.456*	-1.248***	-0.975***
	(-1.369)	(-2.520)	(-2.567)	(-2.010)	(-1.907)	(-4.843)	(-4.133)
$\beta(MKT)$		0.253***	0.249***	0.197*	0.237**	0.268***	0.193*
		(2.645)	(2.615)	(1.709)	(2.494)	(2.738)	(1.649)
$\beta(SMB)$		0.077**	0.079**	0.019	0.079**	0.067*	0.005
		(2.009)	(2.072)	(0.382)	(2.198)	(1.703)	(0.104)
$\beta(HML)$		-0.038	-0.041	-0.073	-0.036	-0.039	-0.066
		(-0.871)	(-0.965)	(-1.352)	(-0.775)	(-0.910)	(-1.113)
$Size$		-0.160***	-0.160***	-0.186***	-0.153***	-0.134***	-0.159***
		(-4.176)	(-4.199)	(-4.752)	(-3.990)	(-3.664)	(-4.349)
BM		-0.001	-0.001	0.222***	0.002	0.007	0.140**
		(-1.282)	(-1.308)	(3.302)	(0.628)	(1.063)	(2.222)
R_t		-0.047***	-0.050***	-0.041***	-0.049***	-0.048***	-0.043***
		(-11.273)	(-11.421)	(-10.217)	(-11.323)	(-11.103)	(-10.694)
$Skew_{1M}$			0.068***				0.102***
			(2.808)				(4.129)
$EIdioSkew$				-0.353***			-0.302**
				(-2.778)			(-2.561)
$CoSkew$					-0.004		-0.011**
					(-0.953)		(-2.307)
$Illiq_{1M}$						0.046***	0.035***
						(4.880)	(3.492)
$Adjusted\ R^2$	0.017	0.061	0.061	0.060	0.064	0.070	0.067
n	5648	3676	3647	3520	3366	3234	3056

Panel B: Value-Weighted Fama-MacBeth Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Constant$	0.546***	1.096***	1.104***	1.523***	1.096***	1.062***	1.531***
	(2.913)	(3.529)	(3.574)	(4.062)	(3.522)	(3.365)	(4.104)
$IVOL$	-0.660	-0.927***	-0.967***	-0.747**	-0.799**	-0.963***	-0.795**
	(-1.422)	(-2.665)	(-2.738)	(-2.241)	(-2.295)	(-2.720)	(-2.287)
$\beta(MKT)$		0.191	0.190	0.105	0.178	0.202	0.096
		(1.442)	(1.433)	(0.649)	(1.348)	(1.501)	(0.593)
$\beta(SMB)$		-0.009	-0.009	-0.081	-0.016	-0.011	-0.087*
		(-0.215)	(-0.216)	(-1.599)	(-0.361)	(-0.252)	(-1.748)
$\beta(HML)$		-0.011	-0.009	-0.031	-0.009	-0.013	-0.027
		(0.150)	(-0.131)	(-0.325)	(-0.126)	(-0.179)	(-0.279)
$Size$		-0.072***	-0.073***	-0.082**	-0.070***	-0.070**	-0.081**
		(-2.688)	(-2.720)	(-2.497)	(-2.608)	(-2.566)	(-2.443)
BM		0.002	0.002	-0.016	-0.002	0.013	-0.026
		(0.596)	(0.620)	(-0.172)	(-0.312)	(0.785)	(-0.284)
R_t		-0.026***	-0.029***	-0.022***	-0.027***	-0.026***	-0.025***
		(-6.174)	(-6.687)	(-4.896)	(-6.173)	(-6.250)	(-5.350)
$Skew_{1M}$			0.092***				0.105***
			(3.074)				(3.308)
$EIdioSkew$				-0.249			-0.237
				(-1.621)			(-1.572)
$CoSkew$					-0.006		-0.016*
					(-0.786)		(-1.959)
$Illiq_{1M}$						0.042	0.001
						(1.601)	(0.096)
$Adjusted\ R^2$	0.027	0.124	0.127	0.142	0.132	0.128	0.153
n	5644	3676	3647	3520	3366	3234	3056

a decrease in expected excess returns ranging from 0.146% (0.389 x 0.374) to 0.467% (1.248 x 0.374) per month. Investigating the same one standard deviation IVOL increase in context of the model from column (7) that incorporates all controls simultaneously, the expected returns decrease by about 0.365% per month. This effect remains economically relevant and proves the negative relationship between IVOL and expected excess returns implied by the IVOL puzzle.

The average coefficients for the remaining variables are mostly in line with their theoretical predictions. Here the average risk premia for all FF3 model risk factor loadings, except for the on $\beta(HML)$, carry a positive sign as predicted by Fama and French (1993). However, it is of note that only $\beta(MKT)$ achieves a positive risk premium that is statistically significant in all regressions with Newey-West robust t-statistics between 1.649 and 2.738. The $\beta(SMB)$ premium is not statistically significant when the models of column (4) or (7) are considered, whereas the one of $\beta(HML)$ is not significant in any of the models. A negative size premium ranging from -0.134 in column (6) to -0.186 in column (4) is verified in all regression models and is furthermore always highly significant at the 1% level where the robust t-statistics lay between -3.664 when controlling for illiquidity and -4.752 in the regression that controls for $EIdioSkew$. In contrast to that, the BM risk premium ranges from 0.222 to -0.001 and is on average only significant at the 1% level when controlling for $EIdioSkew$ or at the 5% level in the complete regression model of column (7) with t-statistics of 3.302 and 2.222 respectively. As predicted by Jegadeesh (1990), I find a negative risk premium for R_t that is significant at the 1% level in all regression models and varies between -0.050 in column (3) and -0.041 when controlling for $EIdioSkew$. In addition, I confirm the negative return premium of expected idiosyncratic skewness found by Boyer et al. (2010) that is statistically significant at a level not less than 5% and the positive illiquidity premium found by Amihud (2002) which is significant at the 1% level in all regressions. $CoSkew$ carries a negative premium consistent with Harvey and Siddique (2000) that is, however, only significant at the 5% level in the complete regression model. Only the estimates on $Skew_{1M}$ do not match the predictions of Barberis and Huang (2008). Here I find a positive premium that varies between 0.068 and 0.102 and is always significant at the 1% level.

Panel B confirms most of the findings from Panel A, however, the IVOL risk premia are higher in magnitude for most of the regression models even though the overall maximum has reduced slightly. Now these premia range from -0.660 in the simple model to -0.967 when controlling for $Skew_{1M}$ and are statistically significant at a level not less than 5%, except for the simple model where it is not significant at any reasonable level. The economic magnitude computed by a one standard deviation increase in IVOL, when holding all other characteristics constant, amounts to a decrease of expected returns ranging from 0.247% to 0.362% per month respectively. In consequence I discover the IVOL puzzle found by Ang et al. (2006) again.

The remaining coefficient estimates are mostly in line with those from Panel A. However, all the FF3 model factor loadings risk premia are no longer significant at any reasonable level except for the one on the $\beta(SMB)$ that is significant at the 10% level in column (7). In addition, the $\beta(SMB)$ premium has switched signs for all models and is, therefore, no longer in line with the predictions of Fama and French (1993). The size-related risk premium decreased in magnitude to values ranging from -0.070 for regressions (5) and (6) to -0.082 in regression (4) with robust t-statistics of -2.566 or -2.608 and -2.497 respectively. It is of note that when controlling for illiquidity, expected idiosyncratic skewness or incorporating all controls, the premium is now only significant at the 5% level. On the other hand, BM is no longer significant in any model and the R_t premium remains negative and highly significant at the 1% level in all settings, however, it has decreased to values between -0.025 and -0.029. Lastly, the negative premium on $EIdioSkew$ and the positive premium on $Illiq_{1M}$ are no longer significant at any reasonable level, no matter which model is considered, whereas the negative premium of $CoSkew$ is now only significant at the 10% level in the complete regression model with an estimated risk premium of -0.016.

When comparing the average fit of the cross-sectional models by means of the adjusted R^2, I find a better fit for all models in Panel B as the adjusted R^2 are largely almost twice as large as the corresponding values from Panel A. This indicates that value-weighting enhances the model fit on average and therefore the results from the VW Fama-MacBeth regressions might be slightly more reliable than those of the EW Fama-MacBeth regression. From the last row of table 2 I conclude that, on average, approximately the same number of stocks is used for the monthly cross-sectional regression estimation, regardless of the weighting scheme employed. This number ranges from 5648 stocks in the simple model to 3056 stocks in the complete regression model, hence allowing, on average, for a valid risk premia estimation.

To sum up the results from Panel A and B, the negative relationship between IVOL and the cross-section of expected returns, implied by the IVOL puzzle, can be found in the risk premium estimates of the EW as well as the VW context. The effect is robust when controlling for skewness preferences and illiquidity measured by $Skew_{1M}$, $EIdioSkew$, $CoSkew$ and $Illiq_{1M}$ respectively and it is always statistically significant except when analyzing the simple regression model. Hence, in the regression-based reference setting, the findings of Ang et al. (2006, 2009) can be verified.

5.3. Portfolio-Based Analysis

This section deals with the replication of the IVOL puzzle using the portfolio-based concept introduced in section 3.3. If the puzzle exists, the portfolio of highest IVOL stocks has to underperform compared to the one with lowest IVOL stocks. Table 3 presents the average portfolio returns, Jensen's time-series alphas relative to the FF3 and the FF6 model, the standard deviation of portfolio excess returns as well as the average size and BM of all stocks within the corresponding portfolio. All stocks traded on the NYSE/AMEX/NASDAQ are

Table 3: Portfolio-Based Results - Reference Results

This table presents the average results from the univariate portfolio sorting procedure as explained in section 3.3 covering all stocks traded on the NYSE/AMEX/NASDAQ. Every month stocks are sorted into quintile portfolios based on their idiosyncratic volatility relative to the FF3 model over the past month that has been calculated by the usage of daily data for the respective month. These portfolios are rebalanced monthly according to the 1/0/1 trading strategy. Portfolio 1 (5) contains all stocks with lowest (highest) IVOL over the past month. The column labeled "5-1" refers to the long-short portfolio that goes long the portfolio of stocks with the highest IVOL and short the respective portfolio of stocks with the lowest IVOL, which is then also rebalanced monthly according to the 1/0/1 trading strategy. In the first four rows I report the average monthly total and excess returns for the portfolios (in percentage terms) and below each return in parenthesis the corresponding Newey and West (1987) adjusted t-statistics where I use 7 lags for the adjustment. Here it is of note that return differences for the long-short portfolio are the same for total and excess returns, wherefore I report them only once in the row for the total returns. SD illustrates the monthly standard deviation of the portfolio excess returns also denoted in percentage terms. The $Size$ and BM report the average natural logarithm of the market capitalization for firms within the portfolio as well as their average book-to-market ratio respectively. The α^{FF3} and the α^{FF6} denote the monthly average of Jensen's time-series alphas relative to the FF3 and the FF6 model that were calculated by equation 1 using the monthly portfolio excess returns over the complete sample horizon. Below each of these alphas I provide the corresponding t-statistics that were again corrected by the procedure of Newey and West (1987) using 7 lags. In Panel A I use the equal-weighted portfolio returns for the analysis, where all firms are getting the same weight. On the other hand, Panel B uses value-weighted portfolio returns that were calculated by weighting the stock returns within the portfolio by their market capitalization which is observable at beginning of the month in order to give higher weights to bigger stocks respectively and therefore diminish the effects that might be explicitly related to small stocks. The complete sample period covers July 1963 to December 2020.

Panel A: Equal-Weighted Portfolio Sorts by Idiosyncratic Volatility

	1	2	3	4	5	5 - 1
Total Return	0.868	1.094	1.158	1.046	0.886	0.018
	(3.893)	(4.510)	(4.169)	(3.167)	(2.216)	(0.064)
Excess Return	0.495	0.721	0.785	0.674	0.514	
	(2.227)	(2.977)	(2.823)	(2.028)	(1.273)	
SD	5.128	6.038	6.877	8.001	9.680	6.786
Size	7.465	7.099	6.320	5.488	4.413	
BM	3.220	4.268	3.518	2.695	2.479	
α^{FF3}	0.018	0.056	0.010	-0.200	-0.455	-0.474
	(0.093)	(0.393)	(0.087)	(-1.614)	(-2.314)	(-2.712)
α^{FF6}	-0.070	0.018	0.079	0.082	0.129	0.199
	(-0.343)	(0.135)	(0.665)	(0.579)	(0.523)	(0.965)

Panel B: Value-Weighted Portfolio Sorts by Idiosyncratic Volatility

	1	2	3	4	5	5 - 1
Total Return	0.804	0.838	0.905	0.648	0.126	-0.678
	(4.146)	(3.911)	(3.492)	(2.067)	(0.350)	(-2.425)
Excess Return	0.432	0.466	0.532	0.275	-0.246	
	(2.219)	(2.178)	(2.053)	(0.872)	(-0.672)	
SD	5.342	5.970	6.880	8.035	9.288	6.659
Size	7.465	7.099	6.320	5.488	4.413	
BM	3.220	4.268	3.518	2.695	2.479	
α^{FF3}	-0.024	-0.108	-0.160	-0.510	-1.145	-1.122
	(-0.192)	(-0.515)	(-1.052)	(-3.244)	(-6.143)	(-6.574)
α^{FF6}	-0.149	-0.165	-0.036	-0.182	-0.599	-0.450
	(-1.714)	(-0.412)	(-0.166)	(-0.807)	(-2.556)	(-2.924)

sorted each month into univariate quintile portfolios based on their past month IVOL calculated from daily data. The column named "5-1" reports the long-short portfolio results. It is of note that I report the long-short monthly average returns just in the row of total returns as the return difference between portfolio 5 and 1 is the same regardless of the return type used for its calculation. Panel A analyzes the EW and Panel B the VW portfolio returns respectively, where in both cases the 1/0/1 trading strategy is used for portfolio formation.

Panel A shows that the EW total (excess) portfolio returns increase monotonically from 0.868% (0.495%) per month for portfolio 1, which contains the lowest IVOL stocks, until portfolio 3 with a average total (excess) returns of 1.158% (0.758%) per month and declines afterwards again to 0.886% (0.514%) per month for portfolio 5 that contains the stocks with the highest IVOL. All total returns are statistically significant at the 1% level, except for the one of portfolio 5 which is only significant at the 5% level. On the other hand, the excess returns for portfolio 2 and 3 are statistically significant at the 1% level with t-statistics of 2.977 and 2.823 respectively, whereas for portfolio 1 and 4 they are only significant at the 5% level and for portfolio 5 they are not significant at any level. Comparing total and excess returns shows that significant results are more likely when analyzing the total returns which might be one reason why

researchers tend to report these in their studies respectively (see e.g. Ang et al., 2006; Bali & Cakici, 2008). Nevertheless, the long-short portfolio return shows that the return difference between portfolio 5 and 1 is only 0.018% per month and not significant at any reasonable level. In line with Bali and Cakici (2008), the return pattern from the IVOL puzzle cannot be detected in my equal-weighted setting regardless of analyzing total or excess returns.

Nonetheless, the previous conclusions change when investigating the alphas instead. The α^{FF3} decrease almost monotonically from portfolio 1 to 5 with values ranging from 0.018% to -0.455% per month respectively, but only the one of portfolio 5 is also statistically significant at the 5% level. Nonetheless, the alpha difference from the long-short portfolio amounts to -0.474% per month, which is not only economically relevant but also statistically significant at the 1% level with a t-statistic of -2.712. So in terms of the alphas relative to the FF3 model, I can verify the existence of an IVOL puzzle as found by Ang et al. (2006), especially when considering the long-short portfolio. Therefore, my results contradict those of Bali and Cakici (2008) who negate its existence in the equal-weighted portfolio context. Their conclusion is only supported by the analysis of the α^{FF6}. Here alphas increase monotonically from -0.070% for portfolio 1 to 0.129 for portfolio 5, albeit none of them are statistically significant at any level. Also the long-short portfolio alpha is no longer significant at any level and is even positive. In conclusion, the FF6 model seems to most appropriately capture the risk dynamics behind the IVOL puzzle as it produces insignificant alphas only. Hence, researchers that decide to compute the α^{FF6} instead of the α^{FF3} can modify their results in a way such that they find no evidence of the puzzle.

The portfolio excess return standard deviations show that portfolio return volatility is monotonically increasing with portfolio IVOL, such that the portfolio with highest IVOL stocks is also the most volatile one. Also the average size of the portfolios exhibits a distinct pattern. Like Ang et al. (2006), I find that stocks with higher IVOL are mostly smaller than low IVOL stocks as the average size decreases across portfolios from 7.465 in portfolio 1 to 4.413 for portfolio 5. In this case, the size effect predicts higher returns for small stocks, which would correspond to higher returns for portfolio 5 compared to portfolio 1 (see Fama & French, 1993). This is indeed observed when looking at the total or excess portfolio returns, but at the same time it contradicts the return pattern implied by the IVOL puzzle. Nevertheless, this size effect is not found in the α^{FF3} as they show lower average returns for portfolio 5 compared to portfolio 1. The average BM increases from 3.220 in portfolio 1 to 4.268 in portfolio 2 and declines afterwards monotonically to 2.479 contradicting the increasing pattern found by Ang et al. (2006). When comparing the BM of portfolio 5 and 1, the value effect predicts lower returns for portfolio 5 as it shows a lower BM relative to portfolio 1. While this effect does not match the EW total or excess portfolio return pattern, it indeed fits to the α^{FF3} pattern. Therefore, the IVOL puzzle found when considering the FF3 alphas might be related to a underlying value effect in the stocks.

Panel B delivers clearer evidence on the existence of the IVOL puzzle, as now all different alphas and the return on the long-short portfolio confirm the underperformance of the highest IVOL portfolio compared to the lowest IVOL one. The VW returns show a similar pattern as the EW ones from Panel A. Here the total (excess) returns increase from 0.804% (0.432%) per month for portfolio 1 to 0.905% (0.532%) for portfolio 3 and afterwards they decline again to 0.126% (-0.246%) per month for portfolio 5. However, the average total return on portfolio 5, as well as the excess returns on portfolios 4 and 5, are not statistically significant at any plausible level, whereas the remaining portfolios show significant returns at a level not less than 5%. The portfolio return difference between portfolio 5 and 1, captured by the long-short portfolio return, illustrates that stocks with high IVOL underperform those of low IVOL by about -0.678% per month which is significant at the 5% level with a Newey-West robust t-statistic of -2.425. In line with Ang et al. (2006), the VW total and excess returns computed here verify the existence of the IVOL puzzle.

Investigating the portfolio alphas delivers the same results. Moving from portfolio 1 to portfolio 5, alphas to all risk-correction models decrease monotonically indicating that the average risk-adjusted performance of the portfolios decreases in IVOL. The only exception is the α^{FF6} that increases from portfolio 2 to 3 but decreases again afterwards. Regardless of the risk-correction model used, all portfolios seem to underperform relative the corresponding risk-correction model, as all alphas in Panel B carry a negative sign. The α^{FF3} ranges from -0.024% for portfolio 1 to -1.145% per month for portfolio 5, whereas the α^{FF6} spans the interval from -0.036% per month for portfolio 3 to -0.599% for portfolio 5. However, only the α^{FF3} of portfolio 4 and 5 are statistically significant at the 1% level with t-statistics of -3.244 and -6.143. In contrast, only portfolio 1 and 5 carry a statistically significant α^{FF6} with robust t-statistics of -1.714 and -2.559. The long-short portfolio alpha is also statistically significant at a level of 1% now, regardless of the risk-correction model employed. It ranges from -0.450% for the α^{FF6} to -1.122% per month for the α^{FF3} with t-statistics of -2.924 and -6.574 respectively. Therefore, I can economically and statistically verify the existence of the IVOL puzzle from these risk-adjusted returns as well. Comparing the alphas of different risk-correction models also delivers evidence on their attractiveness for researchers to integrate them into their study. The fact that long-short portfolio alphas are largest in magnitude when correcting for the FF3 model, no matter whether EW or VW returns are used, might incentivise researchers to use this model for risk-correction in their studies if they seek to discover the IVOL puzzle. This tendency can indeed be observed throughout the literature (see e.g. Ang et al., 2006, 2009; Bali & Cakici, 2008; Boyer et al., 2010).

From the SD the same positive relationship between IVOL and portfolio excess return volatility as in Panel A appears. As portfolios are still sorted similarly to Panel A and only

the return-weighting has been modified, the average portfolio size and BM remain the same. However, as the high IVOL portfolio underperforms the low IVOL one when VW returns are analyzed, this pattern is in line with the predictions of the value effect because high IVOL portfolios have a lower BM on average and should hence have lower returns. As portfolio 5 contains the smallest stocks on average, the size effect would argue for higher returns in the high IVOL portfolio relative to the low IVOL one instead. Nevertheless, this effect seems to be dominated by the value effect, as the size effects might be accounted for by the value-weighting.

In summary, the IVOL puzzle is more apparent in Panel B as here all long-short portfolio performance measures approve its existence. For the EW portfolios of Panel A only the α^{FF3} gives evidence on the existence of the IVOL puzzle. However, the EW long-short portfolio return and its α^{FF6}, albeit not statistically significant, predict higher returns for portfolios with higher IVOL respectively. These results point out that, in contrast to the previous regression-based results, the IVOL puzzle finding in the portfolio-based analysis is sensitive to the choice of the weighting scheme.

6. General Adjustments

After having empirically analyzed the IVOL puzzle with the reference research concepts, I start with adjusting them in a way that is independent of the concrete approach implemented as explained in section 3.4.1. In section 6.1 I modify the way the IVOL is estimated. Section 6.2 then presents all changes related to the sample used for the analysis respectively. All adjustments implemented here are explained in detail in section 3.4.1. Due to space constraints I present only a subset of the regression- and portfolio-based results depicted in tables 2 and 3. However, this selective reporting behavior preserves all information that is necessary for investigation of the IVOL puzzle finding and how it evolves across adjustments.

6.1. Idiosyncratic Volatility-Related

I start with the modification of the way IVOL is estimated by changing the time-window and the data frequency that is used for its computation as well as the underlying risk-correction model towards which it is measured. These adjusted IVOL estimates are then analyzed using the regression- and the portfolio-based approach respectively. For the regression-based method I report the complete model results of column (7) from table 2 that includes all control variables simultaneously.

Table 4 presents the regression-based results where each column corresponds to the findings from a specific IVOL estimate used in the analysis accordingly. The columns labeled $IVOL_{12m}$, $IVOL_{1y}$, $IVOL_{5y}$ depict the results for the IVOLs that are computed relative to the FF3 model but using rolling windows over the past 12 month of daily data as well as the past 1 and 5 years of monthly data respectively. In the remaining columns, $IVOL^{FF6}$, $IVOL^{SY4}$ and $IVOL^{HOU5}$ show the results when IVOL is computed from daily data over the current month but using the FF6, SY4 or the HOU5 model for risk-correction. Panel A displays the EW and Panel B the VW regression results. Panel A shows negative average IVOL risk premia that are statistically significant at the 1% level regardless of the IVOL estimate considered. Since I use the complete regression model for this analysis, these results are robust when controlling for the effects from the control variables introduced previously. The IVOL risk premia estimates fluctuate around -0.238 when using $IVOL_{1y}$ in the regression and -1.189 for $IVOL_{12m}$ with corresponding robust t-statistics of -3.529 and -3.119. Considering the economic significance of the risk premia estimates, a one standard deviation increase in the respective IVOL measure, holding all other variables constant, translates into a reduction of expected returns between 0.281% per month for $IVOL_{1y}$ to 0.391% per month for $IVOL_{12m}$. So no matter which measure is used, the IVOL risk premium remains economically and statistically relevant in the EW Fama-MacBeth regressions and hence verifies the existence of the IVOL puzzle. The standard deviations used for the previous calculation and further summary statistics of the adjusted IVOL estimates can be found in table A.1 of Appendix A (vgl. Anhang, Tabelle A.1). However, it is of note that while the risk premia estimates for all risk-correction models are largely similar and vary from -1.019 for $IVOL^{HOU5}$ to -1.038 for $IVOL^{SY4}$, this is not true for the IVOLs that use different estimation windows and data frequencies. Comparing the IVOL estimates over the same estimation window of one year but with different data frequencies, I find that the risk premium from $IVOL_{12m}$, which uses daily data, is -1.189, whereas the one of $IVOL_{1y}$ that uses monthly data is only -0.238. Nonetheless, when considering the economic significance of these estimates by a one standard deviation increase in the respective IVOL measure, both predict a decrease in expected returns that is close together with values of 0.391% and 0.281% per month. This is most likely explained by the fact that $IVOL_{1y}$ is more volatile than $IVOL_{12m}$ with values of 118.230% and 32.883% as shown in table A.1 of Appendix A (vgl.Anhang, Tabelle A.1). When keeping the data frequency fixed and adjusting the estimation window only, the risk premia estimates hardly change as can be seen by the results from $IVOL_{1y}$ and $IVOL_{5y}$. Hence, in contrast to Bali and Cakici (2008) as well as Spiegel and Wang (2005), the IVOL estimate from monthly data still possess a negative risk premium here. So in line with the findings of Ang et al. (2009), the IVOL puzzle exists in this EW regression setting regardless of the IVOL estimate analyzed. The patterns, as well as the magnitudes, of all other risk premia are almost similar to the reference results of table 2, no matter which IVOL estimate is considered.

Panel B shows that all estimated IVOL risk premia remain negative in sign, but the t-statistics decrease notably when value-weights are used. Only the risk premia from modified risk-correction model IVOL estimates remain significant but only at a level of 5% and with t-statistics around -2.176 for $IVOL^{FF6}$ and -2.570 for $IVOL^{SY4}$. In absolute terms, they are slightly weaker then in the EW context and range from -0.780

Table 4: Regression-Based Results - Idiosyncratic Volatility-Related Adjustments

This table presents the average coefficients from monthly Fama and MacBeth (1973) cross-sectional regressions for individual stocks. Using equation 5, each month I regress the next-month excess firm returns (in percentage points) on a constant, an annualized idiosyncratic volatility measure (in decimals), the stock-specific risk factor loadings relative to the FF3 model over the next month, the end of month firm size defined as the natural logarithm of the market capitalization, the end of month firm book-to-market ratio of equity, the monthly return (in percentage points) computed from the end of the previous month to the current month as well as additional control variables related to further firm characteristics. These firm characteristic variables are the following: $Skew_{1M}$ which is the monthly return skewness measured over daily return data from the current month, $EIdioSkew$ that is the expected idiosyncratic skewness measure from Boyer et al. (2010), $CoSkew$ which is the coskewness measure from Harvey and Siddique (2000) and $Illiq_{1M}$ being the Amihud (2002) illiquidity measure computed over the current month. Instead of the usual IVOL measure, the regression models in each column integrate an idiosyncratic volatility estimate that was computed either over a different time window, with a different data frequency or against an other risk-correction model. More specifically $IVOL_{12m}$, $IVOL_{1y}$, $IVOL_{5y}$ are the idiosyncratic volatilities computed relative to the FF3 model but using a window of the past 12 month of daily data as well as of the past 1 and 5 years of monthly data respectively. $IVOL^{FF6}$, $IVOL^{SY4}$ and $IVOL^{HOU5}$ refer to the IVOL computed based on daily data over the current month but using the FF6, SY4 and the HOU5 model for risk-correction instead. In Panel A I use equal-weighted stock excess returns to run the OLS regression of equation 5. On the other hand, in Panel B I estimate the Fama-MacBeth regression with a weighted-least squares approach, where all individual monthly stock returns are weighted by their current month market capitalization. I report the t-statistics testing the null hypothesis that the average coefficient is equal to zero in parenthesis below each coefficient, where corresponding standard errors are corrected for autocorrelation and heteroskedasticity by implementation of the Newey and West (1987) method using 7 lags. *, ** , and *** indicate significance at the 10%, 5%, and 1% level, respectively. The row "Adjusted R^2" reports the time-series average of the cross-sectional adjusted R^2's. The last row reports the average number of stocks used for the monthly cross-sectional regressions. The sample covers the period from July 1963 until December 2020 and includes all stocks traded on the NYSE/AMEX/NASDAQ. It is shorter when the SY4 model is used in the analysis, which is due a lack of data availability wherefore it then covers the horizon until November 2016 only.

Panel A: Equal-Weighted Fama-MacBeth Regressions

	$IVOL_{12m}$	$IVOL_{1y}$	$IVOL_{5y}$	$IVOL^{FF6}$	$IVOL^{SY4}$	$IVOL^{HOU5}$
$Constant$	1.971***	1.786***	1.884***	1.816***	2.091***	1.835***
	(5.186)	(4.642)	(5.235)	(4.534)	(5.546)	(4.553)
$IVOL$	-1.189***	-0.238***	-0.254***	-1.022***	-1.038***	-1.019***
	(-3.119)	(-3.529)	(-3.188)	(-4.121)	(-4.039)	(-4.169)
$\beta(MKT)$	0.212*	0.198*	0.209*	0.189	0.318***	0.191
	(1.847)	(1.725)	(1.826)	(1.619)	(3.352)	(1.634)
$\beta(SMB)$	0.010	0.010	0.014	0.005	0.021	0.005
	(0.214)	(0.223)	(0.301)	(0.112)	(0.435)	(0.113)
$\beta(HML)$	-0.073	-0.066	-0.071	-0.064	-0.040	-0.065
	(-1.292)	(-1.158)	(-1.237)	(-1.090)	(-0.853)	(-1.102)
$Size$	-0.166***	-0.152***	-0.160***	-0.156***	-0.178***	-0.158***
	(-4.770)	(-4.215)	(-4.503)	(-4.305)	(-4.652)	(-4.321)
BM	0.143**	0.140**	0.112*	0.141**	0.156**	0.140**
	(2.278)	(2.179)	(1.868)	(2.238)	(2.251)	(2.214)
R_t	-0.045***	-0.044***	-0.045***	-0.043***	-0.044***	-0.043***
	(-11.675)	(-11.208)	(-11.691)	(-10.705)	(-11.017)	(-10.644)
$Skew_{1M}$	0.095***	0.092***	0.097***	0.098***	0.103***	0.099***
	(3.563)	(3.182)	(3.436)	(3.917)	(3.790)	(3.950)
$EIdioSkew$	-0.281***	-0.277**	-0.216**	-0.301**	-0.329**	-0.301**
	(-2.631)	(-2.378)	(-2.018)	(-2.544)	(-0.329)	(-2.543)
$CoSkew$	-0.013***	-0.012**	-0.013***	-0.011**	-0.010*	-0.011**
	(-2.683)	(-2.385)	(-2.670)	(-2.303)	(-1.820)	(-2.296)
$Illiq_{1M}$	0.036***	0.029***	0.029***	0.035***	0.033***	0.035***
	(3.539)	(2.947)	(2.953)	(3.487)	(3.194)	(3.492)
$Adjusted\ R^2$	0.069	0.067	0.067	0.067	0.065	0.067
n	3056	3056	3056	3056	3022	3056

Panel B: Value-Weighted Fama-MacBeth Regressions

	$IVOL_{12m}$	$IVOL_{1y}$	$IVOL_{5y}$	$IVOL^{FF6}$	$IVOL^{SY4}$	$IVOL^{HOU5}$
$Constant$	1.414***	1.161***	1.049**	1.508***	1.843***	1.543***
	(3.066)	(2.963)	(2.293)	(4.056)	(5.066)	(4.121)
$IVOL$	-0.689	-0.025	-0.004	-0.780**	-0.959**	-0.831**
	(-1.124)	(-0.262)	(-0.033)	(-2.176)	(-2.570)	(-2.342)
$\beta(MKT)$	0.109	0.070	0.083	0.088	0.160	0.092
	(0.702)	(0.438)	(0.531)	(0.544)	(0.942)	(0.569)
$\beta(SMB)$	-0.085*	-0.084	-0.080	-0.084	-0.076	-0.087*
	(-1.678)	(-1.637)	(-1.603)	(-1.662)	(-1.446)	(-1.710)
$\beta(HML)$	-0.039	-0.017	-0.012	-0.024	0.051	-0.025
	(-0.427)	(-0.183)	(-0.135)	(-0.250)	(0.641)	(-0.256)
$Size$	-0.077**	-0.054	-0.049	-0.079**	-0.097***	-0.081**
	(-2.119)	(-1.636)	(-1.367)	(-2.407)	(-2.782)	(-2.474)
BM	-0.013	-0.009	-0.005	-0.026	-0.026	-0.028
	(-0.151)	(-0.100)	(-0.063)	(-0.285)	(-0.257)	(-0.310)
R_t	-0.027***	-0.027***	-0.027***	-0.026***	-0.026***	-0.025***
	(-5.842)	(-5.811)	(-5.972)	(-5.431)	(-5.205)	(-5.336)
$Skew_{1M}$	0.100***	0.095***	0.099***	0.101***	0.117***	0.104***
	(3.108)	(3.044)	(3.221)	(3.169)	(3.515)	(3.277)
$EIdioSkew$	-0.203	-0.225	-0.207	-0.238	-0.222	-0.239
	(-1.368)	(-1.526)	(-1.442)	(-1.571)	(-1.345)	(-1.585)
$CoSkew$	-0.021**	-0.019**	-0.022**	-0.016*	-0.013	-0.017**
	(-2.407)	(-2.138)	(-2.511)	(-1.924)	(-1.514)	(-1.960)
$Illiq_{1M}$	-0.001	-0.001	0.002	0.000	0.004	0.001
	(-0.096)	(-0.051)	(0.108)	(0.025)	(0.267)	(0.071)
$Adjusted\ R^2$	0.157	0.153	0.156	0.152	0.151	0.152
n	3056	3056	3056	3056	3022	3056

to -0.959 respectively. Matching the results from Bali and Cakici (2008) as well as Rachwalski and Wen (2016), the risk premia estimates decrease remarkably and are no longer statistically significant when IVOL is estimated with a modified data frequency or estimation window. Comparing $IVOL_{12m}$, $IVOL_{1y}$ and $IVOL_{5y}$, the IVOL risk premium increases monotonically from -0.689 for $IVOL_{12m}$ to -0.004 for $IVOL_{5y}$. Using monthly data for IVOL estimation also seems to lead to a risk premium that is not only statistically insignificant but is also very close to zero in absolute terms. This pattern is likewise found when analyzing the economic significance of the IVOL risk premium estimates by a one standard deviation increase in the corresponding measure, holding all other variables constant. The average decrease in expected returns by such an effect amounts to 0.005%, 0.030% and 0.223% per month for $IVOL_{5y}$, $IVOL_{5y}$ and $IVOL_{1y}$ correspondingly up to 0.263%, 0.296% and 0.341% per month for the $IVOL^{FF6}$, $IVOL^{HOU5}$ and $IVOL^{SY4}$. Thus only the IVOL from daily data and these of the different risk-correction models have economically significant risk premia estimates in the VW regression context.

The risk premia pattern of the remaining variables has changed slightly compared to the results from table 2. The $\beta(SMB)$ risk premia remain all negative and similar in magnitude but are now only statistically significant for $IVOL_{12m}$ and $IVOL^{HOU5}$. $\beta(HML)$ carries a positive risk premium when using $IVOL^{SY4}$ that is still not statistically significant. For size, the risk premium remains negative in all cases and similar in magnitude to the reference setting but is no longer significant for $IVOL_{1y}$ and $IVOL_{5y}$ and in contrast significant the 1% level for $IVOL^{SY4}$. While the negative coskewness risk premium increased in significance to a level of 5% for $IVOL_{12m}$, $IVOL_{1y}$, $IVOL_{5y}$ and $IVOL^{HOU5}$, it is no longer significant when using $IVOL^{SY4}$. Also interesting is that when using the $IVOL_{12m}$ and $IVOL_{1y}$, the illiquidity premium is now negative albeit not significant.

The above findings show that, using the EW Fama-MacBeth regressions, a researcher is more likely to detect an IVOL puzzle, no matter which data frequency, estimation window or risk-correction model is used and therefore makes its implementation more attractive to researchers that seek to find the IVOL puzzle in their study. Conversely, when using the adjusted IVOL estimates in the VW Fama-MacBeth regressions, the results less frequently deliver evidence on the existence of the IVOL puzzle, which is why this method is more attractive to researchers that try to negate the presence of the IVOL puzzle. Ang et al. (2009) use also different estimation windows to estimate the IVOL measure and conclude that the existence of the IVOL puzzle is robust to all these estimates. However, it is of note that they implemented an EW Fama-MacBeth regression approach which is, as revealed earlier, more likely to approve this finding and leave the VW results unreported. This again highlights that the IVOL puzzle finding can be influenced by the choice in research design.

Table 5 presents the average total portfolio returns as well as the time-series alphas relative to the FF3 and the FF6 model from the portfolio-based approach for portfolios 1, 5 and the long-short portfolio respectively. Each column corresponds to the portfolio-based results when using one of the modified IVOL measures introduced above for the construction of the corresponding monthly quintile portfolios. Panel A depicts the results of the EW and Panel B those of the VW portfolio returns respectively. In both cases the $1/0/1$ trading strategy is used for portfolio formation.

Panel A shows almost identical results to those from the reference analysis in table 3 which are independent of the IVOL measure analyzed. The long-short portfolio shows positive, albeit not statistically significant, total returns around 0.049% to 0.296% per month for all IVOL measures except for $IVOL^{SY4}$ and $IVOL^{HOU5}$ that display negative total returns instead. But these negative returns are economically small with values of -0.001% and -0.011% per month and are not statistically significant at any reasonable level. The portfolios 1 and 5 total returns are both positive and statistically significant at a level not less than 5%, no matter which IVOL measure is employed. Consequently, from the EW total portfolio returns only, a researcher finds no IVOL puzzle irrespective of the data frequency, estimation window or risk-correction model he uses for IVOL estimation. The α^{FF3} results, on the other hand, do mostly not match those from table 3 showing that the choice of the IVOL estimate influences the alpha pattern of the EW portfolio-based analysis and hence potentially also the conclusions drawn from it. Even though the α^{FF3} of the long-short portfolios are all negative, they are not always statically significant at the 1% level. Only the long-short portfolio α^{FF3} of the $IVOL^{SY4}$ sort that has a return of -0.465% per month with a robust t-statistic of -2.612. Conversely, the $IVOL_{1y}$, $IVOL^{FF6}$ and $IVOL^{HOU5}$ sorts show a negative average long-short α^{FF3} around -0.362% to -0.465% that are significant at the 5% level, whereas those from the $IVOL_{12m}$ and $IVOL_{5y}$ sorts are not significant at any reasonable level. In contrast to the α^{FF3} from the reference setting of section 5.3, here it are mostly the comparatively high and statistically significant α^{FF3} from portfolio 1 that drive the long-short portfolio α^{FF3} and not the statistically significant low alphas from portfolio 5. The positive α^{FF6} from the long-short portfolio in table 3 is robust to all the modified IVOL estimates, as this alpha remains positive, no matter which IVOL estimate is considered. Nevertheless, it is not statistically significant, except for the portfolios that are constructed from the $IVOL_{5y}$ which possesses a long-short α^{FF6} of 0.414% per month with a t-statistic of 2.047 that implies statistical significance at the 5% level. Here portfolio 1 and portfolio 5 seem to drive the long-short α^{FF6} as both are positive and statically significant at a level not less than 5% for all sorts, except for those on $IVOL^{HOU5}$ and $IVOL^{FF6}$ that show a significant alpha for portfolio 5 only. Hence, the positive, albeit statically insignificant, α^{FF6} of the long-short portfolios imply an overperformance of the high IVOL portfolio relative to the low IVOL one independent of the IVOL measured considered which is directly opposite to

Table 5: Portfolio-Based Results - Idiosyncratic Volatility-Related Adjustments

This table presents the average results from the univariate portfolio sorting procedure as explained in section 3.3 covering all stocks traded on the NYSE/AMEX/NASDAQ. Every month stocks are sorted into quintile portfolios based on the idiosyncratic volatility measure depicted in the respective column, where $IVOL_{12m}$, $IVOL_{1y}$, $IVOL_{5y}$ are the idiosyncratic volatilities that are computed relative to the FF3 model but using a window of the past 12 month of daily data as well as of the past 1 and 5 years of monthly data respectively. $IVOL^{FF6}$, $IVOL^{SY4}$ and $IVOL^{HOU5}$, on the other hand, refer to the IVOL computed based on daily data over the current month but using the FF6, SY4 and the HOU5 model for risk-correction instead. These portfolios are rebalanced monthly according to the 1/0/1 trading strategy. In the first row I report the average monthly total returns for the portfolios 1, 5 as well as the long-short portfolio called "5-1" (in percentage points) and in parenthesis below each of these returns I depict the corresponding Newey and West (1987) adjusted t-statistics where I use 7 lags for the adjustment. Portfolio 1 (5) contains all stocks with lowest (highest) IVOL, whereas the long-short portfolio is the one that goes long the portfolio of stocks with the highest IVOL and short the respective portfolio of stocks with the lowest IVOL. This long-short portfolio is also rebalanced monthly according to the 1/0/1 trading strategy. The α^{FF3} and the α^{FF6} in the second and third row denote the monthly average of Jensen's time-series alphas relative to the FF3 and the FF6 model that were calculated by equation 1 using the monthly portfolio excess returns over the complete sample horizon and are again reported for portfolio 1, 5 and the long-short portfolio respectively. Below each of the alphas I provide the corresponding t-statistics that were again corrected by the procedure of Newey and West (1987) using 7 lags. In Panel A I use the equal-weighted portfolio returns for the analysis, where all firms are getting the same weight. On the other hand, Panel B uses value-weighted portfolio returns that were calculated by weighting the stock returns within the portfolio by their market capitalization which is observable at beginning of the month in order to give higher weights to bigger stocks respectively and therefore diminish the effects that might be explicitly related to small stocks. The complete sample period covers July 1963 to December 2020. It is shorter when the SY4 model is used in the analysis, which is due a lack of data availability wherefore it then covers the horizon until November 2016 only.

Panel A: Equal-Weighted Portfolio Sorts

		$IVOL_{12m}$	$IVOL_{1y}$	$IVOL_{5y}$	$IVOL^{FF6}$	$IVOL^{SY4}$	$IVOL^{HOU5}$
Total Return	1	0.871	0.939	0.876	0.863	1.033	0.862
		(4.309)	(4.454)	(4.374)	(3.874)	(5.977)	(3.650)
	5	1.167	1.039	1.212	0.912	1.023	0.861
		(2.718)	(2.586)	(2.975)	(2.284)	(2.530)	(2.078)
	5-1	0.296	0.101	0.337	0.049	-0.011	-0.001
		(0.870)	(0.337)	(1.058)	(0.170)	(-0.035)	(-0.003)
α^{FF3}	1	0.416	0.449	0.410	0.384	0.532	0.396
		(2.462)	(2.413)	(2.481)	(1.663)	(7.094)	(1.639)
	5	0.173	0.086	0.270	-0.051	0.067	-0.062
		(0.796)	(0.436)	(1.294)	(-0.263)	(0.440)	(-0.297)
	5-1	-1.187	-0.362	-0.140	-0.435	-0.465	-0.458
		(0.438)	(-2.038)	(-0.749)	(-2.510)	(-2.612)	(-2.518)
α^{FF6}	1	0.339	0.394	0.340	0.296	0.455	0.303
		(2.054)	(1.963)	(2.008)	(1.097)	(6.434)	(1.113)
	5	0.777	0.580	0.755	0.523	0.715	0.538
		(2.905)	(2.480)	(3.056)	(2.116)	(3.383)	(2.050)
	5-1	0.438	0.186	0.414	0.227	0.260	0.236
		(1.839)	(1.015)	(2.047)	(1.098)	(1.190)	(1.099)

Panel B: Value-Weighted Portfolio Sorts

		$IVOL_{12m}$	$IVOL_{1y}$	$IVOL_{5y}$	$IVOL^{FF6}$	$IVOL^{SY4}$	$IVOL^{HOU5}$
Total Return	1	0.787	0.810	0.756	0.794	0.915	0.826
		(4.167)	(4.292)	(3.975)	(4.124)	(5.852)	(4.107)
	5	0.210	0.639	0.514	0.176	0.214	0.087
		(0.509)	(1.803)	(1.396)	(0.486)	(0.592)	(0.087)
	5-1	-0.577	-0.171	-0.242	-0.618	-0.700	-0.739
		(-1.706)	(-0.616)	(-0.838)	(-2.213)	(-2.374)	(-2.582)
α^{FF3}	1	0.324	0.350	0.285	0.337	0.473	0.375
		(2.571)	(2.720)	(2.301)	(3.089)	(10.401)	(2.648)
	5	-0.788	-0.206	-0.370	-0.724	-0.628	-0.814
		(-3.764)	(-1.146)	(-1.927)	(-4.225)	(-4.398)	(-4.215)
	5-1	-1.112	-0.557	-0.654	-1.061	-1.101	-1.189
		(-5.596)	(-3.304)	(-3.941)	(-6.647)	(-6.447)	(-6.919)
α^{FF6}	1	0.187	0.232	0.140	0.216	0.361	0.241
		(2.098)	(2.467)	(1.415)	(2.721)	(6.931)	(2.431)
	5	-0.120	0.106	-0.040	-0.170	0.014	-0.217
		(-0.579)	(0.644)	(-0.211)	(-0.714)	(0.096)	(-0.832)
	5-1	-0.307	-0.126	-0.180	-0.386	-0.348	-0.458
		(-1.778)	(-0.899)	(-1.344)	(-2.373)	(-2.042)	(-2.868)

the pattern implied by the IVOL puzzle.

The VW portfolio sorts from Panel B show a total return pattern that is almost identical to the one from section 5.3, irrespective of the IVOL measure used. Here the long-short portfolio returns are all negative but only statistically significant for the risk-correction model adjusted IVOL estimates at a level not less than 5% with corresponding returns around -0.618% and -0.739% per month. Among the estimation windows and data frequencies adjusted IVOLs, only the $IVOL_{12m}$ long-short portfolio return is statistically significant at the 10% level and amounts to -0.577% per month. So when using monthly data for the IVOL estimation, the VW total return differences, measured by the long-short portfolio, are not statistically significant at any level. In contrast to table 3, for all IVOL estimates the underperformance of portfolio 5 relative to portfolio 1 appears to be driven by the positive total returns from the low IVOL portfolio that are statistically significant at the 1% level, as those from portfolio 5 are barely significant at any level. When investigating the α^{FF3}, all modified IVOL based sorts confirm the existence of the IVOL puzzle similarly to the results from Panel B of table 3. The long-short portfolios show negative α^{FF3} that are statistically significant at the 1% level for all IVOL measures and fluctuate between -0.557% and -1.189% per month respectively. Also the IVOL puzzle implied negative relationship between IVOL and expected returns becomes apparent from the alpha pattern across portfolios throughout all IVOL sorts, as on average portfolio 5 has negative and portfolio 1 positive α^{FF3} that are statistically significant at a level not less than 5% when excluding the sorts on $IVOL_{1y}$ and $IVOL_{5y}$. Similar to section 5.3, the VW long-short α^{FF6} verify the existence of the IVOL puzzle for all IVOL estimates, except for $IVOL_{1y}$ and $IVOL_{5y}$, as these long-short α^{FF6} are not statistically significant at any plausible level. All other long-short α^{FF6} range from -0.307% for the $IVOL_{12m}$ sort to -0.458% per month for the sort on $IVOL^{HOU5}$ and are always statistically significant at the 5% level, except for the $IVOL_{12m}$ sort where it is only significant at the 10% level. Also, when comparing the α^{FF6} from portfolio 5 and 1, the negative relationship between risk-adjusted expected returns and IVOL appears regardless of the IVOL estimate analyzed. In contrast to the reference results of section 5.3, this pattern is now driven by the positive alphas from portfolio 1 that are statistically significant at a level not less than 5% for all IVOL estimates except for $IVOL_{5y}$.

In summary, the EW reference results from Panel A of table 3 are mostly robust to adjustments in the way IVOL is estimated. As in the reference setting and consistent with the study of Bali and Cakici (2008), the EW total returns and the α^{FF6} show no evidence of an IVOL puzzle, regardless of the IVOL adjustment implemented. Solely the α^{FF3} still confirm the existence of the puzzle which holds for all IVOL sorts, except for those on $IVOL_{12m}$ and $IVOL_{1y}$. The VW portfolio sorts show that the IVOL puzzle finding is sensitive to the choice of the data frequency used for IVOL estimation as all sorts based on an IVOL measure from monthly data are not able to detect the puzzling return pattern in the total returns and the α^{FF6}. This result is in line with Bali and Cakici (2008) as well as Rachwalski and Wen (2016) who also have difficulties detecting an IVOL puzzle for a monthly data based IVOL measure. Only for the VW α^{FF3} the pattern of section 5.3 remains unchanged, showing negative long-short alphas that are statistically significant at the 1% level regardless of the IVOL measured analyzed. When comparing the results from the IVOL estimation adjustments among the portfolio- and the regression-based method conclusions differ notably. In the EW setting, the IVOL puzzle finding is more robust to the usage of different IVOL estimates in the regression-based method as then a statistically significant IVOL risk premium is found regardless of the chosen IVOL estimate. With the portfolio-based analysis it is only found when analyzing the α^{FF3} and excluding the sorts on $IVOL_{12m}$ and $IVOL_{5y}$. However, both approaches deliver similar conclusions in the VW context, as here the IVOL puzzle can only be confirmed statistically for all risk-correction model adjusted IVOLs but not for the IVOLs from different data frequencies or estimation windows.

6.2. Sample-Related

In this section I recompute the regression- and portfolio-based reference analyses from chapter 5, only that I now focus on the corresponding results from specific subsamples. These subsamples are constructed by subdividing the complete data according to the stock price or its size as explained in section 3.4.1.

Table 6 presents the average regression-based results for the different subsamples. In the first column labeled "Price" the complete sample data is subdivided into low, medium high price stocks with cutoffs of 5$ and 10$ and analyzed respectively. The column named "Size" refers to the subsamples that split the data each month into micro, small and big stocks based on the 20th and the 50th percentile of the cross-sectional distribution of market capitalization from all stocks traded on the NYSE. Within each of the subsamples I re-estimate the risk premia using the monthly cross-sectional regressions of equation 5. Then, these risk premia estimates are averaged over time again. The corresponding EW Fama-MacBeth regression results are located in Panel A, whereas the VW ones can be found in Panel B. Panel A shows that the IVOL puzzle finding is robust to all size and price subsamples in the EW regression approach as the estimated IVOL risk premium remains negative and statistically significant at the 1% level in all cases, except for the big stocks where it is only significant at the 10% level. The average IVOL risk premium ranges from -0.796 for low price to -2.402 for medium price stocks in the price subsamples and from -0.657 for big stocks to -1.852 for all small stocks in the size subsamples. As I use the complete regression model for the analysis, none of the corresponding control variables can explain these premia. The economic magnitude of the IVOL risk premia estimates for the price subsamples, measured by the return difference from a one standard deviation change in IVOL while holding all other variables constant, ranges from -0.157% for high

Table 6: Regression-Based Results - Sample-Related Adjustments

This table presents the average coefficients from monthly Fama and MacBeth (1973) cross-sectional regressions for individual stocks. Using equation 5, each month I regress the next-month excess firm returns (in percentage points) on a constant, the annualized idiosyncratic volatility measure of equation 4 calculated from daily data over the current month (in decimals), the stock-specific risk factor loadings relative to the FF3 model over the next month, the end of month firm size defined as the natural logarithm of the market capitalization, the end of month firm book-to-market ratio of equity, the monthly return (in percentage points) computed from the end of the previous month to the current month as well as additional control variables related to further firm characteristics. These firm characteristic variables are the following: $Skew_{1M}$ which is the monthly return skewness measured over daily return data from the current month, $EIdioSkew$ that is the expected idiosyncratic skewness measure from Boyer et al. (2010), $CoSkew$ which is the coskewness measure from Harvey and Siddique (2000) and $Illiq_{1M}$ being the Amihud (2002) illiquidity measure computed over the current month. Instead of the complete sample, the cross-sectional regression of equation 5 are now estimated over subsamples of the data set that matches certain criteria. These criteria can be found in each column accordingly. In the first column named "Price" the complete sample data are subdivided into low, medium high price stocks with cutoffs of 5$ and 10$. The subsample criteria called "Size" refers to the subdivision of the data according to the stock-specific market capitalization. Here the complete data are split each month into micro, small and big stocks based on the 20th and the 50th percentile of the cross-sectional distribution of market capitalization when using all stocks traded on the NYSE. Panel A uses equal-weighted stock excess returns to run the OLS regression of equation 5. On the other hand, in Panel B I estimate the Fama-MacBeth regression with a weighted-least squares approach, where all individual monthly stock returns are weighted by their current month market capitalization. I report the t-statistics testing the null hypothesis that the average coefficient is equal to zero in parenthesis below each coefficient, where corresponding standard errors are corrected for autocorrelation and heteroskedasticity by implementation of the Newey and West (1987) method using 7 lags. *, ** , and *** indicate significance at the 10%, 5%, and 1% level, respectively. The row "Adjusted R^2" reports the time-series average of the cross-sectional adjusted R^2's. n reports the average number of stocks used for the monthly cross-sectional regressions. The last row shows the sample specific time-series mean of the monthly cross-sectional standard deviation of the IVOL estimate (in percentage terms) that was calculated relative to the FF3 model based on daily data over the current month. The sample covers the period from July 1963 until December 2020 and includes all stocks traded on the NYSE/AMEX/NASDAQ.

Panel A: Equal-Weighted Fama-MacBeth Regressions

	Price			Size		
	Low	Medium	High	Microcap	Small	Big
$Constant$	4.857***	2.826***	1.658***	4.165***	2.099***	1.595***
	(7.419)	(5.554)	(4.556)	(7.815)	(3.047)	(3.753)
$IVOL$	-0.796***	-2.402***	-0.834***	-1.172***	-1.852***	-0.657*
	(-3.389)	(-8.743)	(-3.278)	(-5.008)	(-5.797)	(-1.913)
$\beta(MKT)$	0.354***	0.327**	0.094	0.353***	0.160	0.017
	(3.145)	(2.516)	(0.681)	(3.055)	(1.217)	(0.106)
$\beta(SMB)$	0.082	0.041	-0.100**	0.045	-0.023	-0.112**
	(1.327)	(1.016)	(-2.502)	(0.793)	(-0.498)	(-2.456)
$\beta(HML)$	-0.098*	-0.111	-0.047	-0.083	-0.120	0.013
	(-1.726)	(-1.379)	(-0.614)	(-1.593)	(-0.120)	(0.147)
$Size$	-0.939***	-0.297***	-0.111***	-0.696***	-0.148	-0.089**
	(-10.572)	(-4.687)	(-3.544)	(-7.661)	(-1.539)	(-2.512)
BM	0.033	0.174**	0.054	0.114	0.097	0.000
	(0.475)	(2.186)	(0.668)	(1.840)	(0.975)	(-0.003)
R_t	-0.065***	-0.025***	-0.025***	-0.049***	-0.021***	-0.025***
	(-9.049)	(-5.884)	(-6.737)	(-10.398)	(-4.830)	(-5.186)
$Skew_{1M}$	0.076	0.101***	0.089***	0.113***	0.077**	0.085***
	(1.251)	(2.606)	(4.327)	(3.208)	(2.300)	(3.596)
$EIdioSkew$	-0.913***	-0.543***	-0.163	-0.665***	-0.285**	-0.141
	(-4.295)	(-3.915)	(-1.543)	(-4.637)	(-2.005)	(-0.914)
$CoSkew$	-0.016**	0.003	-0.010*	-0.012**	-0.007	-0.017**
	(-2.287)	(0.485)	(-1.787)	(-2.429)	(-1.236)	(-2.139)
$Illiq_{1M}$	0.026**	-0.017	-0.108**	0.030***	0.014	5.394
	(2.515)	(-0.748)	(-2.117)	(2.919)	(0.021)	(0.660)
$Adjusted\ R^2$	0.073	0.069	0.078	0.065	0.076	0.114
n	730	512	1814	1344	758	953
$\sigma(IVOL)$	52.374	28.775	18.822	45.431	21.246	14.135

Panel B: Value-Weighted Fama-MacBeth Regressions

	Price			Size		
	Low	Medium	High	Microcap	Small	Big
$Constant$	3.821***	3.301***	1.537***	3.387***	2.433***	1.509***
	(4.651)	(4.711)	(4.016)	(6.671)	(3.498)	(3.713)
$IVOL$	-1.836***	-2.722***	-0.583	-2.005***	-1.711***	-0.563
	(-6.764)	(-7.627)	(-1.588)	(-7.793)	(-5.334)	(-1.367)
$\beta(MKT)$	0.405***	0.164	0.092	0.371***	0.115	0.102
	(3.413)	(1.134)	(0.554)	(3.250)	(0.854)	(0.597)
$\beta(SMB)$	0.085	0.088	-0.107**	0.042	-0.039	-0.101*
	(1.451)	(1.551)	(-2.149)	(0.892)	(-0.843)	(-1.839)
$\beta(HML)$	-0.088	-0.100	-0.013	-0.087	-0.116	-0.010
	(-1.270)	(-1.016)	(-0.130)	(-1.577)	(-1.423)	(-0.100)
$Size$	-0.518***	-0.341***	-0.082**	-0.432***	-0.199**	-0.079**
	(-5.329)	(-3.943)	(-2.388)	(-5.503)	(-2.072)	(-2.180)
BM	0.115	0.063	-0.042	0.134*	0.068	-0.064
	(1.044)	(0.522)	(-0.415)	(1.939)	(0.656)	(-0.638)
R_t	-0.048***	-0.019***	-0.026***	-0.029***	-0.022***	-0.026***
	(-5.894)	(-3.459)	(-5.002)	(-6.781)	(-4.998)	(-4.783)
$Skew_{1M}$	0.060	0.057	0.110***	0.109***	0.065*	0.110***
	(0.709)	(0.875)	(3.513)	(3.635)	(1.866)	(3.251)
$EIdioSkew$	-0.786***	-0.448**	-0.275*	-0.485***	-0.302**	-0.286
	(-2.950)	(-2.279)	(-1.702)	(-3.299)	(-2.133)	(-1.629)
$CoSkew$	-0.011	0.011	-0.019**	-0.008**	-0.007	-0.018*
	(-1.352)	(1.293)	(-2.092)	(-2.061)	(-1.145)	(-1.814)
$Illiq_{1M}$	0.020**	-0.065	-0.044	0.021**	0.069	4.621
	(2.398)	(-1.514)	(-0.391)	(2.340)	(0.089)	(0.496)
$Adjusted\ R^2$	0.126	0.139	0.163	0.059	0.079	0.174
n	730	512	1814	1345	758	953
$\sigma(IVOL)$	52.374	28.775	18.822	45.431	21.246	14.135

price to -0.691% per month for medium price stocks. In the size subsamples the same effect amounts to -0.093% for the big stocks to -0.532% per month for small stocks. Thus, the negative relationship between IVOL and expected returns remains significant also in economic terms. These results are in line with L. H. Chen et al. (2012) who show that the IVOL puzzle is unlikely caused by market microstructure effects like size, price, liquidity or the security type.

The remaining risk premia estimates show almost the same pattern as the reference results of table 2 with few exceptions that I want to highlight. For $\beta(SMB)$ the premium turns negative in the high price and big stocks subsamples with values of -0.100 and -0.112 respectively that are statistically significant at a level not less than 5% and hence contradict the predictions of Fama and French (1993). Furthermore, the estimated size risk premium is now no longer statistically significant at any reasonable level for the small stock subsample, whereas it is particularly high for low price and microcap stocks with coefficients of -0.939 and -0.695 respectively. The BM premium is now only significant for the medium price subsample, though it generally remains positive. Additionally, in the low price subsample the skewness effect is absent, as the corresponding average risk premium is not significant at any plausible level. It is also interesting, that the effect documented by Boyer et al. (2010) seems to be caused by microstructure effects, as the $EIdioSkew$ premium is highest in magnitude for low price and microcap stocks with values of -0.913 and -0.665 that are also statistically significant at the 1% level and lowest for high price and big stocks with estimates of -0.163 and -0.141 that are not significant at any reasonable level. In addition to that, the coskewness premium turned positive for medium price stocks, although it is not significant at any reasonable level. Lastly, the illiquidity premium turns negative for the medium and high price stocks ranging from -0.017 to -0.108 and is even significant in the high price stock subsample at the 5% level.

The VW regression results of Panel B from table 6 still show negative IVOL risk premia for all subsamples, however, they are not statistically significant at any reasonable level when considering high price or big stock subsamples only. While the low and mid price, as well as the micro and small stocks, show IVOL risk premia around -1.711 and -2.722 that are all significant at the 1% level, they are remarkably lower and insignificant for high price and big stocks with values of -0.583 and -0.563 respectively. Hence, at least in the VW setting small and low price stocks indeed seem to drive the IVOL puzzle, wherefore the results appear to be related to microstructure effects that contradict the findings of L. H. Chen et al. (2012). This pattern is also found when analyzing the economic significance of the risk premia estimates by a one standard deviation increase in IVOL, holding all other variables constant. For the price subsamples the economic effect is the strongest for low price stocks causing a decrease in returns of on average 0.918% per month and the weakest for high price stocks with a associated return decrease of 0.110% per month. Similarly for the size subsamples, the effect ranges from -0.911% per month for the microcap stocks to -0.080% per month for big stocks. Consequently, the economic significance of the results also decreases as higher price or bigger stocks are considered. Hence, focusing on low to medium price or microcap to small stocks might enable researchers to tilt their results towards the discovery of the IVOL puzzle.

Related to the remaining risk premia estimates only a few changes have to be highlighted when compared to the results from Panel B of table 2. In accordance with Fama and French (1993), the $\beta(SMB)$ premium turns positive in the low and medium, as well as microcap stock subsamples, though it is not statistically significant at any plausible level. Also the BM premium has become positive for low to medium price, as well as microcap to small stocks, even though it is only significant for the microcap stock subsample at a level of 10%. Lastly, the illiquidity premium turns negative for medium to high price subsamples contradicting the finding of Amihud (2002), but it is of note that they are not statistically significant at any reasonable level.

In conclusion, the EW and the VW regression-based approach verify the existence of an IVOL puzzle regardless of the price or size subsample analyzed. The VW regressions, however, illustrate that the IVOL puzzle might be caused by microstructure effects, as also argued by Han and Lesmond (2011) as well as W. Huang et al. (2010). This likely holds, as the IVOL risk premium becomes insignificant for high price and big stocks in statistical, as well as economic terms, whereas it is significant in both interpretations for low price and microcap stocks. Hence, researchers that use only low price and microcap stocks in their regression-based analysis are more likely to find the IVOL puzzle than those using stocks with higher price or bigger size.

Table 7 presents the average portfolio total returns as well as the time-series alphas relative to the FF3 and FF6 model of the portfolio-based research concept for portfolios 1, 5 and the long-short portfolio in the previously introduced price and size subsamples. Each column corresponds to a subsample of stocks on the basis of which the monthly quintile portfolios have been constructed and returns as well as alphas have been calculated. Panel A depicts the portfolio-based analysis of the EW and Panel B for the VW portfolio returns respectively. The equal-weighted portfolio sorts from the price subsamples show that the IVOL puzzle finding is robust when looking at the subset of medium and high price stocks but less for the low price stock subsample. Analyzing the average EW long-short portfolio total returns of the price subsamples in Panel A, the positive reference returns of section 5.3 are only found for low price stocks. For the medium and high price stocks, the corresponding average return becomes negative and ranges from -0.210% for high price to -1.144% per month for low price stocks with t-statistics of -5.551 and -1.030. These negative total return differences found by the long-short portfolio are driven by the comparatively high performance of the low IVOL portfolios that is statistically significant at the 1% level in both subsamples. The long-short portfolio α^{FF3} and α^{FF6} show a mainly simi-

Table 7: Portfolio-Based Results - Sample-Related Adjustments

This table presents the average results from the univariate portfolio sorting procedure as explained in section 3.3 covering all stocks traded on the NYSE/AMEX/NASDAQ. Here the portfolio analysis is conducted by subdividing the complete data set into subsamples that are formed based on stock price and stock size as denoted in the respective column and applying the analysis to the subsamples accordingly. In the first column named "Price" the complete sample data are subdivided into low, medium and high price stocks with cutoffs of 5$ and 10$. The subsample criteria called "Size" refers to the subdivision of the data according to the stock-specific market capitalization. Here the complete data is split each month into micro, small and big stocks based on the 20th and the 50th percentile of the cross-sectional distribution of market capitalization when using all stocks traded on the NYSE. Every month stocks from these subsamples are then sorted again into quintile portfolios based on their idiosyncratic volatility relative to the FF3 model over the past month that has been calculated by the usage of daily data for the respective month. These portfolios are rebalanced monthly according to the 1/0/1 trading strategy and the corresponding results for alphas and total returns are calculated. In the first row I report the average monthly total returns for the portfolios 1, 5 as well as the long-short portfolio called "5-1" (in percentage terms) and in parenthesis below each of these returns I show the corresponding Newey and West (1987) adjusted t-statistics where I use 7 lags for the adjustment. Portfolio 1 (5) contains all stocks with lowest (highest) IVOL, whereas the long-short portfolio is the one that goes long the portfolio of stocks with the highest IVOL and short the respective portfolio of stocks with the lowest IVOL. This long-short portfolio is also rebalanced monthly according to the 1/0/1 trading strategy. The α^{FF3} and the α^{FF6} in the second and third row denote the monthly average of Jensen's time-series alphas relative to the FF3 and the FF6 model that were calculated by equation 1 using the monthly portfolio excess returns over the complete sample horizon and are again reported for portfolio 1, 5 and the long-short portfolio respectively. Below each of the alphas I provide the corresponding t-statistics that were again corrected by the procedure of Newey and West (1987) using 7 lags. In Panel A I use the equal-weighted portfolio returns for the analysis, where all firms are getting the same weight. On the other hand, Panel B uses value-weighted portfolio returns that were calculated by weighting the stock returns within the portfolio by their market capitalization which is observable at beginning of the month in order to give higher weights to bigger stocks respectively and therefore diminish the effects that might be explicitly related to small stocks. The complete sample period covers July 1963 to December 2020.

Panel A: Equal-Weighted Portfolio Sorts

		Price			Size		
		Low	Medium	High	Microcap	Small	Big
$TotalReturn$	1	1.316	1.052	0.813	1.020	0.884	0.813
		(3.546)	(3.768)	(3.794)	(3.841)	(3.777)	(4.117)
	5	1.490	-0.092	0.604	0.981	0.381	0.673
		(3.067)	(-0.255)	(1.938)	(2.274)	(1.079)	(2.202)
	5-1	0.175	-1.144	-0.210	-0.039	-0.503	-0.140
		(0.706)	(-5.551)	(-1.030)	(-0.133)	(-1.956)	(-0.641)
α^{FF3}	1	0.217	0.153	-0.002	0.191	0.025	-0.064
		(0.934)	(0.681)	(-0.008)	(0.843)	(0.105)	(-0.498)
	5	0.113	-1.379	-0.555	-0.373	-0.960	-0.506
		(0.355)	(-8.540)	(-4.372)	(-1.540)	(-5.774)	(-3.608)
	5-1	-0.104	-1.533	-0.554	-0.563	-0.986	-0.442
		(-0.443)	(-10.358)	(-4.780)	(-2.631)	(-5.795)	(-3.073)
α^{FF6}	1	0.443	0.132	-0.112	0.149	-0.063	-0.182
		(1.813)	(0.458)	(-0.626)	(0.539)	(-0.222)	(-1.548)
	5	0.870	-1.028	-0.378	0.282	-0.505	-0.131
		(2.374)	(-6.102)	(-3.124)	(0.915)	(-2.731)	(-0.931)
	5-1	0.427	-1.160	-0.267	0.133	-0.442	0.051
		(1.607)	(-8.897)	(-2.749)	(0.513)	(-2.875)	(0.427)

Panel B: Value-Weighted Portfolio Sorts

		Price			Size		
		Low	Medium	High	Microcap	Small	Big
$TotalReturn$	1	1.024	0.901	0.776	1.009	0.890	0.802
		(2.821)	(3.487)	(4.054)	(4.007)	(3.808)	(4.198)
	5	-0.343	-0.297	0.481	-0.256	0.414	0.648
		(-0.708)	(-0.746)	(1.546)	(-0.634)	(1.171)	(2.143)
	5-1	-1.367	-1.198	-0.295	-1.265	-0.475	-0.154
		(-4.388)	(-4.241)	(-1.323)	(-4.431)	(-1.837)	(-0.724)
α^{FF3}	1	-0.108	-0.031	-0.033	0.175	0.016	-0.004
		(-0.409)	(-0.159)	(-0.262)	(0.774)	(0.065)	(-0.038)
	5	-1.810	-1.675	-0.619	-1.648	-0.919	-0.441
		(-5.582)	(-8.243)	(-4.301)	(-8.311)	(-5.415)	(-3.106)
	5-1	-1.702	-1.645	-0.586	-1.824	-0.935	-0.437
		(-6.172)	(-7.983)	(-4.149)	(-9.323)	(-5.460)	(-2.909)
α^{FF6}	1	0.246	0.020	-0.186	0.120	-0.077	-0.132
		0.831	(0.092)	(-1.818)	(0.409)	(-0.243)	(-1.655)
	5	-0.978	-1.145	-0.363	-1.029	-0.452	-0.101
		(-2.866)	(-5.347)	(-2.280)	(-4.575)	(-2.342)	(-0.648)
	5-1	-1.224	-1.165	-0.177	-1.149	-0.375	0.031
		(-4.388)	(-6.297)	(-1.460)	(-5.877)	(-2.449)	(0.237)

lar pattern. Both are negative and statistically significant at the 1% level for medium and high price stocks, whereas low price stocks have a positive α^{FF3} and a negative α^{FF6} that are both not significant at any reasonable level. The alphas for the medium and high price stocks range from -1.533% to -0.267% per month and consequently are also economically relevant while matching the reference results. For both alpha types only the low portfolio 5 alpha is statistically significant at the 1% level and hence drives the negative long-short portfolio alphas for medium and high price stocks. Thus, the price subsamples show that researchers using medium to high price stocks in a EW portfolio-based analysis are more likely to discover a statistically significant IVOL puzzle than those using only low price stocks.

On the other hand, the IVOL puzzle appears more sensitive in the EW portfolio-based analysis of the size subsamples. Even though all size subsamples show negative long-short portfolio total returns now, they are only statistically significant at the 10% level for the small stock subsample. This effect is primarily driven by the high total returns of portfolio 1 as these are all statistically significant at the 1% level in all size subsamples. However, similar to the reference results, the long-short portfolio α^{FF3} is negative and statistically significant for all size subsamples ranging from -0.442% to -0.986% per month with Newey-West robust t-statistics of -3.073 and -5.795. Again the low α^{FF3} of portfolio 5 causes this effect, as it is the only one that is statistically significant at the 1% level for all size subsamples, except for the microcap stocks where it is not significant at any reasonable level. In terms of the α^{FF6}, only the long-short portfolio alpha of the small stocks is negative and statistically significant at the 1% level with a value of -0.442% per month that is also economically viable. Those of the big and microcap stocks match the reference results of section 5.3 and are positive albeit not statistically significant. In addition, no pattern can be found when comparing the α^{FF6} of portfolios 5 and 1 as only the small stock subsample portfolio 5 alpha is statistically significant. Thus, researchers are most likely finding an IVOL puzzle in all return measures using small stocks only or when just computing the α^{FF3} for their EW portfolio-based analysis.

The VW sorts from Panel B find the IVOL puzzle more often than the EW sorts do, even though the finding is still sensitive to the price or size subsample analyzed. In the price subsamples the low and medium price stocks are more likely to show the return pattern implied by the IVOL puzzle than when analyzing the high price stocks. The long-short portfolio total returns are negative in all price subsamples but only statistically significant for low and medium price stocks. They range from -1.198% for medium price to -1.367% per month for low price stocks with t-statistics of -4.241 and -4.388. Thus they are not only statistical significant at the 1% level but are also economically meaningful as in the reference results. Similar to the reference result, the comparatively high total return of portfolio 1 mainly drives this effect, as it is the only one that is statistically significant at a level of 1%, irrespective of the price subsample investigated. Regardless of the price subsample considered, the long-short portfolio α^{FF3} is negative and statistically significant at the 1% level with risk-adjusted returns between -0.589% per month for high to -1.702% per month for low price stocks that are also economically relevant. As in the reference result, the low α^{FF3} of portfolio 5 relative to portfolio 1 seems to cause the negative long-short alpha in all price subsamples as it is the only one that is statistically significant at a level of 1%. The α^{FF6} show the same pattern as the total returns implying a negative long-short portfolio alpha in all price subsamples, although they are only statistically significant for low and medium price stocks. Compared to the reference results, these statistically significant alphas are almost twice as high, amounting to values between -1.165% for medium price to -1.224% per month for low price stocks which are hence also economically significant. However, it is still the low alpha of portfolio 5 that is statistically significant at a level not less than 5% and consequently drives the IVOL effect in the long-short α^{FF6}. Overall the price subsamples reveal a negative relationship between the stock price and the statistical relevance of the IVOL puzzle meaning that researchers who use low or medium price stocks are more likely to prove the existence of the IVOL puzzle than those analyzing high price stocks.

The size subsamples of Panel B illustrate that the IVOL puzzle is also sensitive to the choice of the stock sample when filtering for MTCAP. In general, these subsamples show a negative relationship between stocks' MTCAP and the statistical significance of the IVOL puzzle. The long-short portfolio total returns is negative in all size subsamples but decreases in significance as bigger stocks are used for the analysis. Average long-short portfolio returns increase from -1.265% per month for microcap stocks to -0.154% per month for big stocks with t-statistics of -4.431 and -0.724 which are thus only statistically and economically significant for microcap and small stocks. Therefore, the big stock long-short portfolio total returns match the reference results only in sign but not in statistical significance. Across all size subsamples the high return of portfolio 1 causes this effect again, as it is the only one that is statistically significant at a level not less than 1%. A similar pattern is found for the long-short portfolio α^{FF3} that increases from -1.824% for microcap to -0.473% per month for big stocks and is economically meaningful and statistically significant at the 1% level in all size subsamples. The long-short portfolio α^{FF6} is negative for the microcap, as well as small stocks, matching the reference results but is positive, albeit not significant, for big stocks. It ranges from -1.149% per month for microcap to 0.031% per month for big stocks with t-statistics of -5.877 to 0.237 indicating that only the microcap and small stock α^{FF6} are statistically and economically relevant. In the microcap and small stock subsamples the low portfolio 5 alpha is still statistically significant at a level not less than 5% and thus drives the IVOL puzzle alpha pattern. In summary, the same pattern as in the price subsamples appears implying that researchers using microcap or small stocks are more prone to find an IVOL puzzle than those focusing on stocks with a price higher than 10$.

The previous results indicate that a researcher can affect

the conclusions drawn from his study by a smart choice of the subsample used for his portfolio-based analysis. Especially for VW portfolio sorts, a monotonic pattern shows up entailing that analyzing bigger stocks or those with a higher price makes it harder to discover a statistically significant negative relationship between IVOL and expected stock returns as implied by the IVOL puzzle. Assuming that small and low price stocks are more likely subject to microstructure effects, these results contradict the findings of L. H. Chen et al. (2012) who argue that the IVOL puzzle is robust to such effects. However, regardless of the weighting-scheme and subsample used in the analysis, an IVOL puzzle is almost always found when analyzing the α^{FF3} only. The cross-concept comparison again shows that the IVOL puzzle finding from the EW analysis is more sensitive to the research concept used in the subsample analysis than the VW one, as it mostly delivers the same conclusions across approaches. In the EW regression results the IVOL puzzle exists regardless of the price or size subsample considered, whereas the portfolio-based results draw a slightly different conclusion as there the puzzle is absent in low price and microcap stocks. Consequently the regression-based approach is more likely to find a subsample robust IVOL puzzle than the portfolio-based method. Conversely, the VW results show that the IVOL puzzle exist in all subsamples but becomes less significant when higher prices or larger MTCAP stocks are used for the analysis, regardless of the research concept employed. Hence, for both VW methods the highest price, as well as biggest stock subsamples, show a IVOL puzzle effect that is not statistically significant at any reasonable level except when investigating the α^{FF3} in the portfolio-based approach.

7. Method-Specific Adjustments

This section covers all adjustments that are method-specific and thus only applied to the corresponding research concept. Section 7.1 presents the regression-based and section 7.2 the portfolio-based research concept adjustments respectively.

7.1. Regression-Related

The regression specific adjustments are structured according to the two categories introduced in section 3.4.2 consisting of risk-related control variables and estimation procedure-related adjustments. Section 7.1.1 starts with the adjustments in the risk-related control variables, whereas section 7.1.2 continues with all changes in the way the time-series risk factor loadings as well as the risk premia are estimated.

7.1.1. Risk-Related Control Variables

This section analysis the results from modification of the variables used for systematic or firm-specific risk-correction in the IVOL risk premia estimation of equation 5. For this I use the complete regression model of column (7) from table 2 and exchange the FF3 model risk factor loadings for those relative to the FF6, SY4, HOU5 and the FFA model respectively. Then I integrate further variables into the vector of firm characteristics including the stock's monthly trading volume ($Volume$), the average daily bid-ask spread over the current month ($Spread$) and the maximum daily stock return of the current month ($Maxret$) as proposed by Bali et al. (2011). The corresponding results are depicted in table 8, where Panel A uses the EW and Panel B the VW Fama-MacBeth regression for risk premia estimation.

The results of Panel A show that the IVOL puzzle is robust to the usage of different risk-correction models as well as the inclusion of $Volume$, $Spread$ and $Maxret$ as further firm characteristic control variables in the cross-sectional regression step. Across all the adjustments considered here, the IVOL risk premium estimate remains negative and statistically significant at the 1% level with coefficients ranging from -0.964 when controlling for the HOU5 model to -1.073 when using the FF3 model for risk-correction and incorporating all further firm characteristic control variables. Investigating a one-standard deviation increase in IVOL while holding all other variables constant, these IVOL risk premia translate into a decrease of expected returns between 0.361% to 0.401% per month that are also economically significant. Compared to the complete regression model in Panel A from table 2, the IVOL risk premium is now higher except when systematic risk is corrected by the HOU5 model or $Volume$ is used as firm-specific control. Even though the FF6, the SY4 and the HOU5 model claim to outperform the FF3 model, none of them can explain the IVOL puzzle (see Fama & French, 2018; Hou et al., 2020; Stambaugh & Yuan, 2017). The FFA model results also show that the short-term reversal IVOL puzzle explanation of W. Huang et al. (2010) is unlikely, as the puzzle persists despite the fact that the model explicitly contains a short-term reversal factor that should control for this effect. In addition, I cannot verify the arguments of Bali et al. (2011) and Han and Lesmond (2011) who claim that researchers who account for $Maxret$ or the bid-ask spread in their study could resolve the IVOL puzzle. Instead I find them to have almost no effect on the IVOL risk premia estimates.

The remaining risk premia estimates are nearly identical to the reference results from table 2. The only difference is that, albeit in almost all regressions the $EIdioSkew$ risk premium remains in the same range as in the reference result, it is statistically significant at the 1% level except for the case when the SY4 or $Maxret$ is used for risk-correction. In addition, the $CoSkew$ risk premium is only statistical significant at the 10% level when the SY4 model is used for risk-correction. Among the further firm characteristic control variables only the $Spread$ risk premium is statistically significant at the 10% level with coefficients between between 0.099 when it is added separately to the regression model and 0.106 when all further controls are integrated simultaneously. In contrast, the $Volume$ risk premium is almost zero and not statistically significant at any level, while $Maxret$ carries a positive risk premium that contradicts the findings of Bali et al. (2011) who predict a negative premium instead.

Table 8: Regression-Based Results - Risk-Related Control Variables

This table presents the average coefficients from monthly Fama and MacBeth (1973) cross-sectional regressions for individual stocks. Using equation 5, each month I regress the next-month excess firm returns (in percentage points) on a constant, the annualized idiosyncratic volatility measure of equation 4 calculated from daily data over the current month (in decimals), the stock-specific risk factor loadings relative to a selected risk-correction model over the next month, the end of month firm size defined as the natural logarithm of market capitalization, the end of month firm book-to-market ratio of equity, the monthly return (in percentage points) computed from the end of the previous month to the current month as well as additional control variables related to further firm characteristics. These firm characteristic variables are the following: $Skew_{1M}$ which is the monthly return skewness measured over daily return data from the current month, $EIdioSkew$ that is the expected idiosyncratic skewness measure from Boyer et al. (2010), $CoSkew$ which is the coskewness measure from Harvey and Siddique (2000) and $Illiq_{1M}$ being the Amihud (2002) illiquidity measure computed over the current month. The first four columns show the results from adjustments in the risk-correction model used in the risk premia estimation of equation 5. Here each of these columns refers to a different model towards which the risk factor loadings were computed using daily data over the next month as usual. These models consist of the FF6, the SY4, the HOU5 and the FFA respectively. I do not present the risk premia of the risk-correction models in this table for the matter of brevity. In columns C1 to CA I use the FF3 model for risk-correction again but incorporate additional firm characteristic variables instead which are the following: $Volume$ which is the stock's monthly trading volume, $Spread$ which is the average daily bid-ask spread over the current month and $Maxret$ that refers to the maximum daily stock return of the current month as proposed by Bali et al. (2011). The last column named "CA" then shows the risk premia estimates when all the additional firm characteristic variables are used in the Fama-MacBeth regression simultaneously. Panel A uses equal-weighted stock excess returns to run the OLS regression of equation 5. On the other hand, in Panel B I estimate the Fama-MacBeth regression with a weighted-least squares approach, where all individual monthly stock returns are weighted by their current month market capitalization. I report the t-statistics testing the null hypothesis that the average coefficient is equal to zero in parenthesis below each coefficient, where corresponding standard errors are corrected for autocorrelation and heteroskedasticity by implementation of the Newey and West (1987) method using 7 lags. *, ** , and *** indicate significance at the 10%, 5%, and 1% level, respectively. The row "Adjusted R^2" reports the time-series average of the cross-sectional adjusted R^2's. The sample covers the period from July 1963 until December 2020 and includes all stocks traded on the NYSE/AMEX/NASDAQ. It is shorter when the SY4 model is used in the analysis, which is due a lack of data availability wherefore it then covers the horizon until November 2016 only.

Panel A: Equal-Weighted Fama-MacBeth Regressions

	FF6	*SY4*	*HOU5*	*FFA*	*C1*	*C2*	*C3*	*CA*
Constant	1.833***	2.147***	1.813***	1.873***	1.842***	1.886***	1.805***	1.859***
	(4.585)	(5.603)	(4.537)	(4.755)	(4.439)	(4.694)	(4.760)	(4.805)
IVOL	-1.012***	-0.984***	-0.964***	-0.989***	-0.959***	-1.005***	-1.060***	-1.073***
	(-4.457)	(-3.956)	(-4.180)	(-4.299)	(-4.223)	(-4.245)	(-2.597)	(-2.710)
Size	-0.154***	-0.182***	-0.152***	-0.158***	-0.159***	-0.174***	-0.155***	-0.172***
	(-4.262)	(-4.693)	(-4.199)	(-4.370)	(-3.921)	(-4.733)	(-4.579)	(-4.654)
BM	0.138**	0.148**	0.138**	0.138**	0.143**	0.144**	0.139**	0.146**
	(2.230)	(2.137)	(2.191)	(2.211)	(2.284)	(2.299)	(2.218)	(2.350)
R_t	-0.041***	-0.044***	-0.042***	-0.043***	-0.043***	-0.043***	-0.043***	-0.044***
	(-10.900)	(-11.463)	(-10.808)	(-11.098)	(-10.766)	(-10.684)	(-11.076)	(-11.109)
$Skew_{1M}$	0.100***	0.111***	0.102***	0.106***	0.105***	0.103***	0.133***	0.137***
	(4.199)	(4.195)	(4.247)	(4.396)	(4.211)	(4.177)	(3.618)	(3.720)
EIdioSkew	-0.327***	-0.340**	-0.314***	-0.323***	-0.305***	-0.311***	-0.295**	-0.311***
	(-2.752)	(-2.606)	(-2.645)	(-2.769)	(-2.597)	(-2.659)	(-2.529)	(-2.687)
CoSkew	-0.010**	-0.010*	-0.011**	-0.012**	-0.011**	-0.011**	-0.011**	-0.011**
	(-2.283)	(-1.898)	(-2.401)	(-2.353)	(-2.317)	(-2.195)	(-2.286)	(-2.188)
$Illiq_{1M}$	0.033***	0.032***	0.035***	0.035***	0.036***	0.036***	0.037***	0.037***
	(3.520)	(3.324)	(3.577)	(3.556)	(3.489)	(3.502)	(3.539)	(3.543)
Volume					0.000			0.000
					(-0.994)			(-1.083)
Spread						0.099*		0.106*
						(1.699)		(1.789)
Maxret							0.002	0.001
							(0.110)	(0.094)
Adjusted R^2	0.084	0.069	0.079	0.080	0.068	0.067	0.068	0.069
n	3056	3024	3056	3056	3056	3056	3056	3056

Panel B: Value-Weighted Fama-MacBeth Regressions

	FF6	*SY4*	*HOU5*	*FFA*	*C1*	*C2*	*C3*	*CA*
Constant	1.577***	1.820***	1.531***	1.481***	1.594***	1.556***	1.563***	1.687***
	(4.318)	(5.055)	(4.058)	(4.061)	(3.816)	(4.179)	(4.234)	(4.045)
IVOL	-0.724**	-0.868**	-0.767**	-0.624*	-0.864**	-0.826**	-1.685***	-1.946***
	(-2.271)	(-2.481)	(-2.355)	(-1.882)	(-2.542)	(-2.359)	(-3.133)	(-3.947)
Size	-0.088***	-0.105***	-0.084**	-0.081**	-0.088**	-0.087***	-0.085***	-0.101***
	(-2.657)	(-3.025)	(-2.551)	(-2.478)	(-2.350)	(-2.602)	(-2.607)	(-2.711)
BM	-0.001	0.012	0.022	-0.013	-0.004	-0.018	-0.019	0.012
	(-0.015)	(0.120)	(0.235)	(-0.147)	(-0.040)	(-0.201)	(-0.205)	(0.130)
R_t	-0.030***	-0.029***	-0.031***	-0.037***	-0.026***	-0.025***	-0.029***	-0.030***
	(-6.659)	(-5.632)	(-6.651)	(-8.584)	(-5.602)	(-5.300)	(-5.766)	(-6.057)
$Skew_{1M}$	0.123***	0.131***	0.130***	0.129***	0.106***	0.102***	0.050	0.034
	(4.022)	(3.959)	(4.455)	(4.476)	(3.395)	(3.219)	(1.027)	(0.721)
EIdioSkew	-0.226	-0.225	-0.227	-0.223	-0.218	-0.250*	-0.224	-0.218
	(-1.613)	(-1.438)	(-1.594)	(-1.584)	(-1.460)	(-1.666)	(-1.492)	(-1.478)
CoSkew	-0.016**	-0.014	-0.017**	-0.017**	-0.015*	-0.016*	-0.016*	-0.014*
	(-2.068)	(-1.605)	(-2.142)	(-2.149)	(-1.805)	(-1.901)	(-1.919)	(-1.733)
$Illiq_{1M}$	0.001	0.005	-0.001	-0.005	0.005	0.001	-0.001	0.002
	(0.076)	(0.346)	(-0.091)	(-0.322)	(0.332)	(0.099)	(-0.097)	(0.144)
Volume					0.000			0.000
					(-0.648)			(-0.657)
Spread						0.072**		0.071**
						(2.309)		(2.381)
Maxret							0.052**	0.060***
							(2.088)	(2.660)
Adjusted R^2	0.182	0.164	0.174	0.180	0.162	0.156	0.156	0.169
n	3056	3024	3056	3056	3056	3056	3056	3056

However, the $Maxret$ risk premium found here is not statistically significant.

Moving to the VW regression results of Panel B, the negative IVOL risk premium remains in all regression specifications ranging from -0.624, when controlling risk by the FFA model, to -1.946 when simultaneously controlling for all additional firm-specific control variables. Compared to the reference results, the IVOL risk premium has more than doubled when $Maxret$ is added to the vector of firm characteristics and is still statistically significant at the 1% level, which is at odds with Bali et al. (2011) who claim that the incorporation of $Maxret$ should resolve the IVOL puzzle. However, it appears to have the opposite effect and increases its magnitude instead. On the other hand, when controlling systematic risk by the FFA model, the IVOL risk premium is only significant at the 10% level and decreases in magnitude. The decrease in statistical significance might be related to the inclusion of the short-term reversal factor that accounts for the short-term-related IVOL puzzle explanation of W. Huang et al. (2010). When computing the effect of a one standard deviation increase in IVOL, holding all other variables constant, these risk premia estimates predict a decrease in expected returns between 0.233% and 0.728% per month that appears highly dispersed but remains inside a bandwidth that is economically meaningful.

For the remaining risk premia estimates a few changes need to be highlighted. First, the BM risk premium turns positive when systematic risk is corrected for by the SY4 or the HOU5 model, albeit not statistically significant at any plausible level in both cases. Interestingly the $Skew_{1M}$ risk premium becomes insignificant when $Maxret$ is added to the cross-sectional regression specification. Furthermore, the $Maxret$ risk premium is now statistically significant at the 5% level when added as the only further control variable and even at the 1% level if all additional firm characteristic controls are added to the regression simultaneously. Hence, as both of these coefficients are related to the skewness preferences of investors, $Maxret$ seems to have more explanatory power than $Skew_{1M}$. In addition, the positive $Spread$ risk premium is now statistically significant at the 5% level showing that it has some explanatory power in the VW Fama-Macbeth regression context. Lastly, it is of note that the $Illiq_{1M}$ risk premium turns negative when the HOU5 model, the FFA model or $Maxret$ is used for risk-correction, although it is never statistically significant.

Table 8 shows that the IVOL puzzle finding in the regression-based approach is robust to the inclusion of other systematic risk-correction models or further firm characteristic controls. Whereas for the EW estimation method the results remain similar to those of the reference setting, some changes occur in the VW regressions. When researchers employ the FFA model for risk-correction in the VW context they find only a weakly significant IVOL puzzle, whereas the inclusion of $Maxret$ as further control lets the puzzle appear stronger in absolute terms than in the reference setting.

7.1.2. Estimation Procedure

In this section I analyze the modifications related to the estimation of the first step time-series regressions and the second step cross-sectional regressions as introduced in section 3.4.2. Specifically I address changes in the way the time-series regression of equation 3 is computed to obtain risk factor loadings and the manner in which the risk premia estimates are derived with the cross-sectional regression of equation 5. The corresponding results can be found in table 9. Panel A shows the equal- and value-weighted regression results from the time-series estimation adjustments where the FF3 model risk factor loadings are estimated over rolling windows of 12 and 60 months starting at month $t+1$ as well as over the complete sample horizon as depicted in the respective column. In Panel B I present the results from the cross-sectional regression adjustments where the GLS estimation technique is implemented. Here the stock returns are weighted with a diagonal weighting matrix consisting of the inverse of the estimated stock return variances that are computed over rolling windows from the past 1, 12 and 60 months as well as over the complete sample horizon as indicated in the specific column.

The results of Panel A show that the IVOL risk premium is negative and statistically significant at the 1% level for all risk factor loading estimation windows, no matter whether the equal- or the value-weighted Fama-MacBeth regression procedure is applied. For the EW regressions the IVOL premium ranges from -0.847 for the 60 month window factor loadings to -1.008 when computing them over the complete sample horizon, whereas it is slightly higher in the VW context with values of -1.070 for the 60 months to -1.149 for the 12 months window factor loadings. Using these estimates for the computation of economic significance by a one standard deviation increase in IVOL, holding all other variables constant, the corresponding decrease in expected returns amounts to values between 0.317% and 0.377% per month for the EW as well as 0.400% and 0.430% per month for the VW regression context. So the IVOL puzzle remains economically, as well as statistically, meaningful and is, moreover, comparable in magnitude to the reference results of column (7) in table 2, regardless of the weighting-scheme applied. However, it is of note that 60 months factor loadings lead to the lowest IVOL premium among all estimation windows, irrespective of the weighting-scheme used. Nevertheless, most of the regression specifications show more negative IVOL risk premia when compared to the reference results with exception of the 12 and 60 months factor loading regressions in the EW context that lead to higher IVOL premia instead. Additionally, the adjusted R^2 shows for both weighting-schemes that the model performance decreases as the window used for the factor loading estimation increases. Interestingly the adjusted R^2 for the 12 months factor loadings are higher for both weighting-schemes when compared to the corresponding reference results. This shows that 12 months factor loadings might more precisely capture the underlying risk than the usual 1 months loadings do and hence

Table 9: Regression-Based Results - Estimation Procedure Adjustments

This table presents the average coefficients from monthly Fama and MacBeth (1973) cross-sectional regressions for individual stocks. Using equation 5, each month I regress the next-month excess firm returns (in percentage points) on a constant, the annualized idiosyncratic volatility measure of equation 4 calculated from daily data over the current month (in decimals), the stock-specific risk factor loadings relative to the FF3 model, the end of month firm size defined as the natural logarithm of the market capitalization, the end of month firm book-to-market ratio of equity, the monthly return (in percentage points) computed from the end of the previous month to the current month as well as additional control variables related to further firm characteristics. These firm characteristic variables are the following: $Skew_{1M}$ which is the monthly return skewness measured over daily return data from the current month, $EIdioSkew$ that is the expected idiosyncratic skewness measure from Boyer et al. (2010), $CoSkew$ which is the coskewness measure from Harvey and Siddique (2000) and $Illiq_{1M}$ being the Amihud (2002) illiquidity measure computed over the current month. Panel A shows the results of the Fama-MacBeth regression procedure from adjustments to the way the factor loadings relative to the FF3 model were computed. Instead of computing them only over the next month I now use the next month including the previous 11 or 59 months as well as estimate them over complete sample described as 12M, 60M and FULL in the respective column. These three estimation windows for risk factor loading estimation are implemented in the EW and the VW Fama-MacBeth regression context as shown in the respective column. Panel B reports the results associated with the adjustments in the way the cross-sectional regression of equation 5 is estimated. Here I use a GLS approach where observations are weighted according to a diagonal weighting matrix that consists of the inverse of the estimated stock return variances. The estimation of the return variance is conducted by calculation over rolling windows from the past 1, 12 and 60 months as well as the complete sample while using daily data as shown in the corresponding column. I report the t-statistics testing the null hypothesis that the average coefficient is equal to zero in parenthesis below each coefficient, where corresponding standard errors are corrected for autocorrelation and heteroskedasticity by implementation of the Newey and West (1987) method using 7 lags. *, ** , and *** indicate significance at the 10%, 5%, and 1% level, respectively. The row "Adjusted R^2" reports the time-series average of the cross-sectional adjusted R^2's. The last row reports the average number of stocks used for the monthly cross-sectional regressions. The sample covers the period from July 1963 until December 2020 and includes all stocks traded on the NYSE/AMEX/NASDAQ.

Panel A: Time-Series Regression Step - Different Estimation Windows

	EW			VW		
	12M	60M	FULL	12M	60M	FULL
$Constant$	1.940***	1.847***	2.366***	1.321***	1.254***	2.771***
	(5.214)	(4.964)	(6.121)	(3.281)	(2.796)	(5.507)
$IVOL$	-0.923****	-0.847***	-1.008***	-1.149***	-1.070***	-1.075***
	(-4.844)	(-4.751)	(-5.357)	(-4.050)	(-3.851)	(-3.595)
$\beta(MKT)$	0.102	0.114	1.612***	0.249	0.116	0.577*
	(0.538)	(0.508)	(5.391)	(1.035)	(0.461)	(1.833)
$\beta(SMB)$	0.031	0.000	-0.842***	0.014	0.071	-0.816***
	(0.295)	(0.002)	(-5.047)	(0.126)	(0.548)	(-5.879)
$\beta(HML)$	-0.033	0.002	-0.589**	-0.131	-0.160	-0.642***
	(-0.268)	(0.012)	(-2.552)	(-0.907)	(-0.999)	(-3.800)
$Size$	-0.156***	-0.146***	-0.344***	-0.072**	-0.063*	-0.259***
	(-4.304)	(-3.581)	(-7.213)	(-2.537)	(-1.743)	(-6.968)
BM	0.117**	0.077	0.131***	0.055	0.088	0.093
	(2.285)	(1.561)	(2.717)	(0.710)	(1.007)	(1.147)
R_t	-0.043***	-0.046***	-0.046***	-0.024***	-0.025***	-0.025***
	(-10.456)	(-11.185)	(-11.079)	(-5.351)	(-5.134)	(-5.083)
$Skew_{1M}$	0.110***	0.116***	0.115***	0.126***	0.131***	0.133***
	(4.470)	(4.765)	(4.736)	(4.119)	(3.934)	(4.163)
$EIdioSkew$	-0.336***	-0.275**	-0.383***	-0.193	-0.158	-0.187
	(-3.077)	(-2.534)	(-3.437)	(-1.400)	(-1.099)	(-1.272)
$CoSkew$	-0.011**	-0.008	-0.014***	-0.015*	-0.019**	-0.024***
	(-2.382)	(-1.593)	(-3.115)	(-1.902)	(-2.424)	(-3.130)
$Illiq_{1M}$	0.039***	0.042***	0.043***	0.027	0.034	0.009
	(3.953)	(4.193)	(4.519)	(1.416)	(1.558)	(0.514)
$Adjusted\ R^2$	0.075	0.063	0.058	0.175	0.159	0.145
n	3056	2818	3038	3056	2818	3038

Panel B: Cross-Sectional Regression Step - GLS

	1M	12M	60M	FULL
$Constant$	0.553	1.162***	1.166***	1.339***
	(1.600)	(3.696)	(3.836)	(4.442)
$IVOL$	0.567	-0.818***	-0.842***	-0.549**
	(1.378)	(-2.929)	(-3.325)	(-2.300)
$\beta(MKT)$	0.214*	0.153	0.125	0.149
	(1.796)	(1.265)	(1.017)	(1.201)
$\beta(SMB)$	-0.004	-0.028	-0.015	-0.022
	(-0.105)	(-0.774)	(-0.415)	(-0.629)
$\beta(HML)$	-0.045	-0.045	-0.027	-0.025
	(-0.752)	(-0.704)	(-0.426)	(-0.397)
$Size$	-0.029	-0.061**	-0.059**	-0.092***
	(-1.169)	(-2.540)	(-2.286)	(-3.515)
BM	0.064	0.065	0.041	0.077
	(1.192)	(1.144)	(0.707)	(1.454)
R_t	-0.036***	-0.035***	-0.040***	-0.042***
	(-8.399)	(-9.157)	(-10.745)	(-11.174)
$Skew_{1M}$	0.002	0.084***	0.088***	0.090***
	(0.095)	(5.112)	(5.379)	(4.966)
$EIdioSkew$	-0.128	-0.099	-0.050	-0.027
	(-1.227)	(-0.898)	(-0.444)	(-0.235)
$CoSkew$	-0.011**	-0.014**	-0.016**	-0.014**
	(-1.982)	(-2.600)	(-2.375)	(-2.468)
$Illiq_{1M}$	0.022**	0.021***	0.019**	0.021**
	(2.459)	(2.884)	(2.436)	(2.612)
$Adjusted\ R^2$	0.085	0.071	0.073	0.071
n	3054	3056	2802	3038

researchers should reconsider to use these estimates to improve their model performance. In conclusion I can verify the findings of Z. Chen and Petkova (2012) that the adjustments in the estimation window used for risk factor loading estimation alone, do not influence the IVOL puzzle finding remarkably, as the IVOL risk premia remain negative and statistically significant.

With a few exceptions, the remaining risk premia estimates have hardly changed in comparison to the reference results. The risk premia of the FF3 model risk factor loadings turn statistically significant when the full sample is used for their computation, regardless of the weighting-scheme employed. It is of note, however, that for both weighting-schemes the average risk premia of the full sample $\beta(SMB)$ and $\beta(HML)$ are negative, which contradicts the predictions of Fama and French (1993) who argue for positive premia instead. Furthermore, the negative BM risk premium of the VW regressions from the reference result turns positive here and thus is more in line with the value effect also mentioned by Fama and French (1993). Nevertheless, this premium stays insignificant for all window sizes in the VW context and for the 60 months factor loadings in the EW regression analysis. Panel B shows that weighting stock returns by their estimated variance can have a non-trivial influence on the IVOL risk premium estimate. When the return variance is estimated from rolling windows over the current month, the IVOL risk premium estimate turns positive but also statistically insignificant. Hence, the IVOL puzzle seems to be related to the general stock return variance, as already found in section 5.3 as well as by Barinov and Chabakauri (2020) and accounting for this relationship by, for example, weighting observations as done here, would enable researchers to resolve the IVOL puzzle consequently. However, the remaining columns show that when the weighting is induced by return variance estimates from rolling window estimates covering 12 and 60 months or the complete sample horizon, the negative and statistically significant IVOL premium prevails. In consequence, when the return variance is measured over a wider window than 1 month, the IVOL puzzle persists almost unchanged hinting at a relationship only between IVOL and return variance measured over the same window size. Nevertheless, when using the variance estimates from the complete sample horizon in the weighting matrix, the IVOL risk premium is slightly less significant at a level of 5% and decreased in magnitude to -0.549. The effect of a one standard deviation increase in IVOL, holding all other variables constant, shows that the statistically significant risk premia are also economically meaningful as they predict a decrease in expected returns ranging from 0.205% to 0.315% per month.

Also for the remaining risk premia estimates some differences relative to the reference results have to be highlighted. The $Size$ premium is positive throughout all adjustments here. But interestingly, the level, at which it is statistically significant, increases with the window size used for return variance estimation. When the current month return variances are used in the weighting matrix, the $Size$ premium is statistically insignificant, but if I use the complete sample horizon variance estimates instead, it becomes significant at the 1% level. This indicates that the same relationship between current month return variance and IVOL might also exists for current month return variance and size. A comparable pattern exists for the $Skew_{1M}$ risk premium that also turns statistically insignificant when returns are weighted by their current month variances but is significant at the 1% level in the rest of the cases. In addition, the BM premium is now always positive as in the reference EW complete regression result, albeit not statistically significant. Lastly, the average adjusted R^2 show that the model, which uses the current month return variances in the weighting matrix, has the best fit compared to the other models.

In conclusion the IVOL puzzle is robust to adjustments in the window size used for risk factor loading estimation and mainly also when the GLS estimation technique is employed for estimation of the cross-sectional regressions. Panel A illustrates that, regardless of the choice of the estimation window utilized for computation of the FF3 model risk factor loadings, researchers can only verify the existence of the IVOL puzzle. In contrast, Panel B shows that the estimation of the cross-sectional regression using GLS with a diagonal weighting matrix consisting of the inverse of the current month return variances, enables researchers to argue on the absence of the puzzle respectively. However, if the variances are estimated from rolling windows over 12, 60 months or the complete sample horizon, the IVOL puzzle prevails.

7.2. Portfolio-Related

Finally I analyze the portfolio method-specific adjustments that are introduced in section 3.4. I begin with the modification related to the way the univariate IVOL portfolios are computed and analyzed in section 7.2.1 and afterwards continue with the investigation of the bivariate portfolio sorts in section 7.2.2.

7.2.1. Portfolio Characterization

This section has the objective of finding out whether researchers are able to influence their findings by the way they compute and analyze the univariate portfolios. To do so, as explained in section 3.4, I implement further trading strategies, use different breakpoints for assigning stocks to portfolios, adjust the number of portfolios constructed and also use different benchmarks towards which I compute the time-series alphas. Table 10 depicts the corresponding result of the long-short IVOL portfolio that goes long the portfolio of stocks with the highest IVOL and short the one containing stocks with the lowest IVOL. Here Panel A shows the results of the EW and Panel B those of the VW portfolio sorts.

The average EW total long-short portfolio return of Panel A is always positive and statistically insignificant, regardless of the trading strategy that is used for portfolio formation. Except for an increase in the magnitude of the average total returns, with the highest value of 0.472% per month for the 12/0/12 strategy, the results remain robust to modifications in the trading strategy, wherefore the conclusions drawn here

Table 10: Portfolio-Based Results - Portfolio Characterization

This table presents the average long-short portfolio results from the reference univariate portfolio sorting procedure introduced in section 3.3 that was adjusted as explained in section 3.4.3. Still all stocks traded on the NYSE/AMEX/NASDAQ are covered here. The first column called "Trading strategies" builds the univariate quintile IVOL portfolios based on other trading strategies than the reference 1/0/1 strategy including the 1/1/1, 1/1/12, 1/0/12 and the 12/0/12 strategy respectively. In the column named "Breakpoints" I cover different breakpoints used for portfolio formation other than the usual NYSE/AMEX/NASDAQ ones. NYSE uses only stocks traded on the NYSE to calculate the IVOL breakpoints, whereas "EQUAL MKT" assigns IVOL ranked stocks based on breakpoints such that the stocks in each quintile portfolio make up the same share of the overall market capitalization. The third column is related to the adjustments in the number of univariate portfolios that are constructed. Instead of 5 monthly univariate portfolios I also compute 10 or 15 portfolios and calculate their time-series averages respectively. The last column named "Risk-Correction" refers to the way the risk-adjusted returns in form of the time-series alphas are computed. For this adjustment I compute the alphas for the following additional risk-correction benchmark models: the SY4, the HOU5 and the DANIEL3 model. As this modification is only extending the reference setting for additional alpha estimates, the results from the total returns as well as the α^{FF3} and the α^{FF6} remain unchanged and are therefore left blank. For each of these adjustments, except for those in the column named "Risk-Correction", I compute the average total return, α^{FF3} and α^{FF6} of the long-short portfolio that goes long portfolio 5 that consists of the stocks with the highest IVOL and short portfolio 1 which contains stocks with the lowest IVOL. The α^{FF3} and the α^{FF6} denote the monthly average of Jensen's time-series alphas relative to the FF3 and the FF6 model that were calculated by equation 1 using the monthly portfolio excess returns over the complete sample horizon. Below each of these metrics I depict the corresponding Newey and West (1987) adjusted t-statistic using 7 lags. The last row named "α^X" reports the results of the risk-correction benchmark adjustments for the time-series alphas where the respective models used can be found in the last column. In Panel A I use the equal-weighted portfolio returns for the analysis, where all firms are getting the same weight. On the other hand, Panel B uses value-weighted portfolio returns that were calculated by weighting the stock returns within the portfolio by their MTCAP observable at beginning of the month in order to give higher weights to bigger stocks respectively and therefore diminish the effects that might be explicitly related to small stocks. The complete sample period covers July 1963 to December 2020. It is shorter when the SY4 or the DANIEL3 models are used which is due a lack of data availability and therefore the corresponding analysis cover only the horizon until November 2016 and November 2018 respectively.

Panel A: Equal-Weighted Portfolio Sorts

	Trading Strategies				Breakpoints		# of Portfolios		Risk-Correction		
	1/1/1	*1/1/12*	*1/0/12*	*12/0/12*	*NYSE*	*EQUAL MKT*	*10*	*15*	*SY4*	*HOU5*	*DANIEL3*
Total Return	0.007	0.260	0.226	0.472	0.165	0.238	0.112	0.178			
	(0.025)	(0.906)	(0.784)	(1.417)	(0.685)	(1.088)	(0.338)	(0.510)			
α^{FF3}	-0.482	-0.203	-0.237	-0.028	-0.327	-0.226	-0.480	-0.464			
	(-2.751)	(-1.182)	(-1.376)	(-0.139)	(-2.456)	(-1.925)	(-2.259)	(-2.002)			
α^{FF6}	0.147	0.350	0.333	0.556	0.159	0.205	0.295	0.350			
	(0.718)	(1.838)	(1.727)	(2.412)	(1.132)	(1.610)	(1.182)	(1.329)			
α^X									0.230	0.595	0.383
									(0.896)	(1.824)	(1.410)

Panel B: Value-Weighted Portfolio Sorts

	Trading Strategies				Breakpoints		# of Portfolios		Risk-Correction		
	1/1/1	*1/1/12*	*1/0/12*	*12/0/12*	*NYSE*	*EQUAL MKT*	*10*	*15*	*SY4*	*HOU5*	*DANIEL3*
Total Return	-0.493	-0.201	-0.260	-0.287	-0.173	-0.018	-1.002	-1.216			
	(-1.801)	(-0.771)	(-0.991)	(-0.915)	(-0.778)	(-0.092)	(-3.035)	(-3.512)			
α^{FF3}	-0.945	-0.609	-0.674	-0.780	-0.562	-0.334	-1.509	-1.761			
	(-5.323)	(-4.192)	(-4.599)	(-4.257)	(-4.246)	(-2.761)	(-6.944)	(-7.876)			
α^{FF6}	-0.308	-0.089	-0.134	-0.134	-0.106	0.033	-0.700	-0.913			
	(-2.170)	(-0.783)	(-1.142)	(-0.881)	(-0.837)	(0.310)	(-2.923)	(-4.260)			
α^X									-0.314	0.033	-0.155
									(-1.509)	(0.139)	(-0.808)

match those of the EW reference results from section 5.3. The findings from the α^{FF3} and α^{FF6}, however, seem to be sensitive to the choice of the trading strategy when analyzing the equal-weighted portfolio sorts. Even though the α^{FF3} are negative for all trading strategies, they are not always statistically and economically meaningful. These long-short portfolio α^{FF3} ranges from -0.482% for the 1/1/1 strategy to -0.028% per month for the 12/0/12 strategy with corresponding Newey-West robust t-statistics of -2.751 and -0.139. The 1/1/1 results show that the incorporation of a 1 month waiting period even increases the negative long-short portfolio α^{FF3} in absolute magnitude when compared to the reference results and hence contradicts the explanation of the puzzle based on short-term reversal as proposed by W. Huang et al. (2010). Whereas for the 1/1/12 and 1/0/12 strategies, the α^{FF3} are negative but not statistically significant at any reasonable level. When portfolios are built by the 12/0/12 strategy, on the other hand, the long-short portfolio α^{FF3} is very close to zero and statistically insignificant such that no IVOL puzzle or any systematic effect at all is revealed. In contrast, the α^{FF6} predict an opposite pattern where all alphas are positive and only statistically insignificant for the 1/1/1 strategy. For the remaining strategies, the 12 month holding period implies positive long-short portfolio α^{FF6} ranging from 0.333% to 0.556% per month and that are economically meaningful and statistically significant at a level not less than 10%. Hence, when the FF6 model is utilized for risk-correction, a positive relationship between IVOL and expected risk-adjusted returns remains that is statistically significant if the holding period covers 12 months. This result contradicts the finding of an IVOL puzzle and rather matches the results of Spiegel and Wang (2005) as well as Malkiel and Xu (1997).

The EW reference results, however, appear robust to the breakpoint adjustment considered here. The long-short portfolio total return and α^{FF6} are positive on average for all breakpoints, but none of them is statistically significant at a reasonable level which is similar to the EW results of section 5.3. Also similar is that the α^{FF3} is negative and statistically, as well as economically, significant ranging from -0.327% to -0.226% per month such that the IVOL puzzle can again be detected in the EW context from the α^{FF3} only. Thus in contrast to the findings of Bali and Cakici (2008), the choice of the breakpoints has almost no effect on the EW portfolio sort results when compared to the reference setting. However, when comparing the results of both breakpoints, the α^{FF3} is lower and has a higher t-statistic for the NYSE breakpoints, whereas the total return and the α^{FF6} as well as their t-statistics are higher for the equal market share breakpoints implying that the IVOL puzzle is indeed weaker for the equal market share breakpoints as also found by Bali and Cakici (2008).

The EW long-short portfolio results hardly change, compared to the reference results, when I sort stocks each month into 10 or 15 portfolios respectively. Total returns and the α^{FF6} remain positive and statistically insignificant again and just the α^{FF3} delivers evidence on the IVOL puzzle, as it is negative and statistically significant at the 5% level for 10 and 15 portfolios. The α^{FF3} ranges from an average -0.464% for 15 to -0.480% per month for 10 monthly portfolios which are thus also economically meaningful. Nevertheless, the magnitudes of the total return, the α^{FF3} and the α^{FF6} for both numbers of portfolios is very close to the one in section 5.3, which is why choosing a different number of univariate portfolios in the research study is unlikely to affect the study's conclusions.

In the last column of Panel A I depict the alpha estimates from the EW long-short portfolio relative to further risk-correction models including the SY4, the HOU5 and the DANIEL3 model. All these alphas are found to be positive ranging from an average of 0.230% for the SY4 model alpha to 0.595% per month for the HOU5 model alpha. Even though these appear to be economically meaningful, only the one relative to the HOU5 model is also statistically significant at the 10% level. Hence, researchers that base their arguments only on the HOU5 alpha find a positive relationship between risk-adjusted expected returns and IVOL instead of an IVOL puzzle. For the alphas relative to the SY4 and the DANIEL3 model, conclusions are in line with those drawn from the α^{FF6}, implying the same positive relationship that is, however, not significant at a plausible level. Thus, the choice of the risk-correction model used for computation of the time-series alphas likely influences the conclusions drawn from the study on the existence of the IVOL puzzle.

The results from the VW long-short portfolio seem to be overall more sensitive to the adjustments in the way the portfolios are characterized. With the choice of the trading strategy researchers can influence their findings compared to the reference results of section 5.3. As the total returns, α^{FF3} and α^{FF6} have been negative and statistically significant in the VW reference setting, this result does not hold for all trading strategies. Even though the long-short portfolio total returns are all still negative, just the one of the 1/1/1 strategy is also statistically significant at the 10% level with a average monthly total return of -0.493% that is also economically meaningful. All remaining strategies imply total returns between -0.201% and -0.287% per month that are less than half the magnitude of the reference results and are furthermore even statistically insignificant. The α^{FF3}, on the other hand, remains statistically significant at the 1% level for all trading strategies, although it has reduced in magnitude to average risk-adjusted returns between -0.945% and -0.609% per month that are still economically relevant. In addition, similar to the total returns, the α^{FF6} is only statically significant for the 1/1/1 strategy with an average value of -0.308% per month that is economically relevant and statistically significant at the 5% level. However, compared to the EW portfolio sorts of Panel A, the inclusion of the 1 month holding period does not strengthen the IVOL puzzle effect but instead makes it slightly weaker as the magnitude of all measures shrinks in comparison to the reference results from Panel B of table 3. This fact, as well as the insignificant results from the remaining strategies, indicate that at least in the VW context the correction for short-term reversal helps explaining the IVOL

puzzle, as also pointed out by W. Huang et al. (2010). In conclusion, researchers that exclude the α^{FF3} from their analysis and use trading strategies other than the 1/0/1 or 1/1/1 would not be able discover an IVOL puzzle that is statistically plausible with their VW portfolio-based analysis. However, Ang et al. (2006) claim that their IVOL puzzle finding is robust to the implementation of further trading strategies. This conclusion can be explained by the fact that they focus on the examination of the α^{FF3} in their study which is the only measure that is indeed robust to such adjustments (see Ang et al. (2006)). Nevertheless, their results would likely change when they compute total returns or the α^{FF6} instead. If researchers shift from a 1/0/1 to a 1/1/1 trading strategy they are able to slightly reduce the magnitude of the IVOL puzzle effect.

As found by Bali and Cakici (2008), the IVOL puzzle is sensitive to the choice of the breakpoints used to subdivide stocks into VW portfolios. The long-short portfolio total returns remain negative for both breakpoint types but at the same time become statistically insignificant as well. In addition, the magnitude of the total returns has decreased to -0.173% for the NYSE breakpoints and to an average of only -0.018% per month for the equal market share breakpoints. Therefore, the total return results from the equal market share breakpoints do not only turn statistically insignificant but also become economically negligible. Even though the α^{FF3} remain statistically significant at the 1% level for both breakpoints, their magnitude more than halved to values around -0.562% and -0.334% per month with t-statistics that decrease to -4.246 and -2.761 when compared to the reference results. Hence, although no change in statistical relevance is observable for the α^{FF3}, the decrease in economic magnitude becomes obvious. Even more noticeable is the influence of the choice of the breakpoints when analyzing the α^{FF6}. For the NYSE breakpoints the long-short α^{FF6} is still negative but not statistically significant at any reasonable level, whereas for the equal market share breakpoints the alpha even turns positive albeit remaining statistically insignificant. Hence, researchers that would use the FF6 model for risk-correction and the equal market share breakpoints can almost reverse the IVOL puzzle effect. Conclusively, the choice of the breakpoints has a strong influence on the findings if researchers use total returns or α^{FF6} in their analysis only. The influence on the study is the strongest in terms of changes relative to the reference results when equal market share breakpoints are used for the VW portfolio analysis as also pointed out by Bali and Cakici (2008).

On the other hand, as I increase the number of univariate VW portfolios used for the analysis, the average long-short portfolio total return, α^{FF3} and α^{FF6} as well as the corresponding robust t-statistics grow in magnitude compared to the reference results. Similar to the reference setting, all these measures are negative and statistically significant at the 1% level which also encloses the total returns. In economic terms, the total returns and all the alphas are also plausible, where the total return and the α^{FF3} for the long-short portfolio using 10 and 15 portfolios is in no case higher than -1% per month. Only the long-short portfolio α^{FF6} is slightly larger and varies between an average of -0.700% in case of 10 portfolios and -0.913% per month when 15 portfolios are constructed. Nevertheless, I find that an increase in the number of monthly univariate portfolios strengthens the negative relationship between IVOL and expected cross-sectional returns. Thus in this VW setting researchers can let their results appear more statistically and economically meaningful by increasing the number of univariate IVOL sorted portfolios they compute in their study.

Lastly, the choice of the risk-correction model used for alpha computation also influences the conclusion a researcher would draw from his study. As in the reference setting, the long-short portfolio α^{FF3} and α^{FF6} have been negative and statistically significant at the 1% level which does not apply to the SY4, the HOU5 and the DANIEL3 alphas. The average long-short portfolio alpha relative to the SY4 and the DANIEL3 model are both negative with values of -0.314% and -0.155% per month respectively but both of them are low in magnitude and not statistically significant at any reasonable level. Thus in these cases an IVOL puzzle is detected in signs but cannot be approved statistically in magnitude. In contrast, the long-short portfolio alpha relative to the HOU5 model is positive but not statistically significant at any plausible level and with an average value of 0.033% per month economically negligible. Consequently, the IVOL puzzle finding is not robust to the choice of the risk-correction model used for alpha computation and researchers that use models like the SY4, the HOU5 or the DANIEL3 for that purpose in the VW portfolio sorts would likely not be able to statistically plausible verify the existence of the puzzle.

Conclusively I found out that the EW portfolio sort results are more robust to adjustments in the way the portfolios are characterized than the VW ones. Only the choice of the trading strategy has a notable influence on the results of the EW sorts. Here especially the usage of a longer holding period than 1 month can lead to the conclusion that no systematic effect in the total returns and alphas is found. Conversely, the findings from the VW portfolio sorts are effected by all the adjustments I consider for the category of portfolio characterization. As in the EW setting, a holding period longer than 1 month lets the magnitude of the IVOL puzzle appear mostly statistically insignificant when excluding the analysis of the α^{FF3}. Thus in contrast to the arguments of Ang et al. (2006), the choice of the trading strategy is not irrelevant to the IVOL puzzle finding in the VW context. As stated by Bali and Cakici (2008), the decision on the breakpoints also affects the VW portfolio results because both breakpoints analyzed here show no evidence on the existence of the IVOL puzzle when investigating the total return or α^{FF6}. On the other hand, increasing the number of monthly univariate IVOL portfolios leads to a monotonic increase in the magnitude of the IVOL puzzle in the VW context. Finally, when long-short portfolio alphas are computed towards the SY4, the HOU5 or the DANIEL3 model, no IVOL puzzle can be discovered that is also statistically significant.

7.2.2. Bivariate Portfolio Sorts

Ultimately I investigate if controlling for further firm characteristic effects by bivariate portfolio sorts enables researchers to explain the IVOL puzzle and then also try to determine the sorts they would choose for that purpose. For this I use the dependent portfolio sorting technique and the control variables introduced in section 3.4 to construct the average IVOL portfolios. I compute the total return, the α^{FF3} and the α^{FF6} for average portfolio 1, 5 and the long-short portfolio as well as the corresponding Newey-West robust t-statistics. In table 11 I depict the EW average portfolio sort results in Panel A and those from the VW analysis in Panel B respectively.

The EW bivariately sorted average portfolios from Panel A show that the finding of an IVOL puzzle is predominantly robust to the effect from most of the control variables considered here. Compared to Panel A of table 3, results remain almost unchanged when I control for $Skew_{1M}$, $Volume$ or $CoSkew$. In these cases, I find that the total returns, as well as the α^{FF6}, increases when moving from portfolio 1 to portfolio 5, such that the corresponding measures are positive for the long-short portfolio. However, these positive long-short portfolio total returns and α^{FF6} are still not statistically significant at any reasonable level. Only the corresponding long-short portfolio α^{FF3} are negative with values of -0.384% per month when controlling for skewness and -0.328% per month for the coskewness control that are both economically plausible and statistically significant at the 5% level. Controlling for volume yields a long-short α^{FF6} of -0.493% per month that is economically meaningful and even statistically significant at the 1% level. Hence, researchers that control for the volume effect by Gervais et al. (2001), the monthly return skewness or the coskewness measure of Harvey and Siddique (2000) are not able to modify their findings related to the IVOL puzzle notably and can only verify its existence from the α^{FF3} analysis.

On the other hand, controlling for variables including size, the past months return, expected idiosyncratic skewness, the Amihud (2002) illiquidity measure and the bid-ask spread in the bivariate sorts have only a limited impact on the EW results and often make the IVOL puzzle even more apparent instead of resolving it. For all the bivariate dependent sorts based on these control variables, the analysis of the total returns and the alphas shows the same pattern across average portfolios. In all these cases, the corresponding measure decreases when moving from portfolio 1 to portfolio 5 and thus results in a negative long-short portfolio value, which is consistent with the IVOL puzzle finding. However, only the negative long-short α^{FF3} are also statistically significant even at a 1% level for all controls, whereas the related total returns and α^{FF6} are not significant at any meaningful level. In consequence, even though the IVOL puzzle appears throughout all measures when controlling for the variables mentioned above, the effect is only statistically significant for the α^{FF3}. This pattern found in the α^{FF3} is always driven by a statistically significant underperformance of portfolio 5 relative to portfolio 1 which is similar to the findings of the reference analysis. It is on average the lowest in economic terms when controlling for the past month return with an average long-short portfolio α^{FF3} of -0.477% per month, while it is the highest in the bivariate sorts involving the Amihud (2002) illiquidity measure with an average subsequent month α^{FF3} of -0.832% per month. These α^{FF3} all exceed the corresponding alpha from the reference analysis and are economically meaningful. From this fact, as well as the one that long-short portfolio total returns and α^{FF6} are negative even though not statistically significant for the controls including $Size$, R_t, $EIdioSkew$, $Illiq_{1M}$ and $Spread$, I conclude that researchers using these controls in their EW bivariate dependent portfolio analysis can let the IVOL puzzle appear more plausible than without their consideration.

The only variables that indeed seem to resolve the IVOL puzzle when controlled for in the researchers' EW bivariate portfolio analysis are BM and $Maxret$. Starting with the sorts that control for BM, the total returns show an monotonic increase when moving from portfolio 1 to 5, which is why the long-short portfolio total return is also positive again, even though it is still not significant at any plausible level. Although the negative relationship between IVOL and α^{FF3} persists on average, the long-short portfolio α^{FF3} of -0.229% per month is now no longer statistically significant at any meaningful level as the t-statistic has decreased to -1.229. The typical pattern of an IVOL puzzle even reversed when analyzing the α^{FF6}. Here the α^{FF6} increases monotonically from an average -0.080% for portfolio 1 to 0.356% per month for portfolio 5 resulting in a long-short portfolio α^{FF6} of 0.436% per month that is not only statistically meaningful at the 5% level but is also economically relevant. From this I conclude that the value effect might indeed affect the IVOL puzzle which is also argued by Barinov and Chabakauri (2020) and researchers who are aware of this fact can influence their results by including it in their study. It is of note that the EW regression results from Panel A of table 2 were not able to uncover this relationship which might hint at a non-linear dependence between the value effect and the IVOL puzzle respectively. Controlling for $Maxret$, on the other hand, overturns the IVOL puzzle almost completely, which proves the findings of Bali et al. (2011). An investigation of the EW average portfolios from the dependent bivariate sorts involving $Maxret$ shows that the total returns and alphas monotonically increase when shifting from portfolio 1 to portfolio 5, such that the corresponding long-short portfolio results are also positive for all the measures. The long-short portfolio total return and the α^{FF6} amount on average to 0.408% and 0.481% per month which are not only economically meaningful but also statistically significant at the 1% level and consequently imply a positive relationship between IVOL and expected returns instead of an IVOL puzzle. Whereas for the total returns from portfolio 1 and 5 both are statistically significant at the 1% level, the pattern in the α^{FF6} seems to be driven by the overperformance of the high IVOL portfolio which has a positive α^{FF6} that is statistically significant at the 5% level. Despite the fact that the long-

Table 11: Portfolio-Based Results - Bivariate Portfolio Sorts

This table presents the average results of quintile portfolio 1, 5 and the long-short portfolio from the bivariate portfolio sorting procedure as explained in section 3.4.3 covering all stocks traded on the NYSE/AMEX/NASDAQ. First all stocks are sorted each month into quintile portfolios with respect to their level of the corresponding control variable and afterwards within each of these quintile portfolios the stocks are again sorted into quintile portfolios based on their idiosyncratic volatility computed relative to the FF3 model using daily data over the past month. Afterwards each IVOL quintile portfolios, as well as the long-short portfolio, is averaged over all the quintiles of the control variables to obtain the univariate IVOL quintile portfolios that have been corrected for the effect of the corresponding control variable. For all these bivariately sorted average portfolios I compute the average monthly total returns, the α^{FF3} and the α^{FF6} of portfolios 1, 5 as well as the long-short portfolio (in percentage terms) and in parenthesis below each of these measures I show the corresponding Newey and West (1987) adjusted t-statistics where I use 7 lags for the adjustment. The α^{FF3} and α^{FF6} denote the monthly average of Jensen's time-series alphas relative to the FF3 and the FF6 model again that were calculated by equation 1 using the monthly portfolio excess returns from the respective bivariate portfolios over the complete sample horizon. Each row corresponds to the results of a bivariate portfolio sort where the effect that was controlled for is used as row name. Here I control for the following variables: size that is the natural logarithm of the stock's MTCAP, the firm's BM, the stock's current month return, the Amihud (2002) illiquidity measure calculated over the current month, return skewness over the current month, the expected idiosyncratic skewness measure of Boyer et al. (2010), the coskewness measure of Harvey and Siddique (2000) computed over the recent past 5 years computed, the trading volume of the stock, the monthly average over the daily bid-ask-spreads and the maximum daily return during the current month computed based on the ideas of Bali et al. (2011). These portfolios are rebalanced monthly according to the 1/0/1 trading strategy. In Panel A I use the equal-weighted portfolio returns for the analysis, where all firms are getting the same weight. On the other hand, Panel B uses value-weighted portfolio returns that were calculated by weighting the stock returns within the portfolio by their MTCAP observable at beginning of the month in order to give higher weights to bigger stocks respectively and therefore diminish the effects that might be explicitly related to small stocks. The complete sample period covers July 1963 to December 2020.

Panel A: Equal-Weighted Portfolio Sorts

	Total Return			α^{FF3}			α^{FF6}		
	1	5	5-1	1	5	5-1	1	5	5-1
Size	0.939	0.589	-0.349	0.085	-0.728	-0.813	0.021	-0.213	-0.235
	(3.947)	(1.611)	(-1.423)	(0.412)	(-4.388)	(-5.225)	(0.087)	(-1.074)	(-1.559)
BM	0.932	1.109	0.177	0.003	-0.226	-0.229	-0.080	0.356	0.436
	(4.231)	(2.817)	(0.659)	(0.021)	(-1.105)	(-1.229)	(-0.713)	(1.417)	(2.044)
R_t	0.914	0.830	-0.085	-0.005	-0.482	-0.477	-0.021	-0.051	-0.030
	(4.241)	(2.074)	(-0.328)	(-0.035)	(-2.559)	(-3.199)	(-0.131)	(-0.240)	(-0.198)
$Skew_{1M}$	0.870	0.973	0.103	0.012	-0.372	-0.384	-0.080	0.204	0.283
	(3.920)	(2.460)	(0.363)	(0.059)	(-1.945)	(-2.225)	(-0.355)	(0.841)	(1.388)
EIdioSkew	1.091	0.847	-0.245	0.121	-0.545	-0.666	0.058	-0.028	-0.086
	(4.105)	(2.032)	(-0.937)	(0.768)	(-2.364)	(-3.929)	(0.433)	(-0.086)	(-0.470)
CoSkew	0.955	1.069	0.115	0.053	-0.275	-0.328	-0.016	0.222	0.238
	(4.347)	(2.830)	(0.453)	(0.339)	(-1.469)	(-2.056)	(-0.104)	(0.950)	(1.258)
$Illiq_{1M}$	0.960	0.609	-0.351	0.072	-0.759	-0.832	0.003	-0.171	-0.174
	(4.208)	(1.641)	(-1.350)	(0.371)	(-4.427)	(-5.085)	(0.012)	(-0.804)	(-0.993)
Volume	0.877	0.888	0.011	0.000	-0.493	-0.493	-0.086	0.104	0.190
	(4.104)	(2.234)	(0.038)	(0.001)	(-2.431)	(-2.828)	(-0.464)	(0.411)	(0.945)
Spread	0.889	0.773	-0.115	0.023	-0.540	-0.563	-0.030	-0.120	-0.090
	(3.999)	(2.062)	(-0.456)	(0.135)	(-3.027)	(-3.733)	(-0.170)	(-0.580)	(-0.568)
Maxret	0.913	1.322	0.408	-0.101	0.074	0.175	-0.065	0.416	0.481
	(3.643)	(3.917)	(2.759)	(-0.729)	(0.455)	(1.650)	(-0.335)	(2.164)	(3.848)

Panel B: Value-Weighted Portfolio Sorts

	Total Return			α^{FF3}			α^{FF6}		
	1	5	5-1	1	5	5-1	1	5	5-1
Size	0.802	0.565	-0.237	-0.011	-0.552	-0.540	-0.143	-0.215	-0.072
	(4.167)	(1.848)	(-1.113)	(-0.096)	(-3.963)	(-3.873)	(-1.567)	(-1.359)	(-0.604)
BM	0.794	0.261	-0.533	-0.027	-0.989	-0.962	-0.140	-0.461	-0.321
	(4.067)	(0.744)	(-2.118)	(-0.350)	(-5.308)	(-6.425)	(-2.285)	(-1.674)	(-2.275)
R_t	0.769	0.235	-0.533	-0.092	-1.005	-0.913	-0.143	-0.625	-0.482
	(4.060)	(0.681)	(-2.223)	(-1.203)	(-5.767)	(-6.542)	(-2.282)	(-3.415)	(-3.877)
$Skew_{1M}$	0.787	0.120	-0.667	-0.048	-1.132	-1.085	-0.166	-0.601	-0.435
	(4.058)	(0.331)	(-2.393)	(-0.423)	(-5.940)	(-6.104)	(-2.245)	(-2.428)	(-2.659)
EIdioSkew	0.896	0.453	-0.443	-0.036	-0.872	-0.836	-0.196	-0.480	-0.284
	(3.839)	(1.214)	(-1.855)	(-0.200)	(-2.880)	(-4.981)	(-1.817)	(-0.712)	(-1.919)
CoSkew	0.792	0.318	-0.474	-0.050	-0.913	-0.863	-0.126	-0.546	-0.420
	(4.046)	(0.975)	(-2.157)	(-0.114)	(-5.279)	(-6.209)	(-0.315)	(-2.338)	(-2.995)
$Illiq_{1M}$	0.800	0.530	-0.269	-0.015	-0.619	-0.604	-0.149	-0.268	-0.118
	(4.165)	(1.684)	(-1.221)	(-0.132)	(-4.183)	(-4.193)	(-1.709)	(-1.540)	(-0.944)
Volume	0.785	0.146	-0.639	-0.047	-1.108	-1.061	-0.165	-0.545	-0.380
	(4.081)	(0.404)	(-2.395)	(-0.462)	(-5.770)	(-6.429)	(-2.079)	(-2.105)	(-2.458)
Spread	0.797	0.232	-0.565	-0.028	-0.989	-0.961	-0.040	-0.722	-0.681
	(4.076)	(0.702)	(-2.416)	(-0.282)	(-5.835)	(-6.599)	(-0.467)	(-3.383)	(-4.400)
Maxret	0.756	0.764	0.007	-0.152	-0.280	-0.129	-0.115	-0.293	-0.178
	(3.397)	(2.991)	(0.065)	(-0.991)	(-1.260)	(-1.411)	(-0.672)	(-0.806)	(-1.910)

short portfolio α^{FF3} is also positive with a value of 0.175% per month, it has to be noted that it is rather low in economic terms and only statistically significant at the 10% level. Again the relationship between IVOL and the maximum return is most likely non-linear as the IVOL risk premium from the EW regressions is almost unaffected by the inclusion of $Maxret$ as a further control variable as found in Panel A of table 8.

Panel B displays the results from the VW average dependent sorted portfolios for the same controls as considered in Panel A. When computing the VW average portfolios from bivariate dependent sorts that involve the BM, the past month return, the monthly return skewness, the coskewness measure of Harvey and Siddique (2000), trading volume or the bid-ask spread as control variables, the results from total returns and alphas remain almost unchanged compared to the reference results form Panel B of table 3. The analysis of the long-short portfolio performance in all these sorts shows that portfolio 5 performs worse than portfolio 1, as the long-short portfolio total returns, as well as the alphas relative to both risk-correction models, are negative and statistically significant. The long-short portfolio returns are the lowest in case of the α^{FF3} when controlling for $Skew_{1M}$ with a value of -1.085% per month the highest when controlling for BM with an average α^{FF6} of -0.321% per month. Therefore, the evidence on the IVOL puzzle prevails in all these sorts respectively and is still economically meaningful. It is of note, however, that for all average portfolios considered here the measures are lower in magnitude than compared to their level from the reference analysis of section 5.3. The only exception is the long-short α^{FF6} when controlling for R_t and $Spread$ that are now bigger in absolute terms. Interesting is also that the bivariate dependent sort based on BM was able to resolve the IVOL puzzle in the EW context but is no longer able to do so here. An explanation of that might be that the return weighing by MTCAP resolves the value effect, if the assumption holds that the small firms also have a high BM. In contrast, controlling for $Size$, $EIdioSkew$ or $Illiq_{1M}$ has a limited, albeit notable, effect on the results. All bivariate sorts related to these control variables show the same pattern consisting of a negative long-short portfolio return, regardless of the return type or risk-correction model considered. Just the long-short portfolio α^{FF3} is also always statistically significant. While the long-short portfolio total return and α^{FF6} of these sorts is mostly statistically insignificant at any reasonable level, the α^{FF3} remains statistically significant at the 1% level. Only the bivariate dependent sorts on $EIdioSkew$ also have long-short portfolio total returns and α^{FF6} that are both statistically significant but just at the 10% level. However, compared to the reference result, the economic magnitude of the long-short α^{FF3} decreases to values ranging from -0.836% per month for the sort that controls for $EIdioSkew$ to -0.540% per month when controlling for $Size$. Nevertheless, these values are still economically meaningful and again driven by the low performance of portfolio 5 as the corresponding α^{FF3} is the only one that is statistically significant at a level not less than 1%. Additionally, it is of note that expected idiosyncratic skewness only marginally helps to explain the IVOL puzzle and does not resolve it, as argued by Boyer et al. (2010). In conclusion, researchers who use $Size$, $EIdioSkew$ or $Illiq_{1M}$ as controls in their VW portfolio-based study, are not able to resolve the IVOL puzzle but can achieve a reduction in its magnitude and statistical relevance.

The only variable that is indeed able to resolve the IVOL puzzle almost completely is the maximum daily return measure over the past month. Controlling for $Maxret$ leads to an increasing total return pattern across IVOL portfolios when moving from portfolio 1 to 5 and causes the long-short portfolio total return to be positive with an average value of 0.007% per month, which is statistically insignificant and economically negligible. Despite the fact that the alphas are still decreasing when moving from IVOL quintile portfolio 1 to 5 also yielding negative long-short portfolio alphas, they are only statistically significant at the 10% level in terms of the α^{FF6}. The long-short portfolio alpha amounts to an average value of -0.129% per month for the α^{FF3} and to -0.178% per month for the α^{FF6} that are both economically small. In contrast to the reference results, the alphas from portfolio 1 and 5 are both not statistically significant at any plausible level regardless of the model chosen for risk-correction which is thus another evidence on the explanatory power of $Maxret$. After all, similar to the EW sorts of Panel A, if researchers would use $Maxret$ as a control variable, they are be able to change their findings such that they conclude that the IVOL puzzle found by Ang et al. (2006) is rather implausible.

To sum up the results, it can be said that bivariate depended sorts hardly help researchers to explain the IVOL puzzle. Overall, as also found by Bali et al. (2011), the control associated with the maximum daily return measure is the only variable that could be used by researchers who try to argue against the existence of the IVOL puzzle, regardless of the weighting scheme they would use for the portfolio-based analysis. Researchers that compute EW portfolio returns only could additionally use the book-to-market ratio of equity as a control variable to offset the IVOL puzzle effect. All remaining control variables are not at all or only partly able to explain the IVOL puzzle where the respective conclusion strongly depends on the return type computed and presented in the corresponding analysis. Here it is mostly the α^{FF3} that discovers the IVOL puzzle irrespective of the control or weighting scheme used.

8. Conclusion

With this study I try to shed light on the dispute over the existence of the IVOL puzzle originally found by Ang et al. (2006). I take on an econometric perspective and analyze in how far researchers can draw different conclusions on the existence of the IVOL puzzle only by adjusting their researcher design in a specific manner. As this anomalous finding is most often analyzed by either a regression-based approach that follows the ideas of Fama and MacBeth (1973) or a portfolio-based methodology which involves sorting stocks into portfolios based on their idiosyncratic volatility, I use these two

concepts alongside which I analyze how findings evolve when the research design is modified. As an anchor point I have used the study of Ang et al. (2006) for the portfolio-based and the one of Ang et al. (2009) for the regression-based method as these were among the first to find an IVOL puzzle while using the corresponding research concepts. For the regression-based approach I replicated the IVOL puzzle finding by means of equal-weighted as well as value-weighted Fama-MacBeth regressions. Here both weighting-schemes showed a statistically significant negative IVOL risk premium that is robust when controlling for the current month return skewness, the expected idiosyncratic skewness measure by Boyer et al. (2010), the coskewness measure of Harvey and Siddique (2000) and the Amihud (2002) illiquidity measure. The replication analysis of the portfolio-based approach already showed that the IVOL puzzle finding is more likely to occur when researchers use the value-weighted instead of the equal-weighted portfolio returns in their study. While all measures including excess and total returns as well as risk-adjusted returns relative to the FF3 and the FF6 model prove the existence of the IVOL puzzle in a value-weighted portfolio context, the equal-weighted portfolio returns are solely able to deliver this finding when researchers compute the time-series alphas relative to the FF3 model only. Based on these results I have started the adjustment analysis with the aim to find possible modifications that change the researchers conclusions on the existence of the IVOL puzzle. As general adjustments, that I applied to both research concepts, I have started by modifying the way IVOL is estimated and afterwards adjusted the sample used in the respective analysis. Changing the risk-correction model relative to which IVOL is estimated hardly changes the results researchers would obtain from their study regardless of the research concepts employed. However, those researchers that use the portfolio-based method or the value-weighted Fama-MacBeth regression are able to let the finding of an IVOL puzzle appear statistically insignificant only by switching the data frequency from daily to monthly and using a longer estimation window covering 1 or 5 years when estimating IVOL. The investigation of subsamples constructed based on the stock's price or market capitalization is also shown to influence research findings. Results appear to be particularly sensitive to the sample used when they are obtained from the VW Fama-MacBeth regressions or the VW portfolio sorts, as in these cases the IVOL puzzle effect seems to gradually weaken when stocks with a higher price or bigger market capitalization are used. After implementation of the general adjustments, I start with the analysis of adjustments that focus on the methodological peculiarities of each research concept and reveal how such modifications influence the research findings accordingly. For the regression-based approach I have focused on adjustments in the risk-related control variables as well as the estimation procedure. Results from the inclusion of different systematic risk-correction models or additional firm-specific risk variables appear to be almost similar to the reference results, which is why researchers are typically not able to modify their results by conducting such modifications. Only in the VW Fama-MacBeth regression the usage of a risk-correction model consisting of the FF3 model augmented with a short and long-term reversal factor allows the researchers to slightly mitigate the statistical significance of the puzzle. On the other hand, researchers that modify the estimation procedure of the regression-based approach have only limited ability to change their results. While using different past estimation windows for the computation of risk factor loadings used in the cross-sectional regression does not influence the findings in either the equal- nor the value-weighed context, changing the estimation technique of the cross-sectional regression step by implementation of the GLS technique with a diagonal weighting matrix consisting of the inverse of the stock variances estimated from rolling windows over the current month allows researchers to reverse the IVOL puzzle and obtain a positive risk premium instead that is, however, not statistically significant. Nevertheless, using other windows for the estimation of the stock variance used in the weighting matrix does not change results notably. Then shifting to the portfolio-related adjustments, I modify the way portfolios can be characterized and lastly also use bivariate dependent portfolio sorts to disentangle the effect of the IVOL puzzle that might be due to other firm-related effects. The portfolio characterization adjustments cover modifications in the trading strategy used for portfolio formation, the breakpoints used to subdivide stocks into portfolios, the number of portfolios that are computed as well as further risk-correction models used for alpha computation. Except for the trading strategies that involve a longer holding period than 1 month, researchers are not able to change their results when using EW portfolio sorts. However, the VW portfolio sorts are more sensitive to similar adjustments. Researchers that use trading strategies with longer holding periods than 1 month or breakpoints based on the NYSE stocks or equal market shares are only able to find an IVOL puzzle when they analyze the FF3 model alphas. No IVOL puzzle is found at all, however, by researchers that would compute alphas relative to the SY4, HOU5 or the DANIEL3 model for their VW portfolio sorts. Conversely, an increase in the number of portfolios lets the puzzle appear more present in the VW portfolio analysis. Lastly, I discovered that when researchers control for the maximum daily return over the past month by means of bivariate dependent portfolios sorts, they can conclude that the IVOL puzzle is not statistically relevant or existent at all, regardless of the weighting scheme they use in their study. In addition, those researchers that focus on EW portfolios can also resolve the puzzle when they control for the value effect in their study captured by the BM. All in all, I need to remark that my analysis was only able to investigate a fraction of adjustments that could possibly be employed by researchers to bias their results towards a specific direction. My aim is to draw attention to the fact that the configuration of the researcher design can be important to understand results that contradict theoretically sophisticated predictions and hence trigger a debate in the literature. Furthermore, this study should help to understand why researchers are then also able to provide empirical evidence on contradictory results, even

though they appear to use the same data and methodology.

References

Amihud, Y. (2002, January). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56. Retrieved 2021-07-20, from https://linkinghub.elsevier.com/retrieve/pii/S1386418101000246

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The Cross-Section of Volatility and Expected Returns. *The Journal of Finance*, *61*(1), 259–299.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2009). High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics*, *91*(1), 1–23.

Bali, T. G., & Cakici, N. (2008, March). Idiosyncratic Volatility and the Cross Section of Expected Returns. *Journal of Financial and Quantitative Analysis*, *43*(1), 29–58. Retrieved 2020-12-28, from https://www.cambridge.org/core/product/identifier/S002210900000274X/type/journal_article

Bali, T. G., Cakici, N., & Whitelaw, R. F. (2011, February). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, *99*(2), 427–446. Retrieved 2021-05-19, from https://linkinghub.elsevier.com/retrieve/pii/S0304405X1000190X

Bali, T. G., Del Viva, L., Lambertides, N., & Trigeorgis, L. (2020, November). Growth Options and Related Stock Market Anomalies: Profitability, Distress, Lotteryness, and Volatility. *Journal of Financial and Quantitative Analysis*, *55*(7), 2150–2180. Retrieved 2020-12-21, from https://www.cambridge.org/core/product/identifier/S0022109019000619/type/journal_article

Barberis, N., & Huang, M. (2008, November). Stocks as Lotteries: The Implications of Probability Weighting for Security Prices. *American Economic Review*, *98*(5), 2066–2100. Retrieved 2021-05-10, from https://pubs.aeaweb.org/doi/10.1257/aer.98.5.2066

Barinov, A., & Chabakauri, G. (2020). Idiosyncratic Volatility, Growth Options, and the Cross-Section of Returns. *Working Paper, University of California Riverside*. Retrieved 2020-12-21, from http://www.ssrn.com/abstract=1028869

Black, F. (1972). Capital Market Equilibrium with Restricted Borrowing. *The Journal of Business*, *45*(3), 444–455.

Boehme, R. D., Danielsen, B. R., Kumar, P., & Sorescu, S. M. (2009, August). Idiosyncratic risk and the cross-section of stock returns: Merton (1987) meets Miller (1977). *Journal of Financial Markets*, *12*(3), 438–468. Retrieved 2021-08-06, from https://linkinghub.elsevier.com/retrieve/pii/S1386418109000147

Boyer, B. (2021). *Expected Idiosyncratic Skewness Data.* http://boyer.byu.edu/Research/skewdata2.html. (Accessed: 12.04.2021)

Boyer, B., Mitton, T., & Vorkink, K. (2010, January). Expected Idiosyncratic Skewness. *Review of Financial Studies*, *23*(1), 169–202. Retrieved 2020-12-21, from https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhp041

Cao, J., Chordia, T., & Zhan, X. (2021). The Calendar Effects of the Idiosyncratic Volatility Puzzle: A Tale of Two Days? *Management Science accepted*. Retrieved 2021-07-16, from http://pubsonline.informs.org/doi/10.1287/mnsc.2020.3803

Carhart, M. M. (1997, March). On Persistence in Mutual Fund Performance. *The Journal of Finance*, *52*(1), 57–82. Retrieved 2021-07-15, from https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1997.tb03808.x

Chen, L. H., Jiang, G. J., Xu, D. D., & Yao, T. (2012). Dissecting the Idiosyncratic Volatility Anomaly. *Working Paper, University of Massachusetts*.

Chen, Z., & Petkova, R. (2012, September). Does Idiosyncratic Volatility Proxy for Risk Exposure? *Review of Financial Studies*, *25*(9), 2745–2787. Retrieved 2020-12-25, from https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhs084

Daniel, K. (2021). *Kent Daniel - Data/Appendices.* http://www.kentdaniel.net/data.php. (Accessed: 12.04.2021)

Daniel, K., Hirshleifer, D., & Sun, L. (2020, April). Short- and Long-Horizon Behavioral Factors. *The Review of Financial Studies*, *33*(4), 1673–1736. Retrieved 2021-01-10, from https://academic.oup.com/rfs/article/33/4/1673/5522378

Diether, K. B., Malloy, C. J., & Scherbina, A. (2002, October). Differences of Opinion and the Cross Section of Stock Returns. *The Journal of Finance*, *57*(5), 2113–2141. Retrieved 2021-07-14, from https://onlinelibrary.wiley.com/doi/10.1111/0022-1082.00490

Duarte, J., Kamara, A., Siegel, S., & Sun, C. (2014). The Systematic Risk of Idiosyncratic Volatility. *Working Paper, University of Washington*.

Fama, E. F., & French, K. R. (1992, June). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, *47*(2), 427–465. Retrieved 2021-07-16, from https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F., & French, K. R. (1993, February). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. Retrieved 2021-01-07, from https://linkinghub.elsevier.com/retrieve/pii/0304405X93900235

Fama, E. F., & French, K. R. (2008, August). Dissecting Anomalies. *The Journal of Finance*, *63*(4), 1653–1678. Retrieved 2021-07-20, from https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.2008.01371.x

Fama, E. F., & French, K. R. (2015, April). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. Retrieved 2021-01-09, from https://linkinghub.elsevier.com/retrieve/pii/S0304405X14002323

Fama, E. F., & French, K. R. (2018, May). Choosing factors. *Journal of Financial Economics*, *128*(2), 234–252. Retrieved 2021-01-09, from https://linkinghub.elsevier.com/retrieve/pii/S0304405X18300515

Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *The Journal of Political Economy*, *81*(3), 607–636. Retrieved from http://www.jstor.org/stable/1831028

French, K. R. (2021). *Kenneth R. French - Data Library.* https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. (Accessed: 12.04.2021)

Fu, F. (2009, January). Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics*, *91*(1), 24–37. Retrieved 2021-01-08, from https://linkinghub.elsevier.com/retrieve/pii/S0304405X08001694

Gervais, S., Kaniel, R., & Mingelgrin, D. H. (2001, June). The High-Volume Return Premium. *The Journal of Finance*, *56*(3), 877–919. Retrieved 2021-07-21, from http://doi.wiley.com/10.1111/0022-1082.00349

Han, Y., & Lesmond, D. (2011, May). Liquidity Biases and the Pricing of Cross-sectional Idiosyncratic Volatility. *Review of Financial Studies*, *24*(5), 1590–1629. Retrieved 2020-12-26, from https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhq140

Harvey, C. R., & Siddique, A. (2000, June). Conditional Skewness in Asset Pricing Tests. *The Journal of Finance*, *55*(3), 1263–1295. Retrieved 2021-05-10, from http://doi.wiley.com/10.1111/0022-1082.00247

Hollstein, F., & Prokopczuk, M. (2020). Testing Factor Models in the Cross-Section. *Working Paper, Leibniz Universität Hannover*.

Hou, K., & Loh, R. K. (2016, July). Have we solved the idiosyncratic volatility puzzle? *Journal of Financial Economics*, *121*(1), 167–194. Retrieved 2020-12-20, from https://linkinghub.elsevier.com/retrieve/pii/S0304405X16300137

Hou, K., Mo, H., Xue, C., & Zhang, L. (2020, February). An Augmented q-Factor Model with Expected Growth. *Review of Finance*, *25*(1), 1–41. Retrieved 2021-07-20, from https://academic.oup.com/rof/article/25/1/1/5727769

Hou, K., Xue, C., & Zhang, L. (2015, March). Digesting Anomalies: An Investment Approach. *Review of Financial Studies*, *28*(3), 650–705. Retrieved 2020-12-28, from https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhu068

Hou, K., Xue, C., & Zhang, L. (2021). *global-q Home Page.* http://global-q.org/index.html. (Accessed: 12.04.2021)

Huang, A. G. (2009, June). The cross section of cashflow volatility and expected stock returns. *Journal of Empirical Finance*, *16*(3), 409–429. Retrieved 2020-12-28, from https://linkinghub.elsevier.com/retrieve/pii/S0927539809000036

Huang, W., Liu, Q., Rhee, S. G., & Zhang, L. (2010, January). Return Reversals, Idiosyncratic Risk, and Expected Returns. *Review of Financial Studies*, *23*(1), 147–168. Retrieved 2020-12-20, from https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhp015

Jegadeesh, N. (1990, July). Evidence of Predictable Behavior of Security Returns. *The Journal of Finance*, *45*(3), 881–898. Retrieved 2021-05-10, from http://doi.wiley.com/10.1111/j.1540-6261.1990.tb05110.x

Jegadeesh, N., & Titman, S. (1993, March). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, *48*(1), 65–91. Retrieved 2021-05-08, from http://doi.wiley.com/10.1111/j.1540-6261.1993.tb04702.x

Lehmann, B. N. (1990, February). Fads, Martingales, and Market Efficiency. *The Quarterly Journal of Economics*, *105*(1), 1–28. Retrieved 2021-05-10, from https://academic.oup.com/qje/article-lookup/doi/10.2307/2937816

Lintner, J. (1965). Security Prices, Risk, and Maximal Gains From Diversification. *The Journal of Finance*, *20*(4), 587–615.

Malkiel, B. G., & Xu, Y. (1997, April). Risk and Return Revisited. *The Journal of Portfolio Management*, *23*(3), 9–14. Retrieved 2021-07-14, from http://jpm.pm-research.com/lookup/doi/10.3905/jpm.1997.409608

Malkiel, B. G., & Xu, Y. (2006). Idiosyncratic Risk and Security Returns. *Working Paper, University of Texas at Dallas*. Retrieved 2020-12-28, from http://www.ssrn.com/abstract=255303

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91.

Merton, R. C. (1987, July). A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, *42*(3), 483–510.

Miller, E. (1977). Risk, Uncertainty, and Divergence of Opinion. *The Journal of Finance*, *32*(4), 1151–1168.

Newey, W. K., & West, K. D. (1987, May). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, *55*(3), 703–708. Retrieved 2021-07-16, from https://www.jstor.org/stable/1913610?origin=crossref

Rachwalski, M., & Wen, Q. (2016, December). Idiosyncratic Risk Innovations and the Idiosyncratic Risk-Return Relation. *Review of Asset Pricing Studies*, *6*(2), 303–328. Retrieved 2020-12-28, from https://academic.oup.com/raps/article-lookup/doi/10.1093/rapstu/raw002

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, *19*(3), 425–442.

Shumway, T. (1997, March). The Delisting Bias in CRSP Data. *The Journal of Finance*, *52*(1), 327–340. Retrieved 2021-05-17, from http://doi.wiley.com/10.1111/j.1540-6261.1997.tb03818.x

Spiegel, M., & Wang, X. (2005). Cross-sectional Variation in Stock Returns: Liquidity and Idiosyncratic Risk. *Working Paper, Yale University*.

Stambaugh, R. F. (2021). *Robert Stambaugh's Home Page.* http://finance.wharton.upenn.edu/~stambaug/. (Accessed: 12.04.2021)

Stambaugh, R. F., Yu, J., & Yuan, Y. (2015, October). Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle: Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle. *The Journal of Finance*, *70*(5), 1903–1948. Retrieved 2020-12-29, from http://doi.wiley.com/10.1111/jofi.12286

Stambaugh, R. F., & Yuan, Y. (2017, April). Mispricing Factors. *The Review of Financial Studies*, *30*(4), 1270–1315. Retrieved 2020-12-28, from https://academic.oup.com/rfs/article/30/4/1270/2965095

Tinic, S. M., & West, R. R. (1986, February). Risk, Return, and Equilibrium: A Revisit. *Journal of Political Economy*, *94*(1), 126–147. Retrieved 2021-07-14, from https://www.journals.uchicago.edu/doi/10.1086/261365

Junior Management Science 7(4) (2022) 986-1000

Junior Management Science

journal homepage: www.jums.academy

One User – Two Viewpoints? An Examination of Information Privacy Concerns from the Employee and Consumer Perspective

Marco Wall

Ludwig-Maximilians-Universität München

Abstract

This study aims to answer two questions about the investigation of privacy concerns in the use of video call applications: Firstly, it aims to find out how privacy concerns of users of these technologies differ in the work context from the private context. Secondly, this paper wants to discover the underlying cause for these differing concerns. To answer these questions, focus group interviews were conducted with users who use video call applications in both the business and personal contexts. The results of the qualitative research were presented visually using a thematic map. Users expressed great privacy concerns regarding the control of sensitive data. In addition, work behaviour and employment relationships are becoming more transparent, which raises more concerns. Employees in particular try to protect private, confidential data at work. This paper presents one of the first exploratory findings in the field of privacy research in the workplace.

Keywords: Information privacy; privacy concerns; online privacy; video call applications; workplace privacy.

1. Motivation

The rapid speed of technological development and competition among technology providers has facilitated the adoption of information technology (IT) by individuals and organizations (Moshki & Barki, 2014, p. 2). This digital transformation is leading to a growing number of user data being automatically generated, collected, and evaluated. Due to the increasing transparency of data, the issue of privacy is becoming increasingly important. Recent public opinion polls revealed that "*72 percent*" of the consumers "*are concerned that their online behaviors were being tracked and profiled by companies*" (Consumers-Union, 2008; Smith, Dinev, & Xu, 2011, p. 990). Most of the personal information is being used (by large companies like Google, Microsoft etc.) for tailored advertising, which is also shared with hundreds of affiliated companies. In addition, there is a risk of personal data being lost or stolen, which makes the issue of privacy even more important (e.g. Gomez, Pinnick, & Soltani, 2009, p. 14; Smith et al., 2011, p. 990). According to the Pew Research Center, 74% of participants say it is "*very important*" to be in control of their personal data (Epic, 2020).

Since the concept of information privacy is a latent concept and therefore cannot be measured directly, Information System (IS) scholars use privacy concerns as a proxy for assessing privacy in empirical research (Smith et al., 2011, p. 997). Despite the progress made by previous IS research, the nature of IT-related privacy concerns varies between several studies. Some studies emphasize context-specific privacy concerns (e.g. Jiang, Heng, & Choi, 2013; Xu, Dinev, Smith, & Hart, 2011; Xu, Teo, Tan, & Agarwal, 2012), while other researchers highlight general constructs of privacy concerns related to IT (Hong & Thong, 2013). Previous research has examined privacy concerns related to IT mainly focusing the consumer perspective, only little is known about privacy concerns in the workplace (e.g. Becker, 2018; Connolly & McParland, 2012; Moshki & Barki, 2014). The issue of privacy concerns from an employee perspective has been particularly raised by home office and remote work regulations. Following the outbreak of the Covid-19 pandemic (and government regulations to ensure social distancing) video call applications like Zoom, Skype, Microsoft Teams etc. have become popular solutions to replace face-to-face meetings with virtual meetings (e.g. Kagan, Alpert, & Fire, 2020; Wiederhold, 2020). Nevertheless, there are already some articles which criticize especially the video call application "Zoom" mostly because of its lack of security standards. Specific privacy policies of the application allowed Zoom to gain access to personal data of the users (e.g. Murphy 2020; Wagenseil 2020). In addition, hackers were able to gain access to webcam content and engage in video conference meetings

DOI: https://doi.org/10.5282/jums/v7i4pp986-1000

(Chawla, 2020, p. 2). This current practical example thus also demonstrates the increasing importance of privacy.

This bachelor thesis is therefore going to draw a comparison between privacy concerns of consumers and employees with the use of digital technologies (especially for communication). Due to the lack of literature regarding workplace privacy, the paper will focus more on the concerns in the working environment.

Hence, this paper aims to address the following research questions:

- *RQ1: What are employee privacy concerns and how do they differ from consumer privacy concerns?*
- *RQ2: Why do privacy concerns between both perspectives differ?*

The objective is to obtain relevant findings for research and business in this area, which make the current situation easier to understand and therefore also important for the future of such technologies.

Firstly, important terms such as privacy and privacy concerns are explained, especially in the context of digital technologies, and the current state of research is examined. To gain a deeper understanding of privacy concerns, especially for digital communication applications, semi-structured focus group interviews with experts of these applications were conducted. Afterwards, the results of this evaluation will be discussed. Important limitations and implications will be shown.

Finally, the paper concludes with a presentation of the overall results in compact form and gives a brief outlook on future developments.

2. Theoretical Background

2.1. Terminology / Important Terms

The concept of Informational Privacy is an area of research that is being explored in a variety of research fields, including law, economics, psychology and many others (Bélanger & Crossler, 2011, p. 1018; Pavlou, 2011; Smith, Dinev, & Xu, 2011). This paper aims at exploring the construct of information privacy within the information system domain. The worldwide use of the Internet is increasing continuously. Users' personal data is often collected, stored and analyzed. Therefore, privacy concerns in particular are becoming increasingly important in the digital age (Smith et al., 2011, p. 990). Advances in information and communication technology are opening new forms of exchange, thus also making the issue of privacy more important (Archibald, Ambagtsheer, Casey, & Lawless, 2019, p. 1). The introduction of so-called video call applications (often known as Voice over Internet Protocol (VoIP)-mediated technologies, in the context of this work hereafter limited to the first term) "*allow for real-time interaction involving sound, video, and often, written text*" (Archibald et al., 2019, p. 2; Weiler, Matt, & Hess, 2019). Such technologies therefore replicate features of face-to-face meetings and offer unique advantages, but also challenges, particularly in terms of privacy (Lacono & Brown, 2016, p. 1-3). For better understanding, the most important terms from the field of information privacy research are explained in the following section.

2.1.1. Definition Information Privacy

Privacy covers many areas of human life, is used in many disciplines and is therefore a collective expression (Buck & Dinev, 2020, p. 4232; Solove, 2005), which needs to be defined more precisely and clarified for the context of the use of information systems (Buck, 2018, p. 13). In the literature, privacy is not uniformly defined and is often not precisely delimited with regard to the respective object of investigation (Smith et al., 2011; Solove, 2005). Smith et al. (2011) classify the concept of privacy into the areas of physical privacy and information privacy. While physical privacy addresses the individual and/or the individual's surroundings and private space, information privacy addresses the action and responsibility dimension of the personal information (Smith et al., 2011, p. 990). Furthermore, the term is described as a multidimensional, elastic, and dynamic construct, which must be separated from overlapping concepts like confidentiality, secrecy, anonymity, security, and ethics (Smith et al., 2011, p. 995). Clarke (1999) identifies four dimensions of privacy: "*privacy of a person, personal behavior privacy, personal communication privacy, and personal data privacy*" (Connolly & McParland, 2012, p. 32). Due to the digitalization of information and communication, Bélanger and Crossler (2011) argue that the dimensions "personal communication privacy" and "data privacy" can be merged into the construct of information privacy (Pavlou, 2011, p. 978).

Because of the object of investigation of information systems for communication, this paper focuses on the dimension of information privacy. The definition of information privacy is directly linked to the development of information systems and therefore has different perspectives of definition (e.g. Dinev & Hart, 2006; Krasnova & Kift, 2012; Nissenbaum, 2009). A fundamental discussion is held with the perspective of privacy as a moral or legal right (Warren & Brandeis, 1890). With the perspective on privacy as a state, Westin (1968) and Altman (1975) introduce a way of looking at the individual and situational context of the user. The relationship to other individuals and privacy as a "*state of limited access to information*" (Smith et al., 2011, p. 995) becomes important. Moreover, the control-defined definition of privacy as a scientific discourse is becoming increasingly relevant (Altman, 1975; Smith, Milberg, & Burke, 1996; Westin, 1968). Margulis (1977) describes privacy as *"the control of transactions between person(s) and other(s), the ultimate aim of which is to enhance autonomy and/or to minimize vulnerability."* (Margulis, 1977, p. 10).

The concept of self-determined and autonomous control over the disclosure of personal information is closely related to the development of information technology and information systems (Buck, 2015, p. 107; Dinev & Hart, 2006). As

information systems evolve and expand, the collection, processing, and storage of personal information increases.

Especially when using information systems for communication, users release their personal data for the purpose of obtaining digital services. Individuals have the right to protect themselves from unwanted access in the sense of an intrusion into personal data (Rössler, 2001). This often leads to concerns when using such technologies. These will be examined below in more detail.

2.1.2. Definition Information Privacy Concerns

With increasing digitization and rising use of digital information systems, user concerns regarding information privacy are growing. Since the concept of information privacy is a latent concept and therefore cannot be measured directly, almost all empirical privacy research relies on the concept of information privacy concerns as a proxy for privacy (Smith et al., 2011, p. 997). Due to the broad scope of these concerns, different perspectives and definitions of privacy concerns have been developing in the scientific discourse.

Privacy concerns can be defined as user concerns about a possible future loss of privacy as a result of voluntary or involuntary disclosure of personal data (Dinev & Hart, 2006, p. 65; Reith, Buck, Walther, Lis, & Eymann, 2019, p. 3). This definitional approach is also shared by Malhotra, Kim, and Agarwal (2004) who defines privacy concerns as an individual's subjective view of fairness within the context of information privacy. A more restrictive definition of privacy concerns is used by many researchers, defining them as concerns that users have about the way companies and organizations handle personal data (Buck & Dinev, 2020, p. 4232; Smith et al., 1996, p. 169).

Privacy concerns hence strongly influence the behavioral intentions and the actual behavior of individuals when revealing private information (Pavlou, 2011, p. 982). Therefore, the goal of scientific studies is primarily to analyze the consequences and effects of privacy concerns. In this context, the construct of privacy concerns serves as an antecedent for several behavioral variables, e.g. willingness to disclose personal information (Chellappa & Sin, 2005), intentions to transact (Dinev & Hart, 2006) and information disclosure behavior (Buchanan, Paine, Joinson, & Reips, 2007; Xu et al., 2011, p. 800). IS studies see the construct of privacy concerns mostly as a general construct reflecting the inherent concern of individuals about the possible loss of privacy (e.g. Malhotra et al., 2004; Smith et al., 2011). However, recent legal and social scholars have noted that privacy is maybe more situation-specific than dispositional, making it important to distinguish between general concern and situation-specific concerns (e.g. Margulis, 2003; Solove & Hoofnagle, 2006; Xu et al., 2011, p. 800).

Today's privacy studies investigate the topic of privacy concerns in various contexts, e.g. in e-commerce or in social media. However, empirical research on privacy predominantly uses three constructs for privacy concerns as the base construct. These will be examined in the following chapter "Current State of Research".

2.2. Current State of Research

Previous research investigating the effects of privacy concerns has mainly focused on the intentions of using various types of online services (Bélanger, Hiller, & Smith, 2002, p. 247), disclose of personal information, engagement in e-commerce transactions (Dinev & Hart, 2006), and undertaking online purchases (Pavlou, 2011, p. 979). A major focus of the research lies on the individual level, especially on the consumer side. Bélanger and Crossler (2011) propose that more studies need to be conducted at the group, organizational, and societal level by viewing information privacy as a multilevel concept (Pavlou, 2011, p. 979). Hereby, there are still very few studies on the subject of privacy in the workplace, especially with regard to privacy concerns related to the use of IT (PCIT) (Becker, 2018; Connolly & McParland, 2012; Moshki & Barki, 2014). The following section will focus more on the privacy concern constructs on the consumer side and research concerning privacy in the workplace. In chapter 2.2.3 this paper tries to combine both strands of research and show why it is necessary to empirically investigate privacy concerns, especially in the employee sector.

2.2.1. Privacy in the Consumer Perspective

In empirical privacy research, three main constructs are used for privacy concerns. The Concern for Information Privacy (CFIP) is the first developed and validated construct used to measure information privacy concerns (Smith et al., 1996; Stewart & Segars, 2002). Malhotra et al. (2004) later developed a measuring instrument called the Internet Users' Information Privacy Concerns (IUIPC), which tries to be more specific to the technological conditions of the Internet. With the Mobile Users' Information Privacy Concerns (MUIPC) a construct was developed, which focuses on the privacy concerns which takes into account the environment as well as the specificities of privacy concerns in the context of mobile systems (Buck, 2018, p. 20; Xu et al., 2012). However, all three constructs contain large overlaps in terms of measuring instruments and dimensions. Due to the constant development of different information technologies, however, the number of different situation-specific constructs increases.

Today, one of the most important information privacy concerns in IS research is the construct of Internet Privacy Concerns (IPC). The Internet is one of the most popular mediums through which consumer and organizational data is transmitted, collected and analyzed (Hong & Thong, 2013, p. 276). IPC reflect an individual's perception of user concerns for how personal information is handled by websites. This can be different from users' expectations of how websites should handle users' personal information. Caused by the growing digitalization and the related implementation of digital technologies and new information system, the number of different conceptualizations of the IPC also increased (Hong & Thong, 2013, p. 278). Nevertheless, there are six key dimensions that are most commonly utilized of IPC. They are Collection, Secondary Usage, Errors, Improper Access, Control, and Awareness. Although the field of privacy concerns in the use of video call applications has not yet

been investigated, an attempt is being made to identify these dimensions. The dimensions have the following meanings:

Unauthorized Secondary Use addresses users' concerns that "*personal information is collected for one purpose but is used for another, secondary purpose without authorization from the individual*" (Hong & Thong, 2013, p. 278; Smith et al., 1996, p. 172). The construct *Error* addresses users' concerns about data inaccuracies. It is the degree to "*which a person is concerned that protections against deliberate and accidental errors in personal data collected are inadequate*" (Hong & Thong, 2013, p. 278; Smith et al., 1996, p. 172). *Improper access* describes "*privacy concerns with respect to the perceived threat of unauthorized access by third parties*" (Becker, 2018, p. 3262; Smith et al., 1996, p. 172). The dimension *Collection* describes the "*subjective concern with respect to the accumulation of personal information*" (Becker, 2018, p. 3262; Smith et al., 1996, p. 171). *Control* is the degree to which "*a person is concerned that users do not have adequate control over their personal information*" (Hong & Thong, 2013, p. 278; Malhotra et al., 2004, p. 339). Lastly, *Awareness* is the degree to which a "*person is concerned about users awareness of information privacy practices*" (Hong & Thong, 2013, p. 278; Malhotra et al., 2004, p. 339). The level of analysis of the IPC construct is mainly used with consumer data.

2.2.2. Privacy in the Employee Perspective

Privacy at the workplace is a field of research that has not yet been investigated as extensively as the field of consumer perspective. Due to the questions how privacy is defined in the workplace, scholars have identified three specific organizational contexts where employee privacy matters: First, regarding their information (information privacy), their workplace (work environment privacy) and their ability to work autonomously (autonomy privacy) (Bhave, Teo, & Dalal, 2020, p. 131; Stone & Stone, 1990). Although there is an overlap of the three concepts, the present work focuses on the concept of information privacy. Information privacy in the workplace includes the type of information on employees collected by organizations, the source of the information collection, the purpose of the information collection, and how the information is stored and used (e.g Bhave et al., 2020, p. 137; Smith et al., 1996; Stone & Stone, 1990).

Furthermore, the question arises for which different stakeholders privacy matters. The key stakeholders in workplace privacy research are employees, employers, and the state (Bhave et al., 2020, p. 132). The interests of the individual parties differ here: Employees strive for income and fulfilment; employers desire profit maximization and ensuring stakeholder value; the state wants to safeguard freedom and ensure the rule of law (Budd & Bhave, 2008). However, a strong focus lies on the employment relationship, which is the "*connection between employees and employers through which individuals sell their labor*" (Budd & Bhave, 2010, p. 51). Both stakeholders possess a privacy calculus where they engage in a cost benefit analysis where they weight risks of disclosing information versus withholding (Culnan & Armstrong, 1999; Laufer & Wolfe, 1977). For the employees' privacy calculus, the main risk associated with providing (or accessing) information is the perception of invasion of privacy in connection with the potential loss of control over one's own information (Bhave et al., 2020, p. 133; Stone & Stone, 1990). On the other hand, the key benefit in providing information is the reduction of information asymmetries that exist between them and the employer. Organizational risks when collecting employees' information are a potential invasion of the employees' privacy, the associated detrimental effects as well as the negative impact on employee morale and/or the organization's brand (Bhave et al., 2020, p. 133; Culnan, Smith, & Bies, 1994). The main benefit for the organization when possessing superior information about its employees is making better employment-related decisions and thereby increasing the security of the organization and reducing its legal liability. Both calculus models influence each other and can also be influenced by macro factors (Bhave et al., 2020, p. 134).

In the following section this paper will investigate the employee side and the reasons for possible concerns about their privacy. The issue of privacy concerns from an employee perspective has been particularly raised by home office and remote work regulations. Due to the increased use of digital technologies and new information systems (especially for internal communication), the topic of privacy is becoming more and more important.

2.2.3. Literature Review: Differences in Both Perspectives

There are only very few studies on information privacy research that have been carried out at all levels except the individual level. Investigative literature in the area of privacy concerns about communication technologies is very limited (Smith et al., 2011, p. 1004). Literature in organizational analysis based on communication technologies is outdated. Most of the work is conducted on the individual level, because there are validated concepts that can be used for the research and it is easier to collect and analyze data from a large number of people through surveys or interviews (Bélanger & Crossler, 2011, p 1028). However, despite the high level of research in the field of Information Privacy Concerns, there is only a small effort in developing tools for individuals to protect their information privacy. In addition Bélanger and Crossler (2011) state that most of the studies conducted at the organizational level focuses on information privacy practices as well as instituting appropriate policies. One possible explanation for the lack of organizational research shows that while organizations are interested in understanding the implications of information privacy impacts, researchers have not yet addressed these issues. Collecting information from citizens or consumers in general is easier than encouraging organizations to participate in such surveys.

However, the implementation of new information technologies is also becoming increasingly important in the workplace. Data on employees can be easily analyzed and evaluated, thus increasing the issue of surveillance (Connolly & McParland, 2012). Even if user data is not necessarily col-

lected specifically for monitoring employees, privacy concerns are increasing when using digital technologies. Privacy concerns have the potential to negatively affect organizational productivity and employee morale (Connolly & McParland, 2012, p. 32).

To further advance research in the organizational field, this paper tries to find more insights into the information privacy concerns of employees and to compare them with consumers. The next chapter deals with the chosen methodology for answering the two research questions.

3. Methodology

Since previous research has only examined privacy concerns on the consumer side, a qualitative approach is chosen to explore privacy concerns in the emerging workplace context (Myers, 2009). Interviews will be conducted to ask employees about their privacy concerns and examining how they differ from consumer privacy concerns. Due to the lack of research in the workplace context, the coding of the employee privacy concern is inductive based on the Grounded Theory by Glaser, Strauss, and Strutzel (1968). Consumer privacy concerns are coded inductively and deductively because certain codes are matched with dimensions of existing literature. All codes, factors and dimensions are finally combined inductively and deductively with the iterative thematic analysis approach (Braun & Clarke, 2006, p. 83).

3.1. Choice of Interview Partners

Concerning the choice of interview partners, this paper covers a heterogeneous sample of participants (Patton, 1990, p.169) that use digital video call applications (like Zoom; Skype etc.) in a private as well as in a workplace setting in order to cover their perspectives as a consumer and as an employee. As focus groups are especially well suited to uncover and document the “why” behind opinions, and in obtaining more depth and breadth in the analysis of participants, this paper conducts semi-structured focus group interviews (e.g. Becker, 2018, p. 3263; Morgan, 1996, p. 130-131). In focus group interviews participants have the opportunity to query each other, explain themselves and comment on each other’s experiences (Kitzinger, 1995, p. 299). This research method has been already used to uncover privacy aspects, like privacy concerns about technologies and acceptance of new technologies (Morton, 2014, p. 270). In the selection of interview partners this paper will follows the recommendations of Marshall, Cardon, Poddar, and Fontenot (2013) to collect data until data saturation is achieved.

Due to the psychological complexity of privacy concerns, this paper uses an iterative thematic analysis approach to structure the heterogeneous privacy perceptions into homogeneous themes to compare and analyze the influencing factors on privacy concerns of video call application users (Becker, 2018, p. 3261).

3.2. Data Collection

Three focus groups – are considered to be an adequate number (Morgan, 1996)- with 5-6 participants per group took place, capturing the views of 18 individuals. To ensure participants represent a broad range of experiences and ages, this paper uses opportunistic sampling with peer sampling for all three focus groups. 10 of the interview partners were female, 8 were male. The average age was 38 years. The age of the users ranged from 20 years to 63 years. An overview containing demographic profiles of the participants can be found in Table 1. All participants use video call applications to communicate with various stakeholders.

The group is designed to encourage the participants to interact with each other rather than with the researcher, to allow "*structured eavesdropping*" (Kitzinger, 1995, p. 301). At the beginning of each focus group the researcher provided an overview of the objectives of the study. In addition important terms like privacy and privacy concerns were explained. After questioning demographic data, the usage behavior of video call applications was examined. Next, the researcher attempted to restrict the contribution to reading the following open-end questions out loud, and asking further when required:

1. Do you have privacy concerns when using digital technologies at work?
2. In contrast, what are your privacy concerns in a private setting?
3. How do your privacy concerns differ (when using digital communication technologies) between private use and work use?
4. Why do your privacy concerns differ?

Each focus group lasted approximately from 45 minutes to one hour, 10 minutes asking participants about their demographic data and their usage behavior with the examined technology. Afterwards about 20-30 minutes were spent on questioning the privacy concerns on the employees’ side. It followed with an analysis of how these concerns differed from their private ones. An important subject of the investigation was also the reasons for the differences.

The same questions and procedures were used for each focus group facilitating an investigation into the similarity of the themes discussed across the focus groups (Morton, 2014, p. 272).

3.3. Data Analysis

The interview evaluation is based on the iterative thematic analysis approach to find patterns within the data (Braun & Clarke, 2006, p. 83). This analysis is particularly suitable for sensitive data environments (Patton, 1990) and is frequently employed in IS research (e.g. Becker, 2018, p. 3263; Weiler et al., 2019, p. 5881).

Each focus group interview was held via the video call application Zoom. The use of Zoom as a qualitative data collection tool has already been reviewed in the literature (Archibald et al., 2019). Solely the audio material was

Table 1: Demographic Profile of Participants

Demographics		**Number of Respondents**
Gender	Male	8
	Female	10
Age *(Average)*	Focus Group 20-30	21,8
	Focus Group 30-40	34,4
	Focus Group 40+	54,4
Working Situation *(Average)*	Home Office	65%
	Office	35%
Videocall Applications *(Business Setting)*	Zoom	12
	Microsoft Teams	16
	Cisco WebEx	2
	Skype for Business	2
	Apple FaceTime	1
	KUDO Meetings	1
Industry Sector / Department	Banking Industry	2
	M&A	1
	Consulting	2
	Big Media & Entertainment Company	1
	Luxury Goods:	
	- Off-Trade	2
	- Human Resources	3
	- Logistics	1
	- Marketing	1
	-On-Trade	1
	- Travel Retail	1
	- Management Assistant	1
	- IT	1
	- Customer Service	1

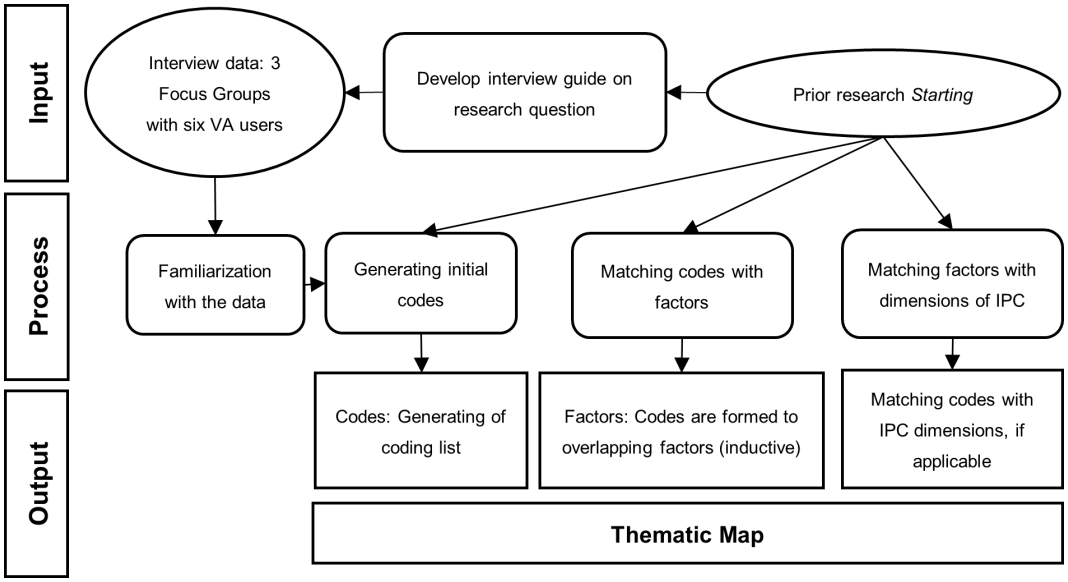


Figure 1: Iterative Thematic Analysis Approach based on Becker (2018)

recorded and first transcribed using "AmberScript", an AI-based software tool that converts audio material into text. The transcript was then repeatedly proofread. Later, the transcripts were uploaded to ATLAS.ti, one of the most widely used software for qualitative data analysis by IS researchers, to generate initial codes by searching for recurring patterns in the raw data (e.g Becker, 2018, p. 3263; Morton, 2014, p. 272). The entire transcript from each focus group was coded to ensure "*each data item has been given equal attention in the coding process*" (Braun & Clarke, 2006, p. 96). The concerns as a consumer are coded both inductively and deductively in order to include existing literature. The generation of employee concerns is inductive. The conceptual basis of the coding process hereby was drawn from already validated privacy concern constructs (e.g. IPC construct (Hong & Thong, 2013), which describes consumer privacy concerns in an online setting).

This paper identifies 56 different codes in the data set. In the next step different codes like "permanent accessibility" were merged with factors like "Performance tracking". Factors were then combined into different dimensions. The process of matching codes with the factors, and then the factors with potential dimensions of already validated privacy concern constructs (e.g. IPC) was accompanied by a constant review of the literature.

4. Results

4.1. Differences in Employee Privacy Concerns to Consumer Privacy Concerns

The first question to be answered is how privacy concerns differ in the use of video call applications as an employee to the use of these applications in private settings. The thematic map (Figure 2) visualizes the results and includes both dimensions derived from the theoretical part of the existing literature on privacy concerns and new dimensions from data collection of the interviews. It is composed of 7 dimensions Control, Collection, Awareness, Errors, Employment, Improper Access, Employment Status, and their related 13 factors. Factors and dimensions will be described in the following section.

4.1.1. Control

The Control dimension covers an individual's concerns that they do not have adequate control over their personal/firm information when using video call applications (cf. Becker, 2018, p. 3266; Malhotra et al., 2004, p. 339). Two factors were analyzed inductively to this dimension.

Dilemma of Usage by Confidential Data: Users express concerns about data control when highly confidential information is involved in communication. It is evident that consumers are switching to alternative technologies to gain more control over data, both in the workplace and in their private lives. In the working context two different types of disuse of video call applications can be distinguished. Firstly, the employer prohibits sharing of certain data on the video call applications. Participants stated that passwords and personal data of employees (especially mentioned in the human resources department) are never shared via the platforms: "*most of the data will be sent by email and not per the data sending tools on Teams"* [P2G1]. Another participant noted: "*the type of documents we have (i.e. human resource department) often prohibits sending them via Zoom*" [P4G2] In this regard, one user stated that there are specific guidelines in which data can be shared with the respective technology*: "I had eLearning in which I have learned what kind of data I am allowed to deliver via which platform. So, a few things I am only allowed to send via email and only encrypted and a few things only via shared folders. In general, I am not allowed to send data via Microsoft Teams or similar application. So, it always has to be encrypted and safe"* [P4G1].

Secondly, users are switching to other technologies they feel more comfortable using and where communication can be better controlled: *"I am super unwilling to send sensitive data over MS Teams, such as an internal price calculation, because it is easier to make mistakes or address someone else incorrectly. I still tend to do important things simply by email, [. . .] because of their clear traceability. Besides, I do not trust the whole thing 100 per cent that it is safe there. Especially with such sensitive data such as customers, prices etc."* [P2G2]. The missing or complicated "*trackability*" [P4G2] of communication in video call applications leads to the use of other technologies such as email.

Also, in the private sphere, participants tend to communicate sensitive messages by phone or in person: *"do not share your passwords or private sensitive data on any of those conversations [. . .]. And I think most of us do not like typing the password via WhatsApp. I called him and told him by phone"* [P3G1]. The use of the telephone as the main communication tool was often mentioned in this context: "*When it really comes to very personal things and serious topics, I prefer to use the telephone"* [P4G3].

Spread of Digital Content: This factor describes the concerns of employees and consumers about the uncontrolled loss or distribution of data using video call applications. Especially when traveling, video call applications are rarely used in the working context to avoid the risk of eavesdropping: *"I would have some concerns with that (i.e. travelling) because I could not guarantee that if I am speaking to someone on the phone or having a call with someone that I am actually taking the measures that their data does not go to someone sitting next to me"* [P3G1]. Another aspect that has been raised in this context is the one-sided control with communication tools like video call application*: "you can only control your outgoing side and you can only take as many precautions as you want, but it's only from your side."* [P3G1]. Concerns about uncontrolled distribution have also been expressed in the private sector: "*WhatsApp is actually more of a communication tool where things get relatively outspoken from time to time [...] there are simply chats where you can't control what the others contribute*" [P8G3]. The last concern that has been made in this context refers to the increased number of different communication technologies that are implemented

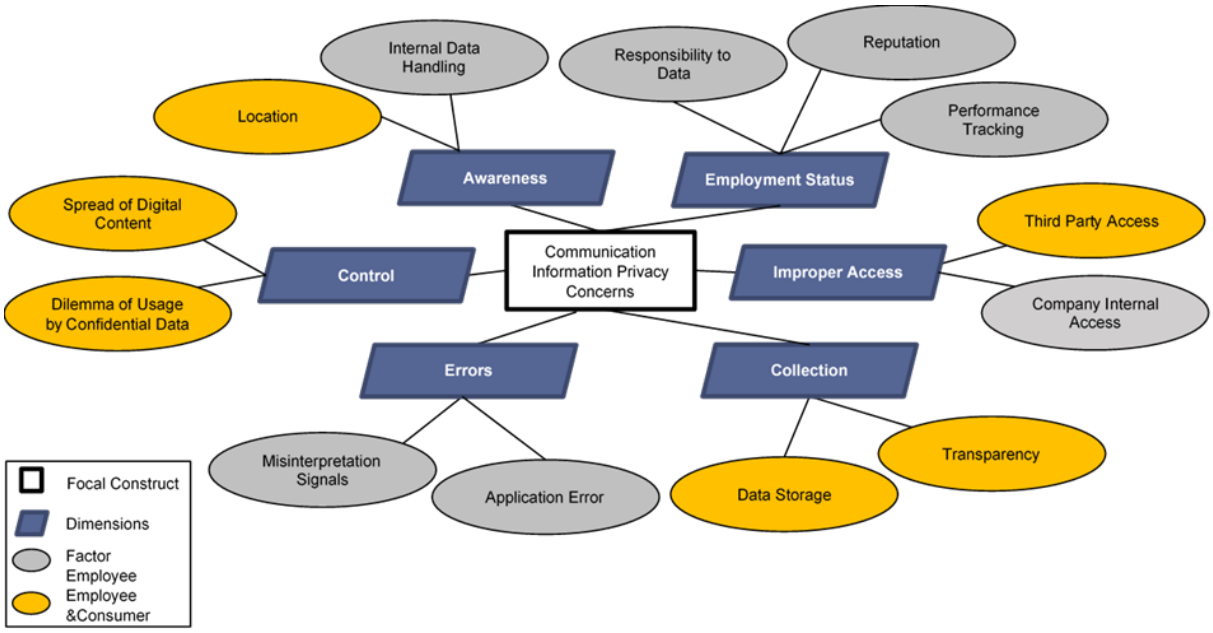


Figure 2: Thematic Map of Privacy Concerns of Employees and Consumers of Communication Technologies

in companies, which tends to lead to a weakening of communication: *"There are simply far too many communication strains and they are being opened up nowadays. And that is a problem, as at some point information gets lost, since you lose the overview, because everyone loses overview"* [P6G2]. Multiple informal forms of communicating and risk of uncontrolled data transfer were described as a "*Chinese-whispers-phenomenon (German Stille-Post-Phänomen)* "[P6G2] among the participants.

4.1.2. Improper Access

The dimension of Improper Access covers individuals' concerns that video call applications do not have adequate measures in place to prevent unauthorized individuals or organizations from accessing personal/firm information or that through the use of such technologies, data may be disclosed to third parties. (cf. Becker, 2018, p. 3265; Smith et al., 1996, p. 173). By analyzing the data, two factors could be found for Improper Access.

Third Party Access: Describes the unauthorized access to personal and firm data by third parties actively or passively while using the applications. This factor was mainly mentioned in the private context, as it is assumed that organizations provide the necessary safety precautions in the workplace: *"If there are difficulties or problem in the security, it is the responsibility of the company to solve them"* [P4G3]. The main concern is that third parties actively eavesdrop or monitor confidential communications, therefore one user mentioned: *"you cannot be sure if there will be a tape at the end (i.e. persons recording meetings). You are never really clear about that, who can see and who cannot."* [P4G2]. Concerns arise in the private sphere when internal family information would be disclosed: *"If my dad tells me some information about his business, it is a private and familiar topic which other people should not know"* [P5G1].

Company Internal Access: This factor describes the concern that personal data could be disclosed within the company to employees or supervisors who should not have access to it. Especially the exchange of private information in a business environment is a sensitive issue, as "*the borders*" between private and professional sphere are often "*floating*" [P8G3].

Users express concerns that private sideline activities made during business meetings may be noticed and picked up by other employees: *„When I get a phone call during a meeting, I am concerned that the conversation will be audible to the other 300 people in the meeting"* [P5G2]. Another user expressed the concern "*that someone is walking through the background or that you have an unfavorable setting with regard to the camera*" [P6G2]. Virtual backgrounds are often used as a security precaution in this context. In addition, the risk of private information being shared between individual employees to other employees/ or the entire team was a common concern: *"Writing stuff about, 'hey, maybe you can make that better or something that like this', what is private information about the team member. And it is not for all the members"* [P5G1].

On the other hand, there is a concern that outside of the work environment, communication about work or other employees may have negative effects when it leads back to the organization: *"when we will have a call or talk privately about work stuff or stuff that happened at work and it would not be great if stuff like that goes out. So, there are concerns"* [P7G1].

4.1.3. Employment Status

The Employment Status dimension includes individuals' concerns that the use of video call applications may have a negative impact on the employment relationship, especially towards the supervisor as well as other employees. This dimension was only mentioned in the context of workplace privacy concerns. Moreover, this dimension was analyzed inductively from the data.

Performance Tracking: This factor describes the increased transparency and traceability by the application, which becomes an increased challenge. Status notifications in applications such as Microsoft Teams and Skype for Business are seen by users as an "*additional obligation"* and access to the individual, as the application "*requires you to be available all the time*"[P8G3]. The transparency of work behavior in particular gives rise to concerns, which are expressed as follows: *"you can see (i.e. related to Microsoft Teams) via the green or yellow or red round indicator, if I am sitting at my computer or if I am in a meeting and so on. . . "* [P4G3].

Reputation: Especially in the young focus group (i.e. 20-30 years), maintaining a good reputation at the workplace is particularly important. Before a video call meeting, security precautions are therefore taken: *"I close all tabs (i.e. other programs) because I do not want to get any one from my colleagues see that I'm doing nonworking stuff in my working time"* [P3G1]. When using the option of screen sharing, people are concerned that private information is visible, which could give a negative impression about the person*: "If I have to do some screen sharing and I have that (i.e. private messenger) open and everybody could read my messages, I would not like that. I want to leave a good impression"* [P2G1]. It is noticeable that one's visible surroundings, especially in video calls, has a strong influence on the reputation. One user felt uncomfortable conducting meetings outside his or her usual workplace: *"I noticed that I did not participate that much in this call because I was in like a restaurant or something. And I did not want the customer to recognize that I am not in a working environment"* [P6G1]. To avoid conclusions about the working environment (increased concern due to home office), users try to have a clean background or even switch off their cameras during meetings.:" *I would avoid sitting in front of by bed, that everyone could see that*" [P4G2]. Other users do not want to provide any information about their environment. "*I do many zoom meetings or other meetings without a camera. I do not want that everyone can see my private background"* [P7G3].

Responsibility to Data: The use of company data in particular leads to increased responsibility, which also gives rise to concerns: *"when I am in the work environment, I am also responsible for the company and the company data and safety."* [P4G1]. In the event of loss of data, this does not only concern a single person but the whole company. This concern is amplified when users use private devices (e.g. private laptop) to communicate, as it may be difficult to maintain the same security measures, while other parties require the same level of responsibility. One user stated: *"when we call an external client of us, for example, or a business partner, because there we mostly call with our private phones, with our private laptops and private data [..] and there I/we or our bank cannot assure that they have the privacy standards which are needed for the sensitive data we have*" [P3G1]. Moreover, communication on company data is shown to cause increased concern among users and therefore more attention is paid to the question *"What kind of information am I processing here?"* [P6G1].

4.1.4. Awareness

The Awareness dimension refers to the individual's concern regarding their lack of awareness of how the use of video call applications utilizes and protects the privacy of personal/firm information (cf. Becker, 2018, p. 3266; Malhotra et al., 2004, p. 339). It is noticeable that users have little knowledge about the security precautions of video call applications made by the application itself or the employer.

Location: An important factor influencing awareness of whether an application perceived as safe is the place of communication or the location of the application. People are unaware of the application when it is managed by companies that are located in countries with controversial government policies: "*There is this concern for me based on the location of the company actually, or the service where the data is stored. I am not really sure if something is safer. I guess that Zoom kind of fixed these concerns. And yeah, it was a lot about media coverage and Zooms new technology, I think based in Asia*" [P6G1].

Internal Data Handling: This factor describes employees' lack of awareness of how organizations deal with internal communication data. One user posed the question*: "Who has access to your Microsoft Team?"* [P5G2]. It is noticeable that users do not know exactly who has access to the communication within the organization. Similarly, a different user poses the question: *"What opportunities do the people or departments we know have to collect data*?" [P7G3] There is a concern that data will be transferred without awareness: *"I am concerned that data may end up where I do not want it to end up"* [P5G2].

4.1.5. Error

The Error dimension refers to an individual's concerns that the application or use of the application does not have adequate measures in place to prevent or correct errors in their personal data (cf. Becker, 2018, p. 3266; Smith et al., 1996, p. 173). This dimension was mainly mentioned in the context of workplace concerns.

Application Error: This factor indicates that users are not sure whether the application or tools of the applications are working properly, and consequently personal data becomes accessible to others. When participating in video calls, it is possible to set your microphone to "mute". The correct function is questioned by a user: *"I have concern that the mute function really works, and I prefer to leave the room during a meeting"* [P5G2]. Even secondary functions such as chatting during a video call are often avoided due to potential errors. Person 4 in Group 2 states: *"Special care must be*

taken with the chat function. I am worried that a private chat will end up in the group...that is why I always look it up three times" [P4G2]. The use of the virtual background is used to avoid giving conclusions about the working environment (see Reputation). This function shows partial errors: "*The virtual background does not work for me. It shines completely through my face and that looks really creepy*" [P3G2].

Moreover, users expressed concerns about using the application when security issues were identified through media reports: "[. . .] *for example, with Zoom, that they had not well functioning privacy [. . .] security measures and that is why I think also when deciding which infrastructure to use, we didn't decide for Zoom"* [P3G1].

Misinterpretation Signals: This factor describes that certain signals (e.g. online status) in the video call applications gives false impressions about work behavior. Especially in video call applications like Microsoft Teams, as well as Skype for Business, every employee can see the presence status of every organization member via a sign symbol. One user states that: when "*the sign is yellow and it may show my team that I'm away and not working, but I am working. And then I have a little bit concerns that they can have a wrong feeling from my working habits*" [P4G1].

4.1.6. Collection

The Collection dimension captures an individual's concern that the application itself of the employer is collecting and storing large quantities of personal data through communication (cf. Becker, 2018, p. 3264; Smith et al., 1996, p. 171). Two factors were related to this dimension.

Data Storage: This factor describes that all communication via the applications is tracked and stored through the application itself as well as through the organization. One user said: "*Everything is tracked. It is not deleted, and someone has access to it."* [P3G2]. Concerns arise that communication will be monitored „*when you make a call, it may be recorded in the background and stored in the cloud*" [P5G2]. The concern that the data collected would be used against a person was particularly pronounced in the workplace: "*if you say something at work that is completely against everything* (i.e. the firm guidelines), *this can lead to dismissal"* [P5G2]. In a private setting people think that: "*in the mass that is generated, in messages, in texts, in pictures, in videos. What role do I play there?"* [P2G2].

Transparency: This factor describes the concern that personal data becomes more transparent using video call applications. Especially in the private sphere, users assume that all communication is stored somewhere in the cloud: "*I think to myself, as soon as I look for something in my mobile phone or anything else, nothing is private anymore. And as soon as you move somehow, it all ends up somewhere"* [P5G2].

4.2. Reasons for Differences in Privacy Concerns

Having identified the different privacy concerns of users in the work context as well as in the private context, the question arises why they differ:

The general trend is that privacy concerns are more widespread in the workplace than in private life. This observation can be found in all focus groups. In the youngest focus group (20-30 years), this is most noticeable, as they describe themselves as "*pretty affirm to digital technology*" [P6G1] and mainly use video call applications (e.g. FaceTime) with friends and communicate on everyday topics: "*when I FaceTime with my friends, I do not have any privacy concerns. We talk about everything and it is not top-secret private data*" [P2G1]. One user stated that the issue of data security is much more questioned in the working context: "*when you are in a business environment, there is this general more data policy stuff going on [...] you think more about it. When I use the digital technology in a business environment, I am more in the "What kind of information am I processing here?". And when I am in the private environment, I do not think about this at all, because I am like private free time. I do not have any privacy concerns*" [P6G1]. In private, the issue of data security is less considered: "*In my private life I think about it much less than I do in my business life."* [P4G2]

A potential loss of company data or incorrect handling of data in working life is considered a greater loss, as data is considered "*more valuable*" [P6G1]. In private life, users only must be aware of their own responsibility for data, while in the workplace company and customer data is also processed. One user summarized that with: "*when I'm talking with friends, I am mainly responsible for my own stuff I say and I feel more safe to handle my own responsibility with my own information, because if something goes outside, then it is my problem and I have to deal with it"* [P4G1]. There is also a concern, especially in the second focus group (30-40 years), that communication data will be screened and used against the individual. Consequences, such as dismissal, increases concerns about privacy at work compared to privacy at home: "*Privately, there are no consequences to be feared, but in the workplace if you send or do something at work and completely against everything and here is your dismissal"* [P5G2]. In private sphere, it is anticipated that mass data in general will be more interesting than individual data: "*I hope that I am one of the 7 billion that nobody cares"* [P7G3]. Interestingly, some of the older participants (40+ years) have more concerns in the private sphere, as there will be no security from the employer: "*I sometimes have the feeling that as an individual I can do so little in private. I can pay attention. I can make my password accordingly complex. But in the end, I do not have the same opportunity to build up protection as a large company"* [P4G3].

Privacy concerns about the application are lower in working life, as employees assume that professional business applications have sufficient security standards. Especially the use of well-known communications applications, which also offer other business solutions in working life, ensure trust: "*In my opinion, I think that MS Teams will be established throughout the business world because everybody is using Microsoft Office already like Excel and Word [. . .] a lot of people trust the Microsoft products in general"* [P6G1]. In addition, people are less concerned about the application, if it is only

used in a business setting: "*I think it (i.e. Microsoft Teams) feels automatically safer for companies because it's already targeting companies and therefore, yeah, it's already of their concerns*" [P7G1].

5. Discussion of Key Findings

This paper conducted a study to identify how information privacy concerns in communication with video call applications differ in a private setting to the workplace. Secondly, this paper analyzed why privacy concerns differ in the various settings. Therefore, three focus groups were interviewed about their concerns and then evaluated in a rigorous iterative thematic analysis (Braun & Clarke, 2006, p. 83). This procedure was used to create interview codes, which were combined into factors and dimensions. The data were visualized in a Thematic Map. The 4 dimensions (Awareness, Error, Collection, Improper Access) were deductively derived from existing literature (Malhotra et al., 2004; Smith et al., 1996). Factors as well as the dimension "Employment Status" were inductively generated from the qualitative data (Glaser et al., 1968). Certain statements could also be assigned to more than one concern based on their message. When analyzing the data, it becomes apparent that the three focus groups, which were classified according to age, have large differences in terms of privacy concerns. Depending on the group, different factors were mentioned in the length and frequency of the discussion, which are structured in more detail below to emphasize the first research questions:

Focus group 20-30 years:

This group is composed mainly of young adults who spend half of their time working, while the other half is still used for studying. In general, it can be seen that young participants in the study are particularly concerned about the Employment Status dimension. The factor Reputation was frequently mentioned, as they want to make a good impression on superiors and other employees in throughout their young professional career. The main concern is that work behavior could be misinterpreted in video call applications. In particular, the employer could draw conclusions about an employee's concentration if, for example, a cell phone is briefly looked at or an e-mail is answered in parallel. For this reason, any secondary activities, such as other open programs, are often closed just before a video call. However, there are also concerns that, if necessary, private messages or information that are not intended for this conversation circle can be read. Responsibility towards data is another factor that was frequently mentioned. There is a concern that company data will be mismanaged, which could again cause the employee relationship to suffer.

After analyzing the data, it can be seen that concerns about control were also frequently named, which is reflected with previous literature (Malhotra et al., 2004, p. 339). Although the interviewed participants use video call applications daily in both the work and private spheres, they switch to other technologies, especially when communicating sensitive data. Existing literature also shows that, although new workplace communication tools have been implemented, e-mail traffic in particular is considered to be the most important one (Chory, Vela, & Avtgis, 2016, p. 26).

In private, young users of the studies show little privacy concerns. Young participants often use video call applications (e.g. FaceTime) to share general topics with their friends and family. It is known that young people have fewer privacy concerns (cf. Culnan et al., 1994; Smith et al., 2011, p. 999). This could result primarily from the fact that young people are more familiar with these communication methods. Only important documents (e.g. passwords) are also shared due to insufficient control by other communication tools. Concerns are expressed here that the data could be used for other purposes.

Focus group 30-40 years:

The second focus group has been in full-time employment for several years. Concerns about employment status are still mentioned, but to a lesser extent. Above all, the factor of company-internal improper access stands out in the frequency of naming. This factor has already often been an important factor/important dimension in other fields of research (e.g. Becker, 2018; Hong & Thong, 2013; Smith et al., 1996). Especially in the work context it is evident that applications like Microsoft Teams are also used for private conversations, where there are concerns that other employees or the company may see them during a call e.g. through screen sharing etc (Clarke, 1999, p. 25). This concern also follows the increase of virtual meetings. Moreover, the concern is that in the worst case, this could lead to dismissal.

Furthermore, the factor of application error is mentioned. The functionality of individual tools and functions is questioned and leads to concerns. The mute function of Zoom is questioned, and participants prefer to leave the room. There are also concerns about control, as communication through video call applications can be poorly tracked. The loss of data is therefore more likely, in the opinion of the users. For communication, this focus group (as well as the focus group 40+) mainly use the video call applications Zoom and Microsoft Teams. The Microsoft Teams application is used primarily for exchanging instant messages (which are increasingly private), while Zoom is used more for scheduled meetings and external customers. However, there is no general trend that privacy concerns are becoming more prevalent in one of the applications mentioned above.

There are also fewer concerns in the private sphere. This focus group explains this mainly stating the fact that they do not have to fear any consequences in their private life that they have in their professional life.

Focus group 40+:

Participants in this focus group have only recently become familiar with communication tools such as video call applications compared to their work careers. Interestingly, privacy concerns per se are hardly mentioned in the last focus group.

Instead, behavioral patterns emerge that allow conclusions to be drawn about concerns.

In the work context, the oldest participants in the study assume that sufficient security standards are provided by the company so that both personal and company data can be secured. Concerns are expressed that all communication is tracked, which brings the collection dimension to the center of attention. In addition, new technologies such as video call applications are seen as an additional challenge, since constant accessibility increases (i.e. performance tracking). In general, the issue of privacy is nevertheless very important for older participants.

In their private lives, they hardly ever use video call applications and clearly prefer to use the telephone or even personal contact. However, personal data as well as privacy issues in general are considered particularly important by older participants. There is a trend that a higher level of privacy regarding data leads the less usage of digital technologies. To underline the second research questions, the reasons for the differences in privacy concerns are investigated below.

Reasons for differences in privacy concerns:

In conclusion, privacy concerns are more apparent in the work domain than in the private domain. This is primarily attributable to the fact that private data must be protected more strongly than when the technologies are used privately. Due to the different relationships with various stakeholders, it is difficult to standardize who can access data in the workspace. In addition, data sharing as well as the issue of data privacy are questioned more often in work life. Concerns about the application are equally low in both areas. The context and the environment are decisive factors influencing concerns (Bansal & Zahedi, 2008; Smith et al., 2011, p. 1002). Nevertheless, it is apparent that privacy concerns are not increasing strongly due to the use of other digital communication applications. Interviewed participants do not feel monitored but rather appreciate the technical capabilities of the communication applications.

6. Implications and Future Research

This study has important theoretical and practical implications. The visualized structure of the thematic map shows the different privacy concerns for the use of video call applications in private and business settings. While there are certainly other factors that influence the concerns, this thematic map provides a good starting point. It indicates that already validated dimensions (e.g. Control, Awareness etc.) of established constructs can be applied to the context of video call applications. The results of this study can therefore contribute to the understanding of privacy concerns in digital communication and identify possible avenues for future research. For example, further research could verify the relationship between the factors developed and the dimensions of the thematic map in a quantitative study (cf. Becker, 2018, p. 3268). Relationships between concerns are already apparent and should be explored through quantitative correlations. Moreover, the results of this study could support efforts to formulate theories to reveal the meaningful interaction between privacy concerns and the perceived benefits and acceptance of communication technologies. The dilemma of using communication technology when transmitting confidential data and privacy concerns about reputational damage in the workplace due to increased transparency show that privacy in the workplace is an important area of research. Further literature should examine other factors of privacy concerns in different technology applications, such as email. The rise of remote work and home office regulations has led to an increase in merging of private and business life, which also raises privacy concerns regarding the working environment. In addition, there is the question of how employers deal with employee concerns and who the appropriate contact person within the company is.

The thematic map can serve as a practical guideline for providers to develop more privacy friendly video call applications in work and private life. This study also shows that control over one's own data is seen as the most important dimension regarding communication privacy concerns. Employers should ensure that employees have more certainty about their control over personal data. Transparency regarding which data is collected and to whom it is accessible should be increased. The implementation of security standards by both the application provider and the employer should be encouraged.

7. Limitations

An obvious limitation of this study was the small sample size of 18 participants who took part in focus groups. Nevertheless, care was taken to ensure that the three focus groups represented a large, diversified sample of the general population in terms of age (i.e., age range from 20-63 years). Moreover, attention was paid to an average group size of approx. 6 participants. Especially when interviewing emotionally charged topics, a group should not be too large in terms of size (Morgan, 1996, p. 146). The ability to facilitate a group-level discussion, which is a strong advantage of the collection method, is also one of its major limitations, because of the danger of dominant personalities controlling a group discussion (Becker, 2018, p. 3269). An attempt was made to include all participants equally in the discussion to avoid a one-sided bias. The first interview with the youngest focus group was conducted in English. Results of the focus group 30-40 years and 40+ had to be translated from German into English, which could potentially lead to changes in what was said. It became apparent that older participants were more comfortable discussing the emotional topic of privacy in their native language. Although the terminology of privacy research was explained at the beginning of each interview, older participants had difficulty understanding it. The result was a shift to other topics and the naming of privacy concerns in the online context.

Many of the questioned participants (12 out of 18 participants) work for the same company, so the number of dif-

ferent applications is limited. Nevertheless, in the selection of these participants, an attempt was made to find as many departments of this company as possible

All group interviews were carried out using the video call application Zoom, which partly disturbed the productivity of the interviews and thus the group dynamics due to technical problems. Nevertheless, it is recognized in literature that Zoom is a suitable medium for collecting qualitative data (Archibald et al., 2019).

Lastly, due to the use of different video call applications in the business as well as the private sector, a comparison of the privacy concerns regarding the application is not always justified.

8. Conclusion

This study attempted to answer two questions regarding the investigation of privacy concerns in the use of video call applications. Firstly, it aimed to find out how the privacy concerns of users of these technologies differ in the work context from the private context. Secondly, this paper wanted to discover the underlying cause for these differing concerns.

Focus group interviews were conducted to answer these questions. Participants were required to use video call applications in both business and personal contexts. The results of the qualitative research were presented visually using a thematic map. Users expressed great privacy concerns regarding the control of sensitive data. In addition, work behavior and employment relationships are becoming more transparent, which raises more concerns. Employees in particular try to protect private confidential data at work.

This paper presents one of the first exploratory findings in the field of privacy research in the workplace.

References

Altman, I. (1975). *The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding*.

Archibald, M. M., Ambagtsheer, R. C., Casey, M. G., & Lawless, M. (2019). Using Zoom Videoconferencing for Qualitative Data Collection: Perceptions and Experiences of Researchers and Participants. *International Journal of Qualitative Methods*, *18*, 1609406919874596.

Bansal, G., & Zahedi, F. (2008). The Moderating Influence of Privacy Concern on the Efficacy of Privacy Assurance Mechanisms for Building Trust: A Multiple-Context Investigation. *ICIS 2008 Proceedings*, 7.

Becker, M. (2018). *Understanding Users' Health Information Privacy Concerns for Health Wearables*.

Bélanger, F., & Crossler, R. E. (2011). Privacy in the Digital Age: A Review of Information Privacy Research in Information Systems. *MIS quarterly*, 1017–1041.

Bélanger, F., Hiller, J. S., & Smith, W. J. (2002). Trustworthiness in Electronic Commerce: The Role of Privacy, Security, and Site Attributes. *The Journal of Strategic Information Systems*, *11*(3-4), 245–270.

Bhave, D. P., Teo, L. H., & Dalal, R. S. (2020). Privacy at Work: A Review and a Research Agenda for a Contested Terrain. *Journal of Management*, *46*(1), 127–164.

Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Buchanan, T., Paine, C., Joinson, A. N., & Reips, U. D. (2007). Development of Measures of Online Privacy Concern and Protection for Use on the Internet. *Journal of the American Society for Information Science and Technology*, *58*(2), 157–165.

Buck, C. (2015). *App-Privacy as an Abstract Value–Approaching Contingent Valuation for Investigating the Willingness to Pay for App Privacy*.

Buck, C. (2018). *Beiträge Zur Untersuchung Der Informationellen Privatheit Im Rahmen Des Experiential Computing*.

Buck, C., & Dinev, T. (2020). Low Effort and Privacy–How Textual Priming Affects Privacy Concerns of Email Service Users. *Proceedings of the 53rd Hawaii International Conference on System Sciences*.

Budd, J. W., & Bhave, D. (2008). Values, Ideologies, and Frames of Reference in Industrial Relations. In *The Sage Handbook of IndustrialRelations* (pp. 92–112). London: Sage.

Budd, J. W., & Bhave, D. (2010). The Employment Relationship. In *Handbook of Human Resource Management* (pp. 51–70).

Chawla, A. (2020). Coronavirus (Covid-19) Zoom application Boon or Bane. *Available at SSRN 3606716*.

Chellappa, R. K., & Sin, R. G. (2005). Personalization Versus Privacy: An Empirical Examination of the Online Consumer's Dilemma. *Information Technology and Management*, *6*(2-3), 181–202.

Chory, R. M., Vela, L. E., & Avtgis, T. A. (2016). Organizational Surveillance of Computer-Mediated Workplace Communication: Employee Privacy Concerns and Responses. *Employee Responsibilities and Rights Journal*, *28*(1), 23–43.

Clarke, R. (1999). Internet Privacy Concerns Confirm the Case for Intervention. *Communications of the ACM*, *42*(2), 60–67.

Connolly, R., & McParland, C. (2012). Dataveillance: Employee Monitoring & Information Privacy Concerns in the Workplace. *Journal of Information Technology Research*, *5*(2), 31–45.

Consumers-Union. (2008). *Consumer Reports Poll: Americans Extremely Concerned About Internet Privacy.* Retrieved 22-11-2020, from http://www.consumersunion.org/pub/core_telecom_and_utilities/006189.html

Culnan, M., Smith, H., & Bies, R. (1994). *Law Privacy and Organizations: The Corporate Obsession to Know V. The Individual Right Not to Be Known*. The Legalistic Organization.

Culnan, M. J., & Armstrong, P. K. (1999). Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. *Organization Science*, *10*(1), 104–115.

Dinev, T., & Hart, P. (2006). An Extended Privacy Calculus Model for E-Commerce Transactions. *Information Systems Research*, *17*(1), 61–80.

Epic. (2020). *Public Opinion on Privacy.* Retrieved 22-11-2020, from https://epic.org/privacy/survey/

Glaser, B. G., Strauss, A. L., & Strutzel, E. (1968). The Discovery of Grounded Theory; Strategies for Qualitative Research. *Nursing Research*, *17*(4), 364.

Gomez, J., Pinnick, T., & Soltani, A. (2009). Knowprivacy.

Hong, W., & Thong, J. Y. (2013). Internet Privacy Concerns: An Integrated Conceptualization and Four Empirical Studies. *MIS Quarterly*, 275–298.

Jiang, Z., Heng, C. S., & Choi, B. C. (2013). Research Note–Privacy Concerns and Privacy-Protective Behavior in Synchronous Online Social Interactions. *Information Systems Research*, *24*(3), 579–595.

Kagan, D., Alpert, G. F., & Fire, M. (2020). Zooming into Video Conferencing Privacy and Security Threats. *arXiv preprintarXiv:2007.01059*.

Kitzinger, J. (1995). Qualitative Research: Introducing Focus Groups. *BMJ*, *311*(7000), 299–302.

Krasnova, H., & Kift, P. (2012). Online Privacy Concerns and Legal Assurance: A User Perspective. *AIS SIGSEC WISP Workshop onInformation Security and Privacy*, 1–23.

Lacono, P., V. L. Symonds, & Brown, D. H. (2016). Skype as a Tool for Qualitative Research Interviews. *Sociological Research Online*, *21*(2), 1–12.

Laufer, R. S., & Wolfe, M. (1977). Privacy as a Concept and a Social Issue: A Multidimensional Developmental Theory. *Journal of Social Issues*, *33*(3), 22-42.

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet Users' Information Privacy Concerns (Iuipc): The Construct, the Scale, and a Causal Model. *Information Systems Research*, *15*(4), 336–355.

Margulis, S. T. (1977). Conceptions of Privacy: Current Status and Next Steps. *Journal of Social Issues*, *33*(3), 5–21.

Margulis, S. T. (2003). Privacy as a social issue and behavioral concept. *Journal of social issues*, *59*(2), 243–261.

Marshall, B., Cardon, P., Poddar, A., & Fontenot, R. (2013). Does Sample Size Matter in Qualitative Research?: A Review of Qualitative Interviews in Is Research. *Journal of Computer Information Systems*, *54*(1), 11–22.

Morgan, D. L. (1996). *Focus Groups as Qualitative Research*. Sage Publications.

Morton, A. (2014). All My Mates Have Got It, So It Must Be Okay: Constructing a Richer Understanding of Privacy Concerns–an Exploratory Focus Group Study. In *Reloading Data Protection* (pp. 259–298). Springer.

Moshki, H., & Barki, H. (2014). *Individuals' It-Related Privacy Concerns: A Two-Phase Cognitive Model*.

Murphy, K. (2020). Why Zoom Is Terrible. *The New York Times*, *23*.

Myers, M. (2009). *Qualitative Research in Business & Management*. Sage Publications.

Nissenbaum, H. (2009). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press.

Patton, M. Q. (1990). *Qualitative Evaluation and Research Methods*. SAGE Publications.

Pavlou, P. A. (2011). State of the Information Privacy Literature: Where Are We Now and Where Should We Go? *MIS quarterly*, 977–988.

Reith, R., Buck, C., Walther, D., Lis, B., & Eymann, T. (2019). *How Privacy Affects the Acceptance of Mobile Payment Solutions*.

Rössler, B. (2001). *Der Wert Des Privaten*. Suhrkamp Frankfurt am Main.

Smith, H. J., Dinev, T., & Xu, H. (2011). Information Privacy Research: An Interdisciplinary Review. *MIS quarterly*, 989–1015.

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information Privacy: Measuring Individuals' Concerns About Organizational Practices. *MIS quarterly*, 167–196.

Solove, D. J. (2005). A Taxonomy of Privacy. *U. Pa. L. Rev.*, *154*, 477.

Solove, D. J., & Hoofnagle, C. J. (2006). Model Regime of Privacy Protection. *U. Ill. L. Rev.*, 357.

Stewart, K. A., & Segars, A. H. (2002). An Empirical Examination of the Concern for Information Privacy Instrument. *Information Systems Research*, *13*(1), 36–49.

Stone, E. F., & Stone, D. L. (1990). Privacy in Organizations: Theoretical Issues, Research Findings, and Protection Mechanisms. *Research in Personnel and Human Resources Management*, *8*(3), 349–411.

Wagenseil, P. (2020). Zoom Security Issues: Here's Everything That's Gone Wrong (So Far). *Tom's Guide*, *11*.

Warren, S. D., & Brandeis, L. D. (1890). The Right to Privacy. *Harvard Law Review*, 193–220.

Weiler, S., Matt, C., & Hess, T. (2019). Understanding User Uncertainty During the Implementation of Self-Service Business Intelligence: A

Thematic Analysis. *Proceedings of the 52nd Hawaii International Conference on System Sciences*.

Westin, A. F. (1968). Privacy and Freedom. *Washington and Lee Law Review*, *25*(1), 166.

Wiederhold, B. K. (2020). *Connecting through Technology During the Coronavirus Disease 2019 Pandemic: Avoiding "Zoom Fatigue*". Mary Ann Liebert.

Xu, H., Dinev, T., Smith, J., & Hart, P. (2011). Information Privacy Concerns: Linking Individual Perceptions with Institutional Privacy Assurances. *Journal of the Association for Information Systems*, *12*(12), 1.

Xu, H., Teo, H.-H., Tan, B. C., & Agarwal, R. (2012). Research Note - Effects of Individual Self-Protection, Industry Self-Regulation, and Government Regulation on Privacy Concerns: A Study of Location-Based Services. *Information Systems Research*, *23*(4), 1342–1363.

Junior Management Science 7(4) (2022) 1001-1031

Junior Management Science

journal homepage: www.jums.academy

Digital Financial Reporting – Consequences and Implications for the Use of the New Medium

Digitalisierung in der Finanzberichterstattung – Folgen und Konsequenzen für den Einsatz des neuen Mediums

Milan Schulze

Heinrich-Heine-Universität Düsseldorf

Abstract

Changes and technological developments in business and society are effects of increasing digitization. New media and technologies within digital business models are gaining importance. The consequences of their usage represent opportunities, challenges, and risks for the digital transformation of companies. The transformation of analog information into a digital format, or the digitization in financial reporting, reaches another level of effectiveness as a new medium. Both the volume of published information and the timely availability of business data continue to increase, not least due to increasing requirements. This poses challenges and risks with regards to the analyzability of this information. The change to digitalized reporting using the new medium must be understood as change management, in that traditional business management thinking and action is digitally oriented and adaptability and consistent willingness to learn are established in a contemporary understanding of roles.

Zusammenfassung

Veränderungen und technologische Entwicklungen in der Wirtschaft und Gesellschaft sind Auswirkungen der zunehmenden Digitalisierung. Neue Medien und Technologien innerhalb digitaler Geschäftsmodelle gewinnen an Bedeutung. Die Folgen und Konsequenzen aus dem Einsatz stellen Chancen, Herausforderungen und Risiken für die digitale Transformation von Unternehmen dar. Die Umwandlung von analogen Informationen in ein digitales Format, bzw. die Digitalisierung in der Finanzberichterstattung erreicht als neues Medium ein weiteres Level der Wirksamkeit. Sowohl der Umfang von veröffentlichten Informationen als auch die zeitnahe Verfügbarkeit von Geschäftsdaten nehmen nicht zuletzt aufgrund steigender Anforderungen weiter zu. Dies birgt insbesondere Herausforderungen und Risiken im Hinblick auf die Auswertbarkeit dieser Informationen. Der Wandel zur digitalisierten Berichterstattung durch den Einsatz des neuen Mediums muss dabei als Veränderungsmanagement verstanden werden, indem traditionelles betriebswirtschaftliches Denken und Handeln digital ausgerichtet und Anpassungsfähigkeit und konsequente Lernbereitschaft in einem zeitgemäßen Rollenverständnis etabliert wird.

Keywords: Digitalisierung; Berichterstattung; ESEF; XBRL.

1. Einleitung

Tiefgreifende Veränderungen und exponentielle technologische Entwicklungen in der Wirtschaft und Gesellschaft sind Auswirkungen der zunehmenden Digitalisierung. Es entstehen neue Geschäftsmodelle, Märkte und Wettbewerbskonstellationen, wobei unternehmerische Entscheidungen mit steigender Unsicherheit und Volatilität durch die Digitalisierung an Komplexität zunehmen. Neue Medien und Technologien innerhalb digitaler Geschäftsmodelle gewinnen an Bedeutung. Die Folgen und Konsequenzen aus dem Einsatz stellen Chancen, Herausforderungen und Risiken für die digitale Transformation von Unternehmen dar (vgl. Caputo, Pizzi, Pellegrini & Dabić, 2021, S. 489ff.). Die Umwandlung von analogen Informationen[1] in ein digitales Format, bzw.

[1] In der Arbeit werden Informationen als verarbeitete (Roh-)Daten verstanden.

DOI: https://doi.org/10.5282/jums/v7i4pp1001-1031

die Digitalisierung in der Finanzberichterstattung (FBE), bei der papierbasierte Geschäftsberichte in neue Formate umgewandelt und dargestellt werden, erreicht als neues Medium[2] ein weiteres Level der Wirksamkeit (vgl. Troshani & Rowbottom, 2021, S. 2f.). In der komplexen Wirtschaft müssen die Akteure zu jedem Zeitpunkt in der Lage sein, zeitnah auf Marktveränderungen und Krisen reagieren zu können, wobei der Zugang und richtige Umgang mit Daten und Informationen diesbezüglich essentielle Wettbewerbsfaktoren darstellen (vgl. Demirkan & Delen, 2013, S. 412). Es liegen vermehrt unstrukturierte Daten in Form von z.B. digitalen Listen und zugleich physischen Gegenständen mit potentiell hohem Informationsgehalt vor, wodurch Unternehmen mit einem digitalen Darwinismus konfrontiert werden, sprich dem Risiko, aufgrund mangelnder Anpassungsfähigkeit und fehlender Lernbereitschaft nicht mit der Digitalisierung und dem Wettbewerb mithalten zu können (vgl. Weißenberger & Bauch, 2017, S. 213; Kreutzer, Neugebauer & Pattloch, 2017, S. 3f.; Smith, 2018, S. 249).

Das Portable Document Format (PDF) steht in den letzten Jahren zum Informationsaustausch von Geschäftsangaben wiederholt im Fokus. Vor drei Jahren publizierten rund 50 Prozent der in einer Studie analysierten Unternehmen einen PDF-Geschäftsbericht (vgl. NeidhartSchön AG, 2018, S. 8). Hierbei stehen heutzutage weitaus mehr Möglichkeiten zur Auswahl, als Informationen auf der Unternehmenswebseite oder im Bundesanzeiger als traditionelles PDF zu veröffentlichen. Vielmehr bieten fortschrittliche Technologien, wie z.B. die Extensible Business Reporting Language (XBRL) die Möglichkeit, den klassischen Finanzbericht auf ein effizientes und digitalisiertes Format zu heben. Das Rechnungswesen eines Unternehmens erreicht damit einen zeitgemäßen und ganzheitlich digitalen FBE-Prozess, während der Berichtsadressat die hohe Informationsdichte seinen Bedürfnissen entsprechend gestalten und effizient verarbeiten kann. Trotz diverser Vorteile einer digitalen FBE zeigten sich Unternehmen zurückhaltend im Hinblick auf die Nutzung neuer digitaler Medien und Technologien und mussten zur Einführung von z.B. digitalen Berichtsprachen wie XBRL zwingend herangeführt werden (vgl. Cordery, Fowler & Mustafa, 2011, S. 84; Mittelbach-Hörmanseder, Hummel, Barrantes & Guggenberger, 2021, S. 23f.). Mittlerweile wird ein Wandel in der Bedeutung neuer digitaler Berichtsformate deutlich – vom statischen Print- und PDF-Bericht hin zur dynamischen Online- und XBRL-Berichterstattung über fortschrittliche Applikationen und digitale Kanäle (vgl. Mittelbach-Hörmanseder, Guggenberger & Barrantes, 2020, S. 47f.).

Sowohl der Umfang von veröffentlichten Informationen, als auch die zeitnahe Verfügbarkeit von produzierten Geschäftsdaten nehmen nicht zuletzt aufgrund steigender Anforderungen weiter zu. Dies birgt insbesondere Herausforderungen und Risiken im Hinblick auf die Auswertbarkeit dieser Informationen. Im Sinne eines effizienten Kapitalmarktes steigt damit analog die Notwendigkeit eines reibungslos auswertbaren Formats für die digitale FBE, um Informationen zeitnah und umfassend in den Markt einpreisen zu können. Es steht in Frage, ob durch die Folgen und Konsequenzen der Digitalisierung in der FBE große und unbeständige Datenmengen, die den Kapitalmarktteilnehmern über diverse Quellen zur Verfügung stehen, effektiver bereit- und dargestellt werden und Transaktionskosten und Verarbeitungskosten reduziert werden, wodurch die Informations- und Markteffizienz ganzheitlich gesteigert werden. Ziel der Arbeit ist es somit, einleitende Ursachen der Digitalisierung in der FBE und anschließende Folgen und Konsequenzen durch den Einsatz neuer digitaler Medien und Technologien, insbesondere die Bestrebungen einer Vereinheitlichung der digitalen FBE in der Europäischen Union (EU) durch XBRL aufzuzeigen und deren Wirkung auf die Informations- und Markteffizienz zu analysieren. Diesbezüglich werden im folgenden Kapitel die Ursachen der Digitalisierung in der FBE aus regulatorischer, phänomenologischer und technologischer Sicht betrachtet. Insbesondere die steigenden Anforderungen an die digitale FBE in Verbindung mit der Hypothese effizienter Märkte (Efficient-Market Hypothesis; EMH), Phänomene wie Big Data und die Entwicklung neuer fortschrittlicher Medien und Technologien führen zu einer anhaltenden Digitalisierung in der FBE. Im anschließenden Kapitel werden die Wirkungen dieser Ursachen anhand von Folgen und Konsequenzen für die digitale FBE erläutert. Darunter fällt der stetig wachsende Umfang der FBE durch die neuen Onlinemedien und steigenden Kapitalmarktforderungen, die Möglichkeit einer erhöhten Datenwertschöpfung und Informationstiefe durch Phänomene wie Big Data und neue fortschrittliche Analysemethoden, vorliegende Vereinheitlichungstendenzen aufgrund einer bis dato heterogenen FBE in der EU und ein neues digitales Rollenverständnis des Rechnungswesens durch die Implementierung von Onlinemedien und Technologien wie z.B. neue digitale Berichtsprachen, Cloud- und Datenbankstrukturen, Robotic Process Automation (RPA) und Künstliche Intelligenz (KI). Mit einer anschließenden Gegenüberstellung möglicher Chancen, Herausforderungen und Risiken aus den Folgen und Konsequenzen für den Einsatz des neuen Mediums wird eine zusammenfassende Würdigung angestoßen und ein ganzheitlicher Blick in die neue digitalisierte Berichterstattung vollzogen. Infolgedessen wird die Arbeit mit einem abschließenden Fazit und Ausblick in mögliche Entwicklungen und Forschungsrichtungen beendet.

Die folgende Abbildung soll einen einleitenden Überblick zu den Ursachen der Digitalisierung in der FBE, die im 2. Kapitel genauer erläutert werden, schaffen, und auf die anschließende Ausarbeitung der Folgen und Konsequenzen für den Einsatz des neuen Mediums vorbereiten (die Ursachen wurden in Hellblau und fettgedruckter Schrift und das Unternehmen bzw. die Adressaten in Hellgrau bzw. Dunkelgrau visualisiert):

[2]Die Begriffe Medium oder Onlinemedium werden als Synonyme für Komponenten der internetbasierten Kommunikation innerhalb der Berichterstattung verstanden. Die Digitalisierung in der FBE beschreibt die Zusammenführung einer Vielzahl von digitalen Medien und Technologien innerhalb der finanziellen Berichterstattung, wobei die grundlegende Internet- bzw. Onlinefähigkeit als übergeordnetes Medium (siehe Titel der Arbeit) bezeichnet wird.

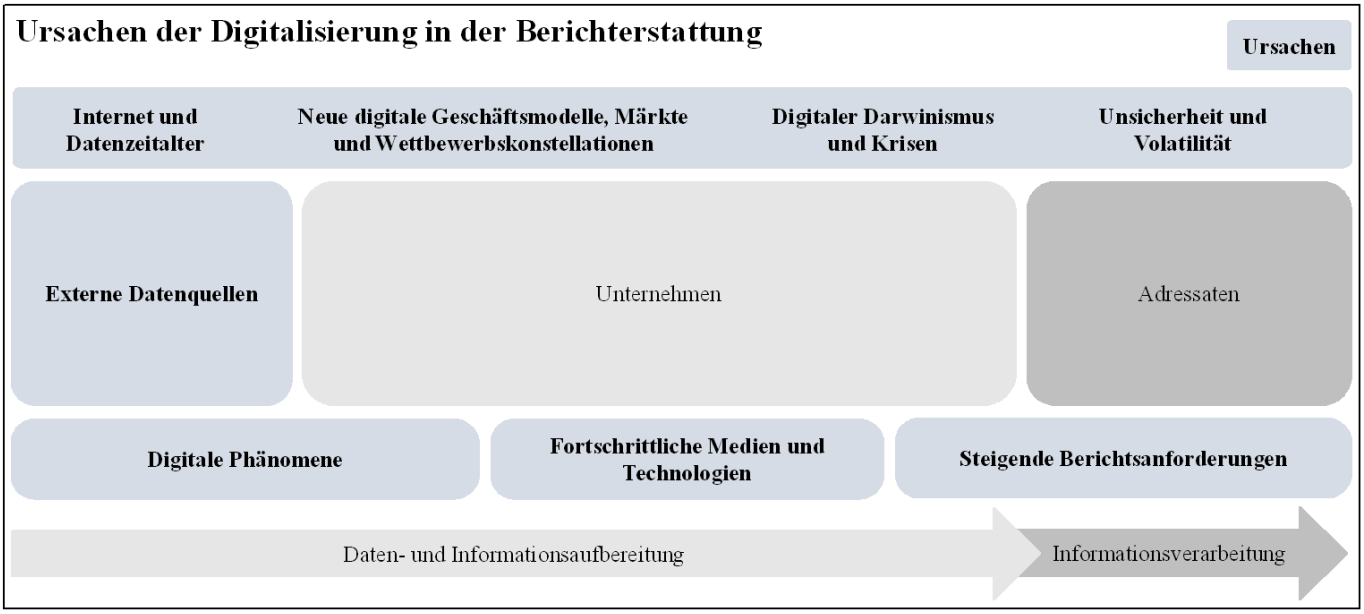


Abbildung 1: Überblick zu den Ursachen der Digitalisierung in der Berichterstattung (eigene Darstellung mit inhaltlicher Anlehnung an Autoren aus Kapitel 1 und 2).

2. Ursachen der Digitalisierung in der Berichterstattung

2.1. Neues Medium und steigende Berichtsforderungen

Die traditionelle FBE als Teilmenge des Rechnungswesens ist grundsätzlich der externen Rechnungslegung zuzuordnen und fokussiert die Sammlung von Geschäftsdaten und der zahlenmäßigen und regelkonformen Abbildung des Realprozesses der Geschäftstätigkeit über die Veröffentlichung von Geschäftsangaben. Diese Veröffentlichung richtet sich vor allem an unternehmensexterne Adressaten wie Kreditinstitute, Investoren, Finanzbehörden oder die Öffentlichkeit. Hierbei sollen veröffentlichte Berichte im Wesentlichen zur Entscheidungsfindung dieser Adressaten beitragen und stellen gleichzeitig die Bemessungsgrundlage für Zahlungsvorgänge, z.B. Dividenden- oder Steuerzahlungen des Unternehmens dar (vgl. Weber & Weißenberger, 2015, S. 27ff.; Wagenhofer & Ewert, 2015, S. 4ff.). Wohingegen aus der aktuellen Forschung deutlich wird, dass neben der traditionellen Definition der Begriff einer digitalen FBE nicht klar abzugrenzen ist. Die Begrifflichkeiten variieren vom „web-based business reporting“ (Beattie & Pratt, 2003, S. 155) bis hin zum „Internet Financial Reporting“ (Debreceny, Gray & Rahman, 2002, S. 372). In der Arbeit wird dahingehend die umfangreiche Definition des International Accounting Standards Committees (IASC) zugrunde gelegt, wonach digitale FBE als „the public reporting of operating and financial data by a business enterprise via the World Wide Web or related Internet-based communications medium“ (IASC, 1999, S. 2) verstanden wird.

Wenn von einer digitalen FBE die Rede ist, sind darüber hinaus Begriffe wie z.B. Multimedialität oder Interaktivität allgegenwärtig. Eine digitale FBE fördert durch den Einsatz verschiedener Onlinemedien die Interaktivität zwischen den Akteuren, sodass digitale Schnittstellen ermöglicht werden (vgl. Barrantes & Sagmüller, 2014, S. 61; Voelzkow, 2019, S. 106). Insbesondere Schnittstellen, die eine Umwandlung von unstrukturierten (Roh-) Daten in ein strukturiertes Format ermöglichen, wodurch das Vergeben einer Bedeutung der Daten erleichtert wird, stehen in ausgeprägter Verbindung mit dem Begriff einer digitalen FBE (vgl. Hoffman & Rodríguez, 2013, S. 86; Valentinetti & Muñoz, 2021, S. 549). Darüber hinaus ist hinzuzufügen, dass neben solchen Begriffen eine digitale Veränderung innerhalb der FBE-Praktik selbst vorzufinden ist, da durch neue digitale Geschäftsmodelle das ausgewiesene Vermögen in Bilanzen vermehrt in die immaterielle Sphäre verlagert wird (vgl. Weißenberger & Bauch, 2017, S. 207f.; Beattie & Pratt, 2003, S. 157).

Derweil hat sich das Financial Accounting Standards Board (FASB) ebenfalls mit der Themeneinordnung und Begriffsdeutung auseinandergesetzt. Grundlegend wird zur Einordnung zwischen der (statischen und dynamischen) Präsentation und dem Inhalt der Geschäftsinformationen unterschieden. Die papierbasierte Berichtspräsentation entspricht der traditionellen Form und setzt sich aus statischen Abbildungen und Texten zusammen, wohingegen bei der dynamischen Präsentationsform z.B. Audio- oder Videodateien, Download-Möglichkeiten und Hypertext Markup Language-(HTML-)Lösungen mit berücksichtigt werden, die den digitalen Aspekt des Begriffes ausmachen (HTML ist eine Auszeichnungssprache für Dokumente, die zur Anzeige in einem Webbrowser bestimmt sind; vgl. FASB, 2000, S. 30f.; Debreceny et al., 2002, S. 372). Diesbezüglich liegen diverse Studien vor, die den komplexen Prozess von der papierbasierten bzw. statischen Berichtspräsentation hin zur dynamischer Veröffentlichung von Berichten in Maschinensprache, was mittlerweile von der Berichtspraxis gefordert wird, ana-

lysiert haben (vgl. Locke, Rowbottom & Troshani, 2018, S. 2006ff.; Troshani, Locke & Rowbottom, 2019, S. 133ff.). Wohingegen neben der reinen Präsentation, die Ausgestaltung des Inhalts zwischen einer schlichten Zusammenfassung von Finanzkennzahlen, bis hin zur Ergänzung weitreichender Geschäftsinformationen neben einer längst ausführlichen FBE variiert (vgl. Debreceny et al., 2002, S. 374ff.). Das IASC hat ein dreistufiges Ausprägungsmodell entwickelt, das den Grad der Nutzung digitaler Berichterstattung von Unternehmen einordnet. In der schwächsten Ausprägungsform wird ein Onlinemedium (z.B. das PDF) rein als zusätzlicher Verbreitungskanal genutzt (vgl. IASC, 1999, S. 48; vgl. Lymer & Debreceny, 2003, S. 105; Anhang A). Inwiefern bei der Veröffentlichung durch das PDF von einer digitalen Berichterstattung die Rede sein kann, bleibt kritisch zu hinterfragen, da signifikante Vorteile von fortschrittlicheren Onlinemedien (wie z.B. die Wiedergabeleistungen von HTML) nicht in Anspruch genommen werden (vgl. NeidhartSchön AG, 2018, S. 7) und die Berichte im Wesentlichen ihre papierbasiert-statische Anordnung und Darstellung beibehalten (vgl. Beattie & Pratt, 2003, S. 161; Troshani & Rowbottom, 2021, S. 2; Anhang A). In einer zweiten Stufe des Modells sorgen Unternehmen dafür, dass die veröffentlichten Informationen z.B. durch Suchmaschinen für die Öffentlichkeit zugänglicher gemacht werden, wohingegen in der dritten Stufe Onlinemedien als Ausgangspunkte für eine Vielzahl fortschrittlicher Lösungen bereitgestellt werden, indem z.B. Analysetools oder andere interaktive Software neben der reinen Datenbereitstellung als erweitere Unterstützung zur Informationsverarbeitung als Informationsplattform zur Verfügung stehen (vgl. IASC, 1999, S. 48; Lymer & Debreceny, 2003, S. 105; Anhang A). Wobei digitale Medien prädestiniert dafür sind, neben den gesetzlichen Bestandteilen, dem Berichtsadressaten zusätzlich geforderte Informationen effizient zu präsentieren, die Verarbeitung zu unterstützen und verbundene Transaktionskosten zu reduzieren (vgl. Ettredge, Richardson & Scholz, 2001, S. 358; Beerbaum, Duncan, McNamara, Ikaheimo & Derichs, 2021, S. 1ff.). Es spielen insbesondere Vereinheitlichungstendenzen eine wachsende Rolle, bei denen ein standardisierter Rahmen im Fokus steht, um Daten einheitlich speichern, für die FBE aufbereiten und anschließend vom Adressaten angemessen verarbeiten zu können (vgl. Bonsón & Escobar, 2006, S. 299f.; Valentinetti & Muñoz, 2021, S. 449). In der zeitlichen Entwicklung in Hinblick auf die Verwendung von Onlinemedien und Technologien für die FBE einzelner Länder wird der zunehmende Einfluss zusätzlich erkennbar. Während die Unternehmenswebseite zur Kommunikation zur Jahrtausendwende verbreitet war, findet zu dieser Zeit eine digitale und harmonisierte FBE ausschließlich in fortschrittlichen Ländern, wie z.B. in Australien oder den Niederlanden, ausführlich statt. Innerhalb Europas war der Einfluss von Onlinemedien und Technologien in den einzelnen Ländern unterschiedlich ausgeprägt (vgl. IASC, 1999, S. 51), was in den letzten zwei Jahrzenten Veränderung erfahren hat und sich in den kommenden Jahren signifikant wandeln wird. Durch die Einführung eines verpflichtenden Berichtsformats innerhalb der EU im Jahr 2020 wurden neue regulatorische Anforderungen geäußert, die die Durchsetzung neuer digitaler Medien und Technologien innerhalb der EU fördern und insbesondere den Adressaten bei der Informationsverarbeitung unterstützen werden (vgl. Kovarova-Simecek, Pellegrini & Aubram, 2018, S. 161).

Die auslösenden Informationsforderungen an die traditionelle und damit digitale FBE ergeben sich aus den gesetzlich festgelegten Mindestausführungen, die vom Markt erwartet werden. Zudem sind die durch die individuellen Adressaten geforderten, nicht gesetzlichen Angaben Gegenstand der freiwilligen Offenlegung (vgl. Brandt & Hütten, 2003, S. 707f.). Diese Informationsfunktion mit dem Hintergrund einer Effizienz der Märkte aufgrund komplexerer Datenstrukturen und steigender Datenmengen tritt in den letzten Jahren zunehmend in den Vordergrund. Gemäß der EMH spiegeln Aktienkurse zu jedem Zeitpunkt alle verfügbaren, also veröffentlichten Informationen vollständig wider. Demzufolge ein Markt in Bezug auf Informationsmengen als effizient bezeichnet wird, wenn der Preis einer Aktie diese Informationsmengen zeitnah reflektiert (vgl. Fama, 1970, S. 383ff.). Vor dem Hintergrund steigender Datenkomplexität und den (kritisch zu beurteilenden) Annahmen der EMH, scheinen Unternehmen in Bezug auf die digitale Veröffentlichung von Informationen vor steigenden Herausforderung zu stehen. Vor allem die Datenaktualität und -integrität, die Relevanz der veröffentlichen Informationen und deren Vergleichbarkeit nehmen durch die Digitalisierung einen neuen Stellenwert für viele Berichtsadressaten ein (vgl. Williams, Scifleet & Hardy, 2006, S. 94).

Es kristallisieren sich im Wesentlichen vier Gruppen von Adressaten heraus, die eine besondere Berücksichtigung benötigen. Darunter zählen Expertengruppen (bestehend aus z.B. institutionellen Investoren und Kreditinstitute), privaten Kleinanleger, der Finanzvorstand und Abschlussprüfer (vgl. Beattie & Pratt, 2003, S. 166f.). Die institutionellen Investoren und Kreditinstitute sind bedeutungsvoll für die berichtenden Unternehmen und erheben einen Anspruch auf eine ausführliche und zeitnahe FBE. Für institutionelle Analysten ist die Effizienz und Qualität der Informationen von Priorität. Infolgedessen fehlende Berichtsqualität durch den Marktmechanismus (in Verbindung mit der EMH) zu verringerter Aufmerksamkeit, einem Mangel an potentiellen Investoren und Liquidität und damit zum Verlust von Unternehmenswert führen kann (vgl. Frank & Brümmer, 2003, S. 40; Losbichler & Otto, 2013, S. 28). Wohingegen private Kleinanleger, die nicht über fachliche Kompetenz, umfangreiches Wissen und notwendige Ressourcen verfügen, in keiner anspruchsstellenden Position stehen. Im Fokus dieser Gruppe liegt der reibungslose Zugang und die Verständlichkeit der teils komplexen Geschäftsangaben (vgl. Weber, Bramsemann, Heineke & Hirsch, 2017, S. 247f.). Hierbei sollte der Einfluss von Kleinanlegern nicht unberücksichtigt bleiben, da die Aufmerksamkeit von Kleinanlegern einen nachweisbar signifikanten positiven Effekt auf die Aktienliquidität auf kurze Sicht hat. Der Effekt scheint über längere Zeithorizonte nicht nachhaltig, die Aktienliquidität trotzdem in einer Periode mit positiver Marktstimmung bei Unternehmen

mit hoher Informationsunsicherheit nachgewiesenermaßen empfindlicher gegenüber der Liquidität von Kleinanlegern ist (vgl. Cheng, Chiao, Wang, Fang & Yao, 2021, S. 180). Es sei außerdem auf die sogenannte soziale Betriebserlaubnis („Social License to Operate“ (Meesters, Wostyn, Van Leeuwen, Behagel & Turnhout, 2021, S. 7)) hinzuweisen, folglich der Frage, ob die Aktivitäten des Unternehmens in den Augen der Gesellschaft als legitim erachtet werden. Der Ausdruck wird mit einer möglichen Missbilligung der Unternehmensaktivitäten und damit zu Widerstand und geschädigten Geschäftsinteressen in Verbindung gebracht, wobei Kleinanleger in der Position stehen, diese Legitimität kollektiv einfordern zu können (vgl. Demuijnck & Fasterling, 2016, S. 675). Ein Finanzvorstand hingegen ist auf der einen Seite gewillt, den Bedürfnissen der Adressaten nachzukommen, auf der anderen Seite ist die Veröffentlichung von Informationen grundsätzlich mit Risiken verbunden, da insbesondere die digitale Offenlegung von z.B. Betriebsgeheimnissen für die Konkurrenz zum Wettbewerbsvorteil werden kann. Der Finanzvorstand ist damit der Herausforderung ausgesetzt, wie steigende Erwartungen unterschiedlicher Adressaten mit den eigenen Interessen des Unternehmens in Einklang gebracht werden können (vgl. Weber et al., 2017, S. 255). Die weitreichende Offenlegung der Geschäftstätigkeit, insbesondere über Onlinemedien, birgt das Risiko, dass dieser frei zur Verfügung stehender, detaillierte Blick ins Unternehmensgeschehen nicht ausschließlich von möglichen Investoren, sondern zugleich von der Konkurrenz zur Entscheidungsunterstützung genutzt wird. Je detaillierter die (digitale) FBE ist, desto nutzenstiftender ist eine ausführliche Durchsicht für den Wettbewerb (vgl. Wallmann 1997, S. 112; Schruff & Kayser, 2002, S. 354f.). Der Konkurrenz werden digitale Informationen kostenlos zur Verfügung gestellt und die Möglichkeit gewährleistet, die Strategien und Ziele mit denen des eigenen Unternehmens zu vergleichen und das operative Geschäft vorteilhaft anzupassen, wonach der Finanzvorstand oftmals zurückhaltend im Hinblick auf die Offenlegung von (insbesondere freiwilligen) Informationen eingestellt ist (vgl. Weber et al., 2017, S. 255). Als vierte und letzte Adressatengruppe sollen Abschlussprüfer im Sinne und zum Schutz der anderen Adressaten agieren, indem durch die Prüfung der Einhaltung der Anforderungen gesetzliche Konformität hinterfragt wird, sodass Attribute wie die Authentizität der Informationen im Vordergrund stehen (vgl. Wagenhofer & Ewert, 2015, S. 410f.). Mit der steigenden Verbreitung von digitalen Geschäftsberichten, komplexen Unternehmensstrukturen und Wettbewerbskonstellationen und einer zugleich steigenden Anzahl unzweckmäßiger (und teilweise schlicht falscher) Informationen, ist eine digitale Weiterentwicklung von Prüfungsmethoden zum Schutz der Adressaten unausweichlich. Die Abschlussprüfung ist in einer Art und Weise anzupassen, sodass z.B. geprüfte von ungeprüften Veröffentlichungen durch ein digitales Prüfsiegel abgegrenzt sind, damit Informationsauthentizität gewährleistet werden kann und Adressaten schließlich abwägen können, ob die vorliegenden Informationen in eine Entscheidung einfließen sollen (vgl. Debreceny & Gray, 1999, S. 335ff.; Jones & Xiao, 2003, S. 147).

Zusammenfassend lässt sich ableiten, dass der Umfang und die Komplexität der traditionellen FBE durch die Internetfähigkeit und gestiegenen Anforderungen der Berichtsadressaten zugenommen haben und Unternehmen gleichzeitig vor der Herausforderung stehen, dass jede veröffentlichte Information auf die digitale Goldwaage gelegt wird und damit eine zeitnahe und risikobehaftete Resonanz auf dem Markt verursachen wird. Andererseits ist es durch die kontinuierliche Zunahme der Anzahl börsennotierter Unternehmen, der Erweiterung des Umfangs externer Veröffentlichungen und der riesigen Datenmengen und -strukturen für Adressaten herausfordernder geworden, die interdependenten (Finanz-)Geschäfte von Unternehmen zu verstehen (vgl. Wen, 2021, S. 2ff.). Mögliche Lösungen für die Herausforderungen der traditionellen FBE in Hinblick auf steigende Anforderungen und den damit einhergehenden wachsenden Informationsumfang liefern unterschiedliche neue digitale Medien und Technologien, wobei ohnehin der Wille zur Verbesserung der Kommunikation als ein wesentlicher Treiber der Digitalisierung in der FBE gesehen wird (vgl. Voelzkow, 2019, S. 4f.; Aaltonen & Penttinen, 2021, S. 66). Diesbezüglich machte bspw. ein weltweit tätiger Softwarekonzern schon vor zehn Jahren deutlich, dass die digitale Berichterstattung an Bedeutung gewinnt und dass das Unternehmen bereit ist, die Entwicklung in Richtung einer fortschreitenden Digitalisierung in der Berichterstattung umzusetzen: „Mit dem Geschäftsbericht vermitteln wir unseren Anlegern - privaten wie institutionellen - nicht nur das Zahlenwerk, sondern auch die Innovationskraft [...]. Dabei spielt die Online-Umsetzung des Geschäftsberichts eine immer größere Rolle. Aus diesem Grund haben wir gerne das Zukunftslabor Geschäftsbericht bei der Suche nach Vorschlägen unterstützt, wie das Online-Reporting Auswege aus der drohenden Überforderung des Kommunikationsinstruments Geschäftsbericht bieten kann“ (Meckel, Fieseler & Hoffmann, 2010, S. 9).

2.2. Big Data und Datenverfügbarkeit

Ausgehend von der grundlegenden steigenden Vernetzung der weltweiten Gesellschaft und Wirtschaft ist unter Big Data ein Datenphänomen zu verstehen, in dem Daten aus diversen Quellen mit ausgeprägtem Volumen, hoher Frequenz und sowohl steigender Vielfalt, als auch Unbeständigkeit gekennzeichnet sind und damit eine Herausforderung für Unternehmen und deren Speicher- und Verarbeitungsmethoden darstellen (vgl. Cockcroft & Russell, 2018, S. 323f.). Hingegen ermöglichen diese Eigenschaften, Entscheidungen auf einer angemessenen Grundlage in Echtzeit zu treffen und die Informationsaufbereitung, Unternehmenssteuerung und Berichterstattung damit datenfokussiert auszurichten (vgl. Bhimani & Willcocks, 2014, S. 479; Krahel & Titera, 2015, S. 409ff.; Vasarhelyi, Kogan & Tuttle, 2015, S. 381ff.). International tätige Unternehmen arbeiten täglich mit mehreren hundert Terabyte an Benutzerdaten[3], wickeln stündlich un-

[3]Ein Terabyte würde bspw. ungefähr einen 6,5 Millionen Seiten langen Bericht im PDF entsprechen.

zählige Kundentransaktionen ab und werden minutenweise mit unüberschaubaren Mengen an Emails und Kommentaren konfrontiert. Darunter steht ein Großteil dieser Daten unstrukturiert zur Verfügung (vgl. Kambatla, Kollias, Kumar & Grama, 2014, S. 2562). Es wird deutlich, dass Unternehmen im Datenzeitalter mit umfangreichen und volatileren Datengegebenheiten konfrontiert werden, Big Data zu einem bedeutsamen digitalen Phänomen der Gegenwart gehört und diese Gegebenheit eine zentrale Rolle bei zukünftigen Veröffentlichungen und darauf aufbauenden Entscheidungen spielt (vgl. H. Wang, Xu, Fujita & Liu, 2016, S. 748). Zudem gewinnen neben Finanzdaten gleichermaßen nicht-finanzielle Daten aus externen Quellen an Bedeutung. Die mehrdimensionale Analyse aus internen und externen (un-)strukturierten Daten kann wertvolle steuerungs- und berichtsrelevante Erkenntnisse für die Geschäftstätigkeit gewährleisten und das Überleben des Unternehmens insbesondere in Krisensituationen sichern. Es verhilft Unternehmen proaktiv und vorausschauend zu agieren und zu planen, da Echtzeitempfehlungen ermöglicht werden (vgl. Assunção, Calheiros, Bianchi, Netto & Buyya, 2015, S. 3; Wamba et al., 2017, S. 357). Viele Unternehmen nutzen vor diesem Hintergrund neben den eigenen internen Datenquellen längst weitaus mehr als fünf externe Datenquellen. Darüber hinaus finden sich wachsende digitale Geschäftsmodelle und Wettbewerbskonstellationen von Datenanbietern zur Datenmonetarisierung und Unternehmen sind bereit, Daten und Informationen in Cloud-basierten Ökosystemen freiwillig zu teilen, um das eigene Geschäft zu verbessern und die Innovationskraft zu stärken (vgl. Business Application Research Center, 2021; Datatrade GmbH, 2021; International Data Corporation, 2021). Durch die wachsende Verfügbarkeit von Echtzeitdaten, die digitale Abwicklung von Geschäftsvorfällen und damit eine angemessenen Informationsgrundlage können Produkte oder Dienstleistungen kundenorientiert zugeschnitten werden und zudem z.B. in (gegenwärtigen) Krisensituationen eine individuelle und transparente Interaktion mit Adressaten, Kunden und Lieferanten ermöglichen (vgl. Provost & Fawcett, 2013, S. 55; Weißenberger & Bauch, 2017, S. 208f.; Javaid et al., 2020, S. 419).

Diesbezüglich sei zu wiederholen, dass das Phänomen Big Data insbesondere durch die gestiegene Vielfalt von externen Quellen, die Unbeständigkeit von Daten durch ständige Aktualisierung und die hohe Datenunsicherheit durch z.B. Inkonsistenzen, Mehrdeutigkeit, Unvollständigkeit oder Manipulationsrisiken getrieben wird (vgl. Gandomi & Haider, 2015, S. 138; Emani, Cullot & Nicolle, 2015, S. 72). Hierbei beeinflusst Big Data die Digitalisierung in der FBE in vielerlei Hinsicht. Es findet ein Wandel statt, wie Geschäftsvorfälle abgewickelt, Jahresabschlüsse erstellt und geprüft werden und das Rollenbild des Rechnungswesens gesehen wird. Es werden multidisziplinäre Teams empfohlen, die unterschiedliches Fachwissen vereinen und die Genauigkeit und Zuverlässigkeit der FBE erhöhen, wobei Buchhalter für das Schaffen von Wissen zu neuen Technologien und Trends im Rechnungswesen zukünftig verantwortlich gemacht werden (vgl. Al-Htaybat & von Alberti-Alhtaybat, 2017, S. 868; Bogasiu & Ardeleanu, 2021, S. 295; Anhang B). Die Verfügbarkeit multidimensionaler Datenquellen, neuer Technologien und damit ausreichender Verarbeitungsmengen erhöhen die Effektivität und Effizienz der Buchhalter und werden das Rechnungswesen maßgeblich verändern und zu Möglichkeiten einer erhöhten Wertschöpfung aus Daten führen (vgl. Gulin, Hladika & Valenta, 2019, S. 432).

2.3. Fortschrittliche Technologien und weitere Trends

Es ist eine Balance zwischen der traditionellen FBE und neuen Onlinemedien und Technologien in der FBE gefragt, wobei ausführliche Kenntnisse zu fortschrittlichen technologischen Entwicklungen zum täglichen Geschäft werden (vgl. Dai, Shi, Meng, Wei & Ye, 2017, S. 975ff.; Gulin et al., 2019, S. 432; Fullana & Ruiz, 2021, S. 78f.). Trotz des längst angebrochenen Extensible Markup Language (XML)-Zeitalters sind die Probleme der semantischen Vereinbarkeit von Geschäftsinformationen weiterhin ungelöst. Da verschiedene Initiativen versuchen herauszufinden, wie Geschäftsangaben angemessen benannt, strukturiert und modelliert werden sollen, ist es elementar einen ganzheitlichen Berichtsansatz im Datenzeitalter zu verfolgen. Es war eine weit verbreitete Annahme, dass das Erstellen von XML-Ausdrücken ausreicht, um Vergleichbarkeit zu erreichen, hingegen entsprach diese Annahme nicht der Realität. XML garantiert nicht, dass die Geschäftsinformationen, die durch Geschäftsprozesse zwischen verschiedenen Unternehmen ausgetauscht werden, von den vorliegenden Systemen gleich interpretiert werden. Das liegt daran, dass die XML-Syntax primär für die Erstellung von weiteren Auszeichnungssprachen geeignet ist, die als Metadaten verwendet werden. Die Sprache selbst ist nicht darauf ausgelegt zu beschreiben, wie Informationen darzustellen sind. Wobei die Semantik diese Lücke abdeckt, indem Daten einheitliche Bedeutungen zugewiesen werden, die sowohl Menschen als auch Maschinen verstehen und verarbeiten können (vgl. Lampathaki, Mouzakitis, Gionis, Charalabidis & Askounis, 2009, S. 1045f.). Ein geeigneter Standard benötigt eine Beschreibungssprache, die einen strukturierten und semantisch einheitlichen Austausch von Informationen ermöglicht. Insbesondere XBRL erfüllt notwendige Funktionen, die die Digitalisierung in der FBE beschleunigen (vgl. Gostimir, 2015, S. 32; Debreceny & Gray, 2001, S. 47ff.; Bonsón, 2001, S. 101ff.). Es lässt sich zwischenzeitlich festhalten, das HTML rein die Optik von Informationen beschreibt, während XML-Weiterentwicklungen wie XBRL der Informationshülle eine substanzielle Bedeutung zuweisen (vgl. Aaltonen & Penttinen, 2021, S. 63).

Die Ausstattung von Gegenständen mit Sensoren zur Verwirklichung eines kontinuierlichen Datenaustausches treibt eine Möglichkeit zur globalen und digitalen Vernetzung. Hierbei steht der Begriff des Internet of Things (IoT), in der eine gemeinsame digitale Infrastruktur geschaffen wird, die nicht ausschließlich von Kontrolle geprägt ist, sondern zugleich den Zustand der verknüpfen Elemente kontinuierlich abrufbar macht und damit Transparenz schafft (vgl. Madakam, Ramaswamy & Tripathi, 2015, S. 164; Valentinetti & Muñoz, 2021, S. 449). Ein Schlüsselkonzept in Verbindung

mit dem IoT ist das erläuterte Phänomen Big Data, da diese Gegebenheit auf dem Zusammenspiel von Faktoren, die zur Produktion von zahlreichen und schnelllebigen Daten aus verschiedenen Quellen führen, basiert. Angemessen berücksichtigt können beide Phänomene in Kombination Chancen in der FBE erzeugen und den Informationsgehalt und die Transparenz von Veröffentlichungen steigern. Darunter haben Analysen gezeigt, dass eine Vielzahl von Aspekten der digitalen FBE vom Einfluss des IoT betroffen sind, die einer kritischen Würdigung zu unterziehen sind. Grundsätzlich scheint eine erhöhte Datenzufuhr und -transparenz durch solche Praktiken vorteilhaft für die FBE (vgl. Valentinetti & Muñoz, 2021, S. 449ff.).

Neben neuen digitalen Berichtssprachen und der fortschreitenden Vernetzung von Computern für einen kontinuierlichen Datenaustausch sind Cloud- und Datenbanksysteme vorteilhafte Anwendungen der Datenverwaltung, die für diese eintretende unbeständige Datenüberlastung geeignet sind und insbesondere Cloudsysteme in kontakteinschränkenden Krisensituationen viele Vorteile zur Virtualisierung und Skalierung mit sich bringen (vgl. Bogasiu & Ardeleanu, 2021, S. 296). Für vielen Unternehmen steigt die Datenmenge schneller, als das die Leistung der vorhandenen Infrastruktur diese angemessen speichern und verarbeiten kann. Vor allem kleine Unternehmen greifen auf Cloud Computing-Systeme von Drittanbietern zurück und erhoffen sich damit Effizienz und Flexibilität, sowie einen verringerten Wartungsaufwand und damit Kostenersparnisse (vgl. Demirkan & Delen, 2013, S. 412f.). Ein darauf aufbauendes Verfahren zum Speichern und Verarbeiten von Daten und Informationen ist die Verwendung von Not only Structured Query Language (NoSQL)-Datenbanken. Eine nicht-relationale Speicherung führt zu steigender Leistungsfähigkeit und eignet sich zur Verarbeitung unstrukturierter Daten (vgl. Assunção et al., 2015, S. 7). Unüberschaubare Datenmengen können auf mehreren dezentralen Servern gespeichert werden und dadurch explorative und prognostizierende Datenverarbeitungen erlauben. Unternehmen nutzen diese Art von Datenbanksystemen, da Daten in Zeiten von Big Data und IoT eine hohe Unbeständigkeit aufweisen, unstrukturiert vorliegen und solche Systeme durch den öffentlichen Zugang günstiger als traditionelle Verfahren sind (vgl. Moniruzzaman & Hossain, 2013, S. 1f.). Darüber hinaus verfügt eine neue Generation von Datenbanksystemen die In-Memory-Technologie. Diese adressiert eine hohe Frequenz und ermöglicht es, Echtzeitanalysen durchzuführen und somit einen Mehrwert aus den vorliegenden (Roh-)Daten zu generieren (vgl. Kataria & Mittal, 2014, S. 109). Diese Technologie gewährleistet eine hohe Zugriffsgeschwindigkeit, indem der Arbeitsspeicher als Verarbeitungsinstanz genutzt wird (In-Memory), wobei die Daten auf Festplatten rein für Sicherheitsaspekte abgelegt werden (vgl. C. P. Chen & Zhang, 2014, S. 15ff.). Schließlich stellt sich die Frage, ob diese Infrastrukturanwendungen zur Verarbeitung von (un-)strukturierten Daten neben Vorteilen der Dezentralisation und Virtualisierung dem Adressaten eine effiziente und umfangreiche Verarbeitung von Informationen ermöglichen.

Die Automatisierungstendenz von RPA nutzt konfigurierte Software, um automatisch bestehende Anwendungen wahrzunehmen, Daten zu beeinflussen, Transaktionen auszuführen und mit anderen Systemen selbstständig zu kommunizieren. Typische Anwendungsgebiete von RPA sind das Öffnen, Lesen und Senden von z.B. E-Mails, das Archivieren, Extrahieren, Aktualisieren, Validieren und Eingeben von Daten über mehrere Anwendungen hinweg (darunter zudem über Formate wie HTML oder XBRL), die Datenverarbeitung und -formatierung und schließlich die Entscheidungsunterstützung (vgl. Jędrzejka, 2019, S. 139f.). Des Weiteren liegen Methoden der Kryptografie vor, die neben einer Automatisierung von bestimmten Prozessen gleichzeitig die Sicherheit und Integrität von Geschäftsinformationen gewährleisten. Durch das Internet befähigte Anwendungen verwenden solche Verfahren, um weltweit und reibungslos Informationen auszutauschen (vgl. Hildner & Danzmann, 2017, S. 385). Solche Anwendungen bilden weitere Lösungen für die Herausforderung der Daten- und Informationsaufbereitung und eröffnen Möglichkeiten einer digitalen Informationsvalidierung. Das resultierende Netzwerk aus verteilten Rechner gewährleistet ein gewisses Maß an Öffentlichkeit, wodurch mögliche Fehler, die zu Systemproblemen führen, verdrängt werden (vgl. Voshmgir, 2016, S. 6f.). Die Blockchain (BC) als ein beispielhaftes Verfahren wird als Verkettung von Transaktionsblöcken verstanden, die einen sicheren Austausch von Informationen gewährleistet. In der BC werden Transaktionen parallel aufgezeichnet und auf dezentralen Datenbanken abgebildet, wodurch eine Legende der durchgeführten Transaktionen inklusive der Möglichkeit eines Prüfungspfades entsteht (vgl. Nakamoto, 2008, S. 1f.; Fuller & Markelevich, 2020, S. 34). Da Geschäftsvorfälle zeitnah zu buchen und validieren sind, ist dieser Prozess mit einer hohen Ressourcenintensivität verbunden, der hingegen ein wesentlicher und notwendiger Bestandteil im Rechnungswesen darstellt (vgl. Cao, Cong, Han, Hou & Yang, 2020, S. 20). Die BC kann hierbei Effizienz und Vertrauen durch das automatische Konsensverfahren schaffen und Datensicherheit und -integrität gewährleisten (vgl. Dai et al., 2017, S. 975ff.). Anschließende Auswirkungen auf andere digitale Medien und Technologien innerhalb der FBE gilt es damit abzuschätzen, wobei die Berichtersteller und Adressaten von den technologischen Möglichkeiten zu profitieren scheinen. Fokussierte Schwerpunkte der gesichteten Studien elaborieren hingegen zu erwartende Auswirkungen der Technologie auf das gesamte Rechnungswesen. Zum einen wird die grundsätzliche Unveränderlichkeit von Daten und Informationen genannt, die durch Transaktionsaufzeichnungen gewährleistet wird und eine erhöhte Transparenz schafft, was in Zeiten von Big Data Vorteile birgt (vgl. Gulin et al., 2019, S. 428; Fullana & Ruiz, 2021, S. 78f.; Min, 2019, S. 9). Zum anderen die Zuverlässigkeit von Daten, die auf der Fähigkeit der BC-Technologie basiert und damit das Potential bietet, die Glaubwürdigkeit von Buchhaltungsdaten und damit veröffentlichen Berichten zu erhöhen und gleichzeitig zuverlässige Beweise für bewusste Manipulation oder unbewusste Fehler zu liefern, wodurch die Abschlussprüfer als einer der primären Adressa-

ten unterstützt werden und schließlich ein kontinuierlicher Erstellungs- und gleichzeitig Prüfungsprozess angestoßen werden kann (vgl. Smith, 2018, S. 243ff.; Fullana & Ruiz, 2021, S. 78f.; Fuller & Markelevich, 2020, S. 34).

Ein weiteres digitales Medium, das neben neuen Speicher- und Validierungsmöglichkeiten und Automatisierungstendenzen für die digitalen FBE zum Vorteil werden kann, ist KI. Dieser Bereich besitzt eine nicht eindeutig einschätzbare Wirkung auf die Gesellschaft, Wirtschaft und menschliche Handlungsweisen (vgl. Fraunhofer IAIS, 2019, S. 7). Als KI ist die Fähigkeit eines Programmes zu verstehen, Aufgaben oder Probleme zu lösen, die grundsätzlich menschenähnlichen Verstand, eine kritische Wahrnehmung und selbstständiges Lernen benötigen. Dazu bedarf es einer Regulierung, die die Zweifel der involvierten Nutzer beseitigen. Um dies zu gewährleisten, muss eine KI-Anwendung entsprechend programmiert worden sein, sodass sie verantwortungsbewusst arbeitet und gleichzeitig verpflichtende Rahmenbedingungen berücksichtigt (vgl. Fraunhofer IAIS, 2019, S. 8). Insgesamt wird zwischen schwacher, starker und überdurchschnittlicher KI unterschieden. Schwache KI-Programme helfen bei Problemlösungen in unverkennbaren und vorbestimmten Umgebungen. Hierzu zählen Anwendungen wie z.B. Service-Roboter (vgl. Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V./ Deutsches Forschungszentrum für Künstliche Intelligenz, 2017, S. 31). Aktuell im Einsatz befindliche KI-Anwendungen sind unter dem Begriff der schwachen KI zusammenzufassen. Wohingegen starke KI umfangreichere Probleme lösen kann, die grundsätzlich ausgeprägte menschliche Fähigkeiten benötigen, während eine überdurchschnittliche KI menschliche Fähigkeiten in vielen Bereichen übertreffen würde (vgl. Huang & Rust, 2018, S. 158; W. Wang & Siau, 2019, S. 62). Machine-Learning (ML) ist eine Teilmenge von KI und bezeichnet Programme, die das selbständige Lernen optimieren und sich somit ohne weitere externe Eingriffe verbessern können, wobei große Datenmengen benötigt werden, damit vordefinierte Anweisungen präzise gelernt und ausgeführt werden (vgl. Bundesministerium für Wirtschaft und Energie, 2019, S. 5; W. Wang & Siau, 2019, S. 63), während Deep Learning (DL) ein Verfahren mit weiterem Detaillierungsgrad darstellt. DL verwendet künstliche neuronale Netze (KNN), deren Struktur sich an dem neuronalen Netzwerk im menschlichen Gehirn orientiert. Mit dieser Struktur können multidimensionale Daten und Muster erkannt und nichtlineare Aufgaben bewältigt werden (vgl. LeCun, Bengio & Hinton, 2015, S. 436), was in Zeiten von Big Data und komplexen Datenstrukturen zu einem Vorteil in der FBE werden kann. Anwendungen für den Einsatz von KI in der FBE sind bspw. das automatisierte und intelligente Einscannen und Buchen von diversen Dokumenten wie z.B. Rechnungen, die Prognose von Inventarbewegungen, die selbstständige Durchführung von Wertminderungstests oder die Unterstützung bei Abstimmungstätigkeiten für z.B. Konzernabschlusszwecke (vgl. Wilson & Sangster, 1992, S. 65; Smith, 2018, S. 242f.; Petkov, 2020, S. 103; Leitner-Hanetseder, Lehner, Eisl & Forstenlechner, 2021, S. 9f.).

Es findet darüber hinaus eine umfassende Veränderung der traditionellen Paradigmen der Informationssystementwicklung statt. Sogenannte No-Code/Low-Code-Entwicklungsumgebungen ermöglichen die Erstellung von Anwendungssoftware über grafische und intuitive Benutzeroberflächen anstelle über traditionelle Programmiersprachen. Mit solchen Werkzeugen können entwicklungsferne Anwender schnell und selbständig neue Applikationen erstellen oder Funktionen hinzufügen, ohne tiefgreifende Programmierungsfähigkeiten zu besitzen, wodurch schließlich der Prozess der FBE eigenständig digitalisiert werden kann (vgl. Chang & Ko, 2017, S. 50ff.; S. Wang & Wang, 2021, S. 1ff.; Talesra & Nagaraja, 2021, S. 346ff.). Die angemessene Verknüpfung solcher neuen Entwicklungstendenzen und den angesprochenen Medien und Technologien kann schließlich zu Möglichkeiten von z.B. Hyperautomatisierung führen. Darunter wird die Automatisierung von Arbeitsschritten mit einem breiten Geschäftsumfang, integriert mit reaktionsfähigen Arbeitskräften in Kombination von RPA mit KI und tiefgreifenden ML-Funktionalitäten, verstanden. Der Einfluss von Hyperautomatisierung auf die digitale FBE sollte unterdessen einer kritischen Würdigung unterzogen werden, da der Begriff heute zu einem der Hauptgründe gehört, die Führungskräfte dazu bewegen, die Unternehmenssteuerung und ihre Managementpraktiken ganzheitlich zu überdenken (vgl. Lasso-Rodriguez & Winkler, 2020, S. 7). Das Überdenken weiterer verwandter Praktiken geht hingegen von den ersten Überlegungen hervor, der Disziplin des Quantum Computings (QC) einen Platz im Rechnungs- und Finanzwesen der Zukunft zu schaffen. Es handelt sich, im Gegensatz zu traditionellen Lösungen, um eine neue Methode zur grundlegenden Ausführung von Programmen mit höherer Geschwindigkeit für Aufgaben, bei denen große und mehrdimensionale Datenmengen in Echtzeit verarbeitet werden müssen, wobei der Einfluss gegenwärtig kaum einzuschätzen ist (vgl. Endeavor Business Media, 2018).

3. Folgen und Konsequenzen digitaler Berichterstattung

3.1. Wachsender Umfang der digitalen Berichterstattung

Technologische Entwicklungen bieten neue Möglichkeiten für die tägliche Aufbereitung und Weiterverarbeitung von Daten und Informationen. Das Rechnungswesen basiert bis dato auf einer periodischen Überprüfung, Veröffentlichung und Analyse von Informationen, wohingegen vermehrt Berichtsadressaten eine kontinuierliche FBE mit Echtzeitfunktionalitäten von Vorteil sehen. Die Implementierung und der Übergang von einem periodischen Berichtswesens zu einer proaktiven und kontinuierlichen Bereitstellung von (nicht-) finanziellen Informationen wird gegenwärtig deutlich vom Markt gefordert (es findet ein Übergang vom Pull- zum Push-Reporting statt, vgl. Mittelbach-Hörmanseder et al., 2020, S. 48) und wird dahingehend eine gewisse Übergangszeit und einige Implementierungsschritte benötigen. Abgeleitet daraus, dass der zeitliche Aspekt der zunehmend digitalen FBE in Form von integrierten Zwischenberichten bereits verkürzt

wurde, ist eine Würdigung der damit verbundenen Folgen und Konsequenzen notwendig (vgl. Jones & Xiao, 2003, S. 136). Während sich Zwischenberichte mittlerweile in der Berichterstattung von Unternehmen flächendeckend etabliert haben, ist die Verbreitung der darauf aufbauenden Echtzeitberichterstattung noch gering. Der Übergang zur kontinuierlichen FBE wird zu einer Verschiebung und Unterbrechung der Prozesse und des aktuellen Rollenverständnisses führen. Insbesondere in neuen digitalen Märkten und Geschäftsmodellen, in denen nahezu kontinuierlich Unmengen an (Roh-) Daten produziert werden und die Adressaten eine zunehmende Vielfalt an Informationen erwarten, erscheinen die derzeit statischen Buchhaltungsprozesse und -verfahren nicht ausreichend (vgl. Smith, 2018, S. 240ff.; Gulin et al., 2019, S. 432).

Wohingegen durch die wachsenden Berichtsanforderungen und damit umfangreicheren Bedürfnissen der Adressaten innerhalb der digitalen FBE die gegensätzliche Tendenz vorzufinden ist, im Sinne der Informationstransparenz und -effizienz den Fokus von Veröffentlichungen rückwirkend auf den ursprünglichen Teil der Finanzangaben zu reduzieren („Reduce to the max“ (Meckel et al., 2010, S. 31ff.)). Diesbezüglich ist an die dritte Stufe des IASC-Ausprägungsmodells zurück zu erinnern, in dem Unternehmen ein Onlinemedium als Ausgangspunkt für eine Vielzahl weiterer Optionen zur Unterstützung der Datenwertschöpfung nutzen. Diese Unterstützung ist aufgrund der gestiegenen Bedürfnisse mittlerweile zur Notwendigkeit geworden. Während der papierbasierte Bericht in der Skalierung unmittelbar an seine Grenzen stößt, bieten integrierte Onlinemedien die Chance die steigenden Informationsbedürfnisse aller Adressaten gezielt zu befriedigen (vgl. Godemann & Herzig, 2012, S. 238). Wobei die Gefahr einer Informationsüberlastung zu berücksichtigen ist und vermieden werden sollte, da die Adressaten des digitalen Geschäftsberichts durch die Fülle an Informationen dem Risiko ausgesetzt werden, den Überblick und die vermittelte Botschaft nicht greifen zu können (vgl. IASC, 1999, S. 48; Barrantes & Sagmüller, 2014, S. 62; Weißenberger & Bauch, 2017, S. 212; Anhang A). Damit ist die Herausforderung gegeben, dass Adressaten nicht eindeutig und zuverlässig zwischen den relevanten und irrelevanten Geschäftsangaben unterscheiden können und schließlich Entscheidungen auf einer unangemessenen und teilweise fehlerhaften Informationsgrundlage treffen (vgl. Menacher, Schütt & Sellhorn, 2017, S. 217; Zülch & Weuster, 2018, S. 102), was dem primären Zweck des Rechnungswesens widerspricht. Es ist eine Reformierung der Berichtsprüfung in Erwägung zu ziehen, sodass insbesondere geprüfte von ungeprüften Informationen abgegrenzt und durch ein Prüfsiegel digital gekennzeichnet wurden, wodurch die geforderte Informationsrelevanz und -authentizität innerhalb der digitalen FBE gewährleistet werden kann (vgl. Jones & Xiao, 2003, S. 147). Demzufolge ist absehbar, dass sowohl Unternehmen, als auch Informationsadressaten sich auf das Wesentliche zu konzentrieren haben. Eine umfangreiche Bereitstellung über Onlinemedien scheint involvierte Parteien zu überfordern, woraus sich ein herausfordernder Balanceakt für Unternehmen im angemessenen Umfang einer digitalen FBE ergibt (vgl. Meckel et al., 2010, S. 31f.).

Der Umfang der FBE wird zudem durch unternehmensspezifische Faktoren und Bedingungen der Umwelt beeinflusst. Dies belegen Studien aus einer unternehmensspezifischen Perspektive, wonach die Unternehmensgröße einen signifikanten Einfluss auf das Ausmaß der FBE hat (vgl. Ashbaugh, Johnstone & Warfield, 1999, S. 251; Debreceny et al., 2002, S. 377ff.; Oyelere, Laswad & Fisher, 2003, S. 30ff.). Erwiesen wird dieser Einfluss z.B. durch verringerte Offenlegungskosten mit zunehmender Unternehmensgröße (vgl. Lang & Lundholm, 1993, S. 250). Darüber hinaus legen Unternehmen ab bestimmten Größenklassen einen gewissen Fokus auf eine angemessene Eigenkapitalquote und haben dadurch zum Ziel, Informationen insbesondere für die Expertengruppe (z.B. für institutionelle Investoren) möglichst transparent und intuitiv zu präsentieren. Dadurch soll der Ruf des Unternehmens gestärkt und Informationsasymmetrien zwischen Management und Aktionären reduziert werden (gemäß der Prinzipal-Agenten Theorie; vgl. Jensen & Meckling, 1976, S. 305ff.; Firth, 1979, S. 273ff.; Verrecchia, 1983, S. 179ff.; Debreceny et al., 2002, S. 371ff.). Wohingegen aus länderspezifischer Perspektive der Umfang der digitalen FBE in den einzelnen Ländern teilweise enorm variierte. Die Ausprägung der Offenlegung von Informationen war und ist vorwiegend geprägt durch rechtliche Gegebenheiten und ob das Regelwerk eines Landes die Implementierung von Onlinemedien im Rechnungswesen fördert bzw. einschränkt (vgl. Deller, Stubenrath & Weber, 1999, S. 355ff.; Ball, Kothari & Robin, 2000, S. 10ff.; Debreceny et al., 2002, S. 376ff.). Wodurch widersprüchliche Anliegen und Konsequenzen in der Zustimmung bzw. Ablehnung neuen Medien und Technologien in der Berichterstattung durch die partizipierenden Akteure erneut ersichtlich werden.

Ferner dieser Gegebenheiten erwarten Adressaten umfangreichere Rahmenwerke, sowie ganzheitliche Darstellungsformate wie z.B. in Form des Integrated Reportings (IR) (vgl. Kretschmer, 2014, S. 169; Engelbrechtsmüller & Kerschbaumer, 2014, S. 11). Eine integrierte Berichterstattung steigert den Informationsumfang, da Geschäftsangaben über das Unternehmen und seine Grenzen hinweg gefordert sind. Abgeleitet von der Unternehmensstrategie und dem Geschäftsmodell sollen Einzelheiten über die gesamte Wertschöpfungskette berichtet werden und gleichzeitig eine Konvergenz des externen und internen Rechnungswesens bewirken (vgl. Weißenberger, 2014, S. 440; De Villiers, Venter & Hsiao, 2017, S. 937ff.). Laut dem International Integrated Reporting Council (IIRC) umfasst die Prüfung eines integrierten Berichts unterschiedliche Teilbereiche, die sich von einem Überblick der Organisation, über die Chancen und Risiken des Geschäftsmodells, bis hin zu möglichen Geschäftsperspektiven erstrecken und von daher den Umfang traditioneller FBE weit überschreiten (vgl. IIRC, 2013, S. 7; Engelbrechtsmüller & Kerschbaumer, 2014, S. 11; Weißenberger, 2014, S. 441). Diesbezüglich steht aus finanzieller Perspektive die Befriedigung der Bedürfnisse der Aktionäre im Vordergrund, wohingegen anderen Adressaten wie z.B.

Lieferanten andere relevante Informationen zu berichten sind. Da die Informationsbedürfnisse der unterschiedlichen Adressaten an Uneinheitlichkeit gewinnen und teilweise im Gegensatz zueinanderstehen, werden Friktionen in Kommunikationsprozessen erkennbar. Darüber hinaus sind neue Interessengruppen in der FBE zu berücksichtigen, wodurch es für Unternehmen zu einer wachsenden Herausforderung wird, allen Adressaten und deren divergierenden Anforderungen gerecht zu werden (vgl. Weißenberger, 2014, S. 445; Voelzkow, 2019, S. 3f.). Unweit des IR berücksichtigen institutionelle Investoren, sowie private Kleinanleger zunehmend Nachhaltigkeitsaspekte in ihren Anlageentscheidungen. Demzufolge Unternehmen z.B. Kriterien in den Bereichen Umwelt und soziale Verantwortung zu entwickeln haben, was die Komplexität und den Umfang der FBE weiter erhöht (vgl. Speich, 2014, S. 210ff.).

Zusammenfassend lässt sich festhalten, dass mehr Informationsanforderungen den je an Unternehmen und deren Wertschöpfung gerichtet werden, neue Stakeholder und Leistungsindikatoren in der FBE zu berücksichtigen sind und ein angemessener Balanceakt zwischen einer umfangreichen und kontinuierlichen Informationsbereitstellung und dem Risiko der Informationsüberlastung von Unternehmen erwartet wird. Dahingehend scheint ein angemessener und effektiver Einsatz von unterschiedlichen digitalen Medien, wie XBRL und Technologien wie z.B. In-Memory-Datenbanken in der FBE diese Herausforderungen bewältigen zu können (vgl. Gerding, 2016, S. 1176f.) und stellen damit erhebliche zukünftige Chancen in der Digitalisierung in der FBE eines Unternehmens dar.

3.2. Erhöhte Datenwertschöpfung und Informationstiefe

Der intrinsische Wert von Daten bedarf im Datenzeitalter einer kritischen Würdigung. Daten haben laut Studien eine geringe Informationsdichte, insbesondere wenn die ursprüngliche Form unverarbeitet genutzt wird. Erst durch die Analyse einer großen Anzahl von (Roh-)Daten kann eine erhöhte Wertschöpfung ermöglicht werden (vgl. Gandomi & Haider, 2015, S. 139). Vorwiegend bei Trends wie Big Data oder IoT ist zu berücksichtigen, dass (Roh-)Daten ohne weitere Verarbeitung keinen Beitrag zur Entscheidungsunterstützung (aus interner bzw. externer Sicht) liefern können (vgl. Tien, 2013, S. 127). Für eine Entscheidungsgrundlage mit fortschrittlichem Datenfokus müssen demnach Verfahren implementiert werden, die mit Fähigkeiten aus den Bereichen Statistik, Data und Process Mining, ML und Visualisierung einhergehen (vgl. LaValle, Lesser, Shockley, Hopkins & Kruschwitz, 2011, S. 22; Fisher, DeLine, Czerwinski & Drucker, 2012, S. 51). Hierbei liegt das grundsätzliche Ziel von Datenanalysen darin, signifikante Erkenntnisse durch Beobachtungen, Messungen und Verarbeitung zu extrahieren, um Fehlerquellen zu identifizieren, Prognosen zu erstellen und die Entscheidungsfindung zu unterstützen, und insbesondere Analysen in Zeiten von Big Data und IoT benötigen fortgeschrittene Analysetechniken (vgl. H. Hu, Wen, Chua & Li, 2014, S. 672).

Neue Verfahren zur Datenanalyse steigern die Wertschöpfung deutlich. Um eine angemessene Auswertbarkeit zu gewährleisten und Mehrwerte zu gewinnen, müssen Daten angemessen extrahiert und für die anschließende Verarbeitung bereinigt und gegebenenfalls zusammengefasst bzw. klassifiziert werden (vgl. Gentsch & Kulpa, 2016, S. 34). Insbesondere Daten aus externen Quellen (deren Anzahl stetig steigt) müssen einer Qualitätssicherung unterzogen und anschließend klassifiziert werden, um die Entscheidungsfindung angemessen unterstützen zu können und keine Fehlschlüsse aus den veröffentlichten Informationen zu ziehen. Im elementaren und aufwendigen Prozess der Datenbereinigung sind Fehlerquellen, sowie Redundanzen zu identifizieren und gegebenenfalls zu beseitigen, falls sonst keine Datenintegrität gewährleistet werden kann (vgl. H. Hu et al., 2014, S. 664f.). Da die Anzahl der Daten exponentiell wächst, werden neue Medien und Technologien erforderlich, um diese unbeständigen und kurzlebigen Daten angemessen aufbereiten und daraus Wert schöpfen zu können (vgl. Gärtner & Hiebl, 2018, S. 10). Aufwendige Modelle für Simulationszwecke und zur Analyse von z.B. Marktrisiken werden durch diese erhöhte Informationstiefe unterstützt. Unterdessen wird die Distanz zwischen der Datenanalyse und dem Treffen einer Entscheidung verringert, wodurch kurzfristiges und agiles Handeln bei Marktveränderungen oder Krisensituationen ermöglicht werden (vgl. Bhimani & Willcocks, 2014, S. 486). Außerdem können durch die Verarbeitung von ganzheitlichen (un-)strukturierten Datenmengen, anstelle von Stichproben, unbekannte Optimierungspotentiale und Manipulationsrisiken aufgedeckt werden, was die Qualität der resultierenden Informationen und damit die Effizienz der digitalen FBE signifikant erhöht und gleichzeitig Abschlussprüfer unterstützt (vgl. Weißenberger & Bauch, 2017, S. 210). Solche explorativen Analysen sind jedoch ausschließlich bei Informationen, die auf bestehenden ökonomischen Transaktionen basierenden, einzusetzen, da ab bestimmten analytischen Informationsmengen die Gefahren von Verzerrungen oder falschen Rückschlüssen vorliegen (vgl. Bhimani & Willcocks, 2014, S. 486f.; Richins, Stapleton, Stratopoulos & Wong, 2017, S. 67f.).

Es liegt eine Vielzahl von Analysetechniken zur Erhöhung der Wertschöpfung aus Daten vor. Eine Vielzahl der unstrukturierten (Roh-)Daten steht in Textform zur Verfügung. Fortschrittliche Textanalysen entnehmen diesen unstrukturierten Texten strukturierte Informationen, können Inhalte zusammenfassend ermitteln und über Social-Media-Kanäle Meinungen und Bewertungen von Kunden analysieren und in die Geschäftsanalyse und FBE einfließen lassen (vgl. H. Chen, Chiang & Storey, 2012, S. 1175f.). Darüber hinaus arbeiten Audioanalysen mit von Natur aus unstrukturierten Audiodateien, wodurch z.B. Callcenter durch aufgenommene Telefonate das Kaufverhalten der Kunden analysieren und mögliche Chancen, sowie Herausforderungen ermitteln können. Wohingegen Videoanalysen Methoden zur Erfassung relevanter Informationen aus unstrukturierten Videodateien beinhalten. Diese unterschiedlichen Analysetechniken erhöhen die Wertschöpfung aus vorliegenden Daten und haben

positive Implikationen auf die Transparenz der Geschäftstätigkeit und damit auf die dadurch anschließend digitalisierte FBE (vgl. Gandomi & Haider, 2015, S. 140ff.). Zur angemessenen Verarbeitung von Datensätzen in Zeiten von Big Data und IoT werden darüber hinaus zunehmend fortschrittliche Data Mining- und ML-Anwendungen herangezogen, um eine effizientere Verarbeitung und höhere Vorhersagegenauigkeit zu gewährleisten. Data Mining identifiziert Muster und Strukturen und kann somit die Unsicherheit und Mehrdeutigkeit von Daten reduzieren. Das Ziel von ML ist die Fähigkeit, eine datenfokussierte und intelligente Entscheidungsfindung zu erlernen. DL-Algorithmen extrahieren hingegen in einem mehrschichtigen Prozess bestimmte Merkmale, Variablen und Muster aus den umfangreichen Daten, bauen damit einen digitalen Erfahrungsschatz auf und erhöhen somit die zukünftige Wertschöpfung aus auftretenden Daten für die FBE (vgl. Kowalczyk & Buxmann, 2014, S. 269; Najafabadi et al., 2015, S. 2f.; Müller, Junglas, Brocke & Debortoli, 2016, S. 294). Inwiefern die erhöhte Datenwertschöpfung einen Beitrag zur Informationsgewinnung und zu Chancen, Herausforderungen und Risiken innerhalb der digitalen FBE führt, ist abschließend kritisch zu würdigen.

3.3. Vereinheitlichung der Berichterstattung in der EU

3.3.1. XBRL und European Single Electronic Format

Die heterogenen Rahmenbedingungen, Formate und Technologien der aktuell praktizierten FBE erschweren eine einheitliche und effiziente Verarbeitung von Informationen für Zwecke der Unternehmensanalyse und -vergleiche. Die Vereinheitlichung der FBE wird diesbezüglich seit geraumer Zeit von vielen Ländern gemeinschaftlich angestrebt. Die USA und China führten bspw. einen einheitlich elektronischen Berichtstandard ein, der die Transparenz der digitalisierten FBE mit Hilfe von XBRL fördert (vgl. Beerbaum & Piechocki, 2016, S. 1f.). Vor über zehn Jahren verlautbarte die verantwortliche Behörde in den USA die Verwendung von XBRL für die Veröffentlichung bestimmter Berichte (z.B. Quartalsberichte; vgl. Liu, Wang & Yao, 2014, S. 71). Daran anschließend erkannte die EU die Notwendigkeit, sowie die Vorteile eines solchen Standards. Zum 01. Januar 2020 sind europäische Unternehmen, die den Kapitalmarkt beanspruchen dazu verpflichtet, unterschiedliche Berichtsbestandteile nach Anweisungen des verlautbarten Formats (gemäß dem European Single Electronic Format (ESEF)) zu erstellen. In diesem Sinne wird für eine angemessene FBE ein schlichtes PDF-Dokument als Informationsmedium nicht mehr ausreichen. Es ist notwendig, fallbezogen entweder HTML oder eine Erweiterung mittels Inline Extensible Business Reporting Language (iXBRL) anzuwenden (namentlich Extensible Hypertext Markup Language (XHTML)) (vgl. Di Fabio, Roncagliolo, Avallone & Ramassa, 2019, S. 33). Es legen Studien nahe, dass die Lesbarkeit von HTML-Berichten durch das XBRL-Mandat verschlechtert wurde, was auf einen parallelen Aufwand der Berichterstellung zurückzuführen ist und mit einer reduzierten Aufmerksamkeit bei der Erstellung von HTML-formatierten Berichten erklärt werden kann, wonach die iXBRL-Erweiterung eine geeignete Zusammenführung darstellt (vgl. Li, Zhu & Zuo, 2021, S. 27f.). Ein iXBRL-Finanzbericht ähnelt dem XBRL-Bericht mit dem Unterschied, dass iXBRL über XHTML dargestellt wird. Damit wird unweit der grundlegenden maschinellen Umschreibung von Informationen eine menschliche Lesbarkeit gleichzeitig gewährleistet (vgl. Basoglu & White Jr, 2015, S. 190; Bravidor & Loof, 2020, S. 178) und vereint damit die Vorteile beider Welten. Das neu eingeleitete Berichtsformat soll mit seinen Funktionalitäten die FBE in Europa digitalisieren und vereinheitlichen. Bei der Umsetzung liegen dennoch Unterschiede in den einzelnen EU-Mitgliedsstaaten vor, sodass es trotz Bemühungen zur Vereinheitlichung zu Unterschieden bei der zusammenführenden Berichtspraxis kommt. Die maßgebende Frage ist, in welchem Land die Offenlegung in XBRL für welche Berichtselemente verpflichtend ist, welche Organisationen involviert sind und welche Vorschriften zugrunde gelegt werden (vgl. Bartolacci, Caputo, Fradeani & Soverchia, 2020). Hierbei sind die Inhalte und Ergebnisse der Hausarbeit „ESEF in der EU - State of the Art und Implementierung“ eingegangen, die im Jahr 2020 dem Lehrstuhl für BWL, insbesondere Wirtschaftsprüfung der Heinrich-Heine-Universität Düsseldorf vorgelegt wurde.

Gegenwärtig ist festzustellen, dass in der FBE diverse heterogene Berichtsformate Verwendung finden. Die gängigen Formate sind das PDF, Textdateien, XML und HTML. Diese Heterogenität der verwendeten Formate stellt im Hinblick auf eine Vereinheitlichung die primäre Herausforderung dar (vgl. Bartolacci et al., 2020). Konkret wurde der Standard entwickelt, um diese heterogene Formatlandschaft zu vereinheitlichen und damit den Anforderungen einer digitalen und globalisierten Wirtschaft und den Kapitalmarktteilnehmer mit hohen Informationsbedürfnissen zu genügen. An das neue Berichtsformat wurden somit erhebliche Erwartungen im Hinblick auf die Verbesserung der Verfügbarkeit von Informationen zu Gunsten der EMH und zur Verringerung des Verwaltungsaufwands bzw. Nützlichkeit von Informationen im Vergleich zu den traditionellen Formaten gestellt (vgl. Baldwin & Trinkle, 2011, S. 12f.; Vasarhelyi, Chan & Krahel, 2012, S. 164). Der Standard verkörpert eine digitale Sprache zur Beschreibung von Informationen, die einen elektronischen, strukturierten und effizienten Informationsaustausch und -vergleich ermöglicht und stellt damit eine Weiterentwicklung der XML-Berichterstattung dar (vgl. Gostimir, 2015, S. 32). Die Entwicklung wurde durch das American Institute of Certified Public Accountants mit Hilfe eines XBRL-Konsortiums angestoßen. Das ursprüngliche Ziel dieses Konsortiums war die Entwicklung einer Sprache, die einen Austausch sowie die Verarbeitung von finanziellen Daten und Informationen unabhängig von verwendeter Software und Systemen sicherstellt (vgl. Nutz & Strauß, 2002, S. 447; Williams et al., 2006, S. 93; Boyer-Wright, Summers & Kottemann, 2010, S. 510).

Das angestrebte Berichtssystem stellt einen offenen, frei zur Verfügung stehenden (Open Source-)Standard dar (vgl. Beerbaum & Piechocki, 2017, S. 3f.). Insbesondere XBRL und dessen Erscheinungsformen erfüllen ausschlaggebende Funktionalitäten, die die Anwender befähigen, Geschäfts-

daten effizient aufzubereiten und umfangreich weiterzuverarbeiten. Als Erweiterung der grundlegenden Beschreibungssprache erlaubt iXBRL zudem, dass sich Berichte von einer maschinellen in eine menschlich auslesbare Form transformieren lassen (vgl. Debreceny & Gray, 2001, S. 47ff.; Bonsón, 2001, S. 101ff.). Ferner stellen diverse Basistaxonomien Verzeichnisse zur Verfügung, die erlauben, Elemente innerhalb der jeweiligen Berichte zu etikettieren. Diese Verzeichnisse orientieren sich an verschiedenen existierenden Rechtsgrundlagen, wie z.B. den International Financial Reporting Standards (IFRS), den Generally Accepted Accounting Principles (GAAP) oder dem deutschen Handelsgesetzbuch (HGB) mit einer daraus resultierenden Kompatibilität von Veröffentlichungen unterschiedlicher Unternehmen (vgl. Bravidor & Loof, 2020, S. 178). Eine daran anschließende Funktion liegt im system- und technologieübergreifenden Austauschformat. Durch Informationsbrüche im Austausch entsteht Ineffizienz, die eine manuelle Eingabe oder Korrektur fehlerhafter Angaben regelmäßig notwendig macht. Der Standard soll hierbei einen von Formatbrüchen befreiten Austausch von Informationen zwischen verschiedenen Akteuren und damit Informationssystemen und Technologien ermöglichen, wodurch eine konsistente und gradlinige Informationskette ermöglicht wird (vgl. Nutz & Strauß, 2002, S. 448; Leibfried, 2010, S. 10; Bravidor & Loof, 2020, S. 178). Somit führt die Etablierung des angestrebten Berichtsstandards in der EU zur Vereinheitlichung der Daten- und Informationsaufbereitung und -verarbeitung und ermöglicht zusätzlich, den gesamten Berichtsprozess und die anschließende Analyse objektiver, kostengünstiger und effizienter zu realisieren. Es wird schließlich eine digitale Transformation der FBE angestoßen, die in einigen Ländern außerhalb der EU bereits langjähriger Pflichtprozess der Berichtspraxis ist (vgl. Bravidor & Loof, 2020, S. 178). Aktuell ermöglichen hingegen viele EU-Mitgliedsstaaten den ansässigen Unternehmen die notwendige ESEF-Umsetzung zu verschieben. Ausschließich Österreich, die tschechische Republik, Slowenien und Deutschland verlangen Rechtswirkung zum 01. Januar 2021 (vgl. Accountancy Europe, 2021) und werden damit erste Einblicke in die europäische digitale FBE der Zukunft ermöglichen. Die proaktive Digitalisierung in der FBE durch einen solchen Standard ist zwar ein Schritt in die richtige Richtung, legt wiederum Herausforderungen und Risiken offen, die in der Natur der verwendeten Formate und partizipierenden Akteure liegen. Im Falle von XBRL ist insbesondere die Fehleranfälligkeit des Etikettierens der Elemente und damit die Validierung der verknüpften Daten herausfordernd und zu würdigen. Infolge solcher Risiken untersuchte die amerikanische Börsenaufsicht das erläuterte Potenzial von iXBRL (vgl. Basoglu & White Jr, 2015, S. 189).

3.3.2. Ausführungen aus rechtlicher Sicht

Aus rechtlicher Sicht nahm sich die EU-Legislative der Notwendigkeit der Entwicklung eines einheitlichen elektronischen Berichtsstandards in der EU an und erließ im Jahr 2013 die einleitende EU-Transparenzrichtlinie (Richtlinie 2013/50/EU). Die Richtlinie adressiert die Notwendigkeit eines einheitlichen Berichtsformats zur Standardisierung des Zugangs, der Informationsverarbeitung und der Vergleichsfähigkeit der veröffentlichten Informationen. Dies wird durch das ESEF unterstützt und betrifft bestimmte Berichtselemente für ausgewählte EU-Mitgliedsstaaten zum Geschäftsjahr 2020 (vgl. Richtlinie 2013/50/EU, Rn. 26; Accountancy Europe, 2021; Beerbaum & Piechocki, 2016, S. 1f.). Die European Securities and Markets Authority (ESMA) wurde dahingehend beauftragt, unter Rücksichtnahme gegenwärtiger und zukünftiger Medien- und Technologieoptionen einen technischen Regulierungsstandard (RTS) auszuarbeiten (vgl. Beerbaum & Piechocki, 2016, S. 1f .; ESMA, 2017, S. 2f.). In diesem Sinne führte die ESMA Tests in Verbindung mit Kosten-Nutzen-Analysen durch, um eine geeignete Herangehensweise und mögliche Optionen zu identifizieren (vgl. ESMA, 2017, S. 7f.; Beerbaum & Piechocki, 2016, S. 2). Im Jahr 2019 wurde ein Vorschlag über die Verordnung 2018/815 von der Europäischen Kommission getätigt und im Amtsblatt der EU der Öffentlichkeit zur Verfügung gestellt. Es wurde verlautbart, dass alle Emittenten des regulierten europäischen Marktes künftig dazu verpflichtet werden, bestimmte Berichtselemente im XHTML-Format zu veröffentlichen, damit involvierte Informationen sowohl maschinell, als auch menschlich verarbeitet werden können. Bestimmte Bestandteile von Konzernabschlüssen nach IFRS sind darüber hinaus mit XBRL zu verknüpfen, was innerhalb des XHTML-Formats durch iXBRL erfolgt (vgl. Di Fabio et al., 2019, S. 36). Innerhalb des Abschlusses sind laut Verordnung in einem ersten Schritt elementare Angaben des Unternehmens, wie z.B. die Firma, der Standort, sowie die Rechtsform zu etikettieren (vgl. Delegierte Verordnung (EU) 2018/815 der Kommission 2018, S. 6). Mit Beginn des Geschäftsjahres 2022 wird die Auszeichnungspflicht nochmals um weitere Elemente ergänzt, um den ausführlichen Anforderungen des Marktes gerecht zu werden. Mit dieser Erweiterung werden primäre Abschlussbestandteile (darunter Elemente der Bilanz, Gewinn- und Verlustrechnung, Kapitalflussrechnung und Eigenkapitalveränderungsrechnung) Teil der digitalen Veröffentlichung (vgl. Delegierte Verordnung (EU) 2018/815 der Kommission 2018, S. 6ff.).

Diesbezüglich wird aus rechtlicher Sicht in der aktuellen Forschung deutlich, dass die Anforderungen und damit viele Projekte primär von europäischen Regulierungsbehörden und Banken ausgehen. Sie sind nicht ausschließlich potenzielle Anwender des Standards, sondern stecken zusätzlich in einer Vorreiterrolle, die den Akteuren einen signifikanten Einfluss auf die Verbreitung des Standards verleiht. Aufsichtsbehörden besitzen die institutionelle Autorität, die Umsetzung des Standards anzuregen bzw. vorzuschreiben und können damit eine einflussreiche Anregung für die Einführung und Weiterenwticklung des Standards geben (vgl. Hucklesby & Macdonald, 2002, S. 54f.; Locke & Lowe, 2007, S. 609). Indessen eine alleinige Anregung nicht auszureichen scheint, da festgestellt wurde, dass Unternehmen zur Einführung einer einheitlichen Berichtssprache teilweise zu zwingen sind (vgl. Cordery et al., 2011, S. 84).

Ergänzend zu den verlautbarten Elementen des Berichts-

standards, haben Länder, die zur Einführung des Standards verpflichtet wurden bzw. sich freiwillig entschieden haben, der vorgegebenen Umgebung der Basistaxonomie weitere Elemente hinzuzufügen, die den nationalen Gegebenheiten entsprechen (vgl. Đurković, Šaponja & Đurković, 2017, S. 13). XBRL International verwaltet unterdessen das gemeinschaftliche Wissen und fördert die Verwendung und weitere Entwicklung von Leitplanken, die sich anschließend in den nationalen Gesetzgebungen und Projekten widerspiegeln. Zudem wird die Offenlegung neu aufgesetzter Taxonomien der nationalen Einrichtungen gewährleistet. Darunter haben diese lokal eingerichteten Konsortien (sogenannte Jurisdiktion) die Pflicht, die Etablierung und Entwicklung eines nationalen Rahmenwerkes, mit Verweisung des nationalen Standards auf die internationale Taxonomie zu fördern (vgl. Florescu & Tudor, 2009, S. 133; XBRL International Inc., 2021b). Nach einer erfolgreichen Ausarbeitung und Validierung werden neu aufgesetzte Elemente als integraler Bestandteil des internationalen Rahmenwerks aufgenommen und der weltweiten Berichtspraxis als Teil der gemeinschaftlichen Open-Source transparent gemacht (vgl. Đurković et al., 2017, S. 13).

Abschließend ist aus rechtlicher Sicht hinzuzufügen, dass Unternehmen den Berichtsstandard zwar überwiegend als obligatorisches Kommunikationsmedium der digitalen FBE zur effizienten Einhaltung rechtlicher Anforderungen verwenden, der Einsatz solcher Sprachen sich zudem zur Verbesserung des internen Rechnungswesens bewährt hat. Innerhalb des internen Rechnungswesens kann z.B. XBRL im Konsolidierungsprozess bei dem Datenexport und -import von Tochtergesellschaften unterstützend eingesetzt werden. Darüber hinaus ist die Anwendung in Verbindung mit dem Beteiligungscontrolling oder einer Balanced Scorecard denkbar (vgl. Kretschmer, 2014, S. 180; Kesselmeyer, 2009, S. 69ff.). Durch solch einen umfangreichen Standard wird dahingehend die Konvergenz des externen und internen Rechnungswesens unterstützt, was zudem im Zeichen der integrierten Berichterstattung steht (vgl. Weißenberger, 2014, S. 440; Hwang, No & Kim, 2021, S. 684). Darüber hinaus unterstützen die Eigenschaften des Standards die Implementierung einer „one version of the truth“ (Simons & Weißenberger, 2009, S. 319), sprich ausgehend von einer einheitlichen und konsistenten Datenbasis agieren zu können und die Integrität der Daten- und Informationsgrundlage und damit der Berichterstattung zu optimieren. Trotz vorteilhafter Auswirkungen auf das interne Rechnungswesen und Datenmanagement des Unternehmens sollten die unterschiedlichen Projekte und Ansätze zur Einführung des Berichtsstandards fokussiert darauf abzielen, dass der Einsatz in Einheit mit den Anforderungen für die obligatorische FBE steht, und gleichzeitig mit regionalen Vorschriften und internationalen Rechnungslegungsstandards (z.B. IFRS) in Einklang ist, um zunächst eine angemessene Herangehensweise aus rechtlicher Perspektive zu gewährleisten, bevor technische Details, Verbundeffekte und unternehmensinterne Folgen und Konsequenzen ausgearbeitet werden (vgl. Florescu & Tudor, 2009, S. 138).

3.3.3. Technische Details zur Vereinheitlichung

Der ursprüngliche Entwurf von XBRL und der einleitende RTS sehen aus technischer Sicht die Hauptfunktion in der Förderung der Effizienz von Geschäftsinformationen. Steigende Effizienz durch eingebettetes Wissen wird durch die jeweilige Taxonomie und nicht innerhalb eines umfassenden Rahmenwerkes verwirklicht (vgl. Đurković et al., 2017, S. 13). Die digitale FBE in einem solchen Standardformat gewährleistet damit die Optimierung der Aufbereitung und Verarbeitung von Daten und Informationen unter Einsatz fortschrittlicher Medien und Technologien (vgl. Florescu & Tudor, 2009, S. 126; Alles & Debreceny, 2012, S. 88). Im Einsatz befindliche Taxonomien liefern ein Regelwerk zur Auslegung von XBRL-Verknüpfungen, die als eine Reihe von voreingestellten Elementen, die von Softwareanwendungen verarbeitet werden, zu interpretieren sind (vgl. Singerová, 2015, S. 136; Đurković et al., 2017, S. 13). Nachdem ein Etikettieren der Informationen stattgefunden hat, können ausschließlich relevante Berichtselemente extrahiert werden, anstatt seitenlange Geschäftsberichte mit hohem Aufwand und dem Risiko von Informationsbrüchen verarbeiten zu müssen (vgl. Bravidor & Loof, 2020, S. 178; Đurković et al., 2017, S. 13). Aus technischer Perspektive wird somit durch XBRL ermöglicht, dass veröffentlichte Informationen eindeutig identifiziert werden können, da bspw. eine Unterscheidung von Bilanzpositionen wie dem Anlage- und Umlaufvermögen digital hinterlegt worden ist. Hierbei weisen die Verknüpfungen nicht ausschließlich aus, ob es sich z.B. um Angaben des Anlagevermögens handelt, vielmehr ergänzen sie das Element mit weiteren Attributen und zeigen daran anschließend auf, ob es sich bei dem markierten Wert z.B. um einen Text, einen positiven oder negativen Wert oder um eine Geldeinheit handelt (vgl. Gostimir, 2015, S. 32f.). Die Verknüpfungen stellen darüber hinaus mehr als die ausschließliche Identifizierung des Inhalts eines Berichtselements dar. Es wird gleichzeitig validiert, ob ein Wert, der sich auf einen bestimmtes Element bezieht (z.B. Anlagevermögen) inhaltlich und strukturell korrekt ist, sprich z.B. innerhalb eines bestimmten Wertebereichs liegt (ausschließlich positive Zahlen bezogen auf das Anlagevermögen). Demzufolge bedeutet die Kodierung von Berichtselementen im XBRL-Format, dass gemäß der eingesetzten Taxonomie ein Bericht in individuelle Elemente gegliedert wird und die verknüpften Informationen des Berichtes nach den Bedürfnissen des Adressaten individuell kombinierbar und elektronisch analysierbar sind (vgl. Đurković et al., 2017, S. 13; Baldwin & Trinkle, 2011, S. 15; Bartolacci et al., 2020).

Für die Funktionsfähigkeit des Etikettierens sind die Taxonomie und die Instanz der XBRL ausschlaggebend. Das Verständnis über den Zusammenhang zwischen dem XBRL-Element bzw. -Taxonomie und dem Fakt/Wert bzw. der XBRL-Instanz wird in der folgenden Abbildung veranschaulicht:

Die XBRL-Instanz konzentriert sich auf Fakten und Werte, die die Ausprägungen der Unternehmensinformationen abbilden. Diese Informationen werden innerhalb der Instanz in einen Kontext gebracht, sprich z.B. einem Geschäftsbereich

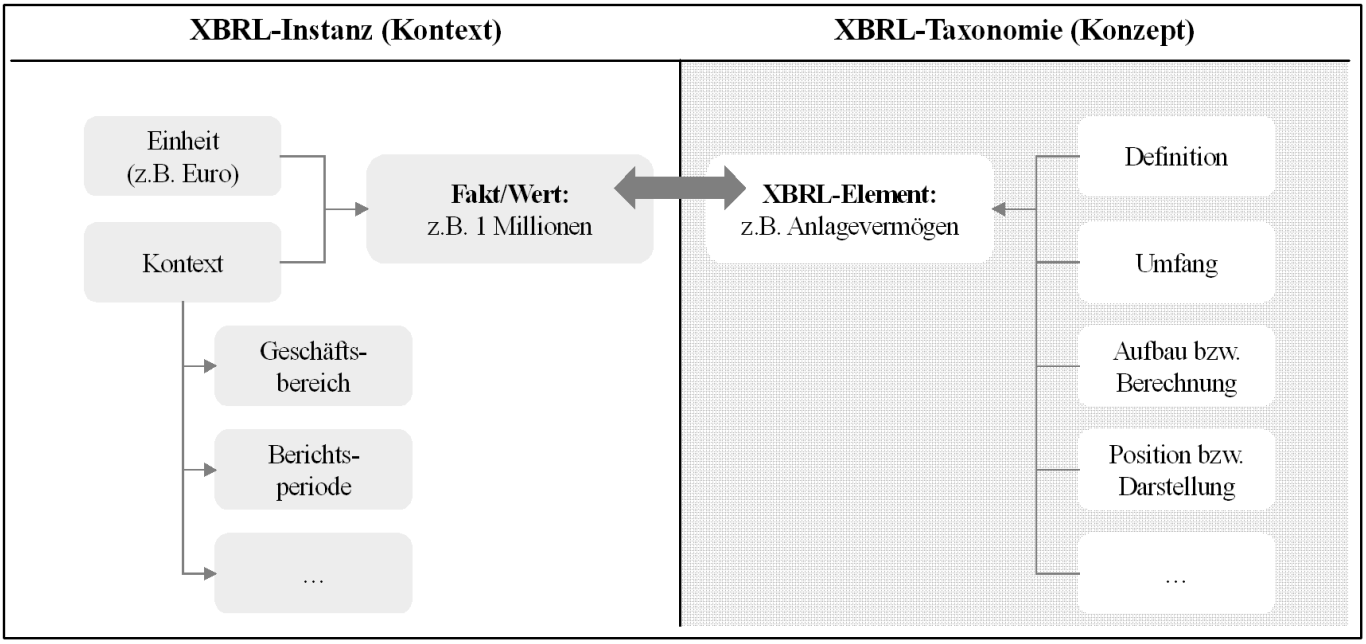


Abbildung 2: Zusammenhang zwischen XBRL-Instanz und XBRL-Taxonomie (eigene Darstellung mit inhaltlicher Anlehnung an Kesselmeyer (2009), S. 70; Pfeile repräsentieren Verbindungen zwischen den einzelnen Elementen in den Abbildungen).

oder einer Berichtsperiode zugeordnet. Zusätzlich ist für das einzelne Element noch eine Einheit festzulegen, die aus finanzieller Perspektive oftmals die geltende Währungseinheit darstellt (vgl. Kesselmeyer, 2009, S. 70). Wohingegen die Taxonomie einem Kontenrahmen ähnelt und das Konzept bildet und damit den Umfang, Aufbau und Position der Informationen festlegt (vgl. Kesselmeyer, 2009, S. 70; Kretschmer, 2014, S. 172). Darüber hinaus wurde von XBRL International ein Global Ledger entwickelt, das ein Rahmenwerk für die Darstellung der Informationen in einem Buchhaltungssystem zusammenfasst. Dazu zählen bspw. Angaben zu Stammdaten von Kunden und Lieferanten, Kontenpläne, Transaktionsdaten und Leistungskriterien (vgl. XBRL International Inc., 2021a; Amrhein, Farewell, Pinsker et al., 2009, S. 147). Durch eine anschließende softwaregeschützte Verarbeitung der Informationen aus unterschiedlichen Datenquellen (z.B. Buchhaltungssystem) bildet sich ableitend aus der Kombination von XBRL-Instanz und -Taxonomie ein Instance Document, das anschließend über Softwareschnittstellen in diverse Formate übersetzt werden kann. Durch das formatübergreifende Instance Document sind Adressaten in der Lage, Informationen effizient nach eigenem Ermessen zu analysieren (vgl. Leibfried, 2010, S. 11). Es wird deutlich, dass die XBRL-Instanz ausschließlich das Grundgerüst der Sprache verkörpert und erst in Verbindung mit der Taxonomie, also einem Informationskonzept, eine angemessene Verarbeitung gewährleistet werden kann (vgl. Kesselmeyer, 2009, S. 71).

Um zu gewährleisten, dass die verwendeten XBRL-Verknüpfungen zu Vergleichungszwecken eingesetzt werden können, ist es erforderlich, einen obligatorischen und einheitlichen Standard bereitzustellen, der die Art und Weise, wie Elemente zu etikettieren sind, festlegt. Dementsprechend hat die ESMA eine Basistaxonomie erarbeitet und eingeführt, die sich an den IFRS inhaltlich orientiert und somit eine Konsistenz in der Verwendung gewährleisten soll (vgl. Beerbaum & Piechocki, 2017, S.3 f.). Es ermöglicht die Erweiterbarkeit der Taxonomie aus den Wurzeln von XML, Daten unternehmensindividuell aussagekräftiger zu gestalten, wobei die Vergleichbarkeit des Instance Documents über Unternehmen hinweg geschwächt werden kann. Aus diesem Grund müssen neu erarbeitete unternehmensindividuelle Elemente mit dem nächstgelegenen Element der Basistaxonomie verankert werden (sogenanntes Anchoring; vgl. XBRL International Inc., 2021c; Aaltonen & Penttinen, 2021, S. 5927).

Buchhaltungsanwendungen sind dabei XBRL-Verknüpfungen fortsetzend einzugliedern, wodurch die Verarbeitung von Buchhaltungsdaten durch das Etikettieren verstärkt profitiert, wobei eine XBRL-Taxonomie und daraus festgelegte Verknüpfungen nicht ausschließlich für Buchhaltungsdaten zu definieren sind (vgl. Baldwin & Trinkle, 2011, S. 15; Đurković et al., 2017, S. 13). Darüber hinaus ist hinzuzufügen, dass der Einsatz von XBRL-Elementen das Aufsetzen und die Vergleichbarkeit einer nicht-finanziellen Berichterstattung verbessert und Nachhaltigkeitskriterien für die Unternehmenssteuerung entwickelt werden können (vgl. Bartolacci et al., 2020), um den wachsenden Umfang insbesondere der Nachhaltigkeits-Berichterstattung innerhalb einer integrierten Berichterstattung unterstützen zu können. Hingegen müssen Buchhaltungslösungen, die ein Unternehmen benötigt, um den XBRL-Standard zu integrieren, neben den individuellen Ausprägungen der Basistaxonomie eine Auto-

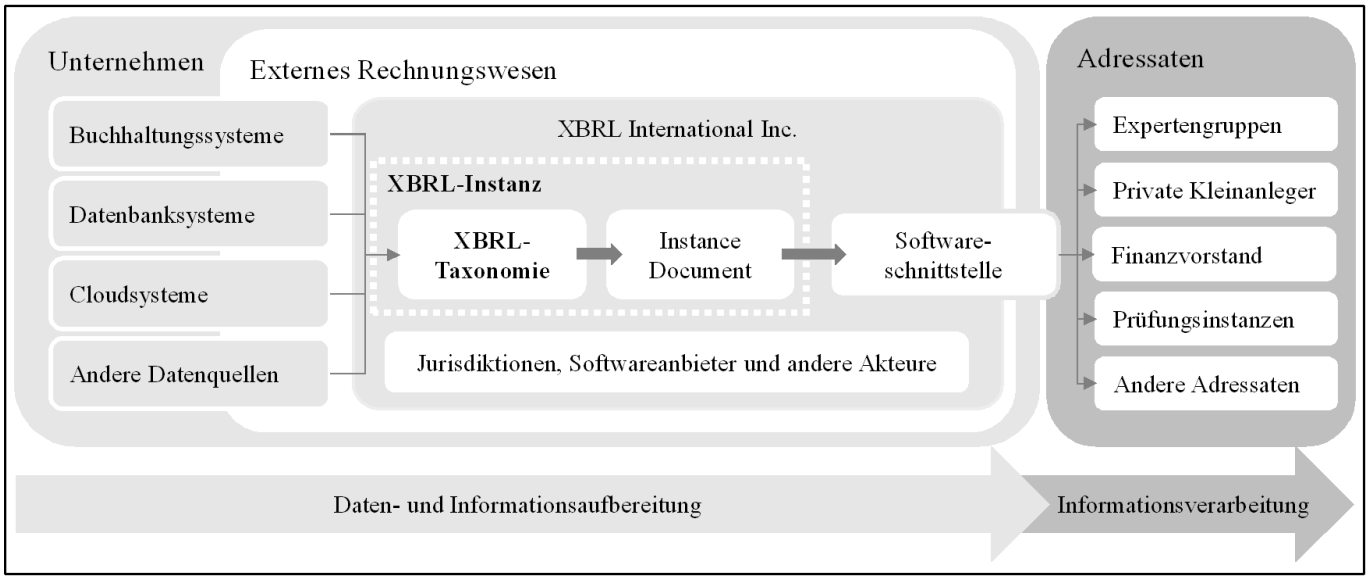


Abbildung 3: XBRL innerhalb der Daten- und Informationslieferkette (eigene Darstellung mit inhaltlicher Anlehnung an Pfister & Venetz, 2009, S. 435; Doolin & Troshani, 2004, S. 96; Kesselmeyer, 2009, S. 70; Für ein Beispiel eines Instance Documents siehe Gostimir, 2015, S. 35).

matisierung der Schaffung, Aufbereitung und Verteilung von insbesondere finanziellen Geschäftsinformationen gewährleisten (vgl. Florescu & Tudor, 2009, S. 134f.).

Abschließend ist sowohl aus rechtlicher, als auch technischer Sicht zu berücksichtigen, dass die Funktionsweise und Implementierung eines einheitlich elektronischen Berichtsstandards neben den rechtlichen und technischen Details primär vom menschlichen Faktor abhängig ist. Neben einer notwendigen Akzeptanz und Rechtfertigung zur Implementierung des neuen Standards inklusive verwendeter Medien und Technologien (dazu empirische Nachweise durch gängige Akzeptanzmodellen, vgl. Felden, 2011; Vasarhelyi et al., 2012; Rawashdeh & Rawashdeh, 2021) wird zunächst ein hoher manueller Aufwand deutlich, der im Prozess des Rechnungswesens mit inhaltlichen Fehlern begleitet wird, gefolgt von kritischen Folgen und Konsequenzen für die Qualität und Zuverlässigkeit der FBE und damit dem Mehrwert durch den Standard (vgl. Florescu & Tudor, 2009, S. 135). Neben solchen traditionellen verhaltensökonomischen Ansätzen steht innerhalb der digitalen Berichterstattung der Begriff des Neuroaccountings im Fokus mit dem Potential, neue Einblicke in die kognitive Verarbeitung von Entscheidungsträgern zu liefern, wenn digitale Rechnungslegungsverfahren und -kontrollen verwendet werden (vgl. Dickhaut, Basu, McCabe & Waymire, 2010, S. 221ff.; Tank & Farrell, 2021, S. 1ff.).

3.3.4. XBRL-Umsetzung und Projekte in der EU

Mit der Etablierung des Berichtsformats seitens der EU wurden die ersten Vorkehrungen getroffen, um zukünftig eine transparente und länderübergreifend vergleichbare FBE in der EU zu gewährleisten. Unter welcher Basistaxonomie und auf welche Berichtselemente bezogen das digitale Format Anwendung erfährt, hängt hierbei von den Ausarbeitungen verschiedener Gremien außerhalb und innerhalb der Mitgliedstaaten ab. Die Einführung und Weiterentwicklung der Sprache und bestimmten Ausprägungen für Zwecke der FBE wurde von einem gemeinnützigen internationalen Konsortium aus einer Vielzahl von Finanzdienstleistungs- und Rechnungslegungsorganisationen, sowie Regulierungsbehörden durchgeführt (vgl. Bonsón, Cortijo & Escobar, 2009, S. 57; Gomaa, Markelevich & Shaw, 2011, S. 155; Durkovic/Saponja/Durkovic 2017, S. 13).

Nach detaillierterer Betrachtung wird deutlich, dass der Berichtstandard weltweit vermehrt wahrgenommen und umgesetzt wird. Der Einsatz variiert von verpflichtenden Umsetzungen börsennotierter Unternehmen, bis hin zu freiwilligen Pilotprojekten von staatlichen Verwaltungseinrichtungen (vgl. Gostimir, 2015, S. 36). Die Förderung der Umsetzung durch ESEF und die Akzeptanz des Standards sind größtenteils dadurch geprägt, dass große Erwartungen, insbesondere an Rahmenwerke für den stark regulierten Finanzsektor gestellt werden. Darüber hinaus wächst die Akzeptanz und Einführungstendenz aufgrund der Kompatibilität zu den anerkannten IFRS (vgl. Gostimir, 2015, S. 32). Diesbezüglich wird die Umsetzung des Standards durch XBRL International als begleitende und beratende Instanz weltweit unterstützt. Die Umsetzung auf nationaler Ebene erfolgt durch die lokal eingerichtete Jurisdiktion (vgl. XBRL International Inc., 2021b). Auf europäischer Ebene nimmt sich XBRL Europe der Etablierung und Entwicklung an. Verschiedene Akteure adressieren unterdessen eine thematische Einbettung des Standards in der EU und unterstützen die jeweiligen nationalen Jurisdiktion bei ihrer Arbeit (vgl. XBRL International Inc., 2021b).

In Belgien nimmt sich XBRL Belgium als nationale Jurisdiktion der Aufsicht und Umsetzung von XBRL an. Die lokale

Basistaxonomie wird auf den nationalen Rechnungslegungsstandards (BE-GAAP) aufbauend ausgeführt (vgl. XBRL International Inc., 2021b). Die Umsetzung von XBRL-Projekten in Belgien unterliegt einer Vielzahl von Organisationen. Die involvierten Berichte erstrecken sich von der Steuerberichterstattung, über die Jahresberichterstattung im weiterem Sinne bis hin zur reinen FBE. Die Jahresberichterstattung richtet sich an Non-Profit-Organisationen, wohingegen die Einreichung der elektronischen Steuerbilanz in XBRL an sämtliche Unternehmen und die allgemeine FBE im XBRL-Format an Kreditinstitute gerichtet ist. In Belgien sind sämtliche Projekte erfolgreich umgesetzt und die Taxonomien verfolgen einen geschlossenen Ansatz, die auf einer einheitlichen Basistaxonomie der Nationalbank Belgiens beruht (vgl. Enachi & Andone, 2015, S. 188; Fourati, Merdassi & Boujelbene, 2018, S. 101). Im deutschen Raum widmet sich hingegen XBRL Germany der Einführung von XBRL. In den letzten Jahren entstand eine Vielzahl von Taxonomien, wie z.B. die XBRL-HGB-Basistaxonomie, sowie weitere Ausprägungen, die bestimmte Industrien oder Branchen repräsentieren, stets in enger Zusammenarbeit mit XBRL Europe (vgl. XBRL International Inc., 2021b). Die XBRL-Implementierung nahmen sich die Deutsche Bundesbank, sowie eine eingetragene Genossenschaft an. Mit dem XBRL-Format zu veröffentlichende Berichte sind z.B. Sonderberichte zu Verbriefungspositionen oder die FBE im engeren Sinne. Darüber hinaus muss die elektronische Steuerbilanz als XBRL-Dokument übergeben werden (vgl. Felden, 2011, S. 162). Hervorzuheben ist, dass sämtliche in Deutschland initiierte Projekte implementiert wurden, die teilweise verpflichtender und optionaler Anwendung unterliegen (vgl. Enachi & Andone, 2015, S. 188f.). Dazu adressiert ein gegenwärtiges Projekt in Deutschland den obligatorischen Austausch von Finanzinformationen an z.B. Banken, die auf Basis der eingereichten Informationen eine Prüfung der Kreditwürdigkeit des Unternehmens vorzunehmen haben. Dieses Vorgehen wird unter Begleitung der deutschen Kreditwirtschaft und Softwaredienstleistern betreut und unter Anwendung einer XBRL-Basistaxonomie digitalisiert, dessen Umstellung seit dem Jahr 2018 umgesetzt wird (vgl. XBRL DE, 2019). XBRL France fördert im regelmäßigen Kontakt zu den französischen Behörden und Wirtschaftsverbänden den XBRL-Standard. Die Aktivitäten in den Bereichen Überwachung, Vernetzung und Kommunikation konzentrieren sich auf die aufsichtsrechtliche Berichterstattung im Banken- und Versicherungswesen und anderer börsennotierter Unternehmen. Beteiligte Organisationen veranstalten zudem Schulungen zu Themen wie z.B. Einführung in die Automatisierung der FBE oder Grundlagen der XBRL-Sprache und ihrer Komponenten (vgl. XBRL International Inc., 2021b). Die Einführung von XBRL steht in Frankreich im Wettbewerb mit einem anderen Standard (dem Electronic Data Interchange for Administration, Commerce and Transport (EDIFACT)), der als stabiler und etablierter Teil der Infrastruktur der steuerlichen Berichterstattung in Frankreich angesehen wird. Allerdings haben Krisenauswirkungen der letzten Jahrzehnte bestehenden Berichtsrahmen in Frage gestellt. Der Druck zu handeln, um weitere Krisen zu verhindern, wurde verstärkt und erlaubt eine mögliche Anpassung mittels XBRL zu rechtfertigen (vgl. Guilloux, Locke & Lowe, 2013, S. 271). XBRL Netherlands ist dagegen eine Initiative, die unter anderem von Organisationen aus den Bereichen Wirtschaft, Regierung und Forschung finanziell und inhaltlich unterstützt wird. Mit mehr als 40 Mitgliedern war das Projekt maßgeblich an der Entwicklung des niederländischen Standard Business Reporting (SBR)-Programms beteiligt. Darüber hinaus werden Softwareanbieter, die XBRL in ihren Anwendungen integrieren, unterstützt und es werden regelmäßig Schulungen abgehalten, um Akteure über die Verwendung und die Vorteile von XBRL zu informieren (vgl. XBRL International Inc., 2021b). Ein besonderer Fokus, neben z.B. Steuerinformationen und der FBE liegt auf der Digitalisierung der Verwaltung der Wasserwirtschaft, indem Quartalsberichte mittels XBRL zu veröffentlichen sind (vgl. Enachi & Andone, 2015, S. 189). Im Vergleich zu den anderen Ländern weist Spanien ein bemerkenswertes XBRL-Engagement auf. Mit über 30 öffentlich verfügbaren Taxonomien hat das Land im Laufe der Zeit viel Erfahrung mit XBRL und ein Repertoire an Informationen in diesem Format aufbauen können. Spanien scheint damit eine Vorreiterrolle der XBRL-Implementierung einzunehmen (vgl. XBRL International Inc., 2021b). Die mit XBRL zu formatierenden Berichte erstrecken sich von der Berichterstattung börsennotierter Unternehmen, über Finanzberichte für Kreditinstitute, bis hin zu Corporate Social Responsibility-Berichten. Indessen trotz der erheblichen Vielfalt an XBRL-Initiativen alle angefangenen Projekte in Spanien umgesetzt wurden (vgl. Enachi & Andone, 2015, S. 189f.).

Zum weiteren Vergleich der XBRL-Projekte stand die Herangehensweise von Enachi und Andone (2015, S. 189) im Fokus. Im Vergleich wurde analysiert, wer die zuständige Jurisdiktion ist und welche Organisationen in den jeweiligen Mitgliedsstaaten die Umsetzung von XBRL begleiten und unterstützen. Zudem wird aufgezeigt, inwiefern lokale Standards oder Taxonomien zur Erweiterung oder Hemmung von XBRL vorhanden sind. Darüber hinaus wurde identifiziert, auf welche Berichtsbestandteile oder Unternehmen XBRL Anwendung findet und ob diese verpflichtender oder freiwilliger Natur ist (vgl. Anhang C). Zwischen den beschriebenen Ländern wird erkennbar, dass sich viele individuelle Umsetzungsansätze entwickelt haben. Länderübergreifend liegt eine Vielzahl von unterschiedlichen Akteuren, die die Einführung von XBRL vorantreiben, vor. Ein Großteil dieser Organisationen sind Regierungsbehörden und Banken und scheinen damit an der Spitze der Umsetzung zu stehen und die Initiative in ihrer autoritären Vorreiterposition angemessen unterstützen zu können, wobei andere Organisationen aus z.B. Wirtschaft, Forschung und Bildung die Umsetzung tatkräftig begleiten. In der Anzahl der Taxonomien zeigen sich die größten Unterschiede. Während in Ländern wie Belgien eine einzige Taxonomie ausgehend von der Nationalbank vorgegeben wird, sind in anderen Ländern wie Spanien über 30 unterschiedliche Taxonomien entwickelt und veröffentlicht worden. Die mit XBRL (verpflichtend oder optional) abzudeckenden Berichtselemente erstrecken sich

von elektronischen Steuerbilanzen (Belgien, Deutschland und Niederlande) und anderen Solvenz-Berichten (Deutschland), über Überschussrechnungen für kleine Unternehmen (Deutschland) und Corporate Social Responsibility-Berichten (Spanien), bis hin zur zweckmäßigen FBE, die länderübergreifend verfolgt wird. In den Niederlanden scheint neben dem fortgeschrittenem SBR-Programm situationsbedingt ein beispielhafter Branchenfokus auf der Wasserwirtschaft zu liegen, wobei in Frankreich weniger branchenfokussiert die allgemein-aufsichtsrechtliche Berichterstattung für bestimmte Unternehmen im Vordergrund steht. Ob diese Elemente einer ganzheitlichen Berichterstattung optional oder verpflichtend im XBRL-Format zu veröffentlichen sind, wird zwischen den Ländern uneinheitlich gehandhabt. In Belgien und Spanien beruht bspw. die FBE mit XBRL grundsätzlich auf verpflichtender Basis, wohingegen in Deutschland diesbezüglich noch eine Option vorlag (vgl. Anhang C). Im Vergleich zu Projekten anderer Staaten, wie z.B. Australien, das bereits vor zwei Jahrzenten als fortgeschritten in der digitalen FBE angesehen wurde, wird die strukturierte Verwendung von Daten und Informationen durch die Einführung eines eigenen SBR-Projekts weiter vorangetrieben (vergleichbar mit der Initiative der Niederlande). Dieser weiterentwickelte Berichtsansatz zielt darauf ab, neben den grundlegenden Forderungen des Berichtsstandards, digitale Meldepflichten zu optimieren, indem australische Organisationen verpflichtende Meldungen an eine zentrale Instanz elektronisch übermitteln, von der aus alle Informationen an den zugehörigen Empfänger verteilt werden, wodurch Zeit- und Kostenersparnisse realisiert werden können (vgl. Markelevich & Riley, 2013, S. 70). Australien verfolgt im Gegensatz zu den EU-Mitgliedstaaten (insbesondere zu Deutschland und Spanien) einen geschlossenen Ansatz ohne direkte Erweiterungen der Basistaxonomie. Diese zunächst einschränkende Wirkung der Individualität einzelner Taxonomien vereinfacht den Umsetzungsprozess und reduziert die Komplexität. Hingegen z.B. branchenspezifische Informationen nicht problemlos erfasst werden können, was letztlich die Qualität und den ursprünglichen Nutzen des Standards negativ beeinflussen kann. Daraus wird ersichtlich, dass eine Balance aus Komplexität und Relevanz vorliegt, die situationsbedingt abzuwägen ist (vgl. Lim, Perrin et al., 2014, S. 41; Zhu & Wu, 2014, S. 359). Es ist im Fall von Australien zu berücksichtigen, dass das erklärte Ziel der XBRL-Implementierung die Effizienzsteigerung durch die Verringerung des administrativen Aufwands des Berichtsprozesses ist. Die Schaffung von Transparenz wird nicht als eines der Hauptziele des australischen Projektes erwähnt. Es wird hingegen deutlich, dass die alleinige Bereitstellung von Geschäftsdaten im XBRL-Format (unabhängig vom verfolgten Zweck) die Transparenz und Interoperabilität von Daten grundsätzlich zwischen verschiedenen Akteuren verbessert (vgl. Y.-C. Chen, 2012, S. 561; Anhang D). Durch den Vergleich verschiedener Projekte wird zudem deutlich, dass die Implementierung des Standards im vollen Gange ist und sich in Zukunft zeigen wird, welche Vorgehensweise sich unter welchen Umständen als vorteilhaft erweisen wird.

Die aus den Folgen und Konsequenzen für den Einsatz des neuen Mediums resultierenden Chancen, Herausforderungen und Risiken, sind Teil des anschließenden Kapitels.

4. Kritische Würdigung der Folgen und Konsequenzen [4]

4.1. Chancen durch den Einsatz des neuen Mediums

Die Nutzung von Onlinemedien in der Berichterstattung ermöglicht nachgewiesenermaßen eine umfangreichere Informationsbereitstellung und eine individuell angepasste Verarbeitung und Darstellungsform von Geschäftsangaben. Traditionell wurden Berichte aufgrund technischer Restriktionen gedruckt oder auf den Webseiten des Unternehmens statisch als z.B. PDF präsentiert. Von diesem traditionell-statischen Verfahren wird durch technologische Entwicklungen eine digital-dynamische Berichterstattung angesteuert, deren Inhalt medien- und technologieübergreifend für die Verteilung, Präsentation und Kommunikation von Informationen und Geschäftsangaben genutzt wird. In erster Linie führt dies für Unternehmen zu einem effizienteren Abschlussprozess, da durch Onlinemedien und Vereinheitlichungs- und Automatisierungstendenzen die manuelle Datenaufbereitung reduziert werden kann. Die automatische Verarbeitung steigert die Effizienz der Geschäftsprozesse, da manuelle Eingaben insbesondere bei Standardabläufen nicht mehr notwendig sind und somit Ineffizienz vermieden werden können (vgl. Pfister & Venetz, 2009, S. 435f.; Weißenberger & Bauch, 2017, S. 210f.; Kovarova-Simecek et al., 2018, S. 163; Bogasiu & Ardeleanu, 2021, S. 296). Diese Aufwandsreduzierung durch effizientere und automatisierte Berichtsprozesse geht zudem mit Kosteneinsparungen einher. Die Chance auf verringerte Transaktionskosten ergibt sich nicht ausschließlich durch erhöhte Prozesseffizienz, sondern insbesondere durch weitere Vorteile von Onlinemedien zur Veröffentlichung, da z.B. Kosten, die mit dem Druck und Versand von papierbasierten Berichten einhergehen ausbleiben. Informationen und Geschäftsberichte können im Datenzeitalter mit nahezu nicht vorhandenen Grenzkosten einer breiten Masse an Adressaten zur Verfügung gestellt werden (vgl. Weißenberger & Bauch, 2017, S. 207). In Zeiten besonderer Umweltsensibilität folgen daraus weitere Verbundeffekte, da aus dem reduzierten Ressourcenverbrauch ein Reputationsgewinn für das Unternehmen resultieren kann und Nachhaltigkeitskriterien unterstützt werden (vgl. Deller et al., 1999, S. 1579f.; Schruff & Kayser, 2002, S. 353).

Durch die Verbindung neuer Onlinemedien in der Berichterstattung tritt der Begriff des „Mash-Up“ (Meckel et al., 2010, S. 39ff.) vermehrt in Erscheinung. Unter Mash-Up wird die Möglichkeit verstanden, bei dem mit Hilfe von unterschiedlichen Medien und Technologien Adressaten in die Lage versetzt werden, relevante Informationen flexibel und

[4] Da in Zukunft die finanzielle Perspektive einer digitalen Berichterstattung nicht von Nachhaltigkeitskriterien deutlich getrennt werden kann, wird für den Rest der Arbeit von Berichterstattung im weiteren Sinne gesprochen.

differenziert zusammenstellen zu können. Es können existierende Inhalte durch neue Darstellungsformen z.B. akustisch oder visuell neu kombiniert und verarbeitet werden (vgl. Meckel et al., 2010, S. 39). Gleichzeitig, im Gegensatz zur medienübergreifenden und umfangreichen Kommunikation, Verteilung und Präsentation von Informationen, entsteht ein Gegenpol mit der Tendenz zum sogenannten „Storytelling“ (Meckel et al., 2010, S. 23ff.). Darunter wird eine Unternehmenskommunikation mit Hilfe von Beschreibungen für die präzise Vermittlung von bestimmten Geschäftsinformationen verstanden. In der Erzählung stehen weder die Geschäfte des Unternehmens, noch deren Dienstleistungen oder Produkte im Vordergrund. Vielmehr wird durch emotionale Erzählungen, z.B. über die Angestellten des Unternehmens, Aufmerksamkeit und Interesse geweckt und mittels dieser Geschichte in Kombination mit medialen Lösungen eine emotionale Verknüpfung zur Berichterstattung hergestellt (vgl. Meckel et al., 2010, S. 23ff.), was einer zielgerichteten Kommunikation gleicht, als der umfangreichen Bereitstellung unter dem Phänomen von Mash-Up. Hierbei ist hinzuzufügen, dass das Wissen von Buchhaltern, die Herangehensweise an Entscheidungen und Werte wie Konservatismus und Verlässlichkeit nicht durch rein statistisch-analytische Vorgehensweisen ersetzt werden dürfen. Insbesondere die digitale Berichterstattung in Zeiten von Big Data erfordert Erfahrung, persönliche Interpretation und ein ausgewogenes Storytelling (vgl. Al-Htaybat & von Alberti-Alhtaybat, 2017, S. 868). Auf der einen Seite kann damit dem Adressaten während der digitalen Berichterstattung eine uneingeschränkte und transparente Informationsverfügbarkeit gewährt werden, in der Freiheitsgrade in der Verarbeitung möglich sind. Wohingegen durch das präzise Storytelling zunächst eine emotionale Verbindung zum Unternehmen aufgebaut wird, bevor eine personalisierte und auf Erfahrung beruhende Berichterstattung stattfindet.

Insbesondere die Erweiterung der Berichterstattung mit (X)HTML-Lösungen unterstützt die dynamische Abbildung eines Geschäftsberichts auf verschiedenen Darstellungsformaten und -größen, wodurch Informationen auf unterschiedlichen Endgeräten abrufbar und verwertbar werden. Es entwickelt sich der Trend einer mobilen Berichterstattung, bei dem der Geschäftsbericht auf Endgeräten flexibel abgebildet und damit eine größere Reichweite ermöglicht wird. Indessen die Verbreitung von Informationen durch die mobile Handhabung in Verbindung mit Social-Media-Kanälen beschleunigt werden kann (vgl. Barrantes & Sagmüller, 2014, S. 68ff.; Mittelbach-Hörmanseder et al., 2020, S. 47; Mittelbach-Hörmanseder et al., 2021, S. 23f.). Die erhöhte Reichweite durch neue Kommunikationskanäle schafft damit eine zeitgemäße Denkweise über Unternehmenskommunikation, was sich schließlich auf das Rollenverständnis des Berichtswesens niederschlägt (vgl. Brennan & Merkl-Davies, 2018, S. 554; Valentinetti & Muñoz, 2021, S. 449; Anhang B). Die dynamische Anpassung von Geschäftsberichten auf verschiedene Bildschirmgrößen begünstigt damit eine zeit- und ortsunabhängige Informationsversorgung der Adressaten und führt zu einem Einklang der veröffentlichten Informationen.

Mit der Integration von Social-Media-Plattformen tritt, neben der mobilen, darüber hinaus die Chance einer interaktiven Berichterstattung ein. Diese Interaktion ermöglicht einen offenen Dialog zwischen Unternehmen und Adressaten und die Möglichkeit, einen persönlichen Austausch in Echtzeit stattfinden zu lassen und durch Feedbackrunden und Kommentaren konstruktive Kritik üben zu können. Darauf aufbauend wird bei angemessener Auslegung die Erstellung von personalisierten Berichten ermöglicht. Das über Social-Media erstellte Profil des Berichtsadressaten wird mit personalisierten Informationen versorgt, die ausschließlich für die individuelle Entscheidungsfindung relevant sind (vgl. Voelzkow, 2019, S. 12; Jones & Xiao, 2003, S. 144). Die erhöhte Informationsvielfalt geht insbesondere in Zeiten von Big Data und IoT mit einer gegenwärtig geringen Inanspruchnahme einher. Emails oder Social-Media-Kommentare von Kunden enthalten wertvolle Hinweise für die zukünftige Geschäftstätigkeit und damit für die Berichterstattung. Eine bedeutsame Schlussfolgerung diesbezüglich ist, dass Big Data in Verbindung mit weiteren Trends das Potenzial besitzt, die aktuellen Lösungen, die durch diverse Onlinemedien und den Social-Media-Plattformen angeboten werden, in die Berichterstattung zu integrieren. Infolgedessen Echtzeit- und Mehrfachdialog ermöglicht werden kann, einschließlich einer ganzheitlichen Mensch-Maschine- und Maschine-Maschine-Interaktion (vgl. Valentinetti & Muñoz, 2021, S. 560).

Neben der Erweiterung der Mobilität, Reichweite und Dialogfähigkeit unterstützt das (X)HTML-Medium in der Berichterstattung zudem die Auffindbarkeit und Verständlichkeit der veröffentlichten Geschäftsinformationen, wodurch insbesondere Kleinanleger bei der Informationsverarbeitung unterstützt werden. Im Gegensatz zur papierbasierten Berichterstattung kann die Bereitstellung mit fortschrittlichen Onlinemedien entsprechend ausgestaltet werden, dass entscheidungsrelevante Informationen zeitnah auffindbar sind (vgl. Barrantes & Sagmüller, 2014, S. 62, Voelzkow, 2019, S. 124f.). Darüber hinaus gewährleistet das (X)HTML-Format die Möglichkeit, die teils komplexen und seitenlangen Geschäftsberichte mit einer erhöhten Verständlichkeit zur Verfügung zu stellen. Dazu kann bspw. eine Verlinkung eines Geschäftsberichts mit einem Glossar als Nachschlagewerk hergestellt werden, womit dem Adressaten Definitionen für unklare Fachbegriffe unmittelbar vorliegen. Weiterhin erhöht der Einsatz multimedialer Elemente die Verständlichkeit, da ein komplexer Geschäftszusammenhang anhand von z.B. Videos greifbarer gemacht werden kann (vgl. Godemann & Herzig, 2012, S. 239). Zusammenfassend führt die Verwendung von Onlinemedien, wie HTML- oder XBRL-Lösungen, für die Berichterstattung zu vorteilhaften Eigenschaften, die mit den Begriffen Verfügbarkeit, Auffindbarkeit und Reichweite erklärt werden können und zu Ressourcenschonung, Vernetzung, Multimedialität, Dialogorientierung, Personalisierung, Anschlusskommunikation und erhöhter Verständlichkeit und Vergleichbarkeit führen (vgl. Barrantes & Sagmüller, 2014, S. 63f.; Anhang D).

Durch die Implementierung einer einheitlichen Berichtsprache folgt eine Standardisierung, Automatisierung und Effizienzsteigerung im Rechnungswesen (vgl. Kretschmer, 2014, S. 169f.). Charakteristisch für ESEF und der Berichtssprache XBRL ist die erhöhte Effizienz von Informationen und eine verbesserte Auswertbarkeit und Analysemöglichkeit für Berichtsadressaten. Digitale Geschäftsberichte können in kürzester Zeit nach persönlich relevanten Informationen gefiltert und mit fortschrittlichen Anwendungen verarbeitet werden, wodurch zudem eine Vergleichbarkeit zwischen verschiedenen Unternehmen ermöglicht wird. Es liegen interaktive Suchfunktionen vor, in denen die vorliegenden Geschäftsinformationen in ihrer Gesamtheit nach bestimmten Begriffen durchsucht und erkennbar gemacht werden können. Daraus resultiert die Möglichkeit, Berichte z.B. nach einer benötigten Kennzahl zu filtern und die verschiedenen Perioden miteinander zu vergleichen, was mit größtmöglicher Transparenz einhergeht. Der Einsatz von Hyperlinks und anderen multimedialen Lösungen macht einen Bericht optisch ansprechender und für Adressaten in gewisser Hinsicht aussagekräftiger. Die Berücksichtigung solcher Such- und Interaktionsfunktionen stärkt die Informationstransparenz und schließlich die Reputation und das Vertrauen in das Unternehmen (vgl. Voelzkow, 2019, S. 116). Folgende Effizienzsteigerungen durch einen vereinfachten Informationsaustausch zwischen unterschiedlichen Softwaresystemen ist eine weitere wesentliche Chance digitaler Berichtssprachen. Die standardisierten Taxonomien unterstützen damit Offenlegungspflichten, da unterschiedliche Rechnungslegungsvorschriften wie HGB oder IFRS kompatibel sind und automatisch implementiert werden können (vgl. Kretschmer, 2014, S. 184f.; Ramin & Kesselmeyer, 2007, S. 568). Es wird darüber hinaus die Ad Hoc-Publizität von Unternehmen begünstigt, da eine unverzügliche Informationsbereitstellung obligatorisch ist, sobald neue Gegebenheiten vorliegen, die unmittelbare Auswirkung auf das Unternehmensgeschehen haben. Solche Fälle werden durch einen standardisierten und effizienten XBRL-Erstattungsprozess unterstützt (vgl. Grüning, 2010, S. 528). Zudem sind die Taxonomien in der Regel in diversen Sprachen vorhanden, wodurch eine reibungslose interne Kommunikation mit ausländischen Tochtergesellschaften stattfinden kann und anschließend die Berichtsqualität und -geschwindigkeit verbessert wird (vgl. Ramin & Kesselmeyer, 2007, S. 568). Insgesamt spielen unternehmensspezifische Angaben, multidimensionale Verlinkungen zwischen den Informationen und Visualisierungsmethoden erhebliche Funktionen in einer angemessenen Auslegung der digitalen Berichterstattung (vgl. Alles & Debreceny, 2012, S. 87; Bravidor & Loof, 2020, S. 138). Es scheint die dynamische Darstellung von Informationen auf Internetseiten von Vorteil, indessen die Verwendung des HTML-Formats und Hyperlinks als unterstützende Onlinemedien eine mehrdimensionale Bereitstellung fördern (vgl. Godemann & Herzig, 2012, S. 238ff.; Beattie & Pratt, 2003, S. 161). Diese Vorteile werden mit dem zusätzlichen Verlinkungscharakter von iXBRL/ XHTML verstärkt, indem neben der Verknüpfung von Informationen auf der Internetseite des Unternehmens mit anderen externen Quellen, zudem die Einbettung von verknüpften Angaben innerhalb des Online-Berichts durchgeführt und gleichzeitig maschinell und menschlich lesbarer gemacht werden. Diesbezüglich wird darüber hinaus das Potential erkannt, dass über die individuellen Erweiterungen der Taxonomien, im Gegensatz zu papierbasiert-erstellten Berichten, verborgene Informationen an den Markt veröffentlicht werden können und damit eine präzisere Informationstiefe und verbesserte Geschäftseinblicke gewährleistet werden (vgl. Beerbaum et al., 2021, S. 18).

Die Herausforderungen von Datenauthentizität und Manipulationstendenzen kann durch eine kontinuierliche und selbständige Prüfung der vorliegenden Informationen mit Hilfe einer implementierten BC überwunden werden. BC wird in zahlreichen Beiträgen als mögliche Alternative bzw. Ergänzung zu aktuellen Buchhaltungssystemen angesehen, wodurch die Buchhaltungsfunktion maßgeblich erweitert und verändert werden kann (vgl. Coyne & McMickle, 2017, S. 110; Kokina, Mancha & Pachamanova, 2017, S. 97; Faccia & Mosteanu, 2019, S. 108ff.). Die BC-Technologie kann die Gefahren bewusster Manipulation oder menschlicher Fehler verringern und gleichzeitig Kosten- und Zeitersparnisse bei dem Prozess der Berichterstattung ermöglichen. Damit hat die Technologie einflussreiches Potenzial und macht es dadurch erforderlich, dass zuständige Mitarbeiter vorhandene Kenntnisse kontinuierlich erweitern und neue Fähigkeiten erlernen, um traditionelle Buchhaltungstätigkeiten und repetitive Abstimmungen auf wertschöpfende Tätigkeiten in den Bereichen Managementunterstützung und Investor Relations verlagern zu können (vgl. Schmitz & Leoni, 2019, S. 338; Fullana & Ruiz, 2021, S. 78f.; Fuller & Markelevich, 2020, S. 37; Pugna & Duţescu, 2020, S. 7ff.). Darüber hinaus findet eine Verbesserung der Qualität der Daten für die digitale Berichterstattung und eine Steigerung der Effektivität und Effizienz von z.B. Erstattungsprozessen aufgrund sinkender Fehlerraten in Datenbanken statt (vgl. Gulin et al., 2019, S. 431; Tan & Low, 2019, S. 314). Die teilnehmenden Akteure scheinen sich durch die Implementierung der BC-Technologie im Rechnungslegungsprozess und damit in der digitalen Berichterstattung auf komplexe und intuitive Themengebiete konzentrieren zu können und ein tieferes Prozessverständnis zur Berichtserstellung aufzubauen. Wobei Tätigkeiten, die z.B. ausgeprägte menschliche Fähigkeiten oder Kreativität innerhalb der Berichterstattung erfordern, weiterhin vom Mensch ausgeführt werden (vgl. Gulin et al., 2019, S. 434f.; Fullana & Ruiz, 2021, S. 78f.). Ansprechend bleibt die Tatsache, dass die Technologie durch den integrierten Mechanismus automatisch Informationsauthentizität zu gewährleisten scheint. Wenn das webbasierte Netzwerk die Ausführung und Buchung von Geschäftsvorfällen selbstständig ausführt und Informationsauthentizität konsequent gewährleisten kann, ist fraglich, ob in Zukunft z.B. Abschlussprüfer in einem ausgeprägten Maße zu involvieren sind, wodurch veröffentliche Geschäftsberichte ohne anschließende Prüfung zeitnah in den Markt eingepreist werden können. Diese Möglichkeit würde im Sinne der EMH stehen, da ein beträchtlicher Ressourcenaufwand für die

Qualität und Validierung der veröffentlichten Berichte aufgebracht wird und damit dem Markt automatisch zeitnah validierte Informationen zur Verfügung gestellt werden. Die Vorteile der BC-Technologie können unter den Begriffen Dezentralisierung, Verifizierung, Resilienz, Automatisierung und angemessene Datenverschlüsselung und -qualität zusammengefasst werden (vgl. Cao, Cong et al., 2020, S. 26f.; Smith, 2018, S. 241f.).

Bestimmte Automatisierungstechnologien halten diesbezüglich Schritt und ermöglichen eine Prozessdurchführung und Entscheidungsfindung mit einem verringerten Aufwand, die ohne menschliche Interaktion effektiver vollzogen werden. Automatisierte Entscheidungen sind somit frei von menschlichen Verzerrungen und bewussten Manipulationen und verbessern damit die Qualität der digital unterstützten Berichterstattung. Die genannten Vorteile von RPA sind Kostenreduktion, erhöhte Prozessgeschwindigkeit und -kontrolle, Resilienz, die Möglichkeit zum Dauerbetrieb und entlastende Auswirkungen auf involvierte Mitarbeiter, da repetitive und mühselige Aufgaben von Software übernommen wird (vgl. Jędrzejka, 2019, S. 141f.).

Durch den fortschrittlichen Einsatz von Methoden aus dem Bereich KI kann zudem eine intelligente Automatisierung von Datenschöpfung und -verarbeitung zu Gunsten einer digitalen Berichterstattung gewährleistet werden. Durch solche Automatisierungstendenzen ändert sich das Rollenbild im Rechnungswesen, da fortschrittliche Fähigkeiten in neuen Fachdisziplinen notwendig und repetitive Aufgaben durch KI-Anwendungen ersetzt werden können, was schließlich positive Effekte auf die Qualität der Informationen hat (vgl. Wilson & Sangster, 1992, S. 74; Lin, 2014, S. 51; Anhang B). Der Wandel wird trotz einiger Herausforderungen als Chance interpretiert, da zusätzliche Qualifikationen und Fähigkeiten erlangt werden und eine digitale Weiterentwicklung des Rechnungswesens und damit der gesamten Organisation ermöglicht wird (vgl. Gänßlen & Losbichler, 2014, S. 146). Buchhalter spielen eine Schlüsselrolle in der Implementierungsphase einer KI-basierten Buchhaltungssoftware. Da es sich bei KI-Methoden teilweise um selbstlernende Systeme handelt, ist es die Aufgabe des Buchhalters, die KI entsprechend der spezifischen Unternehmensbedürfnisse anzupassen. Nach einer erfolgreichen Implementierung wird der Buchhalter in speziellen Fällen eingreifen und dient damit als Schnittstelle zwischen Geschäftstätigkeit und IT-Landschaft (vgl. Leitner-Hanetseder et al., 2021, S. 9). Mit Hilfe von unterschiedlichen KI-Anwendungen kann eine für die Berichterstattung zuständige Abteilung zu einem strategischen Berater des Managements werden und entscheidende Geschäftseinblicke liefern. KI wird insbesondere in Zeiten von Big Data einen besseren Zugang zu Daten in Echtzeit aus einer größeren Anzahl von Quellen ermöglichen und durch angemessenen Einsatz von fortschrittlichen Anwendungen die digitale Berichterstattung weiterentwickeln. Ein Beispiel ist die optische Zeichenerkennung innerhalb der Berichterstattung, die eine unstrukturierte Rechnung selbständig zuordnen, buchen und einen angemessenen Zahlungsvorgang vor dem Hintergrund des Liquiditätsmanagements vorschlagen kann. Die zukünftige Entfaltung solcher Verfahren und Technologien wird, neben dem Abgleich von Bestellnummern und Rechnungsbeträgen, zu einem tieferen Verständnis des Geschäftsmodells und der vorliegenden Informationen führen, was durch erhöhte Datenzugänglichkeit intensiviert wird (vgl. Gulin et al., 2019, S. 432; Leitner-Hanetseder et al., 2021, S. 9).

Die Kombination des Rechnungswesens mit Onlinemedien und Technologien aus dem Bereich XBRL, BC, RPA und KI haben das Potential, die gesamten Prozesse des externen Rechnungswesens und gleichzeitig die beiläufige Abschlussprüfung auf einer autonomeren und selbstständigen Basis durchzuführen. Interne Kontrollsysteme des Unternehmens und extern verpflichtende Prüfungsverfahren werden mit Medien und Technologien ausgestattet sein, die eine kontinuierliche Prüfung der gebuchten Geschäftsvorfälle zeitnah und automatisch vollziehen. Zudem werden durch RPA standardisierte Abschlusserstellungsschritte mit Lösungen, wie der BC-Technologie unterstützt und ausgegliedert. Eine angemessene Kombination wird sich auf die Reduzierung des manuellen Aufwandes und die Steigerung der Effizient, Reibungslosigkeit und Qualität der Informationen innerhalb der digitalen Berichterstattung auswirken und damit diverse Anforderungen unterschiedlicher Adressaten erfüllen (vgl. Fuller & Markelevich, 2020, S. 44; Gulin et al., 2019, S. 505f.; Anhang D). Darüber hinaus zeigt der Trend zur Hyperautomatisierung in Verbindung mit No-Code/Low-Code-Plattformen die wahren Fähigkeiten der zukünftigen software-gestützten und intelligenten Automatisierung, insbesondere im Umfeld des Prozess- und Datenmanagements und wird damit die Digitalisierung in der Berichterstattung verstärkt vorantreiben (vgl. Lasso-Rodriguez & Winkler, 2020, S. 19; Talesra & Nagaraja, 2021, S. 350f.). Wobei sich Folgen und Konsequenzen aus dem Einsatz von fortschrittlichen Disziplinen, wie QC in der Berichterstattung noch nicht präzise deuten lassen.

4.2. Herausforderungen und Risiken des neuen Mediums

Für Unternehmen bestehen neben den Chancen zudem Herausforderungen und Risiken bei der Veröffentlichung von Geschäftsinformationen über Kanäle des neuen Mediums, da Datenauthentizität, Datenintegrität und Manipulationsfreiheit insbesondere im Datenzeitalter nicht problemlos gewährleistet werden können. Die Unternehmensleitung hat auf der einen Seite ein Interesse und steht auf der anderen Seite in der Pflicht, Authentizität und Integrität der Informationen zu gewährleisten und Manipulationen zu vermeiden und damit ein regelkonformes und nicht verzerrtes Abbild der Geschäftstätigkeit offengelegt wird (vgl. IASC, 1999, S. 66; Schruff & Kayser, 2002, S. 355). Die dynamischen Eigenschaften des HTML-Formats, im Gegensatz zur traditionellen statischen Berichterstattung, sind die wesentlichen Gründe für das bestehende Risiko der Manipulationsanfälligkeit, da Informationen in digitalen Darstellungen problemlos (bewusst oder unbewusst) verändert werden können. Ohne angemessene Sicherheitsvorkehrungen können Informationen abgefangen, verändert und gefälscht werden (vgl. Boritz

& No, 2005, S. 11f.; Gerpott & Hoffmann, 2006, S. 371). Im Hinblick auf diese Risiken werden Datenauthentizität und -integrität als wesentliche Pfeiler bei der Haftung von Daten genannt und spielen eine besondere Rolle bei den Anforderungen der Adressaten. Fehlende Datenauthentizität liegt vor, wenn bspw. die veröffentlichten Abschlussangaben im Konflikt mit den vom Abschlussprüfer kontrollierten Gegebenheiten stehen (vgl. Debreceny & Gray, 1999, S. 335ff.; Küting, Dawo & Heiden, 2002, S. 47ff.). Wohingegen das Risiko der fehlenden Integrität von Daten auf den Einsatz des z.B. HTML- oder XBRL-Formats zurückzuführen ist. Berichtsadressaten sind zunehmend der Herausforderung ausgesetzt, dass aufgrund von Verknüpfungen zu externen Quellen durch Links oder andere multimediale Lösungen nicht zu erkennen ist, welche Angaben relevant, vertrauenswürdig und geprüft sind und welche nicht, worunter die Struktur und der Überblick der Berichterstattung leidet (vgl. Debreceny & Gray, 1999, S. 335ff.; Schruff & Kayser, 2002, S. 356; Küting et al., 2002, S. 47ff.; Adlmaier-Herbst & Musiolik, 2017, S. 45f.). Es wird die Herausforderung eines Balanceaktes zwischen einer umfangreichen und vollständigen im Gegensatz zur personalisierten und bestimmten Berichterstattung wiederholt deutlich.

Die größte Herausforderung bei der Erreichung von Vergleichbarkeit ist das fehlende Verständnis auf der Ebene der Geschäftsprozesse und einzelnen Datensätze zwischen Unternehmen. Die Veröffentlichung von Informationen auf Grundlage unterschiedlicher Präsentationsformate, Zwecke und Kontexte macht eine Vergleichbarkeit von Natur aus komplex. Das Fehlen eines gemeinsamen Ansatzes führt automatisch zu erheblichen Vereinheitlichungsproblemen und obwohl Geschäftsinformationen die gleiche semantische Bedeutung haben können, entstehen Herausforderungen durch individuelle Definitionen und/oder Interpretationen (vgl. Lampathaki et al., 2009, S. 1052). Vor diesem Hintergrund bestehen in der Praxis Gestaltungsmöglichkeiten im Erstellen von Geschäftsberichten und der anschließenden Darstellungsform. In Verbindung mit dem Prozess des Jahresabschlusses ist häufig die Rede vom Window Dressing, demzufolge das Rechnungswesen bilanzpolitische Handlungsspielräume nutzt, um die zweckmäßige Außendarstellung zu verbessern und Adressaten damit zu beeinflussen. Dieses allseits bekannte Risiko innerhalb der Berichtspraxis ist insbesondere in der digitalen Berichterstattung zu berücksichtigen, da die berichtenden Unternehmen für die Aufbereitung und Verwaltung der dynamischen Angaben auf Webseiten und in digitalen Berichten selbst verantwortlich sind (vgl. Winter, Saunders & Hart, 2003, S. 309ff.). Es entstehen bei Onlinemedien neue Herausforderungen, da z.B. durch bewusstes Zurückhalten oder Hervorheben einzelner dynamischer Elemente, die Wahrnehmung des Unternehmens in der Öffentlichkeit gesteuert werden kann. Insbesondere anpassbare Grafiken und Tabellen bieten das Risiko einer verzerrten Wahrnehmung. Obgleich visuelle Elemente die Verständlichkeit des Geschäftsberichts grundsätzlich bereichern, besteht ein Risiko, dass die Visualisierung von Informationen in einem Maße eingesetzt wird, sodass grundsätzlich relevante, allerdings unvorteilhafte Angaben in den digitalen Hintergrund gedrängt werden (vgl. Gazdar, 2014, S. 179; KPMG, 2015, S. 9). Solch eine Flexibilität in der Berichterstattung ist in der heutigen volatilen Geschäftswelt zwar unabdingbar, steht hingegen im Konflikt einer wahrheitsgemäßen und fairen Berichterstattung (vgl. Ajekwe, 2021, S. 82). Insbesondere die Visualisierung großer Datenmengen stellt zwar eine einflussreiche, hingegen unerforschte Einsatzmöglichkeit dar. Eine weitreichende Datenbasis ermöglicht es zwar, Muster zu erkennen und darzustellen, wohingegen solche explorativen Analysen ausschließlich auf ökonomisch nachweisbaren Datensätzen und Transaktionen einzusetzen sind, denn ab bestimmten Datenmengen bestehen Risiken von Scheinkorrelationen oder Verzerrungen und damit fehlerhaften Rückschlüssen (vgl. Frankel & Reid, 2008, S. 30; Singh & Best, 2016, S. 58f.; Cockcroft & Russell, 2018, S. 326).

Darüber hinaus ergeben sich für Unternehmen aus der Big Data-Entwicklung weitere Herausforderungen und Risiken, die mit den grundlegenden Charakteristika einhergehen und sich aus prozessualen und verwaltungstechnischen Anforderungen ergeben. Die hohe Geschwindigkeit und Kurzlebigkeit von Daten bergen das Risiko, dass nicht vollständige bzw. veraltete Datensätze bereitgestellt werden. Dies erschwert die Verarbeitung und verzerrt die Wahrnehmung. Daraus resultierend können Datensätze nutzlos werden, wenn der Kontext fehlt oder eine fehlerhafte Verknüpfung vorliegt (vgl. Sivarajah, Kamal, Irani & Weerakkody, 2017, S. 265; Janssen, van der Voort & Wahyudi, 2017, S. 342). Gewisse Datenmengen erschweren zudem die Speicherung trotz fortschrittlicher Datenbanksysteme. Solche Prozessherausforderungen treten in allen Phasen der Verarbeitungskette auf und sorgen für vielfältige Verwaltungsrisiken. Insbesondere stellt sich die Frage der externen Datenhoheit. Dies betrifft vor allem externe Quellen, wie soziale Medien, was in den letzten Jahren zu Skandalen durch Ausnutzung von Nutzerprofilen geführt hat (vgl. Kaisler, Armour, Espinosa & Money, 2013, S. 997ff.; Sivarajah et al., 2017, S. 273f.; Schwegler, 2018). Zudem besteht die Gefahr bei dem Vorhandensein großer Datenmengen zu glauben, unbeeinflussbare Faktoren steuern zu können und damit schließlich einer Kontrollillusion zu unterliegen (vgl. Langer, 1975, S. 311 ff.; Weißenberger & Bauch, 2017, S. 213f.).

Der unmittelbare und personalisierte Dialog durch Interaktionslösungen und Social-Media-Profile stellt neben den Chancen eine Herausforderung dar. Innerhalb kürzester Zeit können sich negative Meinungsbilder über ein Unternehmen verbreiten und damit dem Image des Unternehmens langfristig schaden. Wenn die Aktivitäten des Unternehmens in den Augen der Gesellschaft nicht als legitim erachtet werden und sich eine Missbilligung des Unternehmens digital verselbständigt, kann ein Unternehmen unmittelbar die soziale Betriebserlaubnis verlieren (vgl. Demuijnck & Fasterling, 2016, S. 675; Meesters et al., 2021, S. 7). Negative Bewertungen im Internet halten sich oft dauerhaft und sind schwierig zu verhindern, wenn die kollektive Meinungsmacht von sozialen Medien unberücksichtigt bleibt (vgl. Greven & Lahme, 2014, S. 111). Insbesondere Kleinanleger haben während

der Pandemiejahre 2020 und 2021 einen erleichterten Zugang zum unkomplizierten Handel mit Wertpapieren, mehr Finanzierungsmöglichkeiten und einer höheren Informationsdichte über diverse Social-Media-Plattformen erhalten. Hierbei stellt sich die Frage, ob öffentliche Social-Media-Aktivitäten von Unternehmen, andererseits anonyme Engagements auf Plattformen wie z.B. Reddit den Kleinanlegern helfen, bessere Entscheidungen zu treffen. Es wurde ersichtlich, dass eine hohe Frequenz, ein positiverer Ton von Beiträgen und Kommentaren, mehr Meinungsverschiedenheiten und eine hohe Verbundenheit auf Social-Media-Plattformen zu signifikanten Kursänderungen führen. Das heißt, dass Social-Media-Aktivitäten das Kaufverhalten von Kleinanlegern beeinflussen, sodass die Wirkung solcher Aktivitäten nicht zu ignorieren ist und somit Unternehmen in Zukunft zur angemessenen und zeitgemäßen Berichterstattung Social-Media-Plattformen berücksichtigen sollten. Diesbezüglich wird zugleich an Regulierungsbehörden und politische Entscheidungsträger appelliert, kontinuierlich Kapitalmarktteilnehmer auf Social-Media-Plattformen zu überwachen, da durch solche Aktivitäten und teilweise Falschinformationen Ineffizienz auf dem Markt erzeugt wird, was wiederum die Signifikanz einer angemessenen Implementierung von Onlinemedien in die digitale Berichterstattung von Unternehmen bestätigt (vgl. D. Hu, Jones, Zhang & Zhang, 2021, S. 27f.; Umar, Gubareva, Yousaf & Ali, 2021, S. 6f.; Mittelbach-Hörmanseder et al., 2021, S. 23f.).

Trotz der diversen Chancen entstehen aus der Nutzung von BC Herausforderungen und Risiken. Bisher sind Probleme der technischen Regulierung und der Sicherheit von Daten und Informationen nicht vollständig gelöst. Das Vertrauen in die Technologie ist nicht von einer einzigen Instanz abhängig, sondern von der dezentralen Transaktionsaufzeichnung, den Akteuren und damit der Integrität des gesamten, teils komplexen Netzwerks (vgl. Schmitz & Leoni, 2019, S. 338; Karajovic, Kim & Laskowski, 2019, S. 319). Darüber hinaus liegt die Frage vor, inwieweit BC zukünftig in die Berichtspraxis integriert werden wird. Eine Umfrage aus dem Jahr 2019 verdeutlicht, dass sich lediglich jedes zehnte Unternehmen mit der Einführung von BC im Rechnungswesen beschäftigen. Es scheint bis dato keine signifikante Rechtfertigung für eine Integration der BC in den Geschäftsprozess der Berichterstattung vorzuliegen (vgl. Fuller & Markelevich, 2020, S. 44; PWC, 2019, S. 19). Zudem muss BC erst von einer angemessenen Anzahl an Organisationen akzeptiert und implementiert worden sein, um als neue Technologie im Rechnungswesen flächendeckend wirksam zu werden (vgl. Inghirami, 2018, S. 13; Pugna & Duţescu, 2020, S. 221). Diese Voraussetzung gilt grundsätzlich zudem für die Wahrnehmung der Nutzenstiftung einer digitalen Berichterstattung und wird damit primär durch die Adoptionsraten beeinflusst. Da positive externe Effekte eines Netzwerkes innerhalb der digitalisierten Berichterstattung mit der Anzahl der Anwender wachsen, muss zunächst eine kritische Masse an Unternehmen einen Standard nutzen, bevor die Vorteile in der Berichterstattung wahrgenommen werden (vgl. Troshani & Rowbottom, 2021, S. 1; Đurović, Rep & Dečman, 2021, S. 66). Forscher und Experten der Buchhaltungspraxis sind sich abschließend uneinig, welche Veränderungen und Risiken durch die BC-Technologie zu erwarten sind, was die Beschäftigung mit möglichen Auswirkungen auf die digitale Berichterstattung weiterhin ausausweichlich macht.

Abschließend stehen den Vorteilen, die der Einsatz von KI verspricht, neue Herausforderungen und Risiken gegenüber. Der Einsatz von KI-Anwendungen verschärft das Risiko einer Verzerrung selbständiger Programme. Darüber hinaus tendieren aktuelle ML- und insbesondere KNN-Ansätze häufig dazu, keine Informationen über den Entscheidungsprozess zu gewährleisten und somit in der Form einer Blackbox einen Mangel an Transparenz zu liefern. Dadurch entsteht die Herausforderung bei Anwendern, den teilweise intransparenten Ergebnissen zu vertrauen und der Technologie mit Akzeptanz zu begegnen. Neben technischen Sicherheitsaspekten stellt sich zusätzlich die Frage, welche Rahmenwerke den Einsatz von KI ethisch und rechtlich leiten sollen. Studienergebnisse deuten darauf hin, dass die zunehmende Anzahl von KI-unterstützen Adressaten Unternehmen dazu motiviert, ihre Berichte von vornherein maschinenfreundlicher zu gestalten, was die wachsende Rolle von KI auf den Finanzmärkten und damit auf die Berichterstattung verdeutlicht. Unternehmen steuern bewusst die Wahrnehmung von Berichten, indem z.B. Wörter, die von Algorithmen als negativ wahrgenommen werden, vermieden werden, was womöglich zu einer verzerrten Abbildung der Geschäftstätigkeit führen kann und nicht dem Zweck des Rechnungswesens entspricht (vgl. Cao, Jiang, Yang & Zhang, 2020, S. 28).

4.3. Vom neuen Medium zur digitalisierten Berichterstattung

Aus der kritischen Würdigung der Chancen, Herausforderungen und Risiken[5] der Folgen und Konsequenzen für den Einsatz des neuen Mediums in der Berichterstattung ist abzuleiten, dass Onlinemedien weitaus mehr als neue fortschrittliche Komponenten der digitalen Unternehmenskommunikation darstellen, nämlich bei angemessenem Einsatz in Verbindung mit Medien und technologischen Lösungen eine ganzheitlich digitalisierte und reibungslose Berichterstattung schaffen. Der Übergangsprozess ist fließend zu verstehen und die Wirkung digitaler Ausdrücke auf die Syntax und Semantik der Kommunikation zwischen den Erstellern und Nutzern von Geschäftsberichten zu berücksichtigen. Der Ort und die Methode der Datenerfassung, die gesammelten Variablen, die integrierten Informationen und die Frequenz der Berichterstattung sind einem paradigmatischen Wandel unterworfen, wodurch sich das Rechnungswesen und damit die Berichterstattung grundlegend verändern wird. Das Tempo und die Akzeptanz dieser Veränderungen in der Berichterstattung sind unvorhersehbar, was vor allem an den starren regulatorischen Mechanismen und dem Widerwillen von Unternehmen bei der Nutzung von Onlinemedien zur Veröffentlichung von Geschäftsinformationen liegt. Diesbezüglich

[5] Ein Überblick der ausgearbeiteten Chancen, Herausforderungen und Risiken wurde als Anhang E zusammenstellt.

wird die Rolle von Onlinemedien und Technologien für die Kommunikation zwischen Unternehmen und Interessengruppen verstärkt und es entstehen vorteilhafte Möglichkeiten, eine Informations- und Markteffizienz in Verbindung einer digitalisierten Berichterstattung durch die Interdependenz von z.B. Big Data, BC oder XBRL zu schaffen. Hierbei wurden die ersten Wirksamkeitsüberlegungen unternommen, die Datenintegrität eines XBRL-Berichts durch BC zu unterstützen (vgl. Boixo, Mora & Ruiz, 2019, S. 1ff.). Die Führungskräfte von Unternehmen müssen schließlich anerkennen, dass fortgeschrittene Fähigkeiten und die Implementierung digitaler Lösungen in die Unternehmenstätigkeit zur Notwendigkeit für eine zeitgemäße und effiziente Daten- und Informationsaufbereitung geworden sind. Onlinemedien innerhalb der Berichterstattung in Verbindung mit fortschrittlichen Technologien und daraus resultierenden Tendenzen wie Hyperautomatisierung gehören mittlerweile zu Faktoren, weshalb neue Normen etabliert, traditionelle Managementpraktiken überarbeitet und ein Datenfokus eingerichtet wird. Die Chancen der digitalen Berichterstattung sind bedeutungsvolle Wettbewerbsfaktoren für den zukünftigen Erfolg des Unternehmens und der Mitarbeiter, insbesondere um digitale Rechenschaftspflichten, sowie freiwillige Ansprüche zu erfüllen. Die Wahrnehmung der Chancen birgt zunächst diverse Herausforderungen. Auf der einen Seite ist die digitale Infrastruktur des Unternehmens angemessen zu skalieren oder auszugliedern, wohingegen auf der anderen Seite die Kompetenzen der Belegschaft an die neuen Medien, Technologien und Arbeitsweisen anzupassen sind, um den notwendigen digitalen Rollenwandel stattfinden zu lassen. Hier liegt der Fokus auf digitalisierte und technik-unterstütze Arbeitsweisen und der Erweiterung statistischer und technologischer Fähigkeiten. Mitarbeiter müssen dazu Potentiale der Digitalisierung innerhalb der Berichterstattung erkennen und umsetzen. Mangelndes Wissen und Fähigkeiten können sich negativ auf die digitale Transformation eines Unternehmens auswirken und zum Marktaustritt führen. Hierbei ist Vorsicht walten zu lassen, wenn das Risiko von Informationsüberlastung droht. Falls solch eine vielfältige und umfangreiche Informationsgrundlage geschaffen wird, in der keine Authentizität bzw. Unterscheidung von geprüften von ungeprüften Informationen gewährleistet werden kann, folgen negative Konsequenzen auf beiden Seiten. Unternehmen können insoweit gegensteuern, indem z.B. Investoren proaktiv mit Lösungen versorgt werden, die individuelle Segmentierungen von Informationen ermöglichen und damit irrelevante Informationen nicht in die Entscheidungsunterstützung einfließen zu lassen.

Es ist davon auszugehen, dass sich Unterschiede in der Auslegung und Nutzung digitaler Berichtselemente zunehmend verringern. Institutionelle und private Investoren werden bei zukünftigen Anlageentscheidungen durch die andauernde Internationalisierung der Kapitalmärkte nicht ausschließlich regional ansässige Unternehmen berücksichtigen. Es ist mittlerweile notwendig, sich dem Markt im Hinblick auf Informationseffizienz und -transparenz im internationalen Vergleich anzupassen, um im Wettbewerb auf informierte und anspruchsvolle Investoren zu bestehen, wobei das ESEF einen geeigneten Rahmen für diese Entwicklung vorzugeben scheint. Bei den unterschiedlichen Implementierungsprojekten von ESEF gibt es eine Vielzahl von Interessengruppen, Chancen und Herausforderungen, die die Entwicklung einer digitalisierten und einheitlichen Berichterstattung beeinflussen. Im Fall von XBRL/ESEF gibt es ungeklärte Anliegen in Verbindung mit der technischen Regulierung der Sprache, der Struktur des Berichtsstandards und den Folgen konkurrierender Anforderungen zwischen der internationalen Standardsetzung, der nationalen Ausprägung und anderen regulatorischen Rahmenwerken. Diese Themen wurden im Zuge des Ländervergleichs in einem ersten Schritt adressiert, doch bedürfen diese weitere kritische Abwägung. Es wurde deutlich, dass es viele verschiedene XBRL-Projekte gibt, die sich mit den Möglichkeiten eines einheitlich elektronischen Berichtsstandards beschäftigen und auf diverse Rahmenbedingungen zurückzuführen sind. Auf Anhieb scheint eine solche Vielfalt von Vorteil, hingegen Hindernisse für die Erweiterung und Anwendung des Standards deutlich werden. Wenn die Auslegungen von XBRL-Taxonomien zu unterschiedlich sind, können Adressaten die länderübergreifenden Informationen nicht problemlos vergleichen. Hierbei liegt eine Vielzahl von Weiterentwicklungen vor, die für die Implementierung individueller Vorschriften und Gegebenheiten geschaffen wurden. Durch die Verbreitungsbemühungen und aktive Zusammenarbeit der umfangreichen Gemeinschaft wird der notwendige Kompromiss zwischen Vollständigkeit und Integrität der Ergänzungen erkennbar. Ausufernde Erweiterungen einer Basistaxonomie werden die Herausforderungen zur Vereinheitlichung nicht gänzlich beseitigen können. Stattdessen werden benutzerdefinierte Elemente aus gemeinschaftlich geführten Projekten selektiv in die nächste Fassung der Basistaxonomie verankert und sukzessiv eingebettet und weiterentwickelt. Ein solches Verfahren wird die Kompatibilität der Informationen mit geringfügigen Folgen für die Komplexität und Relevanz des gesamten Standards verbessern und wird sowohl die Gesamtheit, als auch die Integrität der Informationen verbessern und gleichzeitig gewährleisten, dass die veröffentlichten Informationen effizient bearbeitet und zwischen Unternehmen aus unterschiedlichen Ländern vergleichbar sind. Die Nähe anderer Rahmenwerke und Berichtsstandards zur Domäne von XBRL/ESEF ist unterdessen bedeutungsvoll für eine objektive Validierung des Potentials einer Zusammenlegung abweichender Herangehensweisen (wie es z.B. bei den zwei konkurrierenden Standards in Frankreich ersichtlich wurde, vgl. Guilloux et al., 2013). Darunter ist festzustellen, wie geeignet rechtliche Anweisungen zur Gewährleistung eines öffentlichen Standards sind, da vergleichbar zu anderen Open-Source-Projekten die Nützlichkeit solcher frei zugänglichen Rahmenwerke teilweise in Frage gestellt wird (vgl. Locke & Lowe, 2007, S. 610). Durch den Vergleich der EU mit Australien wurde darüber hinaus erneut der Kompromiss zwischen Vereinheitlichung und Nützlichkeit des Standards deutlich. Während die Entwicklung und die Umsetzung des Standards in der EU Freiheitsgrade genießt und z.B. die individuelle Auslegung von branchenspezifischen Taxonomien

ermöglicht wird, liegt in Australien ein geschlossener Ansatz vor, was mit Integrität einhergeht, wohingegen andere Zwecke gefährdet sind. Innerhalb der vorliegenden Projekte und Verfahren ist nicht zu bestimmen, was der richtige Ansatz ist, sondern unter welcher Voraussetzung welches Vorgehen als angemessen angesehen wird.

Es ist zu berücksichtigen, dass der Digitalisierungsgrad der Berichterstattung maßgeblichen Einfluss auf die Bewertung und das Image des Unternehmens zu haben scheint. Der Erfolg und die Legitimität des Unternehmens hängen entscheidend von Vertrauen ab. Die transparente Bereitstellung von Informationen über Onlinemedien stellt hierbei eine große Bandbreite an Möglichkeiten dar. Wenn ein Unternehmen mehr Effizienz durch die Nutzung von digitalen Auszeichnungs- und Berichtssprachen und eine Standardisierung durch das ESEF erreicht, wodurch mehr Transparenz, Informationstiefe und Vereinheitlichung gewährleistet wird, erhöht dies schließlich die Reputation und das Vertrauen in das Unternehmen. Durch die bedingungslose Gewährleistung eines konstruktiven Dialogs mit den Adressaten innerhalb einer mobil-interaktiven Berichterstattung wird eine positive Wahrnehmung weiter unterstützt. Es folgt, dass durch die Nutzung von Onlinemedien in der Berichterstattung zwar das Risiko eingegangen wird, dass das Image des Unternehmens Schaden nimmt und Wettbewerber mit zugänglichen Informationen versorgt werden, hingegen die Chance auf Image- und Vertrauenssteigerung überwiegt und die Resonanz auf dem Kapitalmarkt bei angemessener Umsetzung als positiv interpretiert werden kann. Diesbezüglich wurde aus Studien ersichtlich, dass eine XBRL-basierte Berichterstattung mit Verbesserungen bei einer Reihe von Kapitalmarktindikatoren (darunter z.B. Kapitalkosten) verbunden ist, die stellvertretend für Informations- bzw. Markteffizienz stehen (vgl. Troshani & Rowbottom, 2021, S. 9ff.; Sassi, Othman & Hussainey, 2021, S. 17ff.; Ruan, Liu & Tsai, 2021, S. 13ff.; Anhang D). Diese Evidenz stützt die These, dass die angestrebte digitalisierte Berichterstattung die Informationsmenge, die den Kapitalmarktteilnehmern im Datenzeitalter zur Verfügung steht, effizienter und hochwertiger zur Verfügung stellt und Transaktions- und Verarbeitungskosten reduziert werden, wodurch eine Verarbeitung unterstützt und die Informations- und Markteffizienz gesteigert wird. Schließlich haben weitere Studien, die diese Zusammenhänge zwischen verschiedenen Gruppen von Kapitalmarktteilnehmern untersuchen nahegelegt, dass die Auswirkungen der digitalen Berichterstattung auf Indikatoren insbesondere durch die angemessene Nutzung, als durch das bloße Vorhandensein von Daten beeinflusst werden. Somit ist es nicht ausreichend in Zeiten von Big Data zu agieren, sondern es werden fortschrittliche Methoden und Technologien zur digitalen Datenwertschöpfung unabdingbar (vgl. Troshani & Rowbottom, 2021, S. 15; Anhang D).

Hingegen liegen insgesamt widersprüchliche Ergebnisse vor, ob digital-personalisierte Berichte mit Verbesserungen von Kapitalmarktindikatoren verbunden sind, die auf ihrer Fähigkeit beruhen, detailliertere Informationen zu vermitteln und Informationsasymmetrien zu reduzieren, oder ob sie mit einer Verschlechterung der Indikatoren verbunden sind, die aus einer Verringerung der Vergleichbarkeit von Informationen und damit einem Anstieg der Verarbeitungskosten resultiert (vgl. Troshani & Rowbottom, 2021, S. 15; Liu, Luo & Wang, 2017, S. 48f.; Anhang D). Wohingegen z.B. Bilanzen und Kapitalflussrechnungen eine signifikante und robuste Konvergenz durch die XBRL-Implementierung erfahren haben, was auf eine erhöhte Vergleichbarkeit hindeutet (vgl. Yang, Liu & Zhu, 2018, S. 193ff.; Anhang D). Wiederum liegen Ergebnisse zur Analyse von US-amerikanischer Unternehmen vor, die eine Verringerung der Vergleichbarkeit nach dem XBRL-Mandat wahrgenommen haben. Hierbei ist zu berücksichtigen, dass ein Rückgang der Vergleichbarkeit nicht bedeutet, dass der Standard sein ursprüngliches Ziel verfehlt hat. Es steht zur Frage, ob XBRL bestehende Wechselwirkungen in Geschäftsinformationen und unterschiedlichen Berichten effektiver herausgearbeitet hat. Infolgedessen die Adressaten die Berichte zunächst aufwändiger, dafür schließlich wertstiftender interpretieren und einst verborgene Informationen und wechselseitige Geschäftseinblicke an den Markt veröffentlicht werden können, wodurch die Informationstiefe und -transparenz verbessert wird (vgl. Dhole, Lobo, Mishra & Pal, 2015, S. 43; Beerbaum et al., 2021, S. 18; Anhang D). Die gegenwärtige Umsetzung des ESEF in der EU sollte vor dem Hintergrund solcher Indikatoren und widersprüchlichen Ergebnisse vergangener Studien analysiert werden, um weitere Erkenntnisse hervorzubringen, die anschließend zu beurteilen und in die zukünftig fortgeführte Implementierung zu berücksichtigen sind.

Um einen abschließenden Überblick zu den Ursachen, Folgen und Konsequenzen für den Einsatz des neuen Mediums zu erhalten, folgt auf der nächsten Seite eine zusammenfassende Darstellung der digitalisierten Berichterstattung (die erarbeiteten Folgen und Konsequenzen wurden in Dunkelblau zu den entsprechenden Elementen hinzugefügt; vgl. Abbildung 4).

5. Fazit und Ausblick

Der Wandel zur digitalisierten Berichterstattung durch die Folgen und Konsequenzen für den Einsatz des neuen Mediums muss als individuelles Veränderungsmanagement verstanden werden, indem traditionelles betriebswirtschaftliches Denken und Handeln digital und datenfokussiert ausgerichtet und eine flexible Anpassungsfähigkeit und konsequente Lernbereitschaft in einem zeitgemäßen Rollenverständnis etabliert wird (vgl. Anhang B). Signifikante Faktoren stellen hierbei die Verbesserung der Informationsaufbereitung und -verarbeitung durch Onlinemedien, neue fortschrittliche Technologien und die Vereinheitlichungstendenzen der Berichterstattung in der EU durch ESEF dar. Die Implementierung eines Berichtsstandards fördert die Automatisierung der Informationsbereitstellung und die Effizienz der Verarbeitung, die medien- und technologieübergreifend umsetzbar ist und die individuellen Bedürfnisse diverser Adressaten angemessen befriedigt. Es wird eine einheitliche Struktur zur Aufbereitung und Verarbeitung von Daten

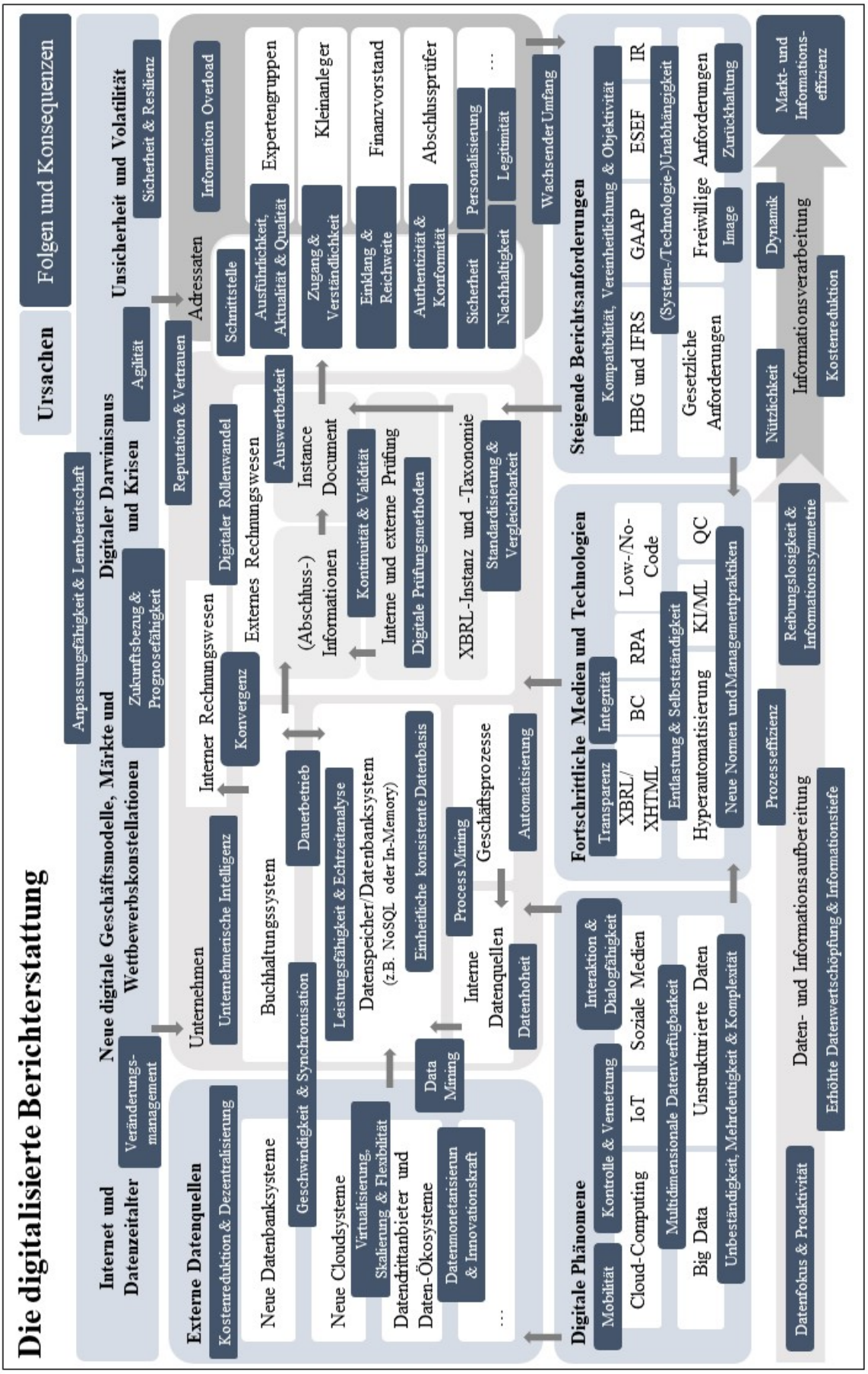


Abbildung 4: Folgen und Konsequenzen für den Einsatz des neuen Mediums (eigene Darstellung mit inhaltlicher Anlehnung an diverse Autoren aus der Arbeit).

und Informationen entlang ganzer Wertschöpfungsketten geschaffen, anstatt ausschließlich neue Rechnungslegungspraktiken einzuführen. Die Interaktionen zwischen dem Unternehmen und individuellen Adressaten und fortschreitende Veränderungen auf traditioneller und digitaler Ebene verschieben das Verständnis der Auswirkungen der Digitalisierung in der Berichterstattung über einen technologischen Darwinismus hinaus. Standardsetzer, Aufsichtsbehörden, Unternehmen und Berichtsadressaten werden von der Digitalisierung in der Berichterstattung beeinflusst, gestalten diese gleichzeitig mit und können zudem bei angemessener Herangehensweise zukünftige Veränderungen und Krisen agiler bewältigen (vgl. Raghavan, 2021, S. 7).

Die EU hat durch die Implementierung von XBRL in der Berichtspraxis einen erfolgreichen Standard für die verpflichtende digitale Berichterstattung kreiert. Die Kommunikation und der Austausch von umfangreichen Informationen auf einheitlicher Art, die Zusammenarbeit zwischen Softwareanwendungen und die automatische Verarbeitung besitzen das Potential, die Berichterstattung auf ein neues Level der Informations- und Markteffizienz zu heben, was sich insbesondere in Zukunft nach erfolgreicher und flächendeckender Umsetzung in der EU zu beweisen hat. Die hierbei ansteigende Verbreitung der involvierten Medien und Technologien stellt unter Beweis, dass die Implementierung des Standards als einheitlicher Rahmen einen eindeutigen Mehrwehrt für die Akteure innerhalb der digitalisierten Berichterstattung liefert. Es findet sich vermehrt Beteiligung an der Erstellung von Taxonomien, der Entwicklung entsprechender Software und damit insgesamt an der Förderung des gesamten Standards. Die Verwendung eines einheitlich elektronischen Berichtstandards ist nachgewiesenermaßen nützlich im Hinblick auf die Heterogenität der Systeme und Formate, den Erstellungs- und Verarbeitungsprozess, die Zielgruppenorientierung und die Vergleichbarkeit der veröffentlichten Informationen und bereichert damit die Digitalisierung in der traditionellen Berichterstattung. Darüber hinaus ist es zukünftig bedeutsam, inwiefern Erfahrungen zu den Auswirkungen von Marktkräften in der EU wie z.B. Kultur oder Politik auf das Bestreben eines einheitlichen Berichtsstandards gesammelt werden können (vgl. Roohani, Xianming, Capozzoli & Lamberton, 2010, S. 144) und welche verhaltensökonomische Rolle der Mensch spielt. Mit stetiger Einführung ist die Berichtspraxis zukünftig verstärkt als eine Informationskette zu betrachten, in der veröffentlichte Informationen nicht lediglich abschließenden Output, sondern darüber hinaus weiterführenden Input in der digitalisierten Berichterstattung darstellen (vgl. Bonsón, 2001, S. 107).

Aus den erarbeiteten Chancen, Herausforderungen und Risiken der Folgen und Konsequenzen für den Einsatz des neuen Mediums wird ersichtlich, dass neue Ansätze und Medien in der Berichterstattung im Datenzeitalter die Konnektivität der Unternehmen und Gesellschaft nutzen, um die Bedürfnisse der Öffentlichkeit und gleichzeitig individueller Adressaten langfristig zu befriedigen. Aus einem angemessenen Einsatz entsteht ein Mehrwert, da es ermöglicht wird, Informationen zeitnah und in unterschiedlichsten Formaten und Kanälen bereitzustellen. Es entsteht gleichzeitig eine maßgeschneiderte und qualitativ hochwertige Berichterstattung für ausgewählte Adressaten, eine leicht zugängliche und verständliche Informationsbereitstellung für die Öffentlichkeit und Kleinanleger, sowie eine zweckorientierte Veröffentlichung für bestimmte Rechenschaftspflichten. Wodurch dem Unternehmen eine ausgewogene intelligente Berichterstattung aller Adressaten gelingt und somit der Zweck des externen Rechnungswesens in Zukunft weiterhin effizient und zeitgemäß erfüllt werden kann. Es lässt sich somit abschließend festhalten, dass durch den angemessenen Einsatz des neuen Mediums große und unbeständige Datenmengen, die den Kapitalmarktteilnehmern zur Verfügung stehen mit reduzierten Kosten und erhöhter Qualität zugänglicher gemacht werden und es liegen reibungslos auswertbare Formate vor, wodurch die Markt- und Informationseffizienz zu Gunsten der EMH gesteigert werden kann. Es wird ersichtlich, dass die langfristigen Chancen die kurzfristigen Herausforderungen und Risiken überwiegen und ein fortschreitender Wandel des Berufsstandes stattfindet, der durch neue Trends und den Einsatz fortschrittlicher Technologien angetrieben wird. Das neue digitale Rollenverständnis mit neu geforderten Fähigkeiten in Zeiten von Big Data, Cloud Computing und IoT sind im vollen Gange den Prozess der Berichterstattung, sowie das Handwerk der Mitarbeiter und Berichtsadressaten von Grund auf nachhaltig zu verändern.

Literatur

Aaltonen, A. & Penttinen, E. (2021). What makes data possible? A sociotechnical view on structured data innovations. In *Proceedings of the 54th hawaii international conference on system sciences* (S. 5922–5931).

Accountancy Europe. (2021). *European Single Electronic Format (ESEF) - Collected guidance.* Zugriff auf https://www.accountancyeurope.eu/publications/esef-guidance (Abfrage: 08.08.2021)

Adlmaier-Herbst, D. G. & Musiolik, T. H. (2017). Digital Storytelling als intensives Erlebnis–Wie digitale Medien erlebnisreiche Geschichten in der Unternehmenskommunikation ermöglichen. In A. Schach (Hrsg.), *Storytelling* (S. 33–59). Springer.

Ajekwe, C. C. (2021). Impact of Flexibility in Accounting on Financial Reporting. *European Journal of Accounting, Auditing and Finance Research*, *9* (3), 74–87.

Al-Htaybat, K. & von Alberti-Alhtaybat, L. (2017). Big Data and corporate reporting: impacts and paradoxes. *Accounting, auditing & Accountability Journal*, *30* (4), 50–873.

Alles, M. & Debreceny, R. (2012). The evolution and future of XBRL research. *International Journal of Accounting Information Systems*, *2* (13), 83–90.

Amrhein, D. G., Farewell, S., Pinsker, R. et al. (2009). REA and XBRL GL: Synergies for the 21st century business reporting system. *The International Journal of Digital Accounting Research*, *9* (15), 127–152.

Ashbaugh, H., Johnstone, K. M. & Warfield, T. D. (1999). Corporate reporting on the Internet. *Accounting Horizons*, *13* (3), 241–257.

Assunção, M. D., Calheiros, R. N., Bianchi, S., Netto, M. A. & Buyya, R. (2015). Big Data computing and clouds: Trends and future directions. *Journal of Parallel and Distributed Computing*, *79*, 3–15.

Baldwin, A. A. & Trinkle, B. S. (2011). The impact of XBRL: A Delphi investigation. *The International Journal of Digital Accounting Research*, *11* (1), 1–24.

Ball, R., Kothari, S. & Robin, A. (2000). The effect of international institutional factors on properties of accounting earnings. *Journal of Accounting and Economics*, *29* (1), 1–51.

Barrantes, E. & Sagmüller, M. (2014). Entwicklungsstand und Perspektiven der Nachhaltigkeitsberichterstattung im Internet. In *CSR und Social Media* (S. 59–75). Springer.

Bartolacci, F., Caputo, A., Fradeani, A. & Soverchia, M. (2020). *Twenty years of XBRL: What we know and where we are going.* Zugriff auf http://eprints.lincoln.ac.uk/id/eprint/42267 (Abfrage: 15.04.2021)

Basoglu, K. A. & White Jr, C. E. (2015). Inline XBRL versus XBRL for SEC reporting. *Journal of Emerging Technologies in Accounting*, *12* (1), 189–199.

Beattie, V. & Pratt, K. (2003). Issues concerning web-based business reporting: An analysis of the views of interested parties. *The British Accounting Review*, *35* (2), 155–187.

Beerbaum, D. & Piechocki. (2017). Besteht ein Konflikt zwischen prinzipienbasierter Standardsetzung und strukturierter elektronischer Berichterstattung mit XBRL. *Research Gate*, *12* (7), 1–15.

Beerbaum, D. O., Duncan, K., McNamara, R., Ikaheimo, S. & Derichs, D. (2021). Digitalization and assurance of IFRS financial reports. *Social Science Research Network 3855045*, 1–38.

Beerbaum, D. O. & Piechocki, M. (2016). On the path to an European Single Electronic Format–ESMA consultation for the IFRS taxonomy of structured electronic reporting. *Social Science Reserch Network*, 1–3.

Bhimani, A. & Willcocks, L. (2014). Digitisation, 'Big Data' and the transformation of accounting information. *Accounting and Business Research*, *44* (4), 469–490.

Bogasiu, I. & Ardeleanu, N. (2021). Advantages and disadvantages of digitalisation in accounting. *EIRP Proceedings*, *16* (1), 294–299.

Boixo, I., Mora, J. & Ruiz, J. (2019). Proof of concept for an XBRL report indexer with integrity and non-repudiation secured by Blockchain using a smart contract: XBRLchain demo. , *44*, 1–12.

Bonsón, E. (2001). The role of XBRL in Europe. *The International Journal of Digital Accounting Research*, *1* (2), 101–110.

Bonsón, E., Cortijo, V. & Escobar, T. (2009). Towards the global adoption of XBRL using International Financial Reporting Standards (IFRS). *International Journal of Accounting Information Systems*, *10* (1), 46–60.

Bonsón, E. & Escobar, T. (2006). Digital reporting in Eastern Europe: An empirical study. *International Journal of Accounting Information Systems*, *7* (4), 299–318.

Boritz, J. E. & No, W. G. (2005). Security in XML-based financial reporting services on the Internet. *Journal of Accounting and Public Policy*, *24* (1), 11–35.

Boyer-Wright, K., Summers, G. & Kottemann, J. (2010). XBRL: Is it time. *Issues in Informing Science and Information Technology*, *7* (7), 509-518.

Brandt, W. & Hütten, C. (2003). Die Finanzberichterstattung international börsennotierter Unternehmen im Licht der jüngsten Entwicklungen–Das Beispiel SAP. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, *55* (7), 707–721.

Bravidor, M. & Loof, J. (2020). XBRL-Reporting in Europa: Eine empirische Analyse des EuroStoxx 50. *Zeitschrift für internationale und kapitalmarktorientierte Rechnungslegung*, *20* (4), 178–180.

Brennan, N. M. & Merkl-Davies, D. M. (2018). Do firms effectively communicate with financial stakeholders? A conceptual model of corporate communication in a capital market context. *Accounting and Business Research*, *48* (5), 553–577.

Bundesministerium für Wirtschaft und Energie. (2019). *Künstliche Intelligenz - Impulse zu einem Megatrend, Initiative Intelligente Vernetzung.* Zugriff auf https://www.bmwi.de/Redaktion/DE/Publikationen/Technologie/kuenstliche-intelligenz-impulse-zu-einem-megatrend.pdf?__blob=publicationFile&v=2 (Abfrage: 21.04.2021)

Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V./ Deutsches Forschungszentrum für Künstliche Intelligenz. (2017). *Entscheidungsunterstützung mit Künstlicher Intelligenz - Wirtschaftliche Bedeu-tung, gesellschaftliche Herausforderungen, menschliche Verantwortung.* Zugriff auf https://www.bitkom.org/sites/default/files/file/import/171012-KI-Gipfelpapier-online.pdf (Abfrage: 26.03.2021)

Business Application Research Center. (2021). *Using internal and external data for decision-making.* Zugriff auf https://bi-survey.com/data-sources (Abfrage: 28.07.2021)

Cao, S., Cong, L. W., Han, M., Hou, Q. & Yang, B. (2020). Blockchain architecture for auditing automation and trust building in public markets. *Computer*, *53* (7), 20–28.

Cao, S., Jiang, W., Yang, B. & Zhang, A. L. (2020). *How to talk when a machine is listening: Corporate disclosure in the age of AI* (Bericht). 1–53: National Bureau of Economic Research.

Caputo, A., Pizzi, S., Pellegrini, M. M. & Dabić, M. (2021). Digitalization and business models: Where are we going? A science map of the field. *Journal of Business Research*, *123*, 489–501.

Chang, Y.-H. & Ko, C.-B. (2017). A study on the design of low-code and no code platform for mobile application development. *International Journal of Advanced Smart Convergence*, *6* (4), 50–55.

Chen, C. P. & Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, *275*, 314–347.

Chen, H., Chiang, R. H. & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, 1165–1188.

Chen, Y.-C. (2012). A comparative study of e-government XBRL implementations: The potential of improving information transparency and efficiency. *Government Information Quarterly*, *29* (4), 553–563.

Cheng, F., Chiao, C., Wang, C., Fang, Z. & Yao, S. (2021). Does retail investor attention improve stock liquidity? A dynamic perspective. *Economic Modelling*, *94*, 170–183.

Cockcroft, S. & Russell, M. (2018). Big data opportunities for accounting and finance practice and research. *Australian Accounting Review*, *28* (3), 323–333.

Cordery, C. J., Fowler, C. J. & Mustafa, K. (2011). A solution looking for a problem: Factors associated with the non-adoption of XBRL. *Pacific Accounting Review*, *23* (1), 69–88.

Coyne, J. G. & McMickle, P. L. (2017). Can blockchains serve an accounting purpose? *Journal of Emerging Technologies in Accounting*, *14* (2), 101–111.

Dai, F., Shi, Y., Meng, N., Wei, L. & Ye, Z. (2017). From Bitcoin to cybersecurity: A comparative study of blockchain application and security

issues. In *2017 4th international conference on systems and informatics (icsai)* (S. 975–979).

Datatrade GmbH. (2021). *Data marketplaces: The ultimate guide 2021*. Zugriff auf https://about.datarade.ai/data-marketplaces (Abfrage: 28.07.2021)

Debreceny, R. & Gray, G. L. (1999). Financial reporting on the Internet and the external audit. *European Accounting Review*, *8* (2), 335–350.

Debreceny, R. & Gray, G. L. (2001). The production and use of semantically rich accounting reports on the Internet: XML and XBRL. *International Journal of Accounting Information Systems*, *2* (1), 47–74.

Debreceny, R., Gray, G. L. & Rahman, A. (2002). The determinants of Internet financial reporting. *Journal of Accounting and Public Policy*, *21* (4-5), 371–394.

Deller, D., Stubenrath, M. & Weber, C. (1999). A survey on the use of the Internet for investor relations in the USA, the UK and Germany. *European Accounting Review*, *8* (2), 351–364.

Demirkan, H. & Delen, D. (2013). Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. *Decision Support Systems*, *55* (1), 412–421.

Demuijnck, G. & Fasterling, B. (2016). The social license to operate. *Journal of Business Ethics*, *136* (4), 675–685.

De Villiers, C., Venter, E. R. & Hsiao, P.-C. K. (2017). Integrated reporting: Background, measurement issues, approaches and an agenda for future research. *Accounting & Finance*, *57* (4), 937–959.

Dhole, S., Lobo, G. J., Mishra, S. & Pal, A. M. (2015). Effects of the SEC's XBRL mandate on financial reporting comparability. *International Journal of Accounting Information Systems*, *19*, 29–44.

Dickhaut, J., Basu, S., McCabe, K. & Waymire, G. (2010). Neuroaccounting: Consilience between the biologically evolved brain and culturally evolved accounting principles. *Accounting Horizons*, *24* (2), 221–255.

Di Fabio, C., Roncagliolo, E., Avallone, F. & Ramassa, P. (2019). XBRL implementation in the european union: exploring preparers' points of view. In *Organizing for the Digital World* (S. 33–47). Springer.

Doolin, B. & Troshani, I. (2004). XBRL: A research note. *Qualitative Research in Accounting & Management*, *1* (2), 93–104.

Emani, C. K., Cullot, N. & Nicolle, C. (2015). Understandable big data: A survey. *Computer Science Review*, *17*, 70–81.

Enachi, M. & Andone, I. I. (2015). The progress of XBRL in Europe–projects, users and prospects. *Procedia Economics and Finance*, *20*, 185–192.

Endeavor Business Media. (2018). *Why and how: Quantum computing.* Zugriff auf https://www.cpapracticeadvisor.com/accounting-audit/article/12432726/why-and-how-quantum-computing (Abfrage: 05.08.2021)

Engelbrechtsmüller, C. & Kerschbaumer, H. (2014). Gestaltungsparameter und Reifegrade im Reporting. *Financial Reporting 2.0: Aktuelle Entwicklungen in der Finanzberichterstattung*, 12–18.

ESMA. (2017). *Final report on the RTS on the European Single Electronic Format.* Zugriff auf https://www.esma.europa.eu/sites/default/files/library/esma32-60-204_final_report_on_rts_on_esef.pdf (Abfrage: 16.04.2021)

Ettredge, M., Richardson, V. J. & Scholz, S. (2001). The presentation of financial information at corporate Web sites. *International Journal of Accounting Information Systems*, *2* (3), 149–168.

Faccia, A. & Mosteanu, N. R. (2019). Accounting and blockchain technology: from double-entry to triple-entry. *The Business & Management Review*, *10* (2), 108–116.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25* (2), 383–417.

FASB. (2000). *Business reporting research project: Electronic distribution of business information.* Zugriff auf https://www.fasb.org/brrp/brrp1.shtml (Abfrage: 15.04.2021)

Felden, C. (2011). Characteristics of XBRL adoption in Germany. *Journal of Management Control*, *22* (2), 161–186.

Firth, M. (1979). The impact of size, stock market listing, and auditors on voluntary disclosure in corporate annual reports. *Accounting and Business Research*, *9* (36), 273–280.

Fisher, D., DeLine, R., Czerwinski, M. & Drucker, S. (2012). Interactions with big data analytics. *Interactions*, *19* (3), 50–59.

Florescu, V. & Tudor, C. G. (2009). The optimization of the internal and external reporting in financial accounting: adopting XBRL international standard. *Annales Universitatis Apulensis: Series Oeconomica*, *11* (1), 126–139.

Fourati, Y. M., Merdassi, S. M. & Boujelbene, N. (2018). The effect of voluntary XBRL adoption on audit fees: Evidence from Belgium. *Journal of Accounting & Finance*.

Frank, R. & Brümmer, M. (2003). Der Kampf um den Investor - Zielgruppengerechte IR als Grundlage der Kapitalbeschaffung. *Going Public „Investor Relations 2003“*, 40–42.

Frankel, F. & Reid, R. (2008). Big data: Distilling meaning from data. *Nature*, *455* (7209), 30–30.

Fraunhofer IAIS. (2019). *Vertrauenswürdig Einsatz von Künstlicher Intelligenz.* Zugriff auf https://www.iais.fraunhofer.de/content/dam/iais/KINRW/Whitepaper_KI-Zertifizierung.pdf (Abfrage: 21.04.2021)

Fullana, O. & Ruiz, J. (2021). Accounting information systems in the blockchain era. *International Journal of Intellectual Property Management*, *11* (1), 63–80.

Fuller, S. H. & Markelevich, A. (2020). Should accountants care about blockchain? *Journal of Corporate Accounting & Finance*, *31* (2), 34–46.

Gandomi, A. & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35* (2), 137–144.

Gänßlen, S. & Losbichler, H. (2014). Big Data-Ein Segen für das Controlling. *White Paper des Internationalen Controller Vereins, Wörthsee*, *8* (4), 144-146.

Gärtner, B. & Hiebl, M. R. (2018). Issues with big data. In *The routledge companion to accounting information systems* (S. 6–172). Routledge.

Gazdar, K. (2014). Zwischen Struktur und Storytelling: Best Practice in Reporting. In *CSR und Reporting* (S. 173–180). Springer.

Gentsch, P. & Kulpa, A. (2016). Mit externen Big Data neue Möglichkeiten erschließen. *Controlling & Management Review*, *60* (1), 32–38.

Gerding, E. F. (2016). Disclosure 2.0: Can technology solve overload, complexity, and other information failures. *Tul. L. Rev.*, *90* (5), 1143–1180.

Gerpott, T. J. & Hoffmann, A. P. (2006). Berichterstattung über immaterielle Vermögenswerte auf Web Sites von Telekommunikationsnetzbetreibern. *KoR–Zeitschrift fü r internationale und kapitalmarktorientierte Rechnungslegung*, *6* (5), 369–78.

Godemann, J. & Herzig, C. (2012). nachhaltigkeitskommunikation. de. *uwf UmweltWirtschaftsForum*, *19* (3), 237–243.

Gomaa, M. I., Markelevich, A. & Shaw, L. (2011). Introducing XBRL through a financial statement analysis project. *Journal of Accounting Education*, *29* (2-3), 153–173.

Gostimir, D. (2015). XBRL standard for financial reporting in Croatia: current state and perspectives. *International Journal of the Society for Advancing Innovation and Research in Economy*, *6* (2), 31–40.

Greven, K. & Lahme, G. (2014). Freiwillige Transparenz führt zum Erfolg. In *Csr und social media* (S. 99–116). Springer.

Grüning, M. (2010). Unternehmenspublizität–Grundlagen und Optimierungsansätze. *Controlling*, *22* (10), 523–530.

Guilloux, V., Locke, J. & Lowe, A. (2013). Digital business reporting standards: Mapping the battle in France. *European Journal of Information Systems*, *22* (3), 257–277.

Gulin, D., Hladika, M. & Valenta, I. (2019). Digitalization and the Challenges for the Accounting Profession. *ENTRENOVA-ENTerprise REsearch InNOVAtion*, *5* (1), 428–437.

Hildner, A. & Danzmann, M. (2017). Blockchain-Anwendungen für die Unternehmensfinanzierung. *Corporate Finance*, *8* (11-12), 385–390.

Hoffman, C. & Rodríguez, M. M. (2013). Digitizing financial reports–Issues and insights: A viewpoint. *The International Journal of Digital Accounting Research*, *13* (1), 73–98.

Hu, D., Jones, C. M., Zhang, V. & Zhang, X. (2021). The rise of reddit: How social media affects retail investors and short-sellers' roles in price discovery. *Social Science Research Network 3807655*, 1–47.

Hu, H., Wen, Y., Chua, T.-S. & Li, X. (2014). Toward scalable systems for big data analytics: A technology tutorial. *IEEE access*, *2*, 652–687.

Huang, M.-H. & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research*, *21* (2), 155–172.

Hucklesby, M. & Macdonald, J. (2002). XBRL: Are you ready to reap the benefits. *Chartered Accountants' Journal of New Zealand*, *81* (4), 54–55.

Hwang, S., No, W. G. & Kim, J. (2021). XBRL mandate and timeliness of financial reporting: The effect of internal control problems. *Journal of Accounting, Auditing & Finance*, *36* (3), 667–692.

IASC. (1999). *Business reporting on the internet.* Zugriff auf https://www.icjce.es/images/pdfs/TECNICA/CO2-IASB/C208-IASB-Estudiosyvarios/BusinessReportingInternet-IASC.pdf (Abfrage: 31.07.2021)

IIRC. (2013). *Consultation draft of the international IR framework.* Zugriff auf https://integratedreporting.org/wp-content/uploads/2013/03/Consultation-Draft-of-the-InternationalIRFramework.pdf (Abfrage: 28.04.2021)

Inghirami, I. E. (2018). *Accounting information system in the time of Blockchain.* Zugriff auf https://www.academia.edu/38813295/Accounting_Information_Systems_in_the_Time_of_Blockchain (Abfrage: 26.04.2021)

International Data Corporation. (2021). *Future of industry ecosystems: Shared data and insights.* Zugriff auf https://blogs.idc.com/2021/01/06/future-of-industry-ecosystems-shared-data-and-insights (Abfrage: 28.07.2021)

Janssen, M., van der Voort, H. & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*, *70*, 338–345.

Javaid, M., Haleem, A., Vaishya, R., Bahl, S., Suman, R. & Vaish, A. (2020). Industry 4.0 technologies and their applications in fighting COVID-19 pandemic. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, *14* (4), 419–422.

Jędrzejka, D. (2019). Robotic process automation and its impact on accounting. *Zeszyty Teoretyczne Rachunkowości* (105 (161)), 137–166.

Jensen, M. C. & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3* (4), 305–360.

Jones, M. J. & Xiao, J. Z. (2003). Internet reporting: Current trends and trends by 2010. In *Accounting forum* (Bd. 27, S. 132–165).

Kaisler, S., Armour, F., Espinosa, J. A. & Money, W. (2013). Big data: Issues and challenges moving forward. In *2013 46th hawaii international conference on system sciences* (S. 995–1004).

Kambatla, K., Kollias, G., Kumar, V. & Grama, A. (2014). Trends in big data analytics. *Journal of Parallel and Distributed Computing*, *74* (7), 2561–2573.

Karajovic, M., Kim, H. M. & Laskowski, M. (2019). Thinking outside the block: Projected phases of blockchain integration in the accounting industry. *Australian Accounting Review*, *29* (2), 319–330.

Kataria, M. & Mittal, P. (2014). Big Data: A review. *International Journal of Computer Science and Mobile Computing*, *3* (7), 106–110.

Kesselmeyer, B. (2009). Potenzial von XBRL für das Beteiligungscontrolling, für Venture Capital Gesellschaften sowie für BSC-Projekte. *Controlling & Management*, *53* (3), 69–76.

Kokina, J., Mancha, R. & Pachamanova, D. (2017). Blockchain: Emergent industry adoption and implications for accounting. *Journal of Emerging Technologies in Accounting*, *14* (2), 91–100.

Kovarova-Simecek, F.-P. M. M., Pellegrini, F.-P. D. T. & Aubram, T. (2018). XBRL im digitalen Reporting-Entwicklungsstand in Europa und in Österreich. *CARF Luzern 2018*, 160–176.

Kowalczyk, M. & Buxmann, P. (2014). Big data and information processing in organizational decision processes. *Business & Information Systems Engineering*, *6* (5), 267–278.

KPMG. (2015). *Information Design in den Geschäftsberichten der Dax 30-Unternehmen.* Zugriff auf http://www.cira.at/fileadmin/downloads/2015/information_design_in_den_geschaeftsberichte_der_dax_30_unternehmen.pdf (Abfrage: 04.05.2021)

Krahel, J. P. & Titera, W. R. (2015). Consequences of big data and formalization on accounting and auditing standards. *Accounting Horizons*, *29* (2), 409–422.

Kretschmer, H. (2014). XBRL: Der lange Weg zur automatisierten Berichterstattung. *Financial Reporting 2.0: Aktuelle Entwicklungen in der Finanzberichterstattung*, 169–186.

Kreutzer, R. T., Neugebauer, T. & Pattloch, A. (2017). *Digital business leadership: Di-gitale Transformation - Geschäftsmodell-Innovation - Agile Organisation - Change-Management*. Wiesbaden, Springer Gabler.

Küting, K., Dawo, S. & Heiden, M. (2002). Internet-Investor Relations – The Impact of the Internet on Corporate Reporting. In C. Scholz & J. Zentes (Hrsg.), *Strategic management: A european approach* (S. 47–67). Wiesbaden: Gabler Verlag. Zugriff auf https://doi.org/10.1007/978-3-322-84457-6_4 doi: 10.1007/978-3-322-84457-6_4

Lampathaki, F., Mouzakitis, S., Gionis, G., Charalabidis, Y. & Askounis, D. (2009). Business to business interoperability: A current review of XML data integration standards. *Computer Standards & Interfaces*, *31* (6), 1045–1055.

Lang, M. & Lundholm, R. (1993). Cross-sectional determinants of analyst ratings of corporate disclosures. *Journal of accounting research*, *31* (2), 246–271.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, *32* (2), 311–328.

Lasso-Rodriguez, G. & Winkler, K. (2020). Hyperautomation to fulfil jobs rather than executing tasks: the BPM manager robot vs human case. *Romanian Journal of Information Technology and Automatic Control*, *30* (3), 7–22.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S. & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, *52* (2), 21–32.

LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning. *Nature*, *521* (7553), 436–444.

Leibfried, P. (2010). Finanzberichterstattung mittels XBRL: Mögliche Erleichterungen durch die Vermeidung von Datenbrüchen. *KPMG's Audit Committee News* (29), 10–13.

Leitner-Hanetseder, S., Lehner, O. M., Eisl, C. & Forstenlechner, C. (2021). A profession in transition: Actors, tasks and roles in AI-based accounting. *Journal of Applied Accounting Research*, 1–18.

Li, X., Zhu, H. & Zuo, L. (2021). Reporting technologies and textual readability: Evidence from the XBRL mandate. *Information Systems Research*, 1–44.

Lim, N., Perrin, B. et al. (2014). Standard business reporting in Australia: Past, present, and future. *Australasian Journal of Information Systems*, *18* (3), 29–51.

Lin, P. P. (2014). What CPAs need to know about Big Data. *The CPA Journal*, *84* (11), 50–55.

Liu, C., Luo, X. R. & Wang, F. L. (2017). An empirical investigation on the impact of XBRL adoption on information asymmetry: Evidence from Europe. *Decision Support Systems*, *93*, 42–50.

Liu, C., Wang, T. & Yao, L. J. (2014). XBRL's impact on analyst forecast behavior: An empirical study. *Journal of Accounting and Public Policy*, *33* (1), 69–82.

Locke, J. & Lowe, A. (2007). XBRL: An (open) source of enlightenment or disillusion? *European Accounting Review*, *16* (3), 585–623.

Locke, J., Rowbottom, N. & Troshani, I. (2018). Sites of translation in digital reporting. *Accounting, Auditing & Accountability Journal*, *31* (7), 2006–2030.

Losbichler, H. & Otto, M. (2013). Finanzmarktkommunikation: Die integrierte Unternehmen-Kapitalmarkt-Perspektive. *Financial Reporting*, 21–44.

Lymer, A. & Debreceny, R. (2003). The auditor and corporate reporting on the internet: Challenges and institutional responses. *International Journal of Auditing*, *7* (2), 103–120.

Madakam, S., Ramaswamy, R. & Tripathi, S. (2015). Internet of Things (IoT): A literature review. *Journal of Computer and Communications*, *3* (05), 164–173.

Markelevich, A. & Riley, T. J. (2013). Embracing and integrating XBRL. *The CPA Journal*, *83* (6), 70.

Meckel, M., Fieseler, C. & Hoffmann, C. P. (2010). *Zukunftslabor Geschäftsbericht.* Zugriff auf https://www.alexandria.unisg.ch/151234/1/Studie_Zukunftslabor.pdf (Abfrage: 22.04.2021)

Meesters, M., Wostyn, P., Van Leeuwen, J., Behagel, J. H. & Turnhout, E. (2021). The Social Licence to Operate and the legitimacy of resource extraction. *Current Opinion in Environmental Sustainability*, *49*, 7–11.

Menacher, J., Schütt, H. H. & Sellhorn, T. (2017). Informationsverarbeitung an Finanzmärkten im Umbruch–zwischen Informationsflut und automatisierter Datenanalyse. *KoR: internationale und kapitalmarktorientierte Rechnungslegung; IFRS*, 217–224.

Min, H. (2019). Blockchain technology for enhancing supply chain resili-

ence. *Business Horizons*, *62* (1), 35–45.
Mittelbach-Hörmanseder, S., Guggenberger, V. & Barrantes, E. (2020). „Digital Shift“ in der Unternehmensberichterstattung. *Zeitschrift für kapitalmarktorientierte Rechnungslegung*, 47–48.
Mittelbach-Hörmanseder, S., Hummel, K., Barrantes, E. & Guggenberger, V. (2021). Digitalization of financial reporting: The preparers' perspective. *SSRN Electronic Journal*, *Nr. 3786100*, 1–46.
Moniruzzaman, A. & Hossain, S. A. (2013). Nosql database: New era of databases for big data analytics-classification, characteristics and comparison. *arXiv preprint arXiv:1307.0191*, *6* (4), 1–14.
Müller, O., Junglas, I., Brocke, J. v. & Debortoli, S. (2016). Utilizing big data analytics for information systems research: challenges, promises and guidelines. *European Journal of Information Systems*, *25* (4), 289–302.
Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R. & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data*, *2* (1), 1–21.
Nakamoto, S. (2008). *Bitcoin. A peer-to-peer electronic cash system, White Paper*.
NeidhartSchön AG. (2018). *Online-Reporting-Perspektiven*. Zugriff auf https://www.financialcommunication.org/fileadmin/webcontent_crifc/Studien/Online_Reports.pdf (Abfrage: 15.04.2021)
Nutz, A. & Strauß, M. (2002). eXtensible Business Reporting Language (XBRL). *Wirtschaftsinformatik*, *44* (5), 447–457.
Oyelere, P., Laswad, F. & Fisher, R. (2003). Determinants of internet financial reporting by New Zealand companies. *Journal of International Financial Management & Accounting*, *14* (1), 26–63.
Petkov, R. (2020). Artificial intelligence (AI) and the accounting function–A revisit and a new perspective for developing framework. *Journal of Emerging Technologies in Accounting*, *17* (1), 99–105.
Pfister, J. & Venetz, S. (2009). Die Anwendung von XBRL in IR. In *Praxishandbuch investor relations* (S. 433–442). Springer.
Provost, F. & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, *1* (1), 51–59.
Pugna, I. B. & Duţescu, A. (2020). Blockchain–the accounting perspective. In *Proceedings of the International Conference on Business Excellence* (Bd. 14, S. 214–224).
PWC. (2019). *Digitalisierung im Finanz- und Rechnungswesen 2019 - Und was sie für die Abschlussprüfung bedeutet.* Zugriff auf https://www.pwc.de/de/digitale-abschlusspruefung/pwc-befragung-digitalisierung-im-finanz-und-rechnungswesen-2019.pdf (Abfrage: 26.04.2021)
Raghavan, K. (2021). Impact of Pandemic and Digital Transformation on Global Accounting Profession. *Journal of Global Awareness*, *2* (1), 1–9.
Ramin, K. P. & Kesselmeyer, B. (2007). XBRL als internetbasierter Standard für die Finanzberichterstattung. *KoR Zeitschrift für internationale und kapitalmarktorientierte Rechnungslegung*, *7* (10), 560–571.
Rawashdeh, B. & Rawashdeh, A. (2021). Factors influencing the usage of XBRL tools. *Management Science Letters*, *11* (4), 1345–1356.
Richins, G., Stapleton, A., Stratopoulos, T. C. & Wong, C. (2017). Big data analytics: opportunity or threat for the accounting profession? *Journal of Information Systems*, *31* (3), 63–79.
Roohani, S., Xianming, Z., Capozzoli, E. & Lamberton, B. (2010). Analysis of XBRL literature: A decade of progress and puzzle. *International Journal of Digital Accounting Research*, *10*, 131–147.
Ruan, L., Liu, H. & Tsai, S. (2021). XBRL adoption and capital market information efficiency. *Journal of Global Information Management (JGIM)*, *29* (6), 1–18.
Sassi, W., Othman, H. B. & Hussainey, K. (2021). The impact of mandatory adoption of XBRL on firm's stock liquidity: A cross-country study. *Journal of Financial Reporting and Accounting*, *19* (2), 299–324.
Schmitz, J. & Leoni, G. (2019). Accounting and auditing at the time of blockchain technology: A research agenda. *Australian Accounting Review*, *29* (2), 331–342.
Schruff, L. & Kayser, M. (2002). Das Internet als Medium zur Publizität von Jahres-und Konzernabschlüssen. In *Electronic business* (S. 347–359). Springer.
Schwegler, P. (2018). *Was Unternehmen nach der DSGVO erwartet.* Zugriff auf https://www.wuv.de/digital/was_unternehmen_nach_der_dsgvo_erwartet (Abfrage: 22.10.2018)
Simons, D. & Weißenberger, B. E. (2009). 'Different costs for different purposes 'vs.' One version of the truth? Zur Konvergenz von externer und interner Rechnungslegung im deutschsprachigen Raum, Teil 1 und 2. *Wirtschaftswissenschaftliches Studium: WiSt*, *38* (8-9), 390–395.
Singerová, J. (2015). XBRL: different approach of utilization. *Procedia Economics and Finance*, *25*, 134–140.
Singh, K. & Best, P. (2016). Interactive visual analysis of anomalous accounts payable transactions in SAP enterprise systems. *Managerial Auditing Journal*, *31* (1), 35–63.
Sivarajah, U., Kamal, M. M., Irani, Z. & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, *70*, 263–286.
Smith, S. S. (2018). Digitization and financial reporting–how technology innovation may drive the shift toward continuous accounting. *Accounting and Finance Research*, *7* (3), 240–250.
Speich, I. (2014). Intangibles: Forderung institutioneller Investoren nach aussagekräftiger extrafinanzieller Berichterstattung. In *Csr und finance* (S. 209–218). Springer.
Talesra, K. & Nagaraja, G. (2021). Low-Code Platform for Application Development. *International Journal of Applied Engineering Research*, *16* (5), 346–351.
Tan, B. S. & Low, K. Y. (2019). Blockchain as the database engine in the accounting system. *Australian Accounting Review*, *29* (2), 312–318.
Tank, A. K. & Farrell, A. M. (2021). Is neuroaccounting taking a place on the stage? A review of the influence of Neuroscience on accounting research. *European Accounting Review*, 1–35.
Tien, J. M. (2013). Big data: Unleashing information. *Journal of Systems Science and Systems Engineering*, *22* (2), 127–151.
Troshani, I., Locke, J. & Rowbottom, N. (2019). Transformation of accounting through digital standardisation: Tracing the construction of the IFRS taxonomy. *Accounting, Auditing & Accountability Journal*, *32* (1), 133–162.
Troshani, I. & Rowbottom, N. (2021). Digital corporate reporting: research developments and implications. *Australian Accounting Review*, 1–20.
Umar, Z., Gubareva, M., Yousaf, I. & Ali, S. (2021). A tale of company fundamentals vs sentiment driven pricing: The case of GameStop. *Journal of Behavioral and Experimental Finance*, *30* (100501), 1–7.
Đurković, E., Šaponja, L. D. & Đurković, O. (2017). Possible avenue of implementation of the XBRL standard for maximum possible efficiency of financial reporting in the Republic of Serbia. *Journal of Strategic Management and Decision Support Systems*, *22* (1), 12-21.
Đurović, M., Rep, A. & Dečman, N. (2021). TOWARDS FULL DIGITIZATION OF THE FINANCIAL REPORTING–OVERVIEW OF THE FINANCIAL REPORTING LANGUAGES. In *Proceedings of feb zagreb international odyssey conference on economics and business* (Bd. 3, S. 55–69).
Valentinetti, D. & Muñoz, F. F. (2021). Internet of things: Emerging impacts on digital reporting. *Journal of Business Research*, *131*, 549–562.
Vasarhelyi, M. A., Chan, D. Y. & Krahel, J. P. (2012). Consequences of XBRL standardization on financial statement data. *Journal of Information Systems*, *26* (1), 155–167.
Vasarhelyi, M. A., Kogan, A. & Tuttle, B. M. (2015). Big data in accounting: An overview. *Accounting Horizons*, *29* (2), 381–396.
Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics*, *5*, 179–194.
Voelzkow, A. (2019). *Unternehmerische Selbstdarstellung im Internet*. Springer.
Voshmgir, S. (2016). *Blockchains, Smart Contracts und das Dezentrale Web. Technologiestiftung Berlin.*
Wagenhofer, A. & Ewert, R. (2015). *Externe Unternehmensrechnung* (3. Aufl.). Springer.
Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J.-f., Dubey, R. & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356–365.
Wang, H., Xu, Z., Fujita, H. & Liu, S. (2016). Towards felicitous decision making: An overview on challenges and trends of Big Data. *Information Sciences*, *367*, 747–765.
Wang, S. & Wang, H. (2021). A Teaching Module of No-Code Business App Development. *Journal of Information Systems Education*, *32* (1), 1–8.
Wang, W. & Siau, K. (2019). Artificial intelligence, machine learning, automation, robotics, future of work and future of humanity: A review

and research agenda. *Journal of Database Management (JDM)*, *30* (1), 61–79.

Weber, J., Bramsemann, U., Heineke, C. & Hirsch, B. (2017). Kapitalmarktkommunikation. In *Wertorientierte unternehmenssteuerung* (S. 233–268). Springer.

Weber, J. & Weißenberger, B. E. (2015). *Einführung in das Rechnungswesen* (9. Aufl.). Stuttgart, Schäffer-Poeschel.

Weißenberger, B. E. (2014). Integrated Reporting: Fragen (und Antworten) aus der Diskussion um die integrierte Rechnungslegung. *Controlling*, *26* (8-9), 440–446.

Weißenberger, B. E. & Bauch, K. A. (2017). Chancen und Risiken der digitalen Transformation für die Rechnungslegung. In *Langfristige perspektiven und nachhaltigkeit in der rechnungslegung* (S. 203–219). Springer.

Wen, Y. (2021). Research on XBRL Financial Reporting Model under Big Data. In *Journal of Physics: Conference Series* (Bd. 1756, S. 1–5).

Williams, S. P., Scifleet, P. A. & Hardy, C. A. (2006). Online business reporting: An information management perspective. *International Journal of Information Management*, *26* (2), 91–101.

Wilson, R. & Sangster, A. (1992). The automation of accounting practice. *Journal of Information Technology*, *7* (2), 65–75.

Winter, S. J., Saunders, C. & Hart, P. (2003). Electronic window dressing: Impression management with websites. *European Journal of Information Systems*, *12* (4), 309–322.

XBRL DE. (2019). *Digitaler Finanzbericht*. Zugriff auf https://de.xbrl.org/news/xbrl-woche-in-frankfurt-2016-ab-jetzt-konnen-sie-sich-registrieren (Abfrage: 20.03.2021)

XBRL International Inc. (2021a). *Global Ledger*. Zugriff auf https://specifications.xbrl.org/spec-group-index-xbrl-gl.html (Abfrage: 31.03.2021)

XBRL International Inc. (2021b). *Jurisdictions*. Zugriff auf https://www.xbrl.org/the-consortium/about/jurisdictions (Abfrage: 14.03.2021)

XBRL International Inc. (2021c). *Using the ESEF rules for anchoring extensions*. Zugriff auf https://www.xbrl.org/guidance/esef-rules-anchoring-extensions/#3-what-is-anchoring (Abfrage: 09.08.2021)

Yang, S., Liu, F.-C. & Zhu, X. (2018). The impact of XBRL on financial statement structural comparability. In *Network, smart and open* (S. 193–206). Springer.

Zhu, H. & Wu, H. (2014). Assessing the quality of large-scale data standards: A case of XBRL GAAP Taxonomy. *Decision Support Systems*, *59*, 351–360.

Zülch, H. & Weuster, C. W. (2018). Change before you have to - Über die Zukunft der Finanzberichterstattung in einer digitalen Welt. *Zeitschrift für internationale und kapitalmarktorientierte Rechnungslegung*, *2*, 102–103.

Junior Management Science 7(4) (2022) 1032-1097

Junior Management Science

journal homepage: www.jums.academy

The Organization of Future Production Work – Requirements and Technical Solution Approaches

Die Organisation der Produktionsarbeit der Zukunft – Anforderungen und technische Lösungsansätze

Jan Felix Csavajda

Universität Stuttgart

Abstract

With a view to the industrial production of the future and Industry 4.0, the focus mostly lies on technology, while organization and the role of humans are less considered. The aim of this paper is to determine organizational requirements and performance indicators for future production. A quantitative empirical study is used to evaluate the relevance from a practical perspective. In addition, a collection of technological solutions illustrates the exemplary practical implementation of the organizational requirements. As a result, an ideal-typical organization of future production is presented. A central finding is that so far little attention is drawn to the areas of organization and humans in the context of Industry 4.0. Modern work and leadership concepts, consequent employee qualification, Lean Management 4.0, improved coordination, connectivity and transparency as well as the use of performance indicators are essential. For the successful implementation of technical Industry 4.0 solutions, the primary establishment of a basic organizational framework is mandatory. This work clarifies which concepts should be in the foreground in the future, also in order to secure competitiveness.

Zusammenfassung

Mit Blick auf die industrielle Produktion der Zukunft und Industrie 4.0 steht vermehrt die Technologie im Fokus, Organisation und die Rolle des Menschen hingegen weniger. Das Ziel der vorliegenden Arbeit ist es, organisationale Anforderungen und Kennzahlen der zukünftigen Produktion zu bestimmen. Anhand einer quantitativen Empirie erfolgt eine Relevanzbewertung aus Praxissicht. Ergänzend zeigt ein Technologielösungsatlas die beispielhafte praktische Umsetzung der organisationalen Anforderungen auf. Als Ergebnis wird eine idealtypische arbeitsorganisatorische Ablauforganisation der zukünftigen Produktion vorgestellt. Eine zentrale Erkenntnis ist, dass die Teilbereiche Organisation und Mensch im Zusammenhang mit Industrie 4.0 bisher unzureichend betrachtet werden. Moderne Arbeits- und Führungskonzepte, eine konsequente Mitarbeiterqualifikation, Lean Management 4.0, eine verbesserte Abstimmung, Vernetzung und Transparenz sowie die Verwendung von Kennzahlen sind essenziell. Für die erfolgreiche Implementierung von technischen Industrie 4.0 Lösungen ist der primäre Aufbau eines organisatorischen Grundrahmens zwingend erforderlich. Diese Arbeit verdeutlicht, welche Konzepte hier zukünftig im Fokus stehen müssen, auch um die unternehmerische Wettbewerbsfähigkeit sicherzustellen.

Keywords: Industrie 4.0; Produktion; Arbeitsorganisation; Anforderungen; Kennzahlen.

1. Einleitung

1.1. Ausgangssituation, Problemstellung und Handlungsbedarf

Mit den drei industriellen Revolutionen der Mechanisierung, Elektrifizierung und der Automatisierung gab es in der Entwicklungsgeschichte der Produktion bereits einige massive Umwälzungen. Die nächste Entwicklungsstufe der Produktion in Form einer zunehmenden Digitalisierung wird als „Industrie 4.0“ bezeichnet.[1] Diese vierte industrielle

[1] vgl. Lucks (2017a), S. 5.

DOI: https://doi.org/10.5282/jums/v7i4pp1032-1097

Fortschrittswelle lässt Unternehmen auf der gesamten Welt über die Verwendung von neuen, digitalen Technologien nachdenken,[2] um weiter zu wachsen und viele Vorgänge zu vereinfachen.[3] Das Thema der Industrie 4.0 (I4.0) als Einzug der Informationstechnik in die Produktion wird schon seit längerer Zeit ausführlich diskutiert.[4] Das behandelte Fachgebiet der vorliegenden Arbeit ist dieser sich aktuell stark verändernde Produktionsbereich. Die beschriebene Veränderung wird durch die hohe Anzahl an Veranstaltungen und Veröffentlichungen zum Thema I4.0 verdeutlicht.[5] Aufgrund des tiefgreifenden Verbesserungspotentials und der Wichtigkeit der Entwicklung hin zu I4.0, ergibt sich ein eindeutiger Bedarf zu einer ausführlichen Diskussion dieses Themas.

Die grundsätzliche Relevanz der Thematik ist als sehr hoch einzustufen. Bereits im Jahr 2015 wies Jean Claude Juncker als damaliger EU Kommissionsvorsitzender gezielt darauf hin, dass die Digitalisierung eine der zentralsten und herausforderndsten Zukunftsfragen für die europäische Politik und Unternehmen darstellt.[6] Zwar hat die deutsche Bundesregierung Industrie 4.0 bereits früh als Zukunftsprojekt auf ihre digitale Agenda gesetzt,[7] doch im Vergleich zu anderen führenden Wirtschaftsnationen wurde der digitale Wandel im Jahr 2016 noch nicht in einem ausreichenden Umfang angesteuert.[8] Damals waren deutsche, produzierende Unternehmen gegenüber der Industrie 4.0 noch zurückhaltender und sahen insgesamt eindeutig mehr Risiken als Chancen, weshalb weitreichende Weckrufe aus der Forschung zur Verdeutlichung der thematischen Dringlichkeit mehr als notwendig waren.[9]

Ein übergreifendes Ziel ist es, durch neue Fortschritte und Erkenntnisse im I4.0 Bereich eine flächendeckende und deutschlandweite Implementierung in produzierenden Unternehmen unabdingbar zu machen.[10] Bis in das Jahr 2018 war eine weitreichende Skepsis existent.[11] In einer Studie des Bundesverbands Informationswirtschaft Telekommunikation und neue Medien e.V. (Bitkom) zeigte sich bereits ein Jahr später eine höhere Akzeptanz: Etwas mehr als jedes zweite Unternehmen setzte auf die Nutzung von Industrie 4.0 Technologien und 74 Prozent der befragten Unternehmen planten, I4.0 Anwendungen einzusetzen. Obwohl sich in den Jahren ab 2015 viel bewegt hat, gaben im Rahmen der Studie fast ein Viertel aller Unternehmen an, den Einsatz von I4.0 Technologien vorerst nicht fördern zu wollen.[12] Dem entgegen steht die Haltung gegenüber der Digitalisierung im Allgemeinen. Diese wird im Großen und Ganzen als beträchtliche Chance angesehen.[13] Aus diesen Gründen ist es in der aktuellen Situation weiterhin notwendig, mithilfe des richtigen Inputs Anpassungen anzustoßen, wie Industrie 4.0 in Deutschland gedacht wird.

Aufgrund dieser Ausgangslage ergibt sich im Konkreten der Handlungsbedarf dieser Arbeit. Die zu bearbeitende Forschungslücke zeigt sich dadurch, dass sich die Folgen der Digitalisierung besonders in der Fertigung von Gütern niederschlagen.[14] Gleichzeitig erfreuen sich innovative Arbeitskonzepte und flexible Arbeitsformen in der Wissensarbeit und im Büroalltag einer weiten Verbreitung. Diese Entwicklungen sind an der Produktionsarbeit bisher jedoch zu großen Teilen vorbeigegangen.[15] Solche Konzepte, und deren Bedeutung im Zusammenspiel mit I4.0 Technologien, werden im Rahmen dieser Arbeit vor einem produktionstechnischen Hintergrund betrachtet. Unter dem Titel „Die Organisation der Produktionsarbeit der Zukunft - Anforderungen und technische Lösungsansätze“ untersucht die vorliegende Arbeit diese Kombination von Industrie 4.0 Aspekten mit arbeitsorganisatorischen und menschenorientierten Ansätzen. Die Durchführung der Forschung erfolgte in Zusammenarbeit mit dem Fraunhofer-Institut für Arbeitswirtschaft und Organisation (IAO) in Stuttgart, wobei die persönliche Motivation des Autors zur Bearbeitung dieses Themas unter anderem auf längeren praktischen Erfahrungen in der Produktions- und Fabrikplanung bei der Daimler AG beruht.

1.2. Zielsetzung, Methodik und Aufbau der Arbeit

Alles in Allem ist das Thema Industrie 4.0 sehr umfangreich und gliedert sich in eine Vielzahl unterschiedlicher Teilbereiche. Der Blickwinkel dieser Arbeit richtet sich weniger auf detaillierte, technologische Inhalte von Industrie 4.0, sondern vielmehr auf die Systembereiche Organisation und Mensch. Diese spielen bei der Betrachtung von Industrie 4.0 bisher fälschlicherweise eine eher untergeordnete Rolle.[16] Weiterhin liegt der Fokus vordergründig in der direkten Produktionsebene sowie deren Ressourcen und Prozessen und nicht in einer übergeordneten Unternehmensstrategiedimension. Zur Abgrenzung zeigt die nachfolgende Abbildung wesentliche Teilbereiche einer digitalen I4.0 Arbeitswelt übersichtlich auf (Abb. 1).

Zusammenfassend werden vorrangig die Arbeitsorganisation, Arbeitsmethoden und der Mensch mit seinen Fähigkeiten behandelt. Der Nutzen und Mehrwert besteht darin, den aktuellen wissenschaftlichen Forschungsstand in diesem Industrie 4.0 Teilgebiet zusammenzutragen und durch eigene Forschungserkenntnisse zu erweitern. Das Hauptziel hierbei ist die Beantwortung mehrerer ausgewählter Forschungsfragen. Nach der Betrachtung allgemeiner, relevanter Grundlagen, fokussiert sich der erste Hauptforschungsteil auf die Frage, mit welchen Kennzahlen die zukünftige Produktion gesteuert wird. Darauf folgt eine Analyse der Anforderungen an die Produktionsorganisation. Im Detail wird evaluiert,

[2] vgl. Lorenz, Küpper, Rüßmann, Heidemann und Bause (2016), S. 3.
[3] vgl. Schnell (2018a), S. 43.
[4] vgl. Manzei (2016), S. 10.
[5] vgl. Schnell (2018a), S. 43.
[6] vgl. Müller und Jentsch (2015), S. 27.
[7] vgl. Manzei (2016), S. 10.
[8] vgl. Kollmann und Schmidt (2016), S. 137.
[9] vgl. Bauernhansl und Dombrowski (2016), S. 5.
[10] vgl. Schlund und Pokorni (2017), S. 163.
[11] vgl. Stief (2018), S. 245.
[12] vgl. A. Berg (2019), S. 2.
[13] vgl. ebenda, S. 13.
[14] vgl. Cole (2017), S. 160.
[15] vgl. Spath et al. (2013), S. 12.
[16] vgl. Kreggenfeld, Prinz und Kuhlenkötter (2016), S. 31f.

Abbildung 1: Modell zur Gestaltung digitaler Arbeitswelten nach Altemeier, Bansmann, Dietrich, Dumitrescu und Nettelstroth (2017).
Quelle: Altemeier et al. (2017), S. 10.

welche Anforderungen an die Produktion der Zukunft sich aus den prognostizierten Herausforderungen der Arbeitswelt ergeben und inwiefern schon länger bestehende oder recht neuartige Organisationskonzepte das Denken der Produktion verändern.

Zur methodischen Bestimmung der Wichtigkeit dieser Kennzahlen und Anforderungen wird eine quantitative Orientierungsumfrage durchgeführt. Durch ein Meinungsbild aus der Praxis erfolgt die Generierung einer Rangfolge der Kennzahlen und theoretischen Konzepte der Produktionsorganisation. Im Zuge dessen stellt sich die Frage, ob die existierenden, arbeitsorganisatorischen Konzepte für die Produktion schon in der breiten Unternehmenspraxis angekommen sind. Der dritte Forschungsaspekt zielt darauf ab, punktuell aufzuzeigen wie die organisationalen Anforderungen in der Praxis umgesetzt werden können und mit welchen innovativen Technologien und I4.0 Anwendungsfällen Shopfloor-Organisationsmethoden unterstützt, optimiert oder gar erst möglich gemacht werden. Hierzu werden technische I4.0 Lösungen in einem Technologielösungsatlas für innovative Produktionsorganisationsformen gesammelt. Als finales Ergebnis wird durch eine auf den jeweils wichtigsten Kennzahlen, Anforderungen und „Best Practices“ basierende, idealtypische Organisationsform ein Lösungsansatz für das Aussehen der zukünftigen Produktion präsentiert. Eine Hypothese ist, dass die Organisation und der Fokus auf den Menschen immer wichtiger wird und Unternehmen dies auch sehen, indem sie diesen Faktoren nach und nach die gleiche Wichtigkeit einräumen wie der reinen I4.0 Technologie. Dies gilt es zu überprüfen.

Des Weiteren wird analysiert, ob das heutige Produktionspersonal dazu geeignet ist mit den zukünftigen Anforderungen und technologischen Neuerungen zurecht zu kommen. Was muss getan werden? Welches Leistungsprofil weist der zukünftige Fabrikarbeiter auf und welche Arbeitsumgebungen und -bedingungen gilt es zu schaffen? Innerhalb des dargelegten, methodischen Vorgehens der Aufgabenbearbeitung wurde zunächst eine systematische Literaturrecherche betrieben. Die quantitative Umfrage dient zusätzlich der Sammlung von eigens erhobenen Daten und ergänzt das Forschungsdesign. Die gewonnenen Daten werden mit Tabellen und Graphen dargestellt und untersucht.

Grundsätzlich erfolgte die Literaturrecherche überwiegend über diverse Literaturdatenbanken wie EconBiz oder Wiso, Google Scholar, Springer Link und durch andere Nachschlagewerke aus den (elektronischen) Bibliotheken der Universität Stuttgart und der Stuttgarter Fraunhofer-Institute. Neben Monografien und Sammelwerken sind Positionspapiere und Forschungsberichte von I4.0 Verbänden und von Bundesministerien häufig verwendete Dokumententypen. Zudem dienen unter anderem Fachartikel aus dem Magazin Industrie 4.0 Management (I40M) und der Zeitschrift für wirtschaftlichen Fabrikbetrieb (ZWF) als Basisliteratur. Zur Recherche wurden hauptsächlich die folgenden Schlüsselwörter im Verbund mit den Zusätzen *4.0 oder *Zukunft genutzt: „Organisation“, „Produktion“, „Arbeit“, „Führung“, „Mitarbeiter“ und „Kompetenz“. Auch die Begriffe „Anforderungen“, „Kennzahlen“, „New Work“ und „Anwendungen“ fanden bei der Quellensuche im Zusammenhang mit Produktion oder Industrie 4.0 Gebrauch.

Der detaillierte Aufbau und Ablauf der Arbeit ist der nachfolgenden Darstellung zu entnehmen. Die Abbildung veranschaulicht die Inhalte der einzelnen Kapitel und liefert einen Überblick über das methodische Vorgehen (Abb. 2).

Zunächst sind die detaillierte Betrachtung und Diskussion einiger grundlegender Thematiken in Kapitel (2.) notwendig.

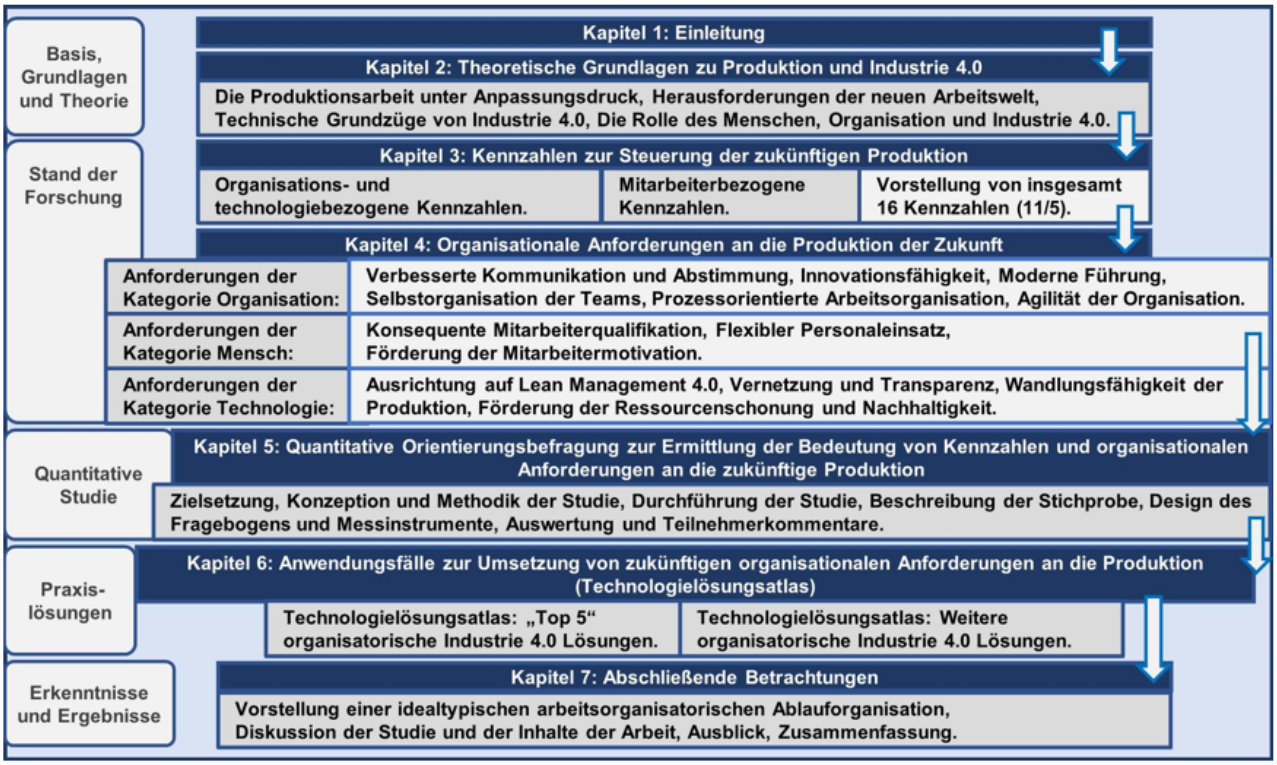


Abbildung 2: Eigene Darstellung - Aufbau und Gang der Arbeit.

2. Theoretische Grundlagen zu Produktion und Industrie 4.0

Als Basis für die detaillierte Betrachtung von produktionsseitigen, organisationalen Anforderungen, zu steuernden Kennzahlen und Industrie 4.0 Lösungsansätzen in den Hauptbestandteilen dieser Arbeit, liefert dieses erste Kapitel theoretische Einblicke und zeigt einleitend den kontextuellen Zusammenhang der Themen Produktion und Industrie 4.0 mit der Organisation und dem Menschen auf.

Anfänglich ist die Definition zentraler Begriffe erforderlich, da diese im weiteren Verlauf häufig verwendet werden und als Verständnisfundament dienen. Der erste Terminus „Produktion“ schließt mehrere Teilbereiche ein.[17] Einer dieser Bereiche ist die Fertigung als „Herstellung materieller Güter unter Einsatz der Ressourcen Material, Energie, Maschinen, Menschen, Kapital, Information und Wissen“.[18] Zusätzlich enthält eine neuere Definition der Produktion die Herstellung immaterieller Güter, die Entwicklung, Konstruktion und Forschung und den gesamten Produktlebenslauf inklusive der Recyclingphasen. Konkret bezieht sich der in dieser Arbeit angesprochene Kontext auf die Fertigung, auch wenn mit Industrie 4.0 verschiedenartige Auswirkungen auf die anderen genannten Bereiche der Produktion einhergehen. Fortlaufend werden die Termini Fertigung und Produktion gleichermaßen verwendet, wobei immer der tatsächliche Wertschöpfungsbereich gemeint ist.[19] Der dritte wesentliche Begriff, ein Synonym, ist der Ausdruck „Shopfloor“, welcher die Wertschöpfungsebene charakterisiert und den Ort, an dem die Wertschöpfungsprozesse eines Unternehmens stattfinden beschreibt.[20] Die vollständige Definition des Begriffs Industrie 4.0 erfolgt aufgrund des größeren Umfangs in einem eigenen Kapitel (2.3).

2.1. Die Produktionsarbeit unter Anpassungsdruck

Zu Beginn steht die Fragestellung im Vordergrund, warum im Produktionsbereich ein hoher Druck zur Anpassung besteht und weshalb die Produktion sich bereits stark verändert hat und auch weiterhin verändern muss.[21] Der Schwerpunkt liegt auf der Veranschaulichung von veralteten Aspekten und Grundbausteinen der Produktion, die nicht mehr den Anforderungen entsprechen und deshalb hinterfragt werden müssen.

Im Allgemeinen sind Fabriken und Produktionssysteme ständigen Veränderungen ausgesetzt.[22] Das circa 100 Jahre alte System der Massenproduktion[23] zur Fertigung einer Variante in großen Mengen wird heutzutage häufig genutzt, um kleinere Mengen vieler Varianten zu produzieren. Das System wird hierbei zweckentfremdet eingesetzt, da sich die Voraussetzungen und Anforderungen an eine Produktion verändert

[17] vgl. Westkämper und Decker (2006), S. 24.
[18] ebenda, S. 24.
[19] vgl. ebenda, S. 24.
[20] vgl. Siepmann (2016), S. 50.
[21] vgl. R. Winkler (2019), URL siehe Literaturverzeichnis.
[22] vgl. Müller und Jentsch (2015), S. 106.
[23] vgl. Schneider (2016), S. 33.

haben. Das Konzept der Massenfertigung ist für eine variantenreiche Produktion nicht ideal, weshalb Produktionssysteme anders ausgelegt und konstruiert werden müssen.[24] Die mittlerweile wichtige Anforderung einer Individualisierung von Produkten wird durch die Massenproduktion ebenfalls nur ungenügend erfüllt.[25]

Seit ihrem Aufkommen wurde die Fließbandfertigung durch zahlreiche Entwicklungen und Technologien vorwärtsgetrieben. Im Umfeld einer klassischen Fließbandfertigung ist eine volle Ausschöpfung des Potentials neuer Technologien nicht möglich. Um auf (Markt-)Veränderungen schnell zu reagieren, muss die Produktion dementsprechend ausgerichtet sein. Es besteht der Anpassungsdruck, sich von den Konzepten der Massenproduktion und Linienorganisation zu lösen. Es sind flexible, kompatible, vernetzte, modular aufgebaute und schnell auf jeweilige Produktfamilien und benötigte Kapazitäten umstellbare Produktionssysteme vonnöten. Grenzenlos verschiebbare Maschinen und Anlagen bieten in der Industrie 4.0 die erforderliche Erhöhung an Mobilität.[26] Ein weiteres zu hinterfragendes Paradigma ist die Aufteilung der Arbeit in repetitive Bearbeitungsschritte mit einem durchgängigen Arbeitstakt. Heute geforderte Tätigkeiten sind längst nicht mehr so eintönig.[27]

Ergänzend sind Unternehmen fortlaufend einem Preisdruck ausgesetzt. Bei der Beantwortung der Frage wie kostensparender produziert werden kann, spielen die Reduktion von Durchlaufzeiten, die Vermeidung von Leerlauf- und Stillstandzeiten,[28] die Termintreue,[29] sowie die Prozessoptimierung eine Rolle. Zudem führt Zeitdruck dazu, Produkte schnell auf den Markt bringen zu müssen, um die unternehmerische Konkurrenzfähigkeit zu sichern.[30] Eine weitere Ineffizienz, die zu einem Aufkommen von Veränderungen führt, ist die Produktivität innerhalb der Produktion. Ohne jegliche Veränderungen zum Beispiel durch Industrie 4.0, ist eine Bewältigung der aufgeführten Probleme nicht in ausreichendem Maße zu schaffen.[31] Dies bezieht sich auch auf das obsolete Konzept der maximalen Planung und Sicherheit, bei welchem alle Produktionsabläufe in maximalem Umfang durchgeplant werden.[32]

Des Weiteren sind Hierarchien mit Meistern und einem strengen Führungsstil zwar bei stupider, erschöpfender Produktionsroutinearbeit zur Aufrechterhaltung von Arbeitsbereitschaften, Disziplin und Leistung des Personals vorteilhaft. Doch mittlerweile muss dieses Konzept durchleuchtet werden. Leistungskontrolle sollte beispielsweise mithilfe von Kennzahl- und Kontrollsystemen erfolgen.[33] Starre und klar definierte Verantwortlichkeiten waren früher in der Produktion sinnvoll.[34] Es lässt sich schlussfolgern, dass hier heutzutage große Anpassungserfordernisse auftreten.

In der nächsten Stufe der Produktion führen Fabrikarbeiter weniger Routinearbeiten durch. Arbeitsaufgaben nehmen an Komplexität zu. Es besteht der Druck ausgebildeten, qualifizierten Mitarbeitern mehr Vertrauen entgegenzubringen und das Konzept der harten Arbeitsteilung zu hinterfragen. Abschließend ist erkennbar, dass die geschilderten Konzepte und Ineffizienzen unter den Voraussetzungen der heutigen Zeit in vielen Umgebungen Schwierigkeiten verursachen, weshalb Produktionsunternehmen an Grenzen der Beherrschbarkeit stoßen.[35] Die Herausforderungen im heutigen Produktionsumfeld müssen mit anderen, neuen Lösungen bewältigt werden.[36] Im weiteren Verlauf der Arbeit soll deutlich gemacht werden, inwiefern Industrie 4.0 als eine solche Antwort auf die verlangte Wandlung und die Beseitigung der Probleme der Produktion fungieren kann.[37] Es stellt sich allerdings weiterhin die Frage, durch welche Faktoren der beschriebene Druck zur stetigen Anpassung der Produktion mitunter ausgelöst wird. Welche Entwicklungen der Arbeitswelt tragen zu verschiedenen Veränderungen in der Produktion bei? Das folgende Teilkapitel liefert hierzu Antworten.

2.2. Veränderungen und Herausforderungen der neuen Arbeitswelt

Bei der Beantwortung der Frage welche Einflüsse und Trends der Arbeitswelt die Gestaltungsfelder Mensch, Technik und Organisation in der Produktion beeinflussen, sind diverse Entwicklungen einzubeziehen.[38] Die digitale Transformation hat einen großen Einfluss auf das wirtschaftliche Geschehen und verändert Unternehmen mitsamt deren Prozessen, Strukturen und Produkten.[39] Zudem gehen mit technologischen Entwicklungen erhebliche Auswirkungen auf die Arbeitsorganisation einher.[40]

Eine beträchtliche Herausforderung der Arbeitswelt ist der globale Wettbewerb. Aufgrund der Globalisierung besteht ein weltweiter Konkurrenzdruck, der die Nutzung von neuen Lösungen unabdingbar macht.[41] Zusätzlich existiert ein verstärkter Druck Innovationen zu erbringen.[42] Neben ökonomischen Entwicklungen verursachen eine Vielzahl von gesellschaftlichen Trends einen starken Anpassungsdruck. Der demographische Wandel ist dabei als ein beispielhafter Faktor zu nennen.[43] Durch diesen kommt es zu einer Verschärfung von ohnehin schon angespannten Wettbewerbssituationen.[44] Dabei rücken der Fachkräftemangel und eine

[24] vgl. Schneider (2016), S. 50.
[25] vgl. Weinreich (2016), S. 186 und Schneider (2016), S. 37.
[26] vgl. R. Winkler (2019), URL siehe Literaturverzeichnis.
[27] vgl. Clement (2006), S. 96f.
[28] vgl. Manzei (2016), S. 11.
[29] vgl. Schneider (2016), S. 37.
[30] vgl. Manzei (2016), S. 11f.
[31] vgl. Ittermann, Niehaus und Hirsch-Kreinsen (2015), S. 26f.
[32] vgl. Schultz-Wild und Lutz (1997), S. 6.
[33] vgl. Clement (2006), S. 96.
[34] vgl. Schultz-Wild und Lutz (1997), S. 6.
[35] vgl. Schneider (2016), S. 35.
[36] vgl. ebenda, S. 37.
[37] vgl. R. Winkler (2019), URL siehe Literaturverzeichnis.
[38] vgl. Nyhuis, Bellmann und Ansari (2015), S. 165.
[39] vgl. Stief (2018), S. 245.
[40] vgl. M. Winkler, Grau und Tilebein (2015), S. 16.
[41] vgl. Abele und Reinhart (2011), S. 11f.
[42] vgl. Nyhuis et al. (2015), S. 168.
[43] vgl. M. Winkler et al. (2015), S. 16.
[44] vgl. Bauer, Schlund, Marrenbach und Ganschar (2017), S. 17.

alternde Belegschaft, deren Arbeitsfähigkeit und Leistungsfähigkeit gesichert werden muss, in den Betrachtungsmittelpunkt.[45] Hinzu kommt, dass mehr Menschen mit Migrationshintergrund Arbeitsplätze in der Produktion einnehmen. Diese Entwicklungen machen Veränderungen in der Produktionsarbeit unabdingbar.

Weiterführend sind neuartige Formen zum Aufbau von Qualifikationen, Kompetenzen und Wissen erforderlich. In diesem Zusammenhang stellt sich die Frage, wie Unternehmen mit den Wirkungen des demographischen Wandels zurechtkommen. Industrie 4.0 kann bei der Bewältigung dieser Herausforderung helfen.[46] Die dynamischen und immer kürzer werdenden Lebenszyklen von Produkten sind in gleichem Maße als eine Herausforderung anzusehen, denn sie animieren Unternehmen zu einem raschen Handeln.[47]

Insgesamt ergeben sich durch die erläuterten Anpassungsgründe, Veränderungen und Entwicklungen konkrete Auswirkungen auf viele Arbeitsplätze. Frey und Osborne (2013) bestimmen den Anteil der in den USA durch die Digitalisierung in Gefahr gebrachten Arbeitsplätze in der Produktion auf nahezu 50 Prozent und gehen von einem ähnlichen Bild für Deutschland aus.[48] In einer weiteren Studie zum zukünftigen digitalisierten Arbeitsleben von Kool und van Est (2015) wird jedoch deutlich, dass es während des Übergangs zu einer veränderten I4.0 Produktion zwar zu einem Verlust von Arbeitsplätzen kommt, zur Kompensation allerdings gleichzeitig neue, andere und hoch technische Arbeitsplätze geschaffen werden.[49] Gut ausgebildete Arbeitskräfte werden von dieser Schaffung profitieren, während gering qualifizierte Produktionsarbeitskräfte zunächst negativ von der Digitalisierung betroffen sind.[50]

Eine beträchtliche Herausforderung der digitalen Transformation besteht darin, durch den Einsatz neuer digitaler Technologien einen größtmöglichen Nutzen zu ziehen. Um dies möglich zu machen müssen Unternehmen ihre Aktivitäten,[51] Technologien und Geschäftsmodelle entsprechend gestalten und neu ausrichten.[52] Zu der Veränderung der Arbeitswelt steuern unterschiedliche zukunftsträchtige Technologien ihren Beitrag bei. Künstliche Intelligenz, maschinelles Lernen, Robotik, mobiles Internet und Cloudservices haben einen großen Einfluss.[53] Die Verbreitung von I4.0 Systemen dient der Anpassung und ist ein Ansatz die Produktionsarbeit zu ändern und weiterzuentwickeln.[54] Industrie 4.0 als übergreifendes Konzept enthält mehrere verschiedene Technologien und Vorstellungen. Doch was genau versteht man unter Industrie 4.0 und welche Kernbestandteile liegen zugrunde?

[45] vgl. Nyhuis et al. (2015), S. 168.
[46] vgl. Abele und Reinhart (2011), S. 19f.
[47] vgl. ebenda, S.15.
[48] vgl. Frey und Osborne (2013), S. 41f.
[49] vgl. Kool und van Est (2015), S. 15ff.
[50] vgl. A. Berg (2019), S. 8.
[51] vgl. Cole (2017), S. 55.
[52] vgl. ebenda, S. 54.
[53] vgl. Klauß und Mierke (2017), S. 23.
[54] vgl. Schumann (2006), S. 72.

2.3. Technische Grundzüge und Inhalte von Industrie 4.0

Der Begriff Industrie 4.0 wird generell zur Beschreibung diverser Vorgänge in der Wirtschaft genutzt. So betrifft er beispielsweise auch Zukunfts- und Digitalisierungsveränderungen in der Medizin und im Verkehr. Der Umfang dieser Arbeit bezieht sich jedoch auf den Industrie 4.0 Teilbereich der industriellen Fertigung.[55] Die entsprechenden Fachbegriffe zur Beschreibung des Einzugs des Internets in die Produktion lauten „Smart Factory“ und „Smart Plant“. In der I4.0 Produktion gewinnen Daten und Dienste an Wertschätzung und sind jetzt essenzielle Bestandteile statt lästiges „Beiwerk“.[56] Zudem ist die Digitalisierung der zu fertigenden Produkte ein zentraler Industrie 4.0 Aspekt („Smart Products“).[57]

In der Industrie 4.0 wird die Produktion grundlegend digitaler und vernetzter.[58] Ziel ist das „Digital Manufacturing“, bei dem der komplette Produktions- und Entstehungsprozess mit sämtlichen vor- und nachgelagerten Wertschöpfungsbereichen virtuell abgebildet wird.[59] Die physische Welt wird digital in eine virtuelle Welt gespiegelt.[60] Es entsteht ein „digitaler Zwilling“ der Realität, der die Fabrik inklusive aller Gebäude abbildet.[61] Dies wird durch die „Smart Factory“ erreicht, ein Netzwerk, in dem Menschen mit intelligenten Maschinen in einer intelligenten Umgebung zusammenarbeiten.[62] Die wesentlichste technische Grundlage ist die komplette Durchdringung der Produktion mit sogenannten cyber-physischen Systemen (CPS). Über sie lassen sich Informationen und Daten über das Produktionsumfeld in Echtzeit ermitteln.[63] In solchen Systemen befinden sich an nahezu allen Orten der Fabrikhalle Sensoren.[64] Zum Beispiel werden Betriebsmittel wie Werkzeugmaschinen, Roboter, Bestandteile des Lagersystems und Transportmittel mit Sensoren und eigenen Computern ausgestattet.[65] Die hergestellten Produkte und die Gebäude erhalten ebenfalls Sensoren. So werden Umgebungsvariablen und Betriebsdaten aufgenommen.[66] Der Mensch ist in dieses hochvernetzte System aller Abläufe komplett eingebunden.[67] Wie Abbildung 3 entnommen werden kann, sind CPS Verbünde vieler verschiedener Komponenten. Durch die vollumfängliche Kombination der physikalischen Welt mit Sensorik- und Aktoriknetzen, Software und Hardware entstehen „eingebettete Systeme“, die Verbindungen, Schnittstellen und Interaktionen zu anderen Systemen, zur Umwelt und zum Menschen zulassen.[68]

[55] vgl. Lucks (2017a), S. 10f.
[56] vgl. Jung (2017), S. 67.
[57] vgl. Zentralverband Elektrotechnik- und Elektronikindustrie e.V. (2015), S. 4.
[58] vgl. Ittermann und Niehaus (2018), S. 36.
[59] vgl. Pötter und Steckenreiter (2017), S. 455f.
[60] vgl. Heidel, Hoffmeister, Hankel und Döbrich (2017), S. 1.
[61] vgl. Bauernhansl und Dombrowski (2016), S. 23.
[62] vgl. Kagermann und Winter (2017), S. 22.
[63] vgl. Mertens, Barbian und Baier (2017), S. 96.
[64] vgl. Cole (2017), S. 162.
[65] vgl. Mertens et al. (2017), S. 96.
[66] vgl. Cole (2017), S. 162.
[67] vgl. Stief (2018), S. 85.
[68] vgl. Broy (2010), S. 21-24.

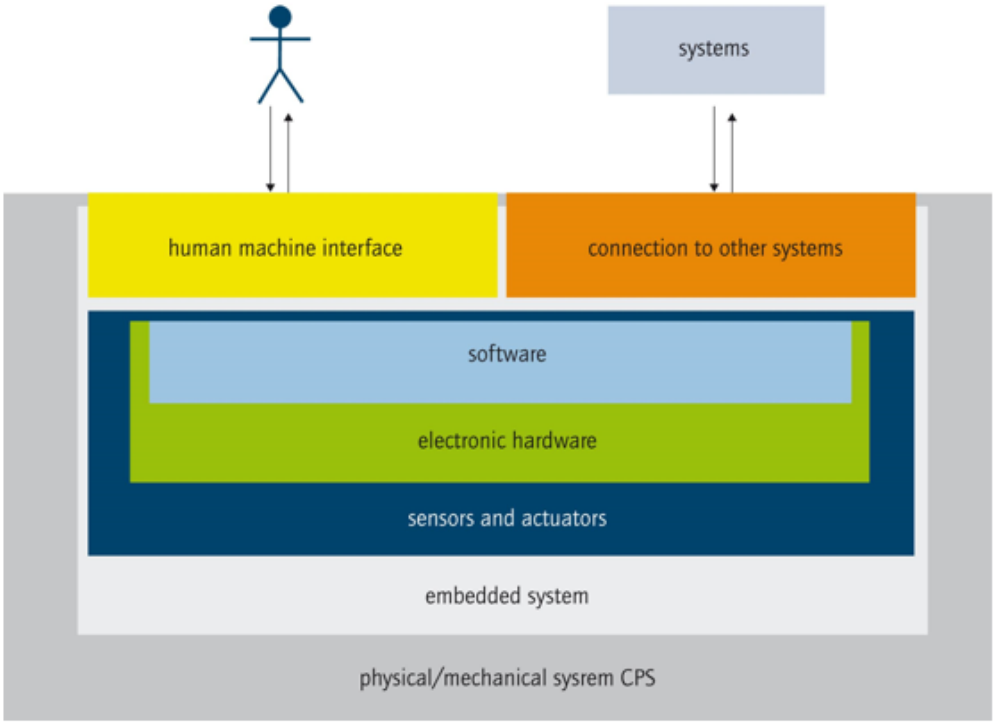


Abbildung 3: System und Prinzip von CPS nach Broy (2010).
Quelle: Broy (2010), S. 24.

Ein notwendiger Bestandteil von Industrie 4.0 ist das „Internet der Dinge“, durch das die Kommunikation aller genannten Systemelemente bewerkstelligt wird.[69] Ergänzend besitzen die „Smart Products“ der Industrie 4.0 ein digitales Gedächtnis, in welchem Fertigungshistorien und kommende Bearbeitungsschritte gespeichert sind. Produktionsanlagen, Roboter und Lager- bzw. Fördersysteme stimmen sich ab und koordinieren, wo Kapazitäten zur Fertigung von Produkten vorhanden sind.[70] Das intelligente Produkt lässt eine eindeutige, durchgängige Identifikation und Lokalisierung in der Fabrik zu.[71] Generell werden I4.0 Methoden erst durch übergeordnete Faktoren, wie Big Data, Datenanalyse[72] und Cloud Computing ermöglicht.[73] Ein endgültiges Ziel der Industrie 4.0 ist eine flexible, kostengünstige „Customized Mass Production“ in der kleine Losgrößen und eine hohe Individualisierung realisierbar sind.[74]

Damit Industrie 4.0 funktioniert, benötigt es eine umfassende Integration in Netzwerken und Produktionssystemen.[75] Die Fähigkeit zur Kommunikation aller Systemkomponenten erfordert die Schaffung geeigneter I4.0 Kommunikations- und Breitbandinfrastrukturen.[76] Die Nutzung von Internet-, Kommunikations-[77] und Datenstandards ist eine weitere Voraussetzung.[78] Der Fachbegriff „Interoperabilität“ steht hier für die nahtlose Zusammenarbeit von Geräten, Systemen und Anwendungen über festgelegte Standards.[79] Für eine Kollaborationsfähigkeit ist es somit erforderlich Schnittstellen zu überwinden. In der I4.0 Produktion müssen alle Akteure eines Wertschöpfungsnetzwerks kooperieren,[80] damit es Übergänge zwischen unterschiedlichen Branchen und Unternehmen gibt.[81] Weiterhin sind Sicherheitskonzepte, der Datenschutz,[82] Haftungsfragen und Datenmissbrauch zu diskutierende Teilaspekte der Industrie 4.0.[83]

Die Verwendung eines Referenzarchitekturmodells, wie die des „Rami 4.0“, soll einen intensiven Datenaustausch und eine durchgängige Kommunikation ermöglichen.[84] Es bietet eine Basis zur Verwirklichung der geforderten Integration verschiedener Systeme.[85] Dieses allgemeine Modell führt die wichtigsten I4.0 Aspekte zusammen, um das grundlegende Ziel der Sicherung des gleichen Verständnisses für alle Nutzer von Industrie 4.0 zu erreichen.[86] Das „Rami 4.0“ Modell und

[69] vgl. Bernsteiner und Bauer (2017), S. 788 und Broy (2010), S. 24.
[70] vgl. Kagermann (2014), S. 244f.
[71] vgl. Heidel et al. (2017), S. 15.
[72] vgl. Schocke (2017), S. 4 und Lucks (2017a), S. 13.
[73] vgl. Schocke (2017), S. 4.
[74] vgl. Heidel et al. (2017), S. 1.
[75] vgl. Banse (2017), S. 126 und Schlund und Pokorni (2017), S. 141.
[76] vgl. Bauernhansl und Dombrowski (2016), S. 23 und Zentralverband Elektrotechnik- und Elektronikindustrie e.V. (2015), S. 9.
[77] vgl. Banse (2017), S. 125.
[78] vgl. Schlund und Pokorni (2017), S. 146.
[79] vgl. Plattform Industrie 4.0 (2019), S. 1f.
[80] vgl. Stief (2018), S. 84.
[81] vgl. Plattform Industrie 4.0 (2018a), S. 3.
[82] vgl. Stief (2018), S.76f., Bauer et al. (2017), S. 19 und Zentralverband Elektrotechnik- und Elektronikindustrie e.V. (2015), S. 11.
[83] vgl. Kagermann et al. (2013), S. 7.
[84] vgl. Hoppe (2016), S. 96f.
[85] vgl. Bernsteiner und Bauer (2017), S. 788f.
[86] vgl. Plattform Industrie 4.0 (2018a), S. 8.

standardisierte Kriterien für I4.0 Produkte sind ergänzend im Anhang einsehbar (Anh. 1 und 2). Insgesamt birgt die Industrie 4.0 viele Herausforderungen. Industrieverbände und die Politik versuchen diese gemeinsam zu bewältigen.[87] Als deutsche Netzwerkinitiative intendiert die „Plattform Industrie 4.0“ die Voranbringung des Themas und arbeitet hierfür mit Verbänden anderer Länder zusammen.[88] Neben solchen konkreten Technologieaspekten besteht bei einem Einsatz von Industrie 4.0 eine Verknüpfung zu grundlegenden Organisationsmodellen.[89]

2.4. Die Rolle des Menschen in der zukünftigen Produktion

Eine erste einleitende Fragestellung des Themenkomplexes Mensch ist es, wie sich die Rolle des Menschen in der I4.0 Produktion verändert und entwickelt. Der Mensch als Mitarbeiter der Zukunftsfabrik nimmt eine immens wichtige Position ein und muss deshalb eine besondere Aufmerksamkeit erhalten.[90] 1991 verdeutlichte Black zunächst, dass die Fabrik der Zukunft weniger Mitarbeiter haben wird.[91] Erste Vorstellungen und Planungen der Industrie 4.0 Produktion gingen sogar von komplett menschenleeren Fabriken aus. Verschiedene Erfahrungen in den letzten Jahren haben dies jedoch eindeutig widerlegt.[92] Der Mensch bildet in Zukunft weiterhin den Mittelpunkt, wobei die eingesetzten Technologien zur Sicherung einer optimalen Unterstützung des Menschen ausgerichtet werden.[93] Das menschliche Agieren in zentraler Rolle wirkt sich vorteilhaft auf die Gewährleistung von Produktivität und Flexibilität in der Produktion aus.[94]

Die veränderte Rolle des Menschen in der Industrie 4.0 zeigt sich insbesondere in seiner neuen Funktion als Überprüfer von Ergebnissen und Abläufen in der Produktion[95] sowie in der Ausfüllung einer übergreifenden Aufsichtsfunktion. Er hat in der Produktion und Montage weniger handwerkliche,[96] manuelle und körperlich belastende Tätigkeiten auszuführen, denn diesen Arbeitsteil erfüllen zunehmend Roboter, die mit dem Menschen kooperieren. Stattdessen übernimmt der Mensch den wissensintensiven Teil, um Maschinen und Anlagen zu überwachen und zu steuern.[97] Durch geringere räumliche Beschränkungen arbeiten flexible Serviceroboter im Team mit dem Menschen.[98] Es bilden sich von starker Kooperation geprägte „Mensch-Maschine-Systeme“ aus.[99]

Resümierend gehört die Zukunft der Produktion Mensch und Maschine gemeinsam.[100] Der Mensch trifft immer noch die Entscheidungen, da er trotz Künstlicher Intelligenz sowie leistungsstarken Maschinen und Technologien in vielen Situationen Vorteile besitzt und überlegen ist. Beispielhaft sind Entscheidungen mit unsicheren Parametern und unklaren Mustern,[101] für die der Mensch mit seinen kognitiven Fähigkeiten optimal gerüstet ist.[102]

Wie bis hierhin erläutert, wandeln sich zum einen die Arbeitsinhalte des Menschen in der industriellen Produktion. Gleichzeitig ändern die komplexen und echtzeitorientierten Produktionssysteme die Erwartungen an Mitarbeiterkompetenzen und Fähigkeiten.[103] Aus diesem Grund müssen produzierende Unternehmen ihre Beschäftigten in der Produktion ausreichend qualifizieren.[104] In vielen Fällen ist das Qualifikationsniveau lediglich für eine einfache Massenproduktion ausreichend und versetzt die Beschäftigten nicht in die Lage, sich stetig weiter zu bilden und weiter zu qualifizieren.[105] Die zur Erfüllung der neuen menschlichen Rolle in der Produktion notwendige konsequente Mitarbeiterqualifikation wird in Kapitel (4.2.1) als Anforderung an die Produktion der Zukunft erfasst und näher beleuchtet.

Zusammenfassend ist die menschliche Arbeit in dem von Veränderung geprägten Produktionsumfeld bereits heute ein wesentlicher Schlüsselfaktor und wird dies auch in Zukunft weiterhin sein.[106] In der I4.0 Produktion sind allerdings neue Formen der Arbeitsgestaltung und Organisation erforderlich.[107] Vor diesem Hintergrund widmet sich das Folgekapitel der wachsenden Bedeutung von Organisation im Produktionskontext.

2.5. Organisation und Industrie 4.0

An dieser Stelle soll die Zusammengehörigkeit von Organisation und Industrie 4.0 genauer betrachtet werden. Die Veränderungsfähigkeit und die ständige Anpassung an Umgebungen sind kritische Erfolgsfaktoren für Unternehmen.[108] Es herrscht die Überzeugung, dass dies auch für die Organisation gilt.[109] Bereits 1986 haben Daft und Huber organisationales Lernen als Anpassung der Organisation an die individuelle Umwelt deklariert.[110] Auch Laloux (2015) setzt seinen Fokus insbesondere auf den Gedanken der Entwicklung von verschiedenen Organisationsformen.[111]

Im Zuge der digitalen Transformation genügen klassische Organisationen mit veralteten und starren Strukturen den neuen digitalen Anforderungen nicht mehr.[112] In der

[87] vgl. Koch, Kuge, Geissbauer und Schrauf (2014), S. 6.
[88] vgl. Plattform Industrie 4.0 (2018a), S. 29f.
[89] vgl. Cole (2017), S. 55.
[90] vgl. Projektträger Karlsruhe (PTKA) des Bundesministeriums für Bildung und Forschung (2017), S. 7.
[91] vgl. Black (1991), S. 3.
[92] vgl. Mertens et al. (2017), S. 23f.
[93] vgl. Pawellek und Schirrmann (2017), S. 120.
[94] vgl. Spath et al. (2013), S. 133.
[95] vgl. Heidel et al. (2017), S. 15.
[96] vgl. Cole (2017), S. 63f.
[97] vgl. Schenk, Haase, Keller und Berndt (2016), S. 131.
[98] vgl. Kagermann und Winter (2017), S. 24.
[99] vgl. Böde et al. (2013), S. 11.
[100] vgl. Spath et al. (2013), S. 45.
[101] vgl. Spath et al. (2013), S. 130.
[102] vgl. Projektträger Karlsruhe (PTKA) des Bundesministeriums für Bildung und Forschung (2017), S. 14.
[103] vgl. Kagermann et al. (2013), S. 6.
[104] vgl. Cole (2017), S. 64.
[105] vgl. Schultz-Wild und Lutz (1997), S. 14.
[106] vgl. Spath et al. (2013), S. 45.
[107] vgl. Projektträger Karlsruhe (PTKA) des Bundesministeriums für Bildung und Forschung (2017), S. 7.
[108] vgl. Westkämper und Löffler (2016), S. 50.
[109] vgl. Zwirner (2004), S. 56ff. und Dichtl und Patermann (2017), S. 200.
[110] vgl. Daft und Huber (1987), S. 2f.
[111] Laloux (2015).
[112] vgl. Klauß und Mierke (2017), S. 137.

Industrie 4.0 muss der Blick auf eine Organisation gerichtet werden,[113] die kollaborativ und übergreifend ausgerichtet ist.[114] Konkret wird die Modifikation der Aufbauorganisation und der Ablauforganisation angesprochen.[115] Die Aufbauorganisation beschreibt die hierarchische Gliederung und reicht in der Produktion von der Abteilungsebene über die Kostenstellen, Arbeitsgruppen und Teams bis zur Mitarbeiterebene nach unten.[116] Sie stellt den Aufbau und die Verantwortlichkeiten im Unternehmen dar.[117] Im Gegensatz dazu umfasst die Ablauforganisation die Planung und Durchführung von Arbeitsinhalten. Ihre Aufgabe ist die zeitliche und räumliche Koordination in der Produktion. Zum Beispiel wird der Einsatz von Ressourcen und der Informationsfluss abgestimmt.[118]

Für die Industrie 4.0 ist die primäre Erarbeitung und Einführung einer grundlegenden Aufbauorganisation essenziell. Entscheidend für den Erfolg ist es, mit den Handlungsfeldern Ablauforganisation und Mitarbeiter zu starten, bevor eine Installation von Technologien erfolgt. Ohne eine vorrangige Entwicklung von Organisation und Menschen wird ein falscher Fokus gesetzt.[119] In vielen Unternehmen ist das technische Know-how vorhanden, jedoch weniger das Wissen über die Wichtigkeit der Veränderung der internen Organisation.[120] Die Organisation der Produktion ist für die geforderten Qualifikationen und das Aussehen der Aufgaben in der Produktion entscheidend.[121]

Für die Organisation in der I4.0 Produktion werden grundsätzlich zwei Organisationsformen verwendet. Ohne konkret auf inhaltliche Details einzugehen, ist es wichtig zu nennen, dass die Mitarbeiterqualifikation sowohl in der sogenannten „polarisierten Organisation“ als auch in der „Schwarmorganisation“ von Bedeutung ist. In jedem Fall wird spezialisiertes Personal mit hohen Handlungsspielräumen benötigt.[122] Es sind umfassende Qualifikationen für das Agieren in I4.0 Organisationsformen notwendig.

Schlussendlich muss die gesamte Organisation des Unternehmens, vor allem jedoch die Produktionsorganisation, auf Industrie 4.0 ausgerichtet werden.[123] Problematisch ist, dass laut einer Bitkom Studie aus dem Jahr 2018 für 68 Prozent der Unternehmen der Einsatz von I4.0 Anwendungen auf technische Prozessverbesserungen abzielt, während sich nur 15 Prozent auf die Schaffung einer flexibleren Arbeitsorganisation fokussieren.[124] Folglich fehlte im Jahr 2018 der Fokus auf die Organisation und den Menschen noch zu großen Teilen. Insgesamt kommen einige neue Konzepte zur Gestaltung der Produktionsorganisation auf.[125] Kapitel (4.) beschäftigt sich mit den organisatorischen Anforderungen an die zukünftige Produktion und stellt mehrerer dieser Konzepte vor. Zunächst erfolgt die Betrachtung wichtiger Kennzahlen zur Steuerung der Produktion.

3. Kennzahlen zur Steuerung der zukünftigen Produktion

Dieses Kapitel zielt darauf ab aufzuzeigen, welche Kennzahlen zur Steuerung einer modernen Produktion verwendet werden können. In einem ersten Schritt werden alle determinierten Kennzahlen übersichtlich dargestellt. Die Bestimmung der Wichtigkeit der ausgewählten organisationalen Steuerungskennzahlen erfolgt durch die quantitative Orientierungsbefragung in Kapitel (5.4). An dieser Stelle werden die Kennzahlen als Verständnisgrundlage zunächst theoretisch erläutert.

Der Auswahl der Indikatoren liegt die konsequente Ausrichtung auf eine Verbesserung von organisatorischen, beziehungsweise menschenbezogenen Themen zugrunde. Ausgewählte, die Technologie betreffende, Kennzahlen sind in die Zusammenstellung integriert, wenn sie in der zukünftigen Produktion ebenfalls eine signifikante Auswirkung auf die Organisation oder den Menschen haben können. Es wird eine Einteilung in die zwei Kategorien „organisations- und technologiebezogene Kennzahlen“ und „mitarbeiterbezogene Kennzahlen“ vorgenommen (Tab. 1).

Diese Kennzahleneinteilung ist aus einer Studie des Karlsruher Instituts für Technologie (KIT) aus dem Jahr 2018 übernommen und leicht adaptiert.[126] Die Tabelle zeigt Kennzahlen, die in dem organisationsbezogenen Kontext dieser Arbeit relevant sind und untersucht werden sollen. Es werden bewusst selten konkrete Formeln erörtert. Im Mittelpunkt steht eher, welche Bedeutung die jeweiligen Kennzahlen haben.

Warum ist bei der Betrachtung von Anforderungen an die zukünftige Produktion auch die Untersuchung entscheidender Kennzahlen relevant? Generell spielen Kennzahlen eine entscheidende Rolle bei der Kontrolle und Überwachung der Produktion. Auch in Zukunft wird sich nichts daran ändern, dass durch sie ein stetiger Verbesserungsprozess vorangetrieben wird.[127] Firmen sind gezwungen die Zusammenstellung ihrer Controllinginstrumente zu überdenken, da mittlerweile andere Aspekte in den Vordergrund rücken und I4.0 Konzepte eine komplexe Auswirkung auf die Wirtschaftlichkeit der Produktion haben.[128] Unternehmen, welche auf I4.0 setzen und ihre Fabriken in „Smart Factories“ verwandeln, sind deshalb zwangsläufig zu einer Adaption ihrer betriebswirtschaftlichen Steuerung, also ihres Kennzahlensystems, gezwungen.[129]

[113] vgl. Duschlbauer, Martin und Saffarnia (2018), S. 11.
[114] vgl. Pawellek und Schirrmann (2017), S. 121.
[115] vgl. Heyse (2018), S. 13.
[116] vgl. Stock (2013), S. 98.
[117] vgl. Lanza, Hofmann, Stricker, Biehl und Braun (2018), S. 47.
[118] Strunz (2012), S. 545.
[119] Dobrzanski und Jungkind (2018), S. 687f.
[120] vgl. Schlund und Pokorni (2017), S. 140f.
[121] vgl. Hirsch-Kreinsen (2018), S. 187.
[122] vgl. Holtgrewe, Riesenecker-Caba und Flecker (2015), S. 26 und Hirsch-Kreinsen und Weyer (2014), S. 26f.
[123] vgl. Reischauer (2015), S. 271ff.
[124] vgl. A. Berg (2018), S. 7.
[125] vgl. Zwirner (2004), S. 56-59.
[126] vgl. Lanza et al. (2018), S. 56.
[127] vgl. Lucks (2017c), S. 286f.
[128] vgl. Schnell (2018a), S. 43.
[129] vgl. Kempkes, Suprano und Wömpener (2018), S. 56f.

Tabelle 1: Eigene Darstellung - Übersicht der ausgewählten Kennzahlen zur Steuerung der zukünftigen Produktion.

Organisations- und technologiebezogene Kennzahlen	Mitarbeiterbezogene Kennzahlen
- Gesamtanlageneffektivität (ÖEE)	
- Qualität: Right First Time/Fehlerrate	
- Schnelligkeit/Durchlaufzeit	
- Bestandshöhen/Reichweiten	- Mitarbeitermotivation und Zufriedenheit
- Liefertreue	- Mitarbeiterengagement
- Instandhaltungskosten	- Personalproduktivität
- Logistikkosten	- Arbeitsschutz und Arbeitssicherheit
- Produktionskosten	- Gesundheitsförderlichkeit und Ergonomie
- Energieverbräuche	
- Wandlungsfähigkeit	
- Produktivität/Produktionsauslastung	

Zunächst ist zur Ausführung möglicher Inhalte eines solchen zukünftigen Kennzahlensystems zu klären, was Kennzahlen überhaupt sind. Kennzahlen oder „Key Performance Indicators“ (KPIs) sind Indikatoren zur Messung einer Leistung.[130] Sie werden in der zugrunde liegenden Literatur als stark verdichtete Messgrößen definiert, die entweder als Verhältniszahlen oder als absolute Zahlen Informationen im Hinblick auf einen bestimmten zahlenmäßig erfassbaren Sachverhalt liefern.[131] Ihre Funktionen liegen darin, die Organisation auf das Wesentliche zu fokussieren, Ziele zu verdeutlichen und Optimierungspotenziale zu identifizieren.[132] Sie informieren über den Stand der Produktion.[133] In der Industrie 4.0 wird das Prozesscontrolling mithilfe von Kennzahlen echtzeitnah durchgeführt und dient der Erkennung von Abweichungen.[134] Zur Überprüfung der Wertentwicklung und der Effekte korrigierender Maßnahmen verlangt der Einsatz von Kennzahlen eine in beliebigen Abständen vollzogene Kontrolle.[135]

Außerdem werden Kennzahlen in der I4.0 Produktion digital über Informationstechnologie (IT) überwacht und ad-hoc gesteuert.[136] Die Verbreitung von Informationen wie Zahlen, Daten und Fakten erfolgt visuell.[137] Dieses Kapitel verdeutlicht die Sinnhaftigkeit des Einsatzes von Kennzahlen in der Industrie 4.0. Die nachfolgenden elf Kennzahlen sind ausgewählte Bestandteile zukünftiger organisationsorientierter Kennzahlensysteme.

3.1. Organisations- und technologiebezogene Kennzahlen

Kennzahl 1: Gesamtanlageneffektivität (OEE)

Zunächst werden die in die erste Gattung von Kennzahlen sortierten organisations- und technologiebezogenen Kennzahlen vorgestellt. Die Gesamtanlageneffektivität wird auch als „Overall Equipment Effectiveness“ (OEE) bezeichnet.[138] Sie beschreibt Verfügbarkeit, Leistung und Qualität von Anlagen und Maschinen.[139] Der Index besteht aus den Komponenten Nutzungsgrad, Leistungsgrad und Qualitätsgrad. Diese drei Komponenten werden durch verschiedene Faktoren beeinflusst. In Bezug auf die Komponente Nutzungsgrad werden beispielsweise Störungen, Rüstzeiten, Einrichtzeiten und Ausfälle in der Berechnung berücksichtigt.[140] Der Leistungsgrad hingegen erfasst unter anderem verringerte Taktgeschwindigkeiten.[141] Wird der Qualitätsgrad betrachtet, fallen zum Beispiel die Menge der Gutteile[142] und Qualitätsverluste durch Nacharbeit und Ausschuss ins Gewicht.[143] Produktionsfehlerquote und Prozessqualität sind entscheidend.[144] Ergänzende Details können nachfolgender Abbildung entnommen werden (Abb. 4).

Durch die Kennzahl OEE kann beobachtet werden, wie effektiv der gegenwärtige Produktionsprozess abläuft. Es lässt sich umgehend erkennen, in welchem Prozess das größte Potenzial zur Verbesserung der Kennzahl steckt.[145]

Kennzahl 2: Qualität - Right First Time/Fehlerrate

Diese Kennzahl beschreibt die Produktqualität und die Fertigungsprozesse in Bezug auf deren Güte und Eignung zur Erfüllung bestimmter Erfordernisse.[146] Dies beinhaltet sowohl Prozessfähigkeit, Prozesssicherheit, als auch das Vorhandensein von Arbeitsqualität. Eine qualitativ hochwertige Fertigung ist transparent und wahrt festgelegte Qualitätsstandards. Es existiert ein etabliertes Qualitätsbewusstsein. Unterziele dieser Kennzahl sind geringe Ausfallraten und niedrige Fehlerquoten.[147] Jede Form von Verschwendung, Ausschuss und Nacharbeit soll vermieden werden, da dies negative Qualitätsauswirkungen mit sich bringt. Bei einer

[130] vgl. Lemmens (2017), S. 58.
[131] vgl. Westkämper und Decker (2006), S. 42.
[132] vgl. Leyendecker und Pötters (2018), S. 56.
[133] vgl. Lanza et al. (2018), S. 23.
[134] vgl. Jeske, Hille und Börkircher (2015), S. 30.
[135] vgl. Leyendecker und Pötters (2018), S. 55.
[136] vgl. Lanza et al. (2018), S. 24f.
[137] vgl. Leyendecker und Pötters (2018), S. 58.
[138] vgl. Strunz (2012), S. 630.
[139] vgl. Bauernhansl und Dombrowski (2016), S. 34.
[140] vgl. Kletti (2015), S. 4 und Strunz (2012), S. 630.
[141] vgl. Strunz (2012), S. 630.
[142] vgl. Stief (2018), S. 88.
[143] vgl. Kletti (2015), S. 4.
[144] vgl. Schlund und Pokorni (2017), S. 149f.
[145] vgl. Leyendecker und Pötters (2018), S. 65.
[146] vgl. Westkämper und Decker (2006), S. 17.
[147] vgl. ebenda, S. 79.

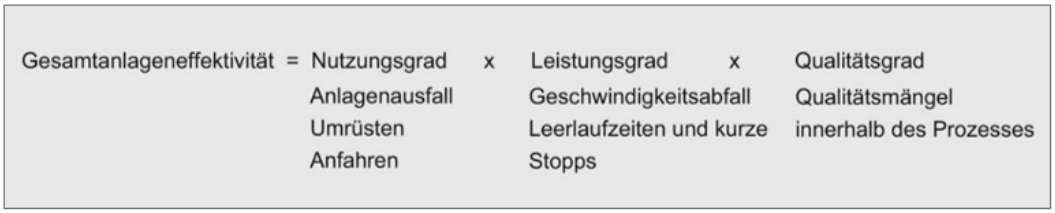


Abbildung 4: Einflussfaktoren auf die OEE nach Strunz (2012).
Quelle: Strunz (2012), S. 630.

Fokussierung auf Qualität wird die Anzahl der korrekt gefertigten Teile der Gesamtproduktion gemessen.[148] Im Zusammenhang mit dieser Definition steht der Produktionsqualitätsaspekt „Right First Time".[149] Teile und Produkte werden „Just in Time" geliefert und produziert.[150] Die „Right First Time" Produktion liefert direkt beim ersten Mal die richtige Qualität und ist auf Anhieb fehlerlos. Diese Kennzahl dient der Beurteilung der Gesamtqualität einer Produktion.[151]

Kennzahl 3: Schnelligkeit/Durchlaufzeit

Die heutige Produktion ist von stetig geringer werdenden Durchlaufzeiten geprägt.[152] Die Durchlaufzeit eines Auftrages wird als Zeitspanne von Beginn der Ausführung, bis zur finalen Auslieferung des Produktes bezeichnet. Sie setzt sich aus der Fertigungsdurchlaufzeit,[153] Bearbeitungszeit, sämtlichen Rüstzeiten und Unterbrechungen des Produktionsflusses in Form von Warte-, Liege- und Transportzeiten zusammen.[154] Außerdem wird die Lagerdauer vor und nach der tatsächlichen Fabrikation einbezogen.[155] Im Kontext von I4.0 kann bei dieser Kennzahl auch nur die voll autonom ablaufende Produktionszeit als Messgrundlage angelegt werden. Die exakte Deklaration und Errechnung der Durchlaufzeit liegt im Ermessen der Unternehmen, eine konsistente Verwendung der Kennzahl ist jedoch ausschlaggebend.[156] Der etwas weiter fassende Begriff der Schnelligkeit inkludiert zusätzlich zu den genannten Zeiten die „Fulfillment Time", bei der die vollständige Auftragsbearbeitungszeit beginnend ab dem ersten Eingang im Unternehmen, bis hin zu späteren After-Sales Aktivitäten berücksichtigt wird.[157]

Kennzahl 4: Bestandshöhen/Reichweiten

Es gibt diverse Gründe die Unternehmen dazu verleiten, mit dieser Größe die Produktion zu lenken. Das liegt unter anderem daran, dass Bestandshöhen häufige Problemverursacher in der Produktion sind.[158] Hohe Bestände bedeuten hohe Kosten[159] und eine unnötige Bindung von Kapital. Das angestrebte Ziel dieser Kennzahl ist eine Bestandsreduzierung. Dies kann Auswirkungen auf die Verringerung der Durchlaufzeit mit sich bringen.[160] Für eine Reduzierung von Bestandsreichweiten wird Kenntnis über die notwendigen Mindestbestände der Produktionsprozesse benötigt. Die Bestandsreichweite misst die absolute Bestandshöhe in Relation zum Verbrauch. Je nach Veränderung des Verbrauches für einen Teil der Produktion wird mit der Festlegung der Bestandshöhen und Bestandsreichweiten eine entsprechende Reaktion zur Anpassung auf veränderte Umstände gegeben.[161] Die Reichweite gibt außerdem an, für welche Dauer der aktuelle Bestand aufgrund von geplanten Absätzen ausreicht.[162]

Kennzahl 5: Liefertreue

Diese Kennzahl ist in der modernen Produktion ebenfalls in Betracht zu ziehen.[163] Die auch als Lieferzuverlässigkeit bezeichnete Messgröße dient zur Ermittlung der Differenzen von geplanten und tatsächlichen Solllieferterminen. Sie berechnet sich aus dem Quotienten der Gesamtanzahl der Aufträge, bei denen der Fertigstellungstermin und Liefertermin eingehalten wurde und einem ausgewählten Betrachtungszeitraum.[164] Es wird der Anteil der pünktlichen Lieferungen an der Zahl der gesamten Lieferungen gemessen.[165] Liefertermintreue heißt, das korrekte Produkt in der erforderlichen Qualität genau zum richtigen Zeitpunkt zu liefern. Dabei ist eine zu frühe Fertigstellung und das Vorziehen von Aufträgen mit unnötigen Lagerkosten verbunden. Eine verspätete Fertigstellung und Auslieferung ist noch unvorteilhafter, da sie diverse kostenintensivere Folgen nach sich zieht.[166] Gruber et al. (2011) stellen diese monetären Auswirkungen mithilfe einer Bewertungsfunktion in einem Zeitverlauf dar und bilden dadurch den Unterschied zwischen zu früher und zu später Fertigstellung ab.[167]

Kennzahl 6: Instandhaltungskosten

Instandhaltungskosten sind die gesamten Kosten für die unterschiedlichen Tätigkeiten der Instandhaltung in der Produktion.[168] Die übergeordnete Zielfunktion strebt eine Mini-

[148] vgl. Dale und Lascelles (2007), S. 125 und Dale (2007), S. 354.
[149] vgl. Huber (2013), S. 120.
[150] vgl. Lowe, Delbridge und Oliver (1997), S. 785.
[151] vgl. Ramezani und Luckow (2019), S. 78.
[152] vgl. C. M. Schlick, Heinen und Frenz (2015), S. 40 und Schlund und Pokorni (2017), S. 149.
[153] vgl. Posluschny (2007), S. 61f.
[154] vgl. Kletti (2015), S. 6 und Leyendecker und Pötters (2018), S. 65.
[155] vgl. Posluschny (2007), S. 62.
[156] vgl. Kempkes et al. (2018), S. 59.
[157] vgl. H. Werner (2014), S. 68.
[158] vgl. Kletti (2015), S. 5.
[159] vgl. Stief (2018), S. 107.
[160] vgl. Kletti (2015), S. 5f.
[161] vgl. Kletti und Schumacher (2014), S. 163f.
[162] vgl. Gabath (2010), S. 159.
[163] vgl. C. M. Schlick et al. (2015), S. 39f.
[164] vgl. Tschandl, Brunner und Wilfinger (2014), S. 79ff.
[165] vgl. Leyendecker und Pötters (2018), S. 62.
[166] vgl. Sengotta (2003), S. 249 und Gruber et al. (2011), S. 71.
[167] vgl. Gruber et al. (2011), S. 71f.
[168] vgl. Strunz (2012), S. 16.

mierung dieser an.[169] Mit der Kennzahl ist es möglich, die Entwicklung der Instandhaltung anlagenspezifisch zu tracken. Alle durch geplante und ungeplante Instandhaltung sowie durch Betriebsunterbrechungen entstehende Kosten werden addiert.[170] Als Beispiele für solche Kostenbelastungen sind an dieser Stelle der Einsatz von Instandhaltungsmaterialien, Werkzeugen und allen möglichen Hilfsmitteln zur Messung und Prüfung zu nennen. Zur Instandhaltung gehören zudem Präventionsmaßnahmen zur Vermeidung von Maschinenausfällen und technischen Störungen durch Wartung und Inspektion.[171] Diese Kennzahl zur Beherrschung des zukünftigen Shopfloors ist nicht die einzige, die auf eine Verringerung von Kosten abzielt.

Kennzahlen 7 und 8: Logistikkosten und Produktionskosten

Generell wird der existierende Druck zur Deckung von Fertigungskosten durch eine hohe Automatisierung und Digitalisierung verstärkt.[172] Neben den Instandhaltungskosten sind somit die Logistik- und Produktionskosten im I4.0 Kontext als Kennzahlen in Frage kommende Kostenarten.[173] Die Größe Logistikkosten nimmt Bezug auf die Wirtschaftlichkeit von logistischen Prozessen. Güter sollen ohne Qualitätsverlust, in einer definierten Menge, zur korrekten Zeit, an einem bestimmten Ort zur Verfügung stehen. Die Messung der dabei auftretenden Kostenfaktoren ist Aufgabe dieser Kennzahl.[174]

Die Kennzahl Produktionskosten besitzt in vielen Fällen höchste Priorität.[175] Eine Senkung der Produktionskosten ist auch in Zukunft eine Möglichkeit zur Steuerung der Produktion, da sich diese Tatsache trotz diverser Neuerungen im Produktionsumfeld nicht ändern wird.[176] Im Detail werden bei der Analyse von in der Produktion anfallenden Kosten unter anderem Preisabweichungen und Mengenabweichungen festgehalten. Exemplarisch ist die Betrachtung von Mehr- oder Minderkosten aufgrund veränderter Einsatzmengen, Fertigungszeiten oder Stücklisten.[177]

Kennzahl 9: Energieverbräuche

Aufgrund von Industrie 4.0 tritt diese Kennzahl im Zusammenhang mit der Einsparung von Energiekosten weiter in den Vordergrund.[178] Eine Möglichkeit ist es, die Produktion hinsichtlich ihres Energieaufwandes und -verbrauches zu optimieren, da hier ein großes Performancesteigerungspotenzial vorhanden ist.[179] Bei der Steuerung dieser Kennzahl sind Ressourcenverbräuche und Energiethemen betreffende Emissionen von Belang.[180] Zu den Energiekosten in der Produktion zählen neben vielen weiteren Aspekten auch die Stromverbräuche für den Betrieb von Anlagen und Maschinen, Heiz- und Kühlkosten der Fertigungsflächen[181] sowie Kosten für Gas und Wasser. Die Energieverbräuche werden durch langzeitiges Erfassen des Bedarfs dieser verschiedenen Faktoren ermittelt.[182]

Kennzahl 10: Wandlungsfähigkeit

Wandlungsfähigkeit ist eine qualitative Eigenschaft der Fabrik, auf jede Art von Veränderung oder Schwankung in der Nachfrage reagieren zu können (siehe Teilkapitel zu Anforderungen an die zukünftige Produktion 4.3.3).[183] Vorerst wird nur die Funktion von Wandlungsfähigkeit in Kennzahlensystemen zur Produktionssteuerung betont. Um den Reifegrad einer praktizierten Wandlungsfähigkeit zu bestimmen und zu bewerten, ist die Nutzung entsprechender Kennzahlen unabdingbar.[184] Dadurch findet eine Bewertung hinsichtlich des Nutzens wandlungsfähiger Systeme und insbesondere der Kosten ungenügender Wandlungsfähigkeit statt. Es können Stellhebel identifiziert werden.[185]

Mithilfe dieser Kennzahl wird außerdem festgestellt, an welcher Position zwischen idealer und maximal möglicher Wandlungsfähigkeit die jeweilige anforderungsgerechte Wandlungsfähigkeit erreicht ist und wie weit das zu prüfende Produktionssystem von diesem Zustand entfernt liegt.[186] Eine genaue Beschreibung der zur Wandlungsfähigkeit gehörigen Teilkennzahlen, kann der Veröffentlichung des Fraunhofer-Institut für System- und Innovationsforschung (2020) „Benchmark Wandlungsfähigkeit“ entnommen werden. Jede Teilkennziffer der Wandlungsfähigkeit bezieht sich dort auf einen der zwei übergeordneten Bereiche „Einsatzintensität und Nutzung von wichtigen Technologien“ oder „organisatorische Konzepte zur Verbesserung der Output-Performance“.[187]

Kennzahl 11: Produktivität/Produktionsauslastung

Die letzte Kennzahl in dieser Kategorie ist die Produktivität.[188] Sie gibt Auskunft darüber, wie gut die zum Einsatz kommenden Produktionsfaktoren im Unternehmen kombiniert und genutzt werden.[189] Die Produktivität wird durch das Verhältnis von Output zu Input definiert.[190] Als Größe für den Output kann die Produktionsleistung herangezogen werden.[191] Der Input wird meist über den Maschineneinsatz oder das verwendete Material angegeben.[192] Ferner wird die Produktivität über verschiedene Teilproduktivitäten konstruiert. Im Kontext dieser Arbeit wird die Kennzahl Produktivität als Zusammenschluss von Material- und Anlagenproduktivität festgelegt. Die Fokussierung der Arbeit auf den

[169] vgl. Strunz (2012), S. 26.
[170] vgl. Leidinger (2017), S. 119f.
[171] vgl. Strunz (2012), S. 22.
[172] vgl. Gottmann (2016), S. 55f.
[173] vgl. Stief (2018), S. 107.
[174] vgl. Reichmann, Kißler und Baumöl (2017), S. 415f.
[175] vgl. Schmitt (2012), S. 123.
[176] vgl. ebenda, S. 127.
[177] vgl. Schnell (2018b), S. 94.
[178] vgl. Stief (2018), S. 107.
[179] vgl. Hohberger und Damlachi (2017), S. 160.
[180] vgl. Roth (2018), S. 134.
[181] vgl. Joos (2012), S. 76.
[182] vgl. Westkämper und Decker (2006), S. 96.
[183] vgl. Steegmüller und Zürn (2017), S. 28f.
[184] vgl. Pachow, Heins, von Bredow, Krebs und Wörn (2008), S. 98.
[185] vgl. Nyhuis (2008), S. 142.
[186] vgl. Berkholz (2008), S. 147.
[187] vgl. Fraunhofer-Institut für System- und Innovationsforschung (2020), S. 1ff., URL siehe Literaturverzeichnis.
[188] vgl. Schmitt (2012), S. 123.
[189] vgl. Leyendecker und Pötters (2018), S. 61.
[190] vgl. Schmitt (2012), S. 128.
[191] vgl. Schnell (2018b), S. 96.
[192] vgl. Schnell (2018b), S. 39. und Schnell (2012a), S. 46.

Menschen und die Organisation hat zur Folge, dass die Personalproduktivität als mitarbeiterbezogene Kennzahl gelistet werden muss. Die Materialproduktivität wird beispielsweise über die Materialausschussquote gemessen, während dies bei der Anlagenproduktivität über den Anlagennutzungsgrad geschieht.[193]

Weiterhin ist die Produktionsauslastung ein Element der Kennzahl Produktivität.[194] Sie wird zum Beispiel über die Aufführung von verfügbarer Kapazität und der tatsächlich genutzten Kapazität determiniert.[195] Für den Fall, dass die Produktivität sehr niedrig ist, kann durch die Mitarbeiter ein Anreiz gegeben werden, die aktuelle Produktionssituation verbessern zu müssen.[196] Welche Kontrollinstrumente zur Überprüfung und Steuerung der Mitarbeiter infrage kommen, wird im nächsten Teilkapitel behandelt.

3.2. Mitarbeiterbezogene Kennzahlen

Mitarbeiter sind grundsätzlich ein erheblicher Potenzialfaktor in Unternehmen, weshalb bei der Produktionssteuerung auch personalbezogene Kennzahlen zu beachten sind.[197] Sie bilden als weiche Faktoren die zweite Gruppe in dieser Kennzahlensammlung.

Kennzahl 12: Mitarbeitermotivation und Zufriedenheit

Bei diesem Indikator ist nicht nur die Zufriedenheit des Personals ausschlaggebend, sondern auch dessen Einsatzbereitschaft und Motivation. Diese Faktoren werden über Einstellungen und Verhaltensweisen der Mitarbeiter abgebildet.[198] Ansatzpunkte zur Bestimmung der Zufriedenheit von Mitarbeitern sind beispielsweise Fehlzeiten und die Fluktuation des Personals.[199] Eine gute Kommunikation, Vertrauen und die Einbeziehung von Mitarbeitern haben positive Effekte, die bei der Nutzung dieser Kennzahl einberechnet werden sollten.[200] Besteht eine hohe Zufriedenheit, ist eine entsprechend hohe Motivation für die Aufgaben in der Produktion zu erwarten.

Über die Erfassung dieser Kennzahl wird die aktuelle Situation und Stimmung der Mitarbeiter aufgenommen.[201] Eine grundlegende Aufnahme und Messung kann anhand der Verbesserungsvorschlagsrate erfolgen.[202] Werden Entscheidungsbefugnisse, Eigenverantwortung sowie interessante Arbeitsaufgaben eingeräumt, verursacht dies ebenfalls motivatorische Ansporn Effekte.[203] All diese Aspekte sollen mit der Kennzahl Mitarbeiterzufriedenheit und Motivation erfasst werden. Mitarbeiterengagement ist eine Folge von Zufriedenheit, benötigt als Grundlage ein hohes Maß an Motivation und ist für die Aufnahme in ein menschenbezogenes Kennzahlensystem ebenfalls geeignet.[204]

Kennzahl 13: Mitarbeiterengagement

Im Vergleich zu der reinen Zufriedenheit und Motivation hat Engagement einen deutlich aktiveren Charakter. Engagement ist durch einen intensiven persönlichen Einsatz, Willen und ein aktives Bemühen, einen wertvollen Beitrag leisten zu wollen, registrierbar.[205] Es resultiert die Bereitschaft, weit über normale Arbeitsanforderungen hinauszugehen[206] und einiges an persönlicher Energie in die Arbeit zu investieren.[207]

Als übergreifende Kennzahl ist das Mitarbeiterengagement mithilfe von Teilkennzahlen messbar. Ein einfaches Zeichen für Engagement ist die Anwesenheit der Mitarbeiter,[208] während die Zeit, die in Schulungen verbracht wird ein stärkeres Zeichen für Engagement setzt.[209] Die verhältnismäßig häufige Leistung von Überstunden ist in gleichem Maße ein starkes Indiz für Engagement.[210] Ergänzende Punkte, die einen Einfluss auf das Engagement und die Performance haben können, sind der Arbeitsplatz, das Maß der Unterstützung im Team, der Führungsstil und die Kultur der Organisation.[211] Bei vorhandenem Engagement ist die Arbeitsqualität signifikant höher.[212]

Kennzahl 14: Personalproduktivität

Die Personalproduktivität als menschenbezogener Teil der Gesamtproduktivität ist eine weitere Größe zur Steuerung und Optimierung der „Smart Factory“.[213] Eine Methode zur Beurteilung der Personalproduktivität ist es, die Produktionsleistung ins Verhältnis zu der Anzahl der eingesetzten Mitarbeiter zu setzten.[214] Im Detail geschieht dies bei dem Fertigungspersonal über das Verhältnis der Produktivzeit zu der Anwesenheitszeit.[215] Die übrigen, indirekt an der Fertigung beteiligten Mitarbeiter werden ebenfalls betrachtet. Bei der Personalproduktivität sind Verlustquellen wie vergeudete Arbeitszeit durch Warten und Suchen hinderlich, sie verringern den resultierenden Wert.[216]

Diese Kennzahl wird zu großen Teilen zur Bestimmung der Arbeitsfähigkeit und Arbeitswilligkeit der Mitarbeiter herangezogen.[217] Im Rahmen von Automatisierung und I4.0 Anwendungen ist eine Kostenreduzierung aufgrund von Zeitersparnis und Aufwandsreduktion bei Mitarbeiteraufgaben durchaus realisierbar.[218] Dies wirkt sich auch auf die Personalproduktivität aus, denn den Mitarbeitern bleibt

[193] vgl. Schnell (2012a), S. 60.
[194] vgl. Lemmens (2017), S. 58 und Schnell (2018b), S. 89.
[195] vgl. Schnell (2018b), S. 37f.
[196] vgl. Leyendecker und Pötters (2018), S. 61.
[197] vgl. Gottmann (2016), S. 56.
[198] vgl. Weber (2006), S. 6.
[199] vgl. Gottmann (2016), S. 59.
[200] vgl. ebenda, S. 57f.
[201] vgl. Horzella (2010), S. 371.
[202] vgl. Schnell (2018b), S. 89.
[203] vgl. Laufer (2013), S. 44.
[204] vgl. Christian, Garza und Slaughter (2011), S. 95f.
[205] vgl. Kahn (1990), S. 700.
[206] vgl. ebenda, S. 718f.
[207] vgl. ebenda, S. 714f.
[208] vgl. Gallup Organization (2017), S. 39.
[209] vgl. ebenda, S.178f.
[210] vgl. Schnell (2018b), S. 89.
[211] vgl. Christian et al. (2011), S. 96f.
[212] vgl. Gallup Organization (2017), S. 39f.
[213] vgl. Sinsel (2020), S. 135.
[214] vgl. Schnell (2018b), S. 97.
[215] vgl. Schnell (2012a), S. 60.
[216] vgl. Sinsel (2020), S. 135f.
[217] vgl. Schnell (2012b), S. 60.
[218] vgl. Stief (2018), S. 85.

nun mehr Zeit für andere, wichtigere Aufgaben, die einen maßgeblicheren Einfluss besitzen. Es benötigt folglich eine Konzentration auf die richtigen Aspekte, um diese Kennzahl hochzuschrauben.[219]

Kennzahl 15: Arbeitssicherheit und Arbeitsschutz

Es ist schwierig die Arbeitssicherheit und den Arbeitsschutz in konkreten Werten auszudrücken. Deswegen werden Hilfsgrößen verwendet.[220] Bei der Kennzahlberechnung fließen Arbeitsunfälle inklusive aller Wegunfälle,[221] Unfälle mit Ausfallzeiten und Unfälle mit Todesfolge ein.[222] Andererseits wird zur Bestimmung die Gesamtheit der Unfallverhütungsmaßnahmen verwendet.[223] Grundsätzlich gehören zum Arbeitsschutz und zur Arbeitssicherheit auch die darauf ausgelegte Gestaltung der Arbeitsstätten, Arbeitsumgebungen und Arbeitsmittel. Mögliche Gefahrenstellen werden systematisch so soweit es geht vom Menschen entfernt, wodurch dieser weitestgehend geschützt ist und sicher arbeiten kann.[224] Bei der Messung des Arbeitsschutzes und seiner Qualität werden unter anderem die Präventionsmaßnahmen betrachtet.[225]

Kennzahl 16: Gesundheitsförderlichkeit und Ergonomie

Die letzte untersuchte Kennzahl zur Steuerung der Produktion blickt auf die ergonomischen Bedingungen[226] sowie auf die gesundheitlichen Beanspruchungen in der Produktion.[227] Als Bestandteil des Arbeitsschutzes konzentriert sich die Gesundheitsförderlichkeit explizit auf Gesundheitsgefahren und eine humane Arbeitsgestaltung.[228] Ergonomie gilt als Ausmaß dafür, inwiefern die Arbeitsumgebung korrekte Bewegungen und Haltungen ermöglicht.[229]

Insgesamt erfolgt eine Betrachtung direkter Belastungen und der Ausführbarkeit, Schädigungslosigkeit, Beeinträchtigungsfreiheit sowie Persönlichkeitsförderlichkeit bei Fertigungsaufgaben.[230] Ein konkretes Beispiel für den Gesundheitsschutz ist die geringe Verwendung von Schad- und Gefahrenstoffen.[231] Weiterhin spielen bei der Betrachtung und Messung dieses Faktors Lärmbelastungen, ein schlechtes Raumklima, ein hohes Stresslevel, zusätzliche Vibrationen und Erschütterungen sowie eine schlechte Beleuchtung eine große Rolle.[232] Die gesundheitsfördernde Arbeitsgestaltung und Arbeitsorganisation dient der Erhaltung der langfristigen Leistungsfähigkeit.[233] Es sind nicht nur potenzielle Schädigungen der physischen Gesundheit zu vermeiden.[234] Bei der Benutzung dieser Kennzahl darf der Schutz vor psychischen Folgeschäden und eine nachhaltige Betreuung der Gesundheit nicht außer Acht gelassen werden.[235] Zusammenfassend misst diese Kennzahl, wie stark gesundheitsförderliche und ergonomische Aspekte in der Produktion ausgeprägt sind.

Schließlich geht es in einer Industrie 4.0 Fabrik darum, die zentralsten KPIs der Fertigung zu verbessern.[236] Unternehmen entwickeln hierzu mit den für Sie relevanten Kennzahlen ein Steuerungssystem und strategisches Instrument, damit die Produktion kontinuierlich optimiert werden kann.[237] Die vorgestellten Kennzahlen sind vor allem für eine auf organisationstechnische Merkmale achtende Produktionssteuerung relevant und müssen im Zusammenhang mit den Anforderungen an die I4.0 Produktion betrachtet werden.

Es ist eine Anforderung an die zukünftige Produktion, dass Unternehmen ihre Kennzahlensysteme individuell und richtig zusammenstellen, um damit eingesetzte I4.0 Anwendungen digital zu steuern. Schlussfolgernd sollten Kennzahlen in der zukünftigen I4.0 Produktion einen umfassenden Einsatz finden,[238] denn sie dienen dem Führen mit konkreten Zahlen, Fakten und Daten und schaffen Transparenz auf dem Shopfloor.[239] Im besten Fall werden dabei, wie in diesem Kapitel dargelegt, einige Kennzahlen verwendet, die das Personal betreffen und miteinschließen. Dieser ersten Anforderung an die Produktion der Zukunft folgen in Kapitel (4.) 14 überwiegend den Menschen und die Organisation betreffende Anforderungen. Am Anschluss erfolgt durch die Auswertung der quantitativen Umfrage in Kapitel (5.) eine konkludierende Untersuchung der Wichtigkeit der Kennzahlen und Anforderungen.

4. Organisationale Anforderungen an die Produktion der Zukunft

Moderne Produktionen stellen neue Anforderungen an die Organisation, durch die dem Anspruch zur Erbringung einer höheren Leistung besser Rechnung getragen werden soll.[240] Ein wesentliches Forschungsziel dieser Arbeit ist es, die essenziellsten organisationalen Anforderungen an die zukünftige Produktion aufzuzeigen. Die ermittelten Anforderungen sind nicht immer eindeutig differenzierbar, da sie sich teilweise überschneiden und bedingen. Wie bereits an anderen Stellen, dient hier die Einteilung in die Kategorien „Organisation“, „Mensch“ und „Technologie“ dazu, eine grundlegende, sinnvolle Zuordnung und Sortierung vorzunehmen. Die Wahl dieser drei Dimensionen auf dem Shopfloor Level wird als sinnvoll erachtet.[241] Alle drei Bereiche sind wesent-

[219] vgl. Schnell (2012a), S. 60f.
[220] vgl. Sengotta (2003), S. 247.
[221] vgl. C. Schlick, Bruder und Luczak (2018), S. 392.
[222] vgl. International Association of Oil & Gas Producers (2014), S. 1.
[223] vgl. C. Schlick et al. (2018), S. 386.
[224] vgl. ebenda, S. 398.
[225] vgl. Bieback und Oppolzer (1999), S. 11.
[226] vgl. Sengotta (2003), S. 247.
[227] vgl. Schaper (2019a), S. 415.
[228] vgl. C. Schlick et al. (2018), S. 386.
[229] vgl. Schaper (2019b), S. 403.
[230] vgl. ebenda, S. 403f. zitiert nach: Hacker und Richter (1984).
[231] vgl. C. Schlick et al. (2018), S. 399.
[232] vgl. Bimos KG (2020a), URL siehe Literaturverzeichnis.
[233] vgl. Schaper (2019a), S. 415.
[234] vgl. Sanders und Lampe (2011), S. 209f.
[235] vgl. Schaper (2019a), S. 421.
[236] vgl. o.V. (2019), S. 35.
[237] vgl. Kempkes et al. (2018), S. 60.
[238] vgl. Ungermann et al. (2019), S. 25f.
[239] vgl. Leyendecker und Pötters (2018), S. 11.
[240] vgl. Adami und Houben (2008), S. 84f.
[241] vgl. Hertle (2018), S. 49 - 53.

liche Teile der industriellen Arbeit und müssen deshalb betrachtet werden.[242]

Es werden Anforderungen behandelt, die laut der zugrundeliegenden Literatur signifikant für die Produktion der Zukunft sind. Die in diesem Kapitel theoretisch erklärten organisationalen Anforderungen sind mögliche Konzepte zur Organisationsgestaltung einer zukünftigen industriellen Produktion. Die Anforderungen aus der Kategorie „Technologie" betreffen neben der Technik auch die Organisation maßgeblich, weshalb sie in dieser Arbeit als organisationale Anforderungen aufgeführt werden. Eine hauptsächliche zu diskutierende und beantwortende Fragestellung ist, welche Shopfloor Organisationsmethoden und Instrumente sich mit dem Blick auf die Faktoren Mensch und Technologie in Verknüpfung mit organisationstechnischen Aspekten für die Produktion ergeben.[243] Dies soll im Folgenden systematisch dargelegt werden. Es folgt zunächst eine Betrachtung der zur Kategorie „Organisation" zugeordneten Anforderungen.

4.1. Anforderungen der Kategorie „Organisation"

4.1.1. Verbesserte Kommunikation und Abstimmung

Eine Art der Organisationsgestaltung der Produktion ist die Fokussierung auf eine verbesserte Kommunikation und Abstimmung. Die Einführung und Nutzung von Industrie 4.0 kann zu einer Verbesserung der Kommunikation auf dem Shopfloor führen.[244] In den operativen Bereichen wird Kommunikation zur Abklärung von Zuständigkeiten und für eine betriebssichere Fertigung als ausgesprochen wichtig eingeschätzt.[245]

Grundsätzlich ist die Kommunikation und Abstimmung in der Produktion keine reine Aufgabe der Managementebene mehr.[246] Der Bedarf an Kommunikation, Kooperation und Abstimmung in der Produktion nimmt im Kontext von I4.0 weiterhin zu.[247] Um diesem Bedarf gerecht zu werden sind die physischen und organisatorischen Gegebenheiten der Produktion entsprechend auszurichten. Die Fähigkeiten der Mitarbeiter sind entscheidend für eine effiziente Kommunikation und Kooperation in der Produktion, die zusätzlich durch diverse IT-Systeme ermöglicht und unterstützt wird.[248] Vor allem die Kommunikationsfähigkeit des Personals ist bei der Umsetzung dieser Anforderung eine unverzichtbare Eigenschaft.[249] Betrachtet man näher, wie die Kommunikation ermöglicht wird, lässt sich die Nutzung von mobilen Kommunikationstechniken identifizieren.[250] Produktionsarbeiter sind über I4.0 Technologien untereinander vernetzt.[251] Der Einsatz von Wearables ist eine von vielen Möglichkeiten hierfür.[252] Das zukünftige Produktionspersonal kann durch eine solche vernetzte Steuerung und Kommunikation von außerhalb der Fabrik arbeiten und ist flexibel einsetzbar.[253]

Ein zentrales Ziel einer verbesserten Kommunikation ist es, Zeiteinsparungen durch unmittelbar verfügbare Informationen zu induzieren. Treten Probleme auf, können Informationen mithilfe von I4.0 Technologien zur sofortigen Problemaufklärung in Form von Bildausschnitten und Erläuterungen schnell an zuständige Stellen und Standorte weitergeleitet werden. Durch diese kommunikative Vernetzung ist eine Realisierung verschiedener organisatorischer Arbeitstechniken in der Produktion möglich.[254] Neben der verbesserten produktionsinternen Kommunikation sind in diese Anforderung externe Abstimmungen und Kommunikationsaufwände mit Lieferanten und Kunden inbegriffen,[255] auf die an dieser Stelle nicht weiter eingegangen wird.

Die Arbeitsorganisation der zukünftigen Produktion ist nicht nur von einer erhöhten Interaktion und Kommunikation unter dem Personal geprägt.[256] Der Mensch kommuniziert in verschiedenen Formen mit Maschinen und Objekten der Produktion.[257] Anlagen, Geräte und Systeme interagieren und verarbeiten direkt Echtzeitdaten. So werden aktuelle Zustände der Produktion kommuniziert. Diese technische Kommunikation trägt verknüpfend dazu bei, in der übergreifenden Organisation eine bessere Kommunikation einzurichten.[258] Durch Technik werden bessere Kommunikationsmöglichkeiten geschaffen, von denen der Mensch wiederum profitiert.[259]

Die Existenz einer einwandfreien Kommunikation ist prägend für eine zukünftige Arbeitskultur,[260] die nun auch in der Produktion Einzug hält.[261] Die Bedeutung von Zusammenarbeit und Kommunikation im Team wächst kontinuierlich.[262] Durch eine effiziente, schnelle und lösungsorientierte Kommunikation sowie eine über Grenzen reichende, direkte Interaktion der Mitarbeiter wird die unmittelbare Arbeitsausführung assistiert.[263] Eine weitere Komponente einer ausgefeilten Kommunikation in der Produktion ist die Verbesserung sämtlicher Kommunikationswege. Diese sollten kurz und transparent sein. Sie werden durch das Produktionslayout bestimmt.[264] Eine kurze Kommunikationsdistanz trägt dazu bei, dass Mitarbeiter besser informiert werden.[265]

Im Spektrum dieser Anforderung steht die Schaffung einer Organisationsstruktur, die eine direkte, zwischenmenschliche Kommunikation sicherstellt.[266] Mit Blick auf die Arbeitsorganisation der Produktion trägt eine gute Kommuni-

[242] vgl. E. Hartmann (2015), S. 11.
[243] vgl. Schultz-Wild und Lutz (1997), S. 190.
[244] vgl. Stich, Gudergan und Senderek (2018), S. 164.
[245] vgl. Böhle und Bolte (2002), S. 97ff.
[246] vgl. ebenda, S. 246.
[247] vgl. Wischmann und Hartmann (2018a), S. 22.
[248] vgl. Will-Zocholl (2011), S. 174f.
[249] vgl. Aßmann (2017), S. 329.
[250] vgl. Schlund und Pokorni (2017), S. 141.
[251] vgl. Spath et al. (2013), S. 57.
[252] vgl. Mättig, Jost und Kirks (2018), S. 70f.
[253] vgl. Jeske et al. (2015), S. 29.
[254] vgl. Spath et al. (2013), S. 57ff.
[255] vgl. Rost, Stölzel und Kozica (2016), S. 11f.
[256] vgl. Matuschek (2017), S. 88.
[257] vgl. Rost et al. (2016), S. 11f.
[258] vgl. Höpner (2019), S. 26.
[259] vgl. Matuschek (2017), S. 88.
[260] vgl. Sass (2019), S. 89.
[261] vgl. Hofert (2018), S. 12.
[262] vgl. Nowotny (2016), S. 74f.
[263] vgl. Mättig et al. (2018), S. 70f.
[264] vgl. Westkämper und Löffler (2016), S. 268.
[265] vgl. Defranceski (2017), S. 149.
[266] vgl. Zwirner (2004), S. 60f.

kation wesentlich dazu bei, dynamisch zu sein und auf Ungeplantes kurzfristig reagieren zu können.[267] Neben einer verbesserten Kommunikation und Abstimmung ist die Innovationsfähigkeit der Produktion und der Produktionsmitarbeiter eine zu untersuchende organisationale Anforderung.

4.1.2. Innovationsfähigkeit

Eine innovationsfähige Produktion benötigt eine innovationsoffene Organisation und verschiedene Rahmenbedingungen als Grundlage. Beispielsweise müssen die Ablauforganisation und die Herstellungsprozesse für Innovationen empfänglich sein.[268] Bei dieser Anforderung besteht die Aufgabe darin, die Produktionsorganisation so auszurichten, dass dem herrschenden Innovations- und Kreativitätsanspruch Rechnung getragen wird.[269] Im Anschluss soll geklärt werden, welche Voraussetzungen im Detail für die Schaffung einer Innovationsfähigkeit in der Produktion erforderlich sind.

Zum einen bezieht sich die Innovationsfähigkeit in der Produktion auf die Innovation von Produkten und zum anderen auf die Innovation von Prozessen,[270] Verfahren,[271] und Technologien. Es geht darum, Verbesserungen an existierenden Produkten zu erzielen und die Produktion und deren Prozesse systematisch zu hinterfragen und zu optimieren. In einer innovationsfähigen Produktion kommen wirkungsvolle Verbesserungen zustande und Mitarbeiter generieren Änderungsideen, die sie einbringen.[272] Sie werden direkt an Veränderungsprozessen beteiligt.[273] Die umfassende Zusammenarbeit zwischen Ingenieuren und Planern sowie den Facharbeitern in der Produktion ist essenziell. Die Mitarbeiter in der Produktion bieten mit kontinuierlichen Innovationsideen Unterstützung bei der fortlaufenden Weiterentwicklung an. Substanziell dafür sind bestimmte Erfahrungen, Fähigkeiten und Eigenschaften.[274]

Demzufolge ist es für eine Innovationsfähigkeit in der Produktion notwendig, die Innovationspotentiale der Mitarbeiter zu erschließen.[275] Der Fabrikarbeiter der Zukunft benötigt innovationsförderliche, persönliche und fachliche Kompetenzen.[276] Er hat die Fähigkeit zum Verständnis digitalisierter, automatisierter Prozesse und ist zu einer Problemanalyse und zu einer frühen Erkennung von Problemsituationen in der Lage.[277] Will die Produktion die Fähigkeit zur Innovation besitzen, ist eine gute Ausbildung der Mitarbeiter bezüglich Industrie 4.0 und Digitalisierung zwingend erforderlich.[278] Die Anforderung der Innovationsfähigkeit hängt somit maßgeblich mit der Anforderung der konsequenten Mitarbeiterqualifikation zusammen (siehe Kapitel 4.2.1).

Ein weiterer Baustein einer innovationsfähigen Produktion ist das systematische Lernen während der Arbeit und am Arbeitsplatz. Der Sinn dahinter ist es, die Bereitschaft zur Beteiligung an Innovationsprozessen in der Produktion zu intensivieren.[279] Auch die Förderung von Motivation und eine gezielte Personalentwicklung sind Wege zur Erhöhung dieser Bereitschaft, so dass das Produktionspersonal 4.0 eine konstruktive Mitgestaltungsrolle einnimmt.[280] Zusätzlich ist eine lernförderliche Arbeitsorganisation unerlässlich, damit die Produktion und die Mitarbeiter innovationsfähig sind.[281] Wenn alle sich bietenden Produktionsoptimierungspotenziale ausgeschöpft werden sollen, partizipieren alle Akteure der innovationsfähigen Produktion an Verbesserungen und Innovationshandlungen.[282] Zum Beispiel helfen sie dabei, Bearbeitungsverfahren zu perfektionieren. Die erhöhte Lernfähigkeit des Personals steigert die Wahrscheinlichkeit, dass Innovationen ausgedacht und im Anschluss umgesetzt werden können.[283]

Schlussendlich ist für das Erreichen einer Innovationsfähigkeit in der Produktion die Gestaltung der I4.0 Umgebung wesentlich. Diese muss so ausgestaltet sein, dass innnovationsfähige Arbeitsumgebungen mit entsprechender Arbeitsatmosphäre vorhanden sind, in denen Menschen die Entfaltung von Innovationsprozessen erleichtert wird. Zudem muss es möglich sein, diese Prozesse grenzüberschreitend an die richtigen Stellen weiterzugegeben.[284] Weiterhin besitzt die innovationsfähige Produktion eine organisatorische Innovationskompetenz. Das bedeutet, dass Einheiten und Personen Rechte zur eigenständigen Umsetzung von Verbesserungen ohne Abstimmung mit der Produktionsleitung zugestanden werden.[285] Laut der zugrundliegenden Literatur ist die Innovationsfähigkeit, genauso wie die moderne Führung, eine wesentliche organisationale Anforderung an die zukünftige Produktion.

4.1.3. Moderne Führung

Eine mehrwertstiftende Organisation mit flachen Hierarchien wird seit schon seit Drucker (1988) für vorstellbar gehalten.[286] Es stellt sich die Frage, ob ein solches organisatorisches Konzept in der Produktion funktioniert und den nächsten Weiterentwicklungsschritt der Produktion begleitet. Die Etablierung und Verwendung einer modernen Führung ist deshalb eine in dieser Arbeit zu untersuchende organisatorische Anforderung an die Produktion der Zukunft.

Industrie 4.0 trägt zu einem Dezentralisierungsschub und einem Abbau von Hierarchie in der Organisation bei.[287] Der

[267] vgl. Böhle und Bolte (2002), S. 43.
[268] vgl. Gausemeier (2007), S. 47.
[269] vgl. Zwirner (2004), S. 59.
[270] vgl. F. Hartmann (2017), S. 27.
[271] vgl. Gausemeier (2007), 47.
[272] vgl. MM MaschinenMarkt (2006), URL siehe Literaturverzeichnis.
[273] vgl. Soder (2017), S. 13.
[274] vgl. Schulze (2000), S. 120f.
[275] vgl. Gronau, Ullrich und Vladova (2015), S. 126.
[276] vgl. ebenda, S. 137f.
[277] vgl. Spath et al. (2013), S. 20f.
[278] vgl. Pfeiffer (2016), S. 40f.
[279] vgl. Stich et al. (2018), S. 164.
[280] vgl. Buhr (2015), S. 20f.
[281] vgl. E. Hartmann (2015), S. 17.
[282] vgl. MM MaschinenMarkt (2006), URL siehe Literaturverzeichnis.
[283] vgl. Wischmann und Hartmann (2018b), S. 2.
[284] vgl. Pfeiffer und Suphan (2018), S. 293f.
[285] vgl. MM MaschinenMarkt (2006), URL siehe Literaturverzeichnis.
[286] vgl. Drucker (1988), S. 6ff.
[287] vgl. Hirsch-Kreinsen und Weyer (2014), S. 21.

Begriff der „New Work“ beschreibt diese grundsätzliche Veränderung der Arbeitswelt.[288] Allerdings wird er bisher selten mit der Produktion in Verbindung gebracht. Es ist jedoch davon auszugehen, dass die Produktion ebenfalls von einer modifizierten Arbeitsstrukturierung profitiert. Inhalte des Konzeptes der „New Work“ sind eine selbstbestimmte Arbeit und eine motivierte, selbstlose Community. Der klassische, alles bestimmende Chef wird substituiert.[289] In der modern geführten Produktion ist trotz flacher Hierarchien eine eindeutige Zugehörigkeit existent. In diesem organisatorischen Modell stehen Vernetzung und Kooperation zunehmend im Fokus.[290]

Die modernen Führungskräfte evolvieren dazu, eher Coach als Entscheider zu sein und Aufgaben wie das Fördern von Teamarbeit und die Stärkung der Gemeinschaft treten in den Vordergrund.[291] Menschen sollen zusammengeführt werden, um einen regen Austausch unter Teammitgliedern voranzutreiben.[292] In Ergänzung dazu ist es eine moderne Führungsaufgabe die Grenzen in Teams deutlich zu machen und Freiheiten der Mitarbeiter zu bestimmen.[293] Führungskräfte müssen trotz ihres geringeren Führungsumfangs stets als Orientierungspfeiler dienen.[294] Sie müssen den Produktionsbeschäftigten mehr Vertrauen entgegenbringen und sich von der Vorstellung lösen, dass diese stets präsent sind (siehe Kapitel 4.2.2).[295]

In dieser Organisationsform ist die Führung generell hochgradig arbeitsteilig.[296] Das Aufgabenspektrum der menschlichen Führungskraft verkleinert und verändert sich durch die konsequente Delegation von Entscheidungen auf untere Ebenen. Es herrscht eine geteilte Führung mit aufgelockerten Strukturen.[297] Aufgabe des „Chefs 4.0“ ist es sicherzustellen, dass seine Mitarbeiter zu einer Selbstführung fähig sind, sodass diese nur Nuancen an Führung benötigen.[298] In einem Umfeld mit moderner Führung entstehen situationsbezogene Hierarchien und weitere Herausforderungen, wie beispielsweise Verantwortung für die aktuelle Organisation zu tragen und Rückmeldungen zu empfangen und zu geben. Diese fordernden Aspekte müssen von der Gemeinschaft gelöst werden, da es keinen allwissenden, eine Richtung vorgebendenden Chef in übergeordneter Position mehr gibt.[299] Mitarbeiter führen rotierend oder werden kurzzeitig in eine Führungsposition gewählt.[300] Die moderne Führung besitzt das Ziel, je nach Situation und Anspruch angepasst zu führen.[301] Die nachfolgende Abbildung zeigt die generellen Stufen und Formen von Führung übersichtlich auf (Abb. 5).

Demnach entwickelt sich Führung weg von einem autoritären Führungsbild. Die aufgelisteten Aspekte der anderen Führungsformen bis einschließlich der „Leaderless Organisation“ kommen für die moderne Führung in der Produktion ebenfalls infrage.

Die gemeinschaftliche Aufgabe der Führung kann nur mit geltenden Werten, Prinzipien und einem adäquaten Mindset aller Beteiligten funktionieren. Andernfalls ist es kompliziert eine moderne Führung zu bewerkstelligen.[302] Wenn eine wesentliche Eigenschaft wie die Kooperationsfähigkeit bei den Mitarbeitern fehlt, kann dies in Führungsproblemen enden.[303] Damit die moderne Führung in der Produktion anwendbar ist, benötigt es bei der Verteilung von Führung auf verschiedene Rollen und mehr Mitarbeiter gleichviel Führung wie es zuvor bei der traditionellen Führung der Fall war. Der Grad an Führung wird nicht verringert, sondern nur anders verteilt.[304] Durch diese flexible Arbeitsorganisation in Form einer veränderten Führung in der operativen Ebene, werden Teile bisher auf der Leitungsebene ausgeführter Steuerungs- und Planungsfunktionen auf untere Ebenen abgegeben. Dies geht mit einer dezentralen Selbstorganisation der Systeme und Mitarbeiter einher,[305] was in Verbindung zu moderner Führung die nächste Produktionsanforderung darstellt.

4.1.4. Selbstorganisation der Teams

Ein arbeitsorganisatorisches Konzept der Zukunft ist die Selbstorganisation der Teams auf dem Shopfloor. In einer mit modernen Strukturen ausgestatteten Produktion bekommen die Mitarbeiter grundlegend mehr Verantwortung und größere Freiheiten.[306] Das Konzept der Selbstorganisation beinhaltet, dass keine zentrale Kontroll- und Koordinationsinstanz mehr vorhanden ist.[307] Eine Veränderung ist die Reduzierung vertikaler Kommunikationsprozesse, da weniger Aufwand in Informationsweitergaben und Absprachen mit übergeordneten Ebenen investiert werden muss.[308] Die Anforderung der Selbstorganisation ist eng mit der modernen Führung verknüpft, da beide Konzepte eine Delegation von Verantwortung auf möglichst tiefe Ebenen umfassen.[309]

Insgesamt übernehmen selbstorganisierte Facharbeiter flexibel und eigenverantwortlich alle verschiedenen Rollen und Aufgaben, die in der Produktion anfallen.[310] Die organisatorische Einheit eines selbstorganisierten Teams erfüllt eine gemeinsame Fertigungsaufgabe, wobei der Handlungs- und Entscheidungsspielraum für alle Beteiligten groß ausfällt. Die Mitglieder der hierarchielosen Gruppe stimmen sich zu jeder Zeit direkt untereinander ab und steuern gemeinsam

[288] vgl. Hackl, Wagner, Attmer und Baumann (2017), S. 3.
[289] vgl. Krügl (2019), S. 236.
[290] vgl. Schwab (2016), S. 92.
[291] vgl. Schomburg, Sobieraj und Kruse (2016), S. 91.
[292] vgl. Liebermeister (2017), S. 193.
[293] vgl. Creusen, Gall und Hackl (2017), S. 105.
[294] vgl. Soder (2017), S. 21.
[295] vgl. Jeske et al. (2015), S. 29.
[296] vgl. Ittermann und Niehaus (2018), S. 52.
[297] vgl. Hofert (2018), S. 98f.
[298] vgl. Möllendorf (2018), S. 159 und Golob und Ullrich (2019), S. 91.
[299] vgl. Munsch (2020), S. 256-260.
[300] vgl. Werther und Bruckner (2018), S. 48.
[301] vgl. M. Berg (2019), S. 120f
[302] vgl. M. Berg (2019), S. 121.
[303] vgl. Hofert (2018), S. 98f.
[304] vgl. H. C. Werner und Gundrum (2019), S. 13.
[305] vgl. Hirsch-Kreinsen und Weyer (2014), S. 21.
[306] vgl. Klauß und Mierke (2017), S. 137.
[307] vgl. Kaltenecker (2018), S. 11.
[308] vgl. Strauß (2019), S. 509.
[309] vgl. ebenda, S. 504.
[310] vgl. ebenda, S. 504f.

Abbildung 5: Verschiedene Formen der Führung nach Strauß (2019).
Quelle: Strauß (2019), S. 506.

die ganze Gruppe sowie die Produktionsabläufe.[311] Daher betreffen selbstorgansierte Handlungen neben einzelnen Arbeitsschritten und Aufgaben organisatorische Veränderungen am System selbst.[312]

Innerhalb einer selbstorganisierten Gruppe hat jeder ein Auge auf das gesamte Arbeitsvorgehen und die Handlungen der anderen.[313] Das Personal arbeitet in automatisierten und intelligenten Prozessen, die zur Beherrschung der Komplexität ebenfalls selbstorganisatorisch gestaltet sind.[314] Zur optimalen Lösung von Arbeitsaufgaben und das Hineinwachsen in die selbstorganisatorische Funktion ist eine Qualifikation der Mitarbeiter unerlässlich (siehe Kapitel 4.2.1).[315]

Das Konzept der Selbstorganisation bietet Freiraum und Flexibilität. Von Mitarbeitern wird dabei Engagement und Initiative gefordert.[316] Zudem ist selbstorganisiertes Lernen unabdingbar.[317] Ein weiterer substanzieller Baustein sind regelmäßige Meetings zur Teamabsprache, denn das Team muss gut eingespielt sein.[318] Der Einsatz von agilen Projektmanagementmethoden und die stetige Visualisierung der Arbeit in der Produktion kann die Zusammenarbeit der selbstorganisierten Teams erleichtern.[319] Sie profitieren von aktuellen Daten, Kennzahlen und dem zeitgetreuen Bild des Produktionsstatus.[320] In Ergänzung dazu werden selbstorganisierte Produktionsmitarbeiter durch Begleitung und Coaching stetig unterstützt.[321]

Die Berücksichtigung fundamentaler Grundwerte wie Offenheit, Transparenz und Feedback hilft dabei, das Funktionieren des Konzeptes Selbstorganisation an sich zu ermöglichen.[322] Zusätzlich ist die Existenz eines klaren Regelwerkes, nach welchem die Mitarbeiter handeln, notwendig.[323] Als weiterer Teil des Rahmens für eine selbstorganisatorische Abwicklung der Produktion müssen Grenzen, Spielregeln und Werte definiert, eindeutig kommuniziert und verbindlich vereinbart werden.[324] Es sind noch einige weitere Festlegungen und Regeln der Zusammenarbeit notwendig, doch auf die genaue Gestaltung dieses Grundgerüsts für die Selbstorganisation soll hier nicht weiter eingegangen werden.[325]

Wenn eine Selbstorganisation der Teams trotz möglicher Hindernisse implementiert wird, ergeben sich temporäre und hochflexible Netzwerke in der Produktion.[326] Ob Unternehmen dieses Shopfloor Organisationskonzept für wichtig erachten ist eine der im weiteren Verlauf zu klärenden Fragen. Zunächst werden mit der prozessorientierten Arbeitsorganisation beginnend weitere Organisationsformen vorgestellt.

4.1.5. Prozessorientierte Arbeitsorganisation

Bei dieser Anforderung an die Produktion der Zukunft orientiert sich die Organisation konsequent am Prozess der Produktion[327] und an sämtlichen weiteren internen Unternehmensprozessen.[328] Die Prozessorientierung strebt eine Überwindung von Barrieren zu anderen Funktionseinheiten des Unternehmens an. Die gesamte Wertschöpfungskette soll konsequent optimiert werden und eine maximale Reduzierung von Schnittstellen stattfinden. Gleichzeitig wird bei dieser Organisationsform eine kleinstmögliche Durchlaufzeit in der Produktion anvisiert.[329] Ein wesentlicher Teilaspekt bei der Betrachtung der Wertschöpfungskette ist es, die Anschlussfähigkeit der eigenen Prozesse an die wichtigen vor-

[311] vgl. Kühl und Kullmann (1999), S. 8-11.
[312] vgl. Matuschek (2017), S. 88.
[313] vgl. Strauß (2019), S. 509f. zitiert nach: Strauß (1996).
[314] vgl. Banse (2017), S. 126.
[315] vgl. Weinreich (2016), S. 154.
[316] vgl. Seel (1995), S. 155.
[317] vgl. Krones, Strauch, Schütze und Müller (2015), S. 92.
[318] vgl. Kaltenecker (2018), S. 5f.
[319] vgl. Lindner und Amberg (2019), S. 33 und Weinreich (2016), S. 154.
[320] vgl. Weinreich (2016), S. 155.
[321] vgl. Geramanis (2020), S. 23.
[322] vgl. Kaltenecker (2018), S. 37.
[323] vgl. Franke (2019), S. 127.
[324] vgl. Weinreich (2016), S. 154.
[325] vgl. Weinreich (2016), S. 154f. und Geramanis (2020), S. 23.
[326] vgl. Ittermann und Niehaus (2018), S. 52.
[327] vgl. Ullrich und Vladova (2015), S. 58.
[328] vgl. C. M. Schlick et al. (2015), S. 53.
[329] vgl. Binner (2005), S. 27ff.

und nachgelagerten Geschäftsprozesse des Kunden sicherzustellen.[330] Was genau die Orientierung auf die internen und externen Unternehmensprozesse ausmacht, ist Gegenstand dieses Teilkapitels.

Das Organisationskonzept der Prozessorientierung benötigt Aspekte der Modernen Führung und der Selbstorganisation. So sind beispielsweise eine weitreichende Zusammenarbeit, Offenheit und mitdenkende Mitarbeiter erforderlich.[331] Bei der Prozessorientierung formiert sich das Personal in selbstgesteuerten, autark und autonom agierenden Teams.[332] Es wird nicht mehr alles von einer Stelle aus gemanagt und der Grad an direkter Führung reduziert sich.[333] Auch in diesem Konzept geht die Übergabe von Verantwortung auf verschiedene Ebenen mit höheren Kompetenzbedarfen und dezentralen Entscheidungen einher.[334]

Mithilfe dieses grundlegenden organisatorischen Aufbaus sollen die Aufgaben einer Prozessoptimierung und dauerhaften Prozesssicherung angegangen werden.[335] Es erfolgt wiederholt eine eingehende Analyse jeglicher Prozesse. Eine solche Analyse resultiert typischerweise in einer visuellen Bewertung, Neubildung und Verbesserung von Prozessen. Diese Kontrollaufgabe geschieht zum Beispiel über Prozessregelkreise, direktes Monitoring und Steuerung unmittelbar auf dem Shopfloor in der Produktion. Ein gewünschter Effekt ist es, die Produktivität in vielen Bereichen zu erhöhen.[336] Eine prozessorientierte Arbeitsorganisation richtet ihren Fokus zusätzlich auf die Existenz von Flexibilität. Deshalb besitzt die Organisation der prozessorientierten Produktion rasch anpassbare Prozesse und flexible Rahmenbedingungen der Arbeit.[337]

Prozessorientierte Organisationsstrukturen erfordern eine prozessbezogene Weiterbildung von I4.0 Fabrikarbeitern. Das bedeutet konkret: Schulungen und Fortbildungen berücksichtigen die speziellen Produktionsprozesse und technologischen Besonderheiten von Unternehmen.[338] Kommunikationsfähigkeit und Teamfähigkeit sind zwei Eigenschaften des Personals, die in einem prozessorientierten Umfeld definitiv stark ausgeprägt sein müssen und folglich ebenfalls in Schulungen trainiert und verbessert werden.[339] Auch hier ist zu erkennen, dass sich die Anforderungen an die Produktion der Zukunft zu einem gewissen Teil überschneiden und eine strikte Kategorisierung nicht gänzlich möglich ist. Es ist festzuhalten, dass eine übergreifende Mitarbeiterorientierung, motivierte Mitarbeiter und ein erfolgreiches Management mitentscheidend für das Funktionieren einer prozessorientierten Produktion sind.[340]

Zusammenfassend ist die konsequente Orientierung der Produktionsorganisation auf den Prozess laut der zugrunde liegenden Literatur durchaus eine wichtige Anforderung an die zukünftige Produktion. In den folgenden zwei Kapiteln wird theoretisch aufgezeigt, was eine agile und transparente Produktionsorganisation charakterisiert und inwiefern diese Konzepte dazu beitragen, die Produktion einen Schritt voran zu bringen.

4.1.6. Agilität der Organisation

Agile Organisationsstrukturen und agile Arbeitsweisen der Mitarbeiter können in der zukünftigen Industrie 4.0 Produktion eine entscheidende Rolle einnehmen.[341] Doch was ist Agilität überhaupt? Agil bedeutet beweglich zu sein. Das Thema Agilität taucht im Zusammenhang mit Organisation und Produktion unter dem Blickwinkel des „Agile Manufacturing“ auf.[342] Agilität bewirkt eine hilfreiche, zielgerichtete und schnelle Anpassung der Organisation in flexibler Art und Weise.[343] Diese Anpassung betrifft auch Prozesse und Strukturen auf dem Shopfloor, um veränderten Umweltsituationen und Gegebenheiten Rechnung tragen zu können.[344]

Eine Eigenschaft agiler Organisationen ist es Chancen, Risikofaktoren und Potentiale zu erkennen und jegliche Handlungen in der Produktion nach diesen Dingen auszurichten. Das Erzielen eines Lerneffektes aus Erfahrungen und Fehlern ist eine weitere essenzielle agile Eigenschaft.[345] Eine Herausforderung ist, dass agiles Arbeiten ein hohes Maß an Selbstführung abverlangt.[346] Die weitgehende Freiheit und der permanente Austausch in der agilen Arbeitsweise sind für Beschäftigte fordernd.[347] Folglich ist die Anforderung einer agilen Organisation eng mit der Anforderung selbstorganisierter Teams verknüpft.

Ergänzende maßgebende Werte für Agilität sind die Kundenorientierung der Organisation und die ständig fortschreitende Verbesserung der Organisation selbst. Ziel ist es, mit weniger Arbeit auf dem Shopfloor mehr Wert zu schaffen.[348] In einer agilen Organisation müssen die Produktionsmitarbeiter agile Haltungen, Fähigkeiten und Kompetenzen mitbringen. Auch die Motivation der Mitarbeiter steuert zu einer Sicherstellung von Agilität bei.[349] Weiterhin lebt eine agile Organisation von einem veränderten, modernen Führungsbild.[350] In diesem Führungsumfeld dient Agilität unter anderem schnellen Entscheidungsprozessen.[351] Eines der ersten expliziten Modelle zur Beschreibung von Agilität hat Parsons (1951) geschaffen. Neben der Adaptionsfähigkeit iden-

[330] vgl. Schatz und Bauernhansl (2017), S. 247.
[331] vgl. Binner (2005), S. 24.
[332] vgl. Kühl (2015), S. 64f.
[333] vgl. Baethge und Baethge-Kinsky (1998), S. 464f.
[334] vgl. Spath, Dworschak und Zaiser (2016), S. 689.
[335] vgl. Baethge und Baethge-Kinsky (1998), S. 465.
[336] vgl. Binner (2005), S. 42 und S. 1000f.
[337] vgl. Baethge und Baethge-Kinsky (1998), S. 466.
[338] vgl. Ullrich und Vladova (2015), S. 70.
[339] vgl. Dorsch-Schweizer und Schwarz (2007), S. 301f.
[340] vgl. Binner (2005), S. 1002f.
[341] vgl. von See und Kersten (2018), S. 8.
[342] vgl. Fischer und Köbler (2018), S. 93.
[343] vgl. M. Berg (2019), S. 120.
[344] vgl. Pfister und Müller (2019), S. 45.
[345] vgl. Hofert (2018), S. 13.
[346] vgl. Pfister und Müller (2019), S. 48.
[347] vgl. ebenda, S. 48f.
[348] vgl. Graf, Gramß und Edelkraut (2017), S. 31f. und Denning (2016), URL siehe Literaturverzeichnis.
[349] vgl. Graf et al. (2017), S. 39.
[350] vgl. Kaltenecker (2018), S. 14.
[351] vgl. Strauß (2019), S. 500.

tifiziert er die Zielverfolgung als Bestandteil von agilen Systemen. Parsons deklariert damit das Vermögen eines Systems, Ziele klar zu definieren und zu verfolgen. Als weitere agile Eigenschaften nennt der Autor die Fähigkeit eines Systems, grundlegende Wertmuster und Strukturen beizubehalten.[352] Wenn die heutige Organisation der Produktion den Anspruch hat agil zu sein, sollte eine Ausprägung dieser Grundzüge und Systemfertigkeiten erfolgen.

Die Produktion nimmt grundsätzlich an Technisierung zu und viele ineinandergreifende I4.0 Lösungen werden genutzt. Das Management von zunehmend mehr Technik muss mit einer geeigneten Methode bewerkstelligt werden, wobei eine agile Einstellung eine dieser Methoden zu sein vermag.[353] Betrachtet man die technische Seite, so ist erkennbar, dass agile Produktionssysteme in der Lage sind, übergreifende organisatorische Veränderungen vorzunehmen. Die agile Wandlungsfähigkeit unterstützt organisatorisch die Erfüllung unerwarteter, kurzfristiger Aufgaben, die außerhalb des normalen Produktionsspektrums liegen (Kapitel 4.3.3). In einer agilen Produktion ereignen sich Anpassungsreaktionen ohne Zeitverlust und Kostenexplosion. Die Anpassung zur Herstellung alternativer neuer Produkte erfolgt schnell.[354]

Ein Ziel der Produktion der Zukunft kann es also sein, den richtigen Grad an Agilität zu finden. Ein Schlüssel zum Erfolg lautet hierbei nicht maximal agil zu sein, sondern genauso agil wie es das Umfeld erfordert. Es ist vorwiegend richtiges Anpassen gefragt, denn fehlerhafte Anpassungen können nachteilige Wirkungen auf das Unternehmen und die Produktion induzieren.[355] Aktuell ist zur Verbesserung der Agilität in der Produktion oftmals die Wandlungsfähigkeit der Technik und der Betriebsmittel vordergründig. Jedoch bringt die beste agile Technik keinen Vorteil, wenn sie innerhalb einer nichtagilen Organisation voller Menschen eingesetzt wird, die weder Kompetenzen noch die angemessene Haltung für ein agiles Handeln mitbringen.[356] Der agilen Organisation kann somit der gleiche Stellenwert zugeschrieben werden, wie einer agilen Fertigung an sich. Nachfolgend richtet sich der Blick auf die Anforderung einer transparenten Produktionsorganisation.

4.1.7. Transparenz in der Organisation

Dieses Kapitel behandelt die letzte Anforderung aus der Kategorie „Organisation“. Die in Kapitel (4.1.3) veranschaulichte Entwicklung der „New Work“ beinhaltet einen umfangreichen Fokus auf Offenheit und Transparenz.[357] Dieses Konzept findet mittlerweile in der Produktion Anklang, weshalb Transparenz im Kontext von I4.0 für die Organisation und deren Entwicklung von Bedeutung ist.[358] Ein Element ist die permanente, transparente Darstellung der Organisation der Produktion für die Mitarbeiter und das Management. Was dies im Detail bedeutet folgt in diesem Teilkapitel.

Bei dieser Anforderung ist eine grundlegende Systemtransparenz unabdingbar.[359] Diese Transparenz lässt sich auf zwei näher zu betrachtende, aber trotzdem zusammengehörige Teilbereiche aufsplitten: Die Transparenz über die Produktionsarbeit und die Transparenz in der Produktionsorganisation. Durch die Transparenz in beiden Bereichen können den Beschäftigten klare Anweisungen erteilt werden.[360] Eine Transparenz in der Produktionsarbeit kann Fortschritte sichtbar machen und darstellen.[361] Dasselbe gilt für Zustandsdaten, die ebenfalls transparent sein sollen.[362] Ziel ist das Erreichen einer völligen Transparenz über den aktuellen und geplanten Zustand von sämtlichen Prozessvorgängen[363] und allen Aktivitäten in der Produktion.[364]

Das aktuell von Unternehmen weitläufig verwendete Instrument des Shopfloor Managements beabsichtigt mithilfe von Visualisierungstafeln Informationen standardisiert, aktuell und aussagekräftig aufzubereiten, so dass Abläufe für jeden Mitarbeiter transparent gemacht werden.[365] Dieser Leitgedanke wird durch die Digitalisierung in der zukünftigen Produktion bestätigt und weiterentwickelt. Dazu werden relevante Informationen an ausgewählten Stellen im Produktionsprozess bereitgestellt. Mittel hierfür sind beispielsweise Assistenzsysteme, Touch-Monitore, Datenbrillen und Smartwatches.[366] Transparenz in jegliche Abläufe zu bringen bewirkt generell eine Verbesserung der grenzüberschreitenden Zusammenarbeit von organisatorischen Einheiten in Unternehmen.[367] Es lässt sich schlussfolgern, dass die Produktion der Zukunft von dieser Wirkung profitieren kann.

Des Weiteren ist die Abbildung der Organisation als digitaler Zwilling eine Möglichkeit Transparenz in der Produktionsorganisation herbeizuführen. In diesem „Digital Twin of Organisation“ kann jeder virtuell die Verrichtung von Tätigkeiten nachverfolgen und sehen, welche Workflows wohin verlaufen.[368] Industrie 4.0 benötigt als Grundlage nicht nur ein digitales Abbild von Objekten innerhalb der Produktion. Eine digitale Abbildung der Organisation ist ebenfalls unerlässlich, da dies unter anderem zu einer optimaleren und umfassenderen Prozessplanung beiträgt. In diesem digitalen Zwilling wird jede Organisationsveränderung in Echtzeit widergespiegelt. Dies bezieht sich sowohl auf unternehmensinterne Prozesse als auch auf externe Prozesse.[369]

Am Ende enthält das digitale Abbild die gesamte Organisation inklusive ihrer Eigenschaften. Die Präsentation für Beschäftigte erfolgt über ausgewählte IT-Lösungen. Darüber

[352] vgl. Fischer und Köbler (2018), S. 93 zitiert nach: Parsons (1951).
[353] vgl. Klauß und Mierke (2017), S. 137f.
[354] vgl. Elkins, Huang und Alden (2004), S. 204f.
[355] vgl. Fischer und Köbler (2018), S. 94f
[356] vgl. ebenda, S. 96.
[357] vgl. Leopold (2019), S. 247.
[358] vgl. Stief (2018), S. 115.
[359] vgl. Hirsch-Kreinsen (2018), S. 189.
[360] vgl. Korge (2017), S. 508f.
[361] vgl. T. Becker (2017), S. 107.
[362] vgl. Schlund und Pokorni (2017), S. 150.
[363] vgl. Plattform Industrie 4.0 (2015), S. 22.
[364] vgl. Nowotny (2016), S. 364.
[365] vgl. Soder (2017), S. 12.
[366] vgl. Tawalbeh, Weißflog und Hopf (2019), S. 16f.
[367] vgl. Spath (2017), S. 6f.
[368] vgl. Dangl (2019), S. 59.
[369] vgl. ebenda, S. 60.

hinaus ist eine Transparenz in der Organisation durch ein digitales Abbild für die Durchführung von Simulationen dienlich. Verschiedene Szenarien, Arbeitsvorgänge, Veränderungen und Entscheidungen können innerhalb des digitalen Abbilds durchexerziert werden.[370]

Eine weitere Komponente einer transparenten Produktionsorganisation ist die Durchsichtigkeit in der Verantwortungs- und Aufgabenverteilung. Es herrscht eine Zunahme an organisationaler Transparenz, wenn geleistete Arbeitsstunden unmittelbar veranschaulicht werden.[371] Die Abbildung der Zuordnung von Mitarbeitern zu Tätigkeitsorten (Mitarbeiterbelegung) sorgt für einen schnellen Überblick über das Personal, ermöglicht einen optimierten Einsatz von Ressourcen und führt Transparenz herbei.[372] Zudem sehen Mitarbeiter die Resultate der eigenen Arbeit direkt.[373] Es wird eine Transparenz über die Arbeitsintensität in Echtzeit ermöglicht.[374] Beispielhaft für Mitarbeitertransparenz ist außerdem, dass jederzeit kontrollierbar ist, welche Person verantwortlich für einen Fehler ist. Der Mitarbeiter wird „gläsern“ und kontrollierbar.[375]

Transparenz dient generell als Basis für Entscheidungen. Diese können sicherer getroffen werden, je mehr Informationen vorliegen.[376] Aufgrund dessen ist die zielgruppenspezifische Aufbereitung von Informationen wichtig.[377] In einer transparenten Organisation müssen Kommunikationsverläufe und Kommunikationsergebnisse für alle Organisationsmitglieder zugänglich sein. Weiterhin kann durch eine Transparenz ein Lernen der gesamten Organisation erzielt werden.[378] Es lässt sich schlussfolgern, dass diese Grundsätze auch auf die Shopfloor Organisation übertragbar sind. Insgesamt wird der organisationalen Anforderung einer transparenten Organisation in der Produktions- und Industrie 4.0 Literatur eher weniger Beachtung geschenkt. Ein weitaus intensiveres behandeltes Thema ist die konsequente Mitarbeiterqualifikation in der Produktion.

4.2. Anforderungen der Kategorie „Mensch“

4.2.1. Konsequente Mitarbeiterqualifikation

Die Weiterentwicklung des Produktionspersonals für die veränderten Aufgaben bei der Anwendung von Industrie 4.0 ist eine der größten Herausforderungen für alle Unternehmen.[379] Die Arbeit wird anspruchsvoller und nimmt an Komplexität und Vernetzung zu.[380] Die einführenden Kapitel bezüglich der Rolle des Menschen und zur Verknüpfung von Organisation und Industrie 4.0 reißen hinführend an, dass das System der zukünftigen Produktion Fachkräfte mit einem breit aufgestellten Qualifikationsprofil benötigt. Die Fachkräfte agieren als steuernde und entscheidungsmächtige Experten.[381] Sie sind dafür verantwortlich die I4.0 Umgebung zu lenken und bei Störungen einzugreifen. In der zukünftigen Produktion gibt es weniger reine Facharbeiter.[382] Das Management der Produktion wird von hochqualifizierten Mitarbeitern betrieben.[383] Diese Anforderung erfasst die nötigen Fähigkeiten und mögliche Qualifizierungswege im Detail. Generell enthält das Verständnis von Kompetenz die beiden Teilaspekte Wissen und Qualifikation.[384] Wissen schließt Fertigkeiten mit ein.[385] Der Begriff Kompetenz umfasst somit neben den fachlichen Schlüsselqualifikationen übergreifende Fähigkeiten.[386] In einer I4.0 Umgebung ist der Ausbau mehrerer Kompetenzen notwendig. Die Einnahme der Überwachungsrolle verlangt ein hohes Komplexitätsverständnis.[387] Es ist Fachkenntnis in Bezug auf den Umgang mit innovativen I4.0 Technologien gefordert.[388] Ein generelles Verständnis für Interaktionen von Maschinen,[389] die Fähigkeit zur Filterung und Analyse von Prozessinformationen[390] und die Fähigkeit zur schnellen Einarbeitung in unbekannte Abläufe sind in der Produktionsarbeit zunehmend notwendige Aspekte.[391] Wichtige personale Kompetenzen sind ein ganzheitliches Denken,[392] analytische Fähigkeiten[393] und eine eigenständige Lösung auftretender Probleme.[394]

Neue Konzepte der Arbeitsorganisation verlangen überfachliche Kompetenzen und soziale Attribute.[395] Es benötigt Kreativität,[396] Kooperationsfähigkeit,[397] Teamfähigkeit[398] und Kommunikationsfähigkeit.[399] Ein selbstbestimmtes Handeln,[400] die Eignung zur Selbstorganisation[401] und Innovationsoffenheit sind gleichfalls wichtige Eigenschaften.[402] Industrie 4.0 Fabrikarbeiter müssen fähig sein, formell und informell zu lernen.[403] Die Bereitschaft zum lebenslangen Lernen nimmt eine Schlüsselposition ein.[404]

[370] vgl. Dangl (2019), S. 59.
[371] vgl. T. Becker (2017), S. 107
[372] vgl. Dichtl und Patermann (2017), S. 210.
[373] vgl. Wischmann und Hartmann (2018b), S. 2.
[374] vgl. Butollo und Engel (2015), S. 35.
[375] vgl. Windelband und Dworschak (2018), S. 71.
[376] vgl. Kletti (2015), S. 3.
[377] vgl. Lanza et al. (2018), S. 40f.
[378] vgl. Probst und Büchel (1994), S. 21.
[379] vgl. Killian, Nendel, Markert und Riedel (2019), S. 53 und Lorenz et al. (2016), S. 3.
[380] vgl. Matuschek (2017), S. 88.
[381] vgl. Windelband und Dworschak (2018), S. 77.
[382] vgl. Spath, Dworschak, Zaiser und Kremer (2015), S. 117f.
[383] vgl. Matuschek (2017), S. 88.
[384] vgl. Heyse (2018), S. 46.
[385] vgl. M. Becker (2005), S. 4
[386] vgl. Heyse (2018), S. 47.
[387] vgl. Graf et al. (2017), S. 108.
[388] vgl. C. M. Schlick et al. (2015), S. 53 und Krones et al. (2015), S. 89.
[389] vgl. Ahrens und Spöttl (2018), S. 190.
[390] vgl. Graf et al. (2017), S. 108.
[391] vgl. Kagermann und Winter (2017), S. 30.
[392] vgl. Rost et al. (2016), S. 20.
[393] vgl. Matuschek (2017), S. 89 und Rost et al. (2016), S. 16.
[394] vgl. Stich et al. (2018), S. 159.
[395] vgl. Klauß und Mierke (2017), S. 137 und Krones et al. (2015), S. 89.
[396] vgl. Frey und Osborne (2013), S. 45 und Heidel et al. (2017), S. 11.
[397] vgl. Klauß und Mierke (2017), S. 137, Rost et al. (2016), S. 18 und Stich et al. (2018), S. 159.
[398] vgl. Kagermann und Winter (2017), S. 30.
[399] vgl. Kagermann und Winter (2017) und Stich et al. (2018), S. 159.
[400] vgl. Stich et al. (2018), S. 159.
[401] vgl. Kagermann und Winter (2017), S. 30.
[402] vgl. Rost et al. (2016), S. 23.
[403] vgl. Stich et al. (2018), S. 159.
[404] vgl. Schlund und Pokorni (2017), S. 139 und Müller und Jentsch (2015), S. 107.

Mit der in Zukunft wesentlichen Methodenkompetenz wird ein Verständnis über die Zusammenhänge von operativen Kennzahlen angesprochen.[405] Ein weiterer Baustein der Qualifizierung ist der umfangreiche Aufbau von IT-Kompetenzen.[406] Diese enthalten ein grundlegendes Verständnis für Daten und deren Analyse.[407] Wissen in Bezug auf die Arbeit in virtuellen Umgebungen ist essenziell.[408] Der IT-Kompetenzaufbau bezieht sich nicht nur auf die Fabrikarbeiter, sondern auf Mitarbeiter in verknüpften Unternehmensbereichen, die produktionsbezogene Wissens- und Planungsarbeit vollbringen. Folglich sind IT-Fähigkeiten zukünftig nicht nur in der direkten Wertschöpfung von hoher Bedeutung.[409] Genaueres zu wichtigen Daten- und Softwarefähigkeiten der Angestellten in der Industrie 4.0 ist im Anhang zu diesem Kapitel zu finden (Anh. 3).

Aktuell ist die Aus- und Weiterbildung der eigenen Belegschaft der vielversprechendste Weg der Mitarbeiterqualifikation, da noch selten fertig ausgebildete I4.0 Fachkräfte am Arbeitsmarkt zu finden sind.[410] Die verschiedenen Qualifizierungsmaßnahmen bauen systematisch auf das Erfahrungswissen der Fabrikarbeiter auf.[411] Ein Durchführungsweg der Mitarbeiterqualifikation ist die arbeitsplatznahe Qualifikation.[412] Bei der „On the Job" Qualifizierung erfolgt das Lernen in kleinen Einheiten und richtet sich nach dem individuellen Leistungsniveau.[413] Für diese Lehrmethode werden Assistenzsysteme lernförderlich gestaltet und Arbeitsvorgänge, Aufgaben und Prozessschritte visuell abgebildet.[414] Es ergibt sich ein zielgerichtetes Lernen als Training mit Präsenz und direkten Arbeitserklärungen.[415]

Die I4.0 Qualifizierungsmaßnahmen Coaching und Mentoring zielen auf eine individuelle Kompetenzförderung ab. Zudem können Unternehmen ihre Mitarbeiter über Workshops, Seminare und Informationsveranstaltungen schulen.[416] Diese nicht direkt an Arbeitsplätzen stattfindenden Fortbildungen bilden die Gruppe der „Off the Job" Maßnahmen. „Near the Job" Maßnahmen werden in Lehrwerkstätten vollzogen.[417] Beispielsweise bieten Forschungsinstitutionen in Lernfabriken verschiedene Demonstratoren an, die zur Schulung genutzt werden können.[418] Auch mit einzelnen Methodenräumen und Arbeitsstationen erfolgt dort eine I4.0 Weiterbildung.[419]

Neben den Fortbildungsmethoden spielt in Zukunft die grundlegende Ausbildung von Fachkräften eine zentrale Rolle. In der Produktionsarbeit ergeben sich drei Qualifizierungsstufen mit unterschiedlichen Kompetenzanforderungen. Auf der obersten Ebene sind akademische Grade vorhanden. Es folgen die mittleren Fachkräfte mit Ausbildungen und Fortbildungsabschlüssen (z.B. Meister). Auf der untersten Schicht befinden sich Ungelernte und einfache Werker.[420] Die klassischen Produktionsberufe (z.B. Industriemechaniker) werden durch neuartige Qualifikationsprofile wie die des Produktionstechnologen, Applikationsexperten oder Prozessmanagers ersetzt.[421] Aufgaben bezüglich der Produktionsorganisation werden von allen Qualifikationsebenen durchgeführt.[422] Schlussendlich sind die Anforderungen der zukünftigen Produktion ohne eine konsequente Mitarbeiterqualifikation auf keinen Fall zu bewerkstelligen.[423]

4.2.2. Flexibler Personaleinsatz

Das Konzept des flexiblen Personaleinsatzes steht aktuell noch weniger stark im Fokus. Ein flexibler Arbeitskräfteeinsatz ist eine Möglichkeit der Organisationsgestaltung, die unter anderem zur Bewältigung der steigenden Bedarfsabhängigkeit der Produktion beiträgt.[424] Je nach Bedarf und Situation steigen oder fallen die Anforderungen in der Produktion erheblich. Die Kapazitätsbedarfe der Produktion in Bezug auf das Personal variieren nicht nur wöchentlich, sondern oftmals auch sehr kurzfristig während eines Tages,[425] oder von Stunde zu Stunde. Bei einer solchen Konfrontation mit Schwankungen jeglicher Art müssen Unternehmen dazu in der Lage sein, immer die richtigen und passenden Kapazitäten bereitzustellen. Ein starres System versagt an dieser Stelle, weshalb ein flexibles Personaleinsatzsystem verlangt wird.[426]

Der erste Baustein des flexiblen Personaleinsatzes ist die Reaktion auf unvorhergesehene Kapazitätsbedarfsschwankungen. Diese werden durch veränderte Markt-, Produktions-, Kunden-, und Mitarbeiteranforderungen sowie Auftrags-[427] und Absatzschwankungen hervorgerufen.[428] Zudem ist eine aktive und tagesaktuelle Steuerung der Personalkapazität Inhalt des Konzeptes, damit jederzeit die genau passende Anzahl an Mitarbeitern zur Verfügung steht. Schichtanpassungen, Schichtbesetzungen und Arbeitszeiten sollen durch eine digitale, intelligente Kommunikation via Smartphones, Applikationen und IT-Systemen besprochen und abgestimmt werden. Dies ermöglicht eine Selbstbestimmung und Selbststeuerung der Mitarbeiter.[429] Sie entschei-

[405] vgl. Deuse, Achenbach und Lenze (2015), S. 251.
[406] vgl. Matuschek (2017), S. 83f.
[407] vgl. Ahrens und Spöttl (2018), S. 190.
[408] vgl. Kagermann und Winter (2017), S. 30.
[409] vgl. Lorenz et al. (2016), S. 6.
[410] vgl. Ullrich und Vladova (2015), S. 57.
[411] vgl. Dumitrescu und Marquardt (2017), S. 649f.
[412] vgl. Matuschek (2017), S. 88.
[413] vgl. Kagermann (2014), S. 244.
[414] vgl. Senderek (2018), S. 94f.
[415] vgl. Kagermann und Winter (2017), S. 30f.
[416] vgl. Schlund und Pokorni (2017), S. 152 und IHK Darmstadt Service GmbH (2020), URL siehe Literaturverzeichnis.
[417] vgl. Ullrich und Vladova (2015), S. 67.
[418] vgl. IHK Darmstadt Service GmbH (2020), URL siehe Literaturverzeichnis.
[419] vgl. Stich et al. (2018), S. 156.
[420] vgl. Spath et al. (2015), S. 116.
[421] vgl. Borch und Zinke (2008), S. 44f.
[422] vgl. C. M. Schlick et al. (2015), S. 40.
[423] vgl. ebenda, S. 54.
[424] vgl. Kempter, Jost und Künz (2020), S. 229.
[425] vgl. Spath et al. (2013), S. 5f.
[426] vgl. ebenda, S. 71ff.
[427] vgl. Fraunhofer-Institut für Arbeitswirtschaft und Organisation (2020d), S. 8f., URL siehe Literaturverzeichnis.
[428] vgl. Spath et al. (2013), S. 72.
[429] vgl. Fraunhofer-Institut für Arbeitswirtschaft und Organisation (2020d), S. 9, URL siehe Literaturverzeichnis.

den eigenständig an welchem Ort und zu welcher Zeit ihr Arbeitseinsatz stattfindet. Durch dieses Modell der Arbeitsorganisation kommt es zu einer hohen Flexibilisierung von Arbeitszeiten. Die Verantwortung des Personals steigt.[430] Der Aufwand der Personaleinsatzplanung wird mithilfe dieser selbstständigen Organisationsübernahme durch das Personal maßgeblich reduziert.[431]

Für den flexiblen Personaleinsatz können verschiedene Methoden eingesetzt werden, wozu auch die Verleihung von Mitarbeitern zählt. Mitarbeiter werden in verschiedenen Organisationseinheiten[432] oder an anderen Produktionslinien eingesetzt.[433] Die Existenz dieser Einsatzflexibilität ist ein bedeutender Faktor. Zum Konzept des flexiblen Personaleinsatzes gehört neben der räumlichen und aufgabentechnischen Flexibilität auch der Austausch von Mitarbeitern über mehrere Standorte hinweg.[434] Es muss eine Eignung für diesen flexiblen Einsatz vorhanden sein (Mitarbeiterqualifikation).[435] Grundsätzlich ist der flexible Personaleinsatz eng mit der Selbstorganisation der Teams verknüpft. Die Produktionsbeschäftigten arbeiten in ortsunabhängigen und vernetzten Teams sowie in (internationalen) Projekten. Eine Anwesenheit im Betrieb ist bei vielen Tätigkeiten nicht mehr zwingend erforderlich und die Arbeit erfolgt weniger in getakteten Schichtlängen.[436]

Zusätzlich stehen Fragen der „Work-Life-Balance“ im Vordergrund.[437] Darunter fallen Überlegungen, das Privatleben der Produktionsmitarbeiter in Einklang mit deren Arbeitsleben zu bringen.[438] Es ergeben sich Verbesserungen für viele Mitarbeiter, da es möglich wird, je nach individueller Lebenssituation, am besten passende Einsatzorte und Einsatzzeiten auszuwählen. Somit werden persönliche Bedürfnisse berücksichtigt.[439] In Ergänzung dazu finden in dieser Form der Arbeitsorganisation flexible Beschäftigungsverhältnisse Anwendung.[440] Bei nachteiligen Schwankungen ermöglicht es ein flexibler Personaleinsatz korrigierende Handlungen in Form von Arbeitszeitverringerungen und Leiharbeit einzuleiten. Der generelle Einsatz von Teilzeit, erweiterten Arbeitszeitkonten und Gleitzeit in der Produktion dient außerdem einer schnellen Personalkapazitätsanpassung. Bei kurzzeitig erhöhten Bedarfen können Zusatzschichten und Überstunden geleistet werden. Prinzipiell sind auch Vertrauensarbeitszeit, variable Bezahlung, Zeitarbeit und Stellen auf Minijob Basis denkbare und praktikable Mittel der zukünftigen Produktion.[441]

Bei einem flexiblen Einsatz von Produktionspersonal werden verschiedene Assistenzsysteme zur Verbesserung der Personaleinsatzplanung genutzt. Dabei werden individuelle Qualifikationen, Verfügbarkeiten, bevorzugte Arbeitszeiten und auch ergonomische Aspekte berücksichtigt.[442] Eine Intention der Verwendung eines flexiblen Personaleinsatzes ist es, die Flexibilität von Produktionstechnik besser auszuschöpfen. Eine personalseitige Flexibilität kann eine technologische Flexibilität unterstützen (Kapitel 4.3.3).[443] An all diesen Punkten zeigen sich die vielseitigen Ausgestaltungsmöglichkeiten des flexiblen Personaleinsatzes. Fortführend widmet sich das nächste Teilkapitel der Motivation von Mitarbeitern.

4.2.3. Förderung der Mitarbeitermotivation

Die letzte Anforderung an eine zukünftige Produktion in der Teilkategorie „Mensch“ ist die Förderung der Mitarbeitermotivation. Motivation beinhaltet die Bereitschaft und die Beweggründe für bestimmte Verhaltensweisen und Handlungen.[444] Für die Motivation von Produktionspersonal bedeutet diese allgemeine Definition, dass in jeder Situation die richtigen Handlungsanreize vorhanden sind und das richtige Verhalten zur Erledigung der Arbeitsaufgaben zum Vorschein kommt.[445]

Das Personal hat innerhalb der Produktion ein hohes Leistungslevel zu erreichen. Mit einer personalseitigen hohen Motivation zur Leistungserbringung kann diesem Anspruch Rechnung getragen werden. Betrachtet man die Organisation der Produktion der Zukunft, so sind der dauerhafte Erhalt und die stetige Förderung der Mitarbeitermotivation zur Festigung organisatorischer Strukturen zu diskutieren.[446] Das Motivationslevel muss aufrechterhalten werden, weil umfangreiche Kompetenzen zur Erfüllung von fordernden Organisationsaufgaben notwendig sind.[447] Obwohl die Arbeit von erhöhter Komplexität geprägt sein wird, sind die Arbeiter wesentlich freier in dem was sie tun. Arbeitgeber müssen deshalb dafür sorgen, die Motivation und den Antrieb der Mitarbeiter hoch zu halten.[448] Unter diesen Gesichtspunkten ist die individuelle Mitarbeiterqualifikation eng mit der Mitarbeitermotivation verwurzelt. Ist dem Mitarbeiter der Nutzen seiner Arbeit bewusst[449] und weiß er um seine Eignung die jeweilige Aufgabe zu bewältigen, ist er grundsätzlich mit einer motivierteren Arbeitshaltung zugange.[450] Sind die Auswirkungen der eigenen Arbeit und die Wichtigkeit des eigenen Arbeitsplatzes bekannt, hat dies gleichermaßen einen positiven Einfluss auf die Motivation.

Allein schon die Produktionsorganisation und ihr Aufbau an sich sind motivationsbeeinflussend. Durch Unsicherhei-

[430] vgl. Niesing (2013), S. 12f.
[431] vgl. Spath et al. (2013), S. 85.
[432] vgl. ebenda, S. 82.
[433] vgl. Senderek (2018), S. 92.
[434] vgl. Spath et al. (2013), S. 77.
[435] vgl. Holtgrewe et al. (2015), S. 27.
[436] vgl. Jeske et al. (2015), S. 29.
[437] vgl. Jeske et al. (2015), S. 29.
[438] vgl. Spath et al. (2013), S. 73.
[439] vgl. Zimmermann (2017), S. 61.
[440] vgl. Matuschek (2017), S. 87.
[441] vgl. Fraunhofer-Institut für Arbeitswirtschaft und Organisation (2020c), S. 2, URL siehe Literaturverzeichnis.
[442] vgl. Börkircher et al. (2018), S. 13.
[443] vgl. Holtgrewe et al. (2015), S. 27.
[444] vgl. Schröder (2019), S. 11.
[445] vgl. Adami und Houben (2008), S. 86.
[446] vgl. ebenda, S. 84f.
[447] vgl. Graf et al. (2017), S. 119.
[448] vgl. F. Becker (2019), S. 2
[449] vgl. Graf et al. (2017), S. 119.
[450] vgl. Rosenstiel (2007), S. 57.

ten, ein ungenügendes Verständnis der Organisation sowie die nicht vorhandene Kenntnis von Entscheidungswegen und Entscheidungskriterien fällt die Motivation geringer aus.[451] Für die Erreichung einer hohen Mitarbeitermotivation muss dem entgegengewirkt werden. Zusätzlich schlägt sich eine erhöhte Motivation in dem Faktor Arbeitsbereitschaft nieder.[452] Je nach Motivation und Einstellung differiert die Arbeitsleistung,[453] die Geschwindigkeit der Arbeit, oder das Verhalten während der Schichtübergabe und in den Pausen. Der zweite Faktor zur Beschreibung von Effekten der Motivation ist die Einsatzbereitschaft. Sie zeigt sich exemplarisch im Grad der Bereitschaft zur Verantwortungsübernahme[454] und im kontinuierlichen Anstreben von Verbesserungen durch Mitarbeiter.[455] Positive Wirkungen von Mitarbeitermotivation zeigen sich in einer Reduzierung von Unfallzahlen und Fehlverhalten.[456] Der dritte Teilbereich der Motivation ist die Lernbereitschaft.[457] Zum einen betrifft dies die eigenen Lernprozesse.[458] Zum anderen zeigt sich Lernbereitschaft in der Wahrnehmung von Qualifikationsmaßnahmen und deren anhaltendem Erfolg.[459]

Eine Möglichkeit zur Beeinflussung und Steuerung von Mitarbeitermotivation ist es, verschiedene Anreizsysteme und Belohnungskonzepte zu schaffen. Es können positive Anreize Anwendung finden, jedoch sind auch negative Anreizfaktoren zum Beispiel in Form von monetären Einbußen denkbar, auch wenn hier die praktische Handhabung eher schwierig erscheint. Soziale Anreizsysteme, die Reputation oder Leistung messen und vergleichen, sind weitere denkbare Mittel, um Mitarbeiter in der Produktionsumgebung zu motivieren.[460] Die Gewährung von Arbeitsplatzsicherheit ist ein weiterer Motivator für Mitarbeiter.[461] Zufriedenheit und Motivation sind des Weiteren durch einen hohen Teamzusammenhalt,[462] eine verbesserte Kommunikation und eine ausgedehnte Kooperation zu erreichen.[463] Attraktivität zwecks Mitarbeitermotivation kann durch die Veränderung des Arbeitsumfeldes geschaffen werden.[464] Es ist wichtig, dass Mitarbeiter in die Gestaltung der Arbeitsumgebung einbezogen werden, da sie dadurch motivierter sind und Veränderungen leichter annehmen. Motivationsfördernde Maßnahmen betreffen zum Beispiel Zeitpläne und die Arbeitszeitverteilung.[465]

Inhalte der Konzepte der modernen Führung und der Selbstorganisation stellen wesentliche Komponenten eines solchen Arbeitsumfeldes mit motivationssteigernden Effekten dar.[466] Durch die Einräumung von Autonomie und Verantwortlichkeiten entwickelt sich die Motivation positiv.[467] Die Motivation von Mitarbeitern hat ergänzend zentrale Einflüsse auf die Kennzahlen zur Steuerung der Produktion und kann selbst als Kennzahl herangezogen werden.[468] Die Förderung der Mitarbeitermotivation ist letztendlich eine wichtige, zu diskutierende, menschenbezogene Anforderung an die zukünftige Produktion.

4.3. Anforderungen der Kategorie „Technologie“

4.3.1. Ausrichtung auf Lean Management 4.0

Die Ausrichtung der Produktion auf Lean Management in Kombination mit der Veredelung durch Industrie 4.0 ist eine technologische Anforderung an die Produktion der Zukunft. Schlanke Prozesse und Industrie 4.0 widersprechen sich nicht, sondern können sich sinnstiftend ergänzen.[469] Bei einer Umgestaltung der Produktion zur „Smart Factory“ werden existierende schlanke, prozesssichere Produktionssysteme digital aufgewertet.[470] Der eindeutige Fokus dieser Anforderung liegt auf der Darstellung der Möglichkeiten zur Anpassung von Lean Management durch I4.0 Konzepte. Die grundlegenden Prinzipien des Lean Managements werden deshalb nur grob erläutert.

Eine schlanke Produktion verfolgt das Ziel einer exzellenten Qualität und Produktivität.[471] Die wesentlichste Säule des Lean Ansatzes ist die Übertragung von Verantwortung für die permanente Verbesserung von Prozessen auf die Mitarbeiter in der Produktion. Verschwendungsvermeidung, eine geringe Fehlerrate und ein optimaler Material- und Arbeitseinsatz sind weitere Bestandteile des Lean Managements. Mitarbeiter sollen als eigenständige, systematische Problemlöser agieren.[472] Basis des „Toyota Production Systems“ sind standardisierte, stabile und beherrschbare Prozesse. Das „Jidoka“ Element des Systems fixiert sich darauf, Fehler und Defekte einzugrenzen und diese rapide zu eskalieren. Ein visuelles Management dient dazu, Abweichungen von gesetzten Standards aufzuspüren. Die „Just in Time“ Produktion ist eine weitere Systemkomponente.[473] Neben den Prinzipien „One Piece Flow“, „Kanban“ und „Ordnung und Sauberkeit“ gibt es noch weitere Lean Werkzeuge, auf welche an dieser Stelle nicht weiter eingegangen werden kann.[474] Doch wie sind diese Lean Ansätze mit neuartigen digitalen Möglichkeiten in Einklang zu bringen? Die Kombination von Industrie 4.0 und Lean Management wird als Lean 4.0 bezeichnet. Industrie 4.0 bietet das Potential die Anforderungen an eine Lean Management Produktion hinsichtlich null

[451] vgl. Adami und Houben (2008), S. 88f.
[452] vgl. ebenda, S. 85.
[453] vgl. Ullrich und Vladova (2015), S. 58 und F. Becker (2019), S. 2
[454] vgl. Adami und Houben (2008), S. 85.
[455] vgl. Antoni (2017), S. 164.
[456] vgl. F. Becker (2019), S. 2.
[457] vgl. Adami und Houben (2008), S. 85.
[458] vgl. Strauß (2019), S. 509.
[459] vgl. Adami und Houben (2008), S. 85.
[460] vgl. Schuh, Fabry, Schmitz-Urban und Siegers (2015), S. 651.
[461] vgl. Adami und Houben (2008), S. 89.
[462] vgl. F. Becker (2019), S. 94f.
[463] vgl. Soder (2017), S. 12f.
[464] vgl. Adami und Houben (2008), S. 84f.
[465] vgl. Schröder (2019), S. 30.
[466] vgl. F. Becker (2019), S. 103 und Sass (2019), S. 10f.
[467] vgl. Orpen (1997), S. 193 und S. 204.
[468] vgl. Stich et al. (2018), S. 177.
[469] vgl. Metternich, Adolph et al. (2017), S. 193.
[470] vgl. ebenda, S. 196.
[471] vgl. Meudt, Hartmann und Metternich (2018), S. 6.
[472] vgl. Metternich, Adolph et al. (2017), S. 193-196.
[473] vgl. Meudt et al. (2018), S. 7f. zitiert nach: Ohno (1988).
[474] vgl. Fischer und Köbler (2018), S. 53.

Fehler und einer hohen Prozessqualität zu erfüllen.[475] Wenn eine Ausrichtung auf Lean 4.0 vollzogen wird, ergeben sich schlanke Wertschöpfungsketten, die durch innovative Organisationskonzepte und Technologien geformt und unterstützt werden.[476]

Bei der Verwendung von I4.0 Ansätzen wird Mitarbeitern durch die technologische Unterstützung vergleichsweise der Druck einer schnellen Problemlösung und unmittelbaren Analyseerkenntnis genommen.[477] Zusätzlich wird mit der Förderung einer Selbstoptimierung der Produktion durch I4.0 Technologie der Leitgedanke einer stetigen Verbesserung durch reaktive Mitarbeiter angegangen. Die Lean Philosophie des Strebens nach jederzeitigem Wissen über Material und Ort wird in der Industrie 4.0 durch Echtzeitdaten maßgeblich erleichtert.[478] Durch Datenkorrelationen können Ursachen für Probleme und Abweichungen in dem realen Prozess gefunden werden.[479] Das schlanke Produktionssystem wird durch die Informationsbereitstellung in Echtzeit über den ganzen Wertstrom hinweg organisatorisch erleichtert und optimiert.[480] Die Umsetzung der Lean Philosophie einer hohen Produktionsprozessqualität erfolgt mithilfe des hohen Grades an Überwachung durch Sensoren, sowie durch intelligente Produkte und Maschinen.[481]

Die Verbindung von Lean und Industrie 4.0 hegt resümierend sehr große Synergiepotenziale[482] und Produktivitätspotentiale.[483] Ergänzende Chancen durch Lean 4.0 sind exemplarisch das optimale Auslasten von Montagelinien, die frühzeitige Erkennung von Engpässen, die Verringerung von Durchlaufzeiten,[484] ein besserer Informationsfluss und eine hohe Transparenz.[485] Nachfolgend sind ergänzende Punkte zur Integration und Kombination der beiden Ansätze Lean Management und Industrie 4.0 abgebildet (Abb. 6). Zusätzliche Informationen sind bei Meudt et al. (2018) zu finden.[486]

Eine Lean Management Produktion ist eine notwendige organisatorische Basis für die Einführung von Industrie 4.0.[487] Ist diese Basis vor der Implementierung von technischen I4.0 Lösungen vorhanden, liegt ein gutes, abgestimmtes, fehlerfreies Prozessgerüst zugrunde. Dadurch wird die Gefahr minimiert, eine problembehaftete Produktion mit I4.0 Lösungen zu überdecken, welche dann nicht wie geplant funktionieren können.[488] Neben der Ausrichtung der Produktion auf Lean 4.0 ist eine ausgiebige Vernetzung und Transparenz eine Anforderung an die Organisation der zukünftigen Produktion.

[475] vgl. Glück (2015), S. 39f.
[476] vgl. Neuhaus (2018), S. 55f.
[477] vgl. Losch (2016), URL siehe Literaturverzeichnis.
[478] vgl. Meudt et al. (2018), S. 9.
[479] vgl. Metternich, Müller, Meudt und Schaede (2017), S. 346.
[480] vgl. Meudt et al. (2018), S. 9.
[481] vgl. Kaufmann (2015), S. 27f.
[482] vgl. Goschy und Metternich (2016), S. 13.
[483] vgl. Lanza und Nyhuis (2016), S. 25.
[484] vgl. Meudt et al. (2018), S. 14ff.
[485] vgl. ebenda, S. 18f.
[486] Ebenda.
[487] vgl. Fischer und Köbler (2018), S. 53.
[488] vgl. Jeske et al. (2015), S. 30.

4.3.2. Vernetzung und Transparenz

Die Vernetzung in der industriellen Fertigung steigt stetig an.[489] Im Kontext der Produktion bezieht sich Vernetzung auf alle wissensbezogenen, organisatorischen und technischen Abläufe zwischen verschiedenen Leistungseinheiten sowie auf eine Kooperation von allen Beteiligten. Die Vernetzung begrenzt sich nicht nur auf interne Bereiche, sondern überschreitet Unternehmensgrenzen.[490] Partner, Kunden, Lieferanten und Berater fungieren infolge der Vernetzung als interne Unternehmensteile.[491] Der Begriff einer horizontalen Vernetzung wird für die Beschreibung der Vernetzung innerhalb einer Unternehmensebene und über Unternehmensgrenzen hinweg verwendet, während die vertikale Vernetzung die Ebenen übergreifende Vernetzung im Unternehmen definiert.[492]

Eine prinzipielle Zielsetzung der Vernetzung in der Produktion ist die Installation von kurzen Wegen in Bezug auf physische Bewegungen, Wissen und Information, um zeitnah auf Störungen und Abweichungen jeder Größe zu reagieren.[493] Die Vernetzung des gesamtheitlichen Produktionsprozesses ermöglicht eine transparente Darstellung des Produktionszustandes im Sinne eines Echtzeitabbildes. Dank dieser Transparenz ist es möglich, jederzeit in die mit Echtzeitinformationen abgebildete Wertschöpfung einzugreifen und diese zu steuern. Die Durchgängigkeit von IT-Systemen ist eine der Voraussetzungen dafür, dass Produktionsstandorte, Anlagen, Maschinen, Computer und Datenbanken digital vernetzt sein können und miteinander interagieren.[494] Durch die Verkettung von jeglichen CPS werden eine Vielzahl an Daten gewonnen, sodass im Anschluss intelligente Geräte Informationen für Beschäftigte liefern.[495] So können mobile Geräte in einer „Smart Factory“ genutzt und zusätzlich in das vernetzte System integriert werden, wodurch beispielsweise auch Mobilfunkgeräte, Laptops oder Tablets ihren Teil zu der umfassenden Vernetzung beitragen. Die verbesserte Transparenz durch die Vernetzung vieler verschiedener Komponenten unterstützt die Mitarbeiter bei ihren Aufgaben in der Produktion. Sie erhalten aktuelle Zustands- und Prozessinformationen, welche es ihnen erlauben, jederzeit unmittelbar zu kommunizieren und operative oder managementbezogene Aktionen einzuleiten.[496]

Transparenz führt dazu, dass zu jeder Zeit Orte und Bewegungen von allen möglichen Dingen (z.B. Mitarbeiter, Teile, Lagerbehälter) in der Produktion verfolgt werden können. Bei einer starken Vernetzung und Transparenz wird der Materialfluss bis außerhalb der Fabrik verfolgt. Es werden Verbrauchsdaten und Prozessparameter gesammelt, analysiert und zuverlässig geplant.[497] Die durchgängige Lieferung von

[489] vgl. Aßmann (2017), S. 321.
[490] vgl. Westkämper und Löffler (2016), S. 128.
[491] vgl. Cole (2017), S. 60.
[492] vgl. Pawellek und Schirrmann (2017), S. 121.
[493] vgl. Westkämper und Löffler (2016), S. 131.
[494] vgl. Pawellek und Schirrmann (2017), S. 121.
[495] vgl. Geisberger und Broy (2012), S. 124.
[496] vgl. Westkämper und Löffler (2016), S. 173.
[497] vgl. Hoffmann (2017), S. 58f.

Abbildung 6: Integration der Digitalisierung in den klassischen Lean Ansatz nach Meudt et al. (2018).
Quelle: Meudt et al. (2018), S. 10.

Informationen zu Prozessen dient der Dokumentation und sichert eine transparente Ressourcennutzung.[498] Zusammenfassend schafft eine Verwendung smarter Komponenten zur Abbildung des gesamten Betriebs und dessen Zustands eine hohe Transparenz und beabsichtigt eine weitläufige Optimierung von Abläufen, Kosten und einzelnen Komponenten der Produktion.[499] Die Transparenz wird benötigt um Effizienz zu sichern und eine stetige Rückmeldung aus der Produktion zu erhalten. Dadurch wird zudem die Steuerung von Kennzahlen unterstützt.[500]

Ein ergänzender Schwerpunkt von Vernetzung und Transparenz ist es, das gesamte Management der Produktion realitätsnah zu vollziehen. Im Wesentlichen sollen Entscheidungen immer auf Basis der richtigen Informations- und Datengrundlage getroffen werden und somit Verbesserungen bewirken. Ein allumfassendes Systemmonitoring unterstützt zügige Reaktionen auf Störungen und Abweichungen. Zeitraumbeobachtungen fördern in diesem vernetzten, transparenten System die Gewinnung von Erkenntnissen zum Verhalten der Produktion. Somit können Planungen detaillierter ausfallen.[501] Ein weiteres Ziel ist es, Beschäftigte aus unterschiedlichen Qualifikationsebenen lernfördernd zu verbinden.[502]

Eine umfangreiche Vernetzung und Transparenz kann, wie in diesem Kapitel aufgezeigt, einige positive Wirkungen haben. Schlussendlich ist die vernetzte und transparente Produktion eine bereits länger diskutierte technologische Anforderung, die von Industrie 4.0 affinen Produktionsunternehmen anvisiert werden kann. Dasselbe trifft auch auf die nächste zur Debatte stehende Anforderung zu. Die Wandlungsfähigkeit der Produktion ist die dritte organisationale Anforderung in der Kategorie „Technologie“.

4.3.3. Wandlungsfähigkeit der Produktion

Die Wandlungsfähigkeit ist das Vermögen einer Fabrik, mit niedrigem Aufwand aktiv strukturelle Veränderungen zu vollziehen, wenn interne und externe Auslöser dies erforderlich machen.[503] Eine kontinuierliche Anpassung wird erst durch anpassbare Gegebenheiten, Strukturen und Prozesse ermöglicht.[504] Der überwiegend automatisierte, schnelle physische Umbau der Fertigung und die Anpassung des Produktionslayouts sind typisch für eine wandlungsfähige Fabrik.[505] Anpassungsmaßnahmen können beispielsweise Kapazitäten und Produktionsvolumina betreffen,[506] aber auch die Durchführung veränderter Aufträge und die Herstellung vielfältiger, andersartig konfigurierter Produkte sind eine Stärke der wandlungsfähigen Produktion.[507]

Modularität, Mobilität, Skalierbarkeit, Kompatibilität und die Universalität stellen relevante Elemente zur Beschreibung einer wandlungsfähigen Produktion dar. Durch eine Veränderung dieser als „Wandlungsbefähiger“ bezeichneten Faktoren, lassen sich Flexibilitätskorridore bedarfsweise reaktiv oder proaktiv verschieben. Die Fähigkeit zu dieser Verschiebung ist als Wandlungsfähigkeit einer Fabrik definiert (Abb. 7).[508]

Flexibilität beinhaltet die Erhöhung oder Erniedrigung von Anforderungen innerhalb eines Korridors, beispielswei-

[498] vgl. Westkämper und Löffler (2016), S. 161.
[499] vgl. Lucke, Defranceski und Adolf (2017), S. 81.
[500] vgl. Kletti (2015), S. 3f.
[501] vgl. Westkämper und Löffler (2016), S. 170f.
[502] vgl. Ittermann und Niehaus (2018), S. 52.
[503] vgl. Wiendahl (2002), S. 123.
[504] vgl. Westkämper (1999), S. 131.
[505] vgl. Arbeitsgruppe Forschung und Innovation der Plattform Industrie 4.0 (2016), S. 12.
[506] vgl. Das (2001), S. 4154.
[507] vgl. Dobrowolski (2016), URL siehe Literaturverzeichnis.
[508] vgl. Heinen, Rimpau und Wörn (2008), S. 25ff.

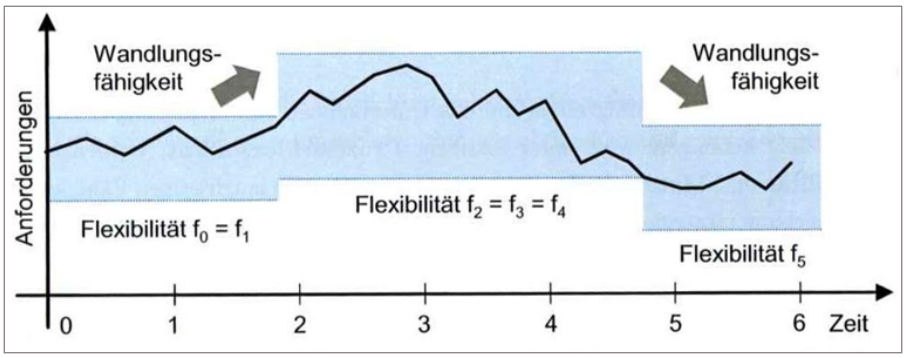


Abbildung 7: Abgrenzung von Flexibilität und Wandlungsfähigkeit (Originalquelle Zäh und Reinhart (2005) adaptiert von Heinen et al. (2008)).
Quellen: Zäh und Reinhart (2005), S. 4 und Heinen et al. (2008), S. 25.

se in Bezug auf Qualität oder Menge. Mithilfe von Wandlungsfähigkeit kann zu bestimmten Zeitpunkten eine Erweiterung oder Schrumpfung dieser Flexibilität anvisiert werden.[509] Wenn sich Leistungen außerhalb des Flexibilitätskorridors befinden, sind diese allein mit Flexibilität nicht ausreichend zu erfüllen. Bei diesem Zustand ist die Wandlungsfähigkeit eines Produktionssystems gefragt. Jedoch ist diese Wandlungsfähigkeit durch zeitliche und finanzielle Grenzen limitiert.[510] Die Flexibilität stellt sowohl für die wandlungsfähige Produktion als auch für die nächste Stufe der agilen Produktion eine essenzielle Voraussetzung dar (siehe Anh. 4).[511] Außerdem greift die Anlagen- und Kapazitätsflexibilität eine flexible Justierung bezüglich Losgrößen, Variantenvielfalt und Prozessschrittreihenfolge auf.[512] Flexibilität ist demnach also ein „Rüstzeug“ der Produktion.[513] Im Endeffekt ist es ein Ziel, die flexible Anpassung so kostengünstig wie möglich durchzuführen.[514]

Die Modularität als Industrie 4.0 Komponente der Wandlungsfähigkeit ist der modulare Aufbau der Produktion mit interoperablen und intelligenten Modulen. Aufgrund von Schnittstellen und einer hohen Kompatibilität der Module sind unterschiedliche Konfigurationen und funktionsspezifische Kombinationen verschiedener Anlagen und Maschinen realisierbar.[515] Die separaten Einheiten des Fertigungssystems sind standardisiert und einfach auszuwechseln. Auf der softwaretechnischen Seite ermöglicht die Kompatibilität eine Vernetzung aller Anlagen und Einheiten. Dadurch wird ein verknüpfender Austausch zwischen diesen garantiert. Der Parameter Universalität kennzeichnet, in welchem Maße Anpassungsfähigkeiten von Arbeitern und Maschinen vorliegen.[516] Weitere Elemente sind die Mobilität als unbegrenzte räumliche Bewegbarkeit aller Dinge in der Produktion sowie die Skalierbarkeit als das Vermögen von Objekten zur technologischen und räumlichen Erweiterung und Reduzierung.[517]

Eine wandlungsfähige Produktion lässt sich einfach und schnell je nach den Rahmenbedingungen, wie zum Beispiel der Marktsituation und den aktuellen Kundenwünschen, umbauen.[518] Diese Rahmenbedingungen vermögen auch rechtlicher und gesellschaftlicher Natur zu sein.[519] Ein auf solche äußere Faktoren reagierender Umbau der Produktion erfolgt selbstorganisiert und reibungslos.[520] Zur Verdeutlichung der Thematik ist es beispielsweise eine konkrete wandlungsfähige Maßnahme einen Maschinenausfall aufzufangen, indem eine selbstentscheidende Umleitung von Materialen auf andere Anlagen stattfindet.[521] Relevante Zielgrößen, die mit einer wandlungsfähigen Produktion intendiert werden, sind die Erhöhung der Gesamtauslastung der Fertigung und die Steigerung der Lieferfähigkeit.[522]

Insgesamt sind bei der Wandlungsfähigkeit menschliche Fähigkeiten mitentscheidend, denn der Mensch prägt den Wandel der Produktion durch seine intelligente Anschauung und Kreativität maßgeblich mit.[523] Wandlungsfähige Systeme stoßen an ihre Grenzen, wenn die Integration und Einbeziehung der Mitarbeiter, die den Wandel vollziehen und sich in den vielzählig veränderten Strukturen orientieren müssen, ausbleibt. Diesbezüglich wird an dieser Stelle auf einen ergänzenden Beitrag in der Zeitschrift für wirtschaftlichen Fabrikbetrieb von Heinen, Hirsch und Nyhuis (2009) hingewiesen. Dieser zeigt fortführend auf, inwiefern Wandlungsfähigkeit mit Mitarbeiterkompetenz verknüpft ist.[524]

Die wandlungsfähige Produktion der Zukunft ist nicht nur flexibel, sie ist zur Selbstoptimierung geeignet, fähig zu

[509] vgl. Zäh und Reinhart (2005), S. 4.
[510] vgl. Upton (1994), S. 73f.
[511] vgl. Wiendahl (2002), S. 126.
[512] vgl. Stief (2018), S. 100.
[513] vgl. Sharma und Ali (2010), S. 210.
[514] vgl. Gibson et al. (2008), S. 714.
[515] vgl. Arbeitsgruppe Forschung und Innovation der Plattform Industrie 4.0 (2016), S. 12f.
[516] vgl. Heinen et al. (2008), S. 25ff.
[517] vgl. Steegmüller und Zürn (2017), S. 29.
[518] vgl. Arbeitsgruppe Forschung und Innovation der Plattform Industrie 4.0 (2016), S. 12.
[519] vgl. Westkämper (1999), S. 131f.
[520] vgl. Dumitrescu und Marquardt (2017), S. 641.
[521] vgl. Stief (2018), S. 80.
[522] vgl. Arbeitsgruppe Forschung und Innovation der Plattform Industrie 4.0 (2016), S. 12f.
[523] vgl. Westkämper, Zahn, Balve und Tilebein (2000), S. 25f.
[524] Heinen et al. (2009).

lernen und passt sich jederzeit dem rasanten technologischen Wandel an.[525] Industrie 4.0 Technologien und Lösungen tragen in diesem Szenario dazu bei, die Wandlungsfähigkeit zu verbessern.[526] Mit Blick auf den I4.0 Bereich betrifft die Flexibilität als Vorstufe einer wandlungsfähigen Produktion neuartige Komponenten, wie etwa die schnelle und flexible Darstellung von Daten und Informationen auf Datenbrillen, Bildschirmen, Tablets oder sonstigen Wearables.[527] Die reale Fabrik wird in eine I4.0 Fabrik transformiert, um die ständige Adaption der Produktion zu begünstigen und zu optimieren.[528] Schlussendlich sollen die beschriebenen Aspekte und Bestandteile der Wandlungsfähigkeit durch Industrie 4.0 unterstützt und verbessert werden. Es ist eine Anforderung an die Produktion der Zukunft wandlungsfähig zu sein, doch die Förderung der Ressourcenschonung und Nachhaltigkeit ist laut der zugrundeliegenden Literatur eine gleichfalls bedeutsame technisch-organisatorische Anforderung.

4.3.4. Förderung der Ressourcenschonung und Nachhaltigkeit

In den vergangenen Jahren hat sich die Nachhaltigkeit in der Produktion zu einer wichtigen Strategie entwickelt. In einer Welt, in der globale Erwärmung, Klimawandel und Umweltverschmutzung stark diskutierte Themen sind, müssen Unternehmen Antworten und Lösungen finden.[529] Die Rohstoffverknappung ist ein Faktor der Unternehmen zum Umdenken anleitet,[530] da dieser mit Preissteigerungen einher geht.[531] Nachhaltigkeit wird von Unternehmen angetrieben und in vielen Fällen von Kundenseite gewünscht.[532]

Die Nachhaltigkeit im Kontext der Produktion beruht auf mehreren Säulen.[533] Die ökonomische Nachhaltigkeit stellt Finanzaspekte in den Vordergrund. Infrastruktur, Gebäude und Betriebsmittel zählen als finanzielle Ressourcen, bei denen ein kostensparender und nachhaltiger Ansatz verfolgt werden kann.[534] Der Auftrag der sozialen Nachhaltigkeit ist es, mitarbeiterbezogene Aspekte nachhaltig auszurichten. Hier geht es beispielsweise um die Erhaltung und Schaffung von Arbeitsplätzen, einen fairen Umgang mit dem Personal und eine soziale Chancengleichheit. Zu dieser Säule gehört zudem die nachhaltige Gestaltung der Produktionsarbeitsplätze.[535] Die dritte Säule der Nachhaltigkeit zielt auf ein ökologisch verantwortliches Handeln ab.[536] Zu dieser Form der Nachhaltigkeit gehört ein ausschweifender Blick auf die Umwelt,[537] zum Beispiel in Form der Verwendung erneuerbarer Energiequellen und umweltfreundlicher Ressourcen. Eine auf ökologische Nachhaltigkeit achtende Herstellungskette ist komplett umweltfreundlich konzipiert und lässt so wenig Abfall wie möglich zurück.[538]

Fortführend setzt sich eine nachhaltige Produktion mit der Herkunft, dem Transport, der Verarbeitung und der Entsorgung von Ressourcen auseinander. Anlagen und Maschinen müssen dazu fähig sein möglichst umweltschonend zu arbeiten.[539] Innerhalb der Produktionsprozesse sind zu einem großen Anteil Materialspar- und Effizienzpotentiale vorhanden, die genutzt werden können. Wenn Ausschuss und Überproduktion auf einem geringen Niveau sind, deutet dies auf eine effiziente und nachhaltige Produktion hin. Grundlage für die Ressourcenschonung und die effiziente Ressourcennutzung ist der produktive Ressourceneinsatz nach dem Maximalprinzip, bei dem mit gegebenem Input ein möglichst hoher Output realisiert wird. Ressourcenschonung betrifft den sparsamen Umgang mit Ressourcen. Rohstoffe,[540] Zeit, Personal, Arbeit,[541] Fernwärme, Gas[542] und Wasser sind beispielhafte Ressourcenarten, deren Einsparung lohnend sein kann.[543]

Die Energieeffizienz ist ein weiteres wichtiges Merkmal einer nachhaltigen Produktion.[544] Sämtliche Produktionsanlagen sollen so energieeffizient wie möglich arbeiten.[545] Industrie 4.0 Anwendungen können bei einer Energiesteuerung assistieren, indem beispielsweise Energiedaten der Produktion gesammelt und ausgelesen werden.[546] Das Ermitteln von Schwachstellen und Energiesparpotenzialen durch eine vernetzte Steuerungstechnik und Maschinen in Echtzeit ist eine Möglichkeit Einfluss auf die Energieeffizienz zu nehmen.[547] Die intelligente Nutzung von oftmals ungenutzter Abwärme oder überschüssiger Energie kann den Fabrikbetrieb gleichfalls nachhaltig prägen.[548] Die „Smart Factory“ der Zukunft besitzt durch eine intelligente Kommunikation mit der ganzen Fabrik und den Gebäuden weiteres Potential zur Einsparung von Energie, wobei die Gebäudesteuerung bezüglich Heizung, Klima und Lüftung sowie der Ausgleich von Energiespitzen Angriffspunkte für eine Energieverbrauchsregelung darstellen.[549]

Anhand dieser Sachverhalte lässt sich erkennen, dass Industrie 4.0 und Nachhaltigkeit durch gemeinsame Werte prinzipiell eng miteinander verbunden sind. Alle in Kapitel (4.) dargelegten organisationalen Anforderungen an die zukünftige Produktion sind laut des aktuellen Forschungsstandes von Bedeutung. Die Untersuchung der Relevanz

[525] vgl. Dobrowolski (2016), URL siehe Literaturverzeichnis.
[526] vgl. Lanza et al. (2018), S. 15.
[527] vgl. Stief (2018), S. 80f.
[528] vgl. Riffelmacher, Kluge und Westkämper (2009), S. 30.
[529] vgl. Westkämper und Löffler (2016), S. 202f.
[530] vgl. Ramsauer (2013), S. 9.
[531] vgl. ebenda, S. 12.
[532] vgl. Stief (2018), S. 79.
[533] vgl. Westkämper und Löffler (2016), S. 203.
[534] vgl. Ramsauer (2013), S. 9.
[535] vgl. Bimos KG (2020b), URL siehe Literaturverzeichnis.
[536] vgl. Wiendahl, Reichardt und Nyhuis (2014), S. 148.
[537] vgl. Stief (2018), S. 79f.
[538] vgl. Bimos KG (2020b), URL siehe Literaturverzeichnis.
[539] vgl. ebenda.
[540] vgl. Ramsauer (2013), S. 9f.
[541] vgl. Stief (2018), S. 87.
[542] vgl. Abele und Reinhart (2011), S. 168.
[543] vgl. Manzei (2016), S. 12.
[544] vgl. Laub (2013), S. 6.
[545] vgl. Abele und Reinhart (2011), S. 168.
[546] vgl. Stief (2018), S. 79.
[547] vgl. Laub (2013), S. 6.
[548] vgl. Abele und Reinhart (2011), S. 168.
[549] vgl. Ramsauer (2013), S. 8.

dieser Produktionsanforderungen aus den Kategorien „Organisation“, „Mensch“ und „Technologie“ in der Praxis ist ein Bestandteil der nachfolgenden quantitativen Orientierungsbefragung.

5. Quantitative Orientierungsbefragung zur Ermittlung der Bedeutung von Kennzahlen und organisationalen Anforderungen der zukünftigen Produktion

Die Kapitel (2.) bis (4.) zeigen die aktuelle Situation der Produktionsarbeit und wesentliche vorstellbare Zukunftskonzepte auf. Die vorliegende Studie intendiert ein umfassendes Bild der Wichtigkeit der theoretisch erklärten Steuerungskennzahlen und Produktionsanforderungen abzuzeichnen. Konkret sollen folgende Forschungsfragen beantwortet werden: Welches sind die fünf zentralen Kennzahlen zur Steuerung der zukünftigen Produktion und wie setzt sich die Gruppe der für Unternehmen bedeutsamsten Produktionsanforderungen zusammen? Hierfür erfolgt zunächst die Beschreibung der Studienmethodik und des Vorgehens bei der Durchführung der Datenerhebung, bevor auf wesentliche Inhalte und Erkenntnisse eingegangen wird. Eine konkrete Zielsetzung ist es, dadurch den Forschungsstand zu erweitern und eine Orientierung zu schaffen, in welchen organisatorischen Rahmen Industrie 4.0 eingebettet werden sollte. Es erfolgt auch eine Integration von teilweise technologiebezogenen Kennzahlen und Produktionsanforderungen in die Studie, um einen Vergleich zu den organisations- und menschenbezogenen Komponenten zu gewährleisten.

5.1. Konzeption und Methodik der Studie

Innerhalb dieses Abschnitts werden alle genutzten Methoden zur Findung von Antworten auf die Forschungsfragen mithilfe der Studie reflektiert. Grundsätzlich wird bei einer empirischen Überprüfung von theoretischen Aussagen und Konzepten eine Unterscheidung zwischen quantitativen und qualitativen Verfahren vorgenommen.[550] Zur Überprüfung der theoretischen Überlegungen wird in dieser Studie eine quantitative Vorgehensweise verwendet,[551] da sich diese Art der Befragung für eine breite Erfassung möglichst vieler Erfahrungen, Meinungen und Einstellungen eignet und eine Darstellung von Zusammenhängen in konkreten Zahlenwerten ermöglicht.[552] Ergänzend handelt es sich hier aufgrund einer relativ überschaubaren Stichprobengröße und der Zielsetzung eine grundsätzliche Orientierung zu liefern um eine Orientierungsbefragung. Außerdem wurde die quantitative Studie als standardisierte schriftliche Befragung mit einem Papierfragebogen durchgeführt.[553] Eine Messung von Einstellungen kann entweder explizit oder implizit erfolgen. Bei der Verwendung von expliziten Einstellungsmaßen werden die Befragten direkt darum gebeten, über ihre Einstellung nachzudenken und diese im Anschluss zu berichten.[554] Daher wurde die Einstellung der Probanden zu zukünftig wichtigen Kennzahlen und Anforderungen in dieser Arbeit explizit erhoben.

Beschreibung der Stichprobe: Allgemein bildet eine Stichprobe ein verkleinertes, strukturell ähnliches Abbild einer Grundgesamtheit ab. Im weiteren Verlauf werden die Stichprobenelemente mit „n“ angegeben.[555] Die Studie erfolgt als Stichprobenziehung, da eine Erfassung der Grundgesamtheit in diesem Fall keineswegs möglich ist. Zudem ist eine Vollerhebung aus zeitlichen, monetären, aufwandstechnischen und organisatorischen Gründen nicht sinnvoll.[556] Eine Teilerhebung, wie sie hier durchgeführt wurde, hat zudem in vielen Fällen den Vorteil qualitativ besser zu sein.[557] Die dem Autor zugängliche Umfragestichprobe ist grundsätzlich eine bewusste Auswahl von Personen, die Kenntnisse und Erfahrungen im Bereich Produktion und Industrie 4.0 besitzen. Die Teilnehmer haben in ihren Funktionen in verschiedenen Unternehmen und Forschungs- oder Bildungseinrichtungen mit diesem Themengebiet zu tun. Die Auswahl der Umfrageteilnehmer erfolgte vorausgewählt über das Firmen- und Kontaktnetzwerk des Fraunhofer IAO. Details hierzu folgen in der Studiendurchführung, jedoch muss an dieser Stelle bereits erwähnt werden, dass die Stichprobenziehung innerhalb dieser Vorauswahl als zufällig beschrieben werden kann. Es fand demnach keine explizite Auswahl nach bestimmten soziodemographischen Merkmalen statt (Quotenauswahl).[558]

In dieser Befragung wurde eine relativ kleine Stichprobe von circa 50 Probanden angestrebt. Aufgrund der Tatsache, dass sich die Gewinnung von Studienteilnehmern als schwierig herausstellte, ergab sich letztendlich eine niedrigere Zahl an Probanden (n=34). Die tatsächliche Stichprobengröße unterscheidet sich weiterhin leicht in den einzelnen Teilen der Umfrage. Außerdem stammen die Probanden aus 24 in ihrer Größe variierenden Einrichtungen. Die exakte Aufteilung ist in Abbildung 8 dargestellt.

Fast 70 Prozent aller befragten Unternehmen und Institutionen lassen sich der Kategorie der Großunternehmen zuordnen, während 20 Prozent der Probanden in mittelgroßen Einrichtungen arbeiten. Kleine Unternehmen sind die am seltensten angesprochene Größenkategorie (12%). Abbildung 9 veranschaulicht in welchen Branchen die befragten Unternehmen und Institutionen tätig sind. Die Teilnehmergruppe der Befragung ist über einige Branchen gestreut, weshalb ein repräsentatives Bild aus verschiedenen Wertschöpfungsketten vorhanden ist. An dieser Stelle war es nicht möglich Angaben in Prozent zu machen, da viele der Einrichtungen mehreren Branchen zuzuordnen sind.

Die am stärksten vertretenen Gruppen sind die Metall- und Elektroverarbeitung (N=11) und die Branche der produzierenden Unternehmen (N=9). Viele Unternehmen stam-

[550] vgl. Mayer (2009), S. 15.
[551] vgl. Brosius, Haas und Koschel (2016), S. 1.
[552] vgl. ebenda, S. 3ff.
[553] vgl. Mayer (2009), S. 59.
[554] vgl. Haddock und Maio (2014), S. 213.
[555] vgl. Brosius et al. (2016), S. 61.
[556] vgl. Kromrey, Roose und Strübing (2016), S. 254f.
[557] vgl. Schumann (2006), S. 82.
[558] vgl. Brosius et al. (2016), S. 67ff.

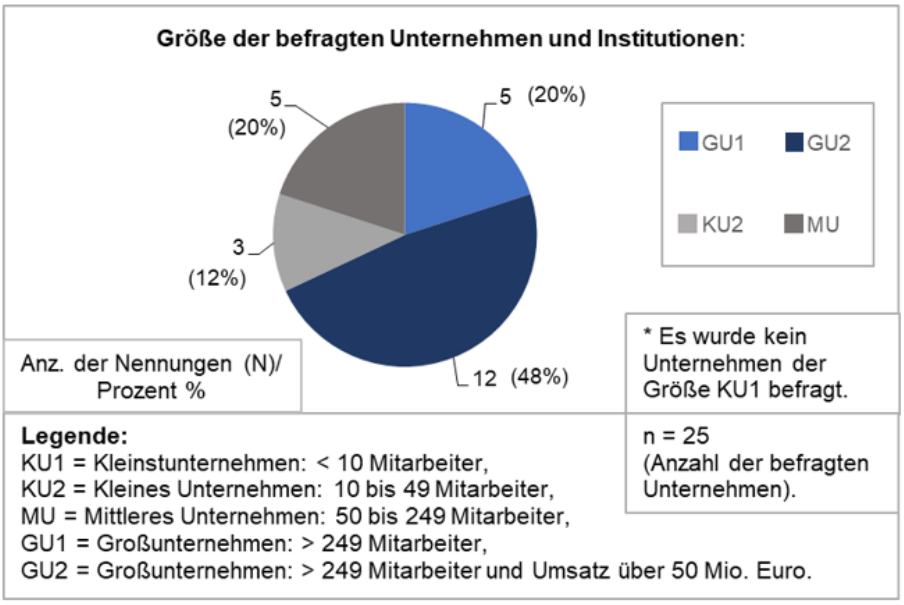


Abbildung 8: Studie - Größe der befragten Unternehmen und Institutionen.

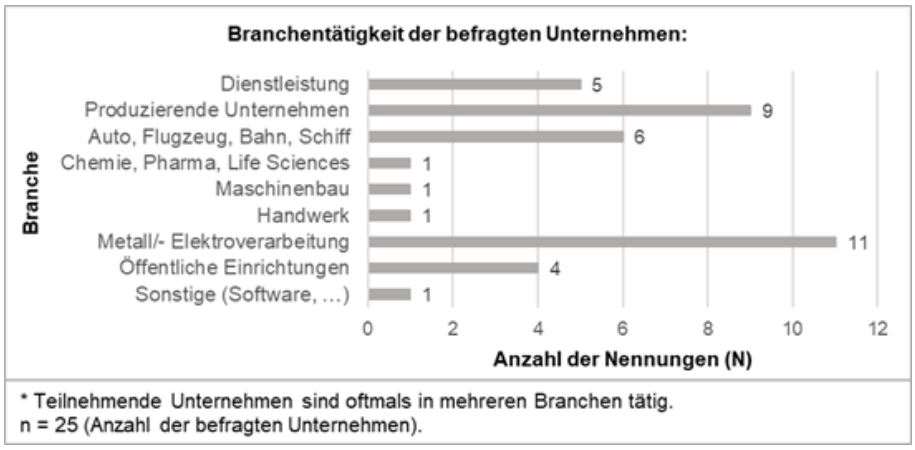


Abbildung 9: Studie - Branchentätigkeit der befragten Unternehmen und Institutionen.

men aus den Bereichen Auto, Flugzeug, Bahn und Schiff (N=6). Die Dienstleistungsbranche (N=5) und der Sektor der öffentlichen Einrichtungen (N=4) sind stark vertreten. Der restliche Anteil verteilt sich gleichmäßig auf sonstige Branchen. Die Stichprobe wird durch den Tätigkeitsbereich der befragten Personen näher beschrieben (Abb. 10).

Anhand von Abbildung 10 ist eine überwiegende Tätigkeit der Probanden in den Bereichen der Produktion und Auftragsabwicklung zu erkennen (N=14). Ein Großteil der Probanden betreibt angewandte Forschung, hat eine Lehrtätigkeit inne, oder bewegt sich in sonstigen nicht explizit aufgelisteten Bereichen (N=8). Ansonsten fällt die Verteilung auf die diversen Tätigkeitsbereiche recht homogen aus. Eine vollständige Beschreibung der Umfrageteilnehmer inklusive aller Umfragedaten ist in tabellarischer Form im Anhang zu finden (Anh. 5). Zugleich muss erwähnt werden, dass bei dieser Studie zur Einholung von Einschätzungen zu zukünftig wichtigen Produktionsanforderungen und Kennzahlen ein Querschnittsdesign vorliegt. Dies bedeutet, dass die Messung und Datenerhebung einmalig waren, nicht wiederholt wurden, und anhand einer Stichprobe stattfanden.[559]

Durchführung der Studie: Der Untersuchungszeitraum erstreckte sich über zwei Monate (November und Dezember 2019). Dieser Zeitraum wurde benötigt, um ausreichend Probanden für die Befragung zu rekrutieren. Die Probennahme erfolgte am Fraunhofer IAO in Stuttgart. Hauptsächlich konnten dazu zwei Eventtage genutzt werden. Die Durchführung der Studie an diesen Tagen wurde entsprechend vorbereitet.

Ergänzend wurden Firmen und Experten per E-Mail kontaktiert, die an dem praxisorientierten Forschungsprojekt ARENA2036 mitwirken.[560] Außerdem konnten bestehen-

[559] vgl. Schirmer, Blinkert und Buchen (2009), S. 172.
[560] vgl. ARENA2036 e.V. (2020), URL siehe Literaturverzeichnis.

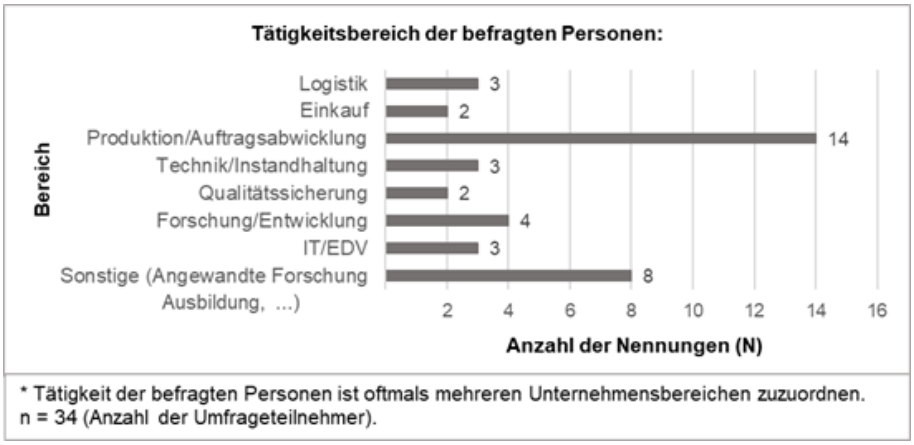


Abbildung 10: Studie - Unternehmenstätigkeitsbereich der Probanden.

de reguläre Meetings im Arbeitsalltag der Betreuerin dieser Arbeit genutzt werden, um entsprechende I4.0 Fachleute Fragebögen ausfüllen zu lassen. Die letzte Quelle von Probanden sind persönliche Kontakte des Autors zu Experten regionaler Großunternehmen. Sie wurden über verschiedene Wege kontaktiert, um hochwertigen Input zu der Zukunft der Produktion zu erhalten.

Der größte Anteil an Teilnehmern konnte über zwei von dem Fraunhofer IAO und dem Fraunhofer Institut für Produktionstechnik und Automatisierung (IPA) durchgeführte Veranstaltungen gewonnen werden. Die erste größere Datenerhebung erfolgte am Freitag den 29. November 2019 während des „Open Lab Days“ im Future Work Lab in Stuttgart. An diesem Tag wurden in dem Innovationslabor der Fraunhofer Institute anhand von Demonstratoren Szenarien der Zukunftsproduktion vorgestellt. Seminare, Workshops und Führungen ergänzten die Veranstaltung.[561] Im Laufe dieses Tages konnten einige zufällig ausgewählte Teilnehmer des Events zur Ausfüllung eines Fragebogens gewonnen werden.

Am 12. Dezember 2019 fand ein zweiter, größerer Veranstaltungstag der Stuttgarter Fraunhofer Institute statt. Dieser wurde ebenfalls zur Erhebung der grundlegenden Daten für diese Arbeit genutzt. Die Initiative „Innovationsnetzwerk - Produktionsarbeit 4.0“ ist eine Kooperation, bei der das Fraunhofer IAO den Austausch mit verschiedenen Partnerunternehmen und Institutionen vorantreibt.[562] Der insgesamt zwölfte Workshop dieses Projektes zur Gestaltung und Implementierung von „Produktionssystemen 4.0“ konnte in gleichem Maße genutzt werden, um Studienteilnehmer zu gewinnen. Die folgenden Wochen bis Anfang 2020 wurden genutzt, um weitere von den beiden größeren Erhebungstagen unabhängige Rückmeldungen einzuholen und ausgefüllte Fragebögen zu bekommen. Generell bearbeiteten die Probanden den Fragebogen selbstständig und wurden mithilfe von textuellen Hinweisen durch die einzelnen Fragen geleitet. Der Fragebogen inklusive aller seiner Bestandteile wird im nächsten Kapitel näher erläutert.

5.2. Design des Fragebogens, Messinstrumente und Auswertungsmethoden

In diesem Kapitel wird im Detail auf die Konzeptionierung und das Design des verwendeten Fragebogens inklusive der genutzten Messinstrumente eingegangen. Bevor die Kapitel (5.3) bis (5.5) die Auswertung der erhobenen Daten liefern, erfolgt in diesem Abschnitt abschließend die Erklärung der Datenauswertungsmethoden.

Insgesamt wurde bei der Formulierung der einzelnen Teilfragen des Fragebogens darauf geachtet, mehrdeutige Formulierungen zu vermeiden und jegliche Grundsätze der Fragenformulierung einzuhalten.[563] Weiterhin sind bei der Gestaltung die Kriterien der Verständlichkeit und Eindeutigkeit berücksichtigt worden.[564] Als eine grundlegende Basis für den prinzipiellen Aufbau und die Gestaltung des Fragebogens dient das Design einer Studie von Stief (2018) zum Nutzen und zur Akzeptanz von Industrie 4.0 Anwendungen.[565] Außerdem fand nach einer vorangestellten Literaturrecherche eine mehrstufige iterative Verbesserung des Fragebogens statt.

Der Fragebogen wurde überwiegend mit geschlossenen Fragen konzipiert. So konnten die Probanden inhaltlich geführt und ihre Antworten schnell ausgewertet werden.[566] Bei solchen Fragen wird eine begrenzte Anzahl an Antwortmöglichkeiten geboten. In Ergänzung dazu gestatten offene Fragen eine beliebige Äußerung zu einem Bereich.[567] Dieses Fragenelement wurde zu einem geringen Anteil ebenfalls in das Fragebogendesign integriert. Der Fragebogen inklusive aller Messbestandteile ist im Anhang einsehbar (siehe Anh. 6).

[561] vgl. Fraunhofer-Institut für Arbeitswirtschaft und Organisation und aucobo GmbH (2020), S. 1f, URL siehe Literaturverzeichnis.
[562] vgl. Fraunhofer-Institut für Arbeitswirtschaft und Organisation (2019), S. 3 - 20, URL siehe Literaturverzeichnis.
[563] vgl. Schumann (2006), S. 62 - 67.
[564] vgl. Brosius et al. (2016), S. 97.
[565] vgl. Stief (2018), S. 49 - 122.
[566] vgl. Brosius et al. (2016), S. 86f.
[567] vgl. ebenda, S. 85.

Die Befragung gliedert sich grundsätzlich in drei unterschiedliche Teilerhebungen und zwei Themenblöcke. Zu Beginn einer Befragung ist es üblich, zuerst Charakteristika der Teilnehmer zu erfassen.[568] Dies wurde in dieser Studie übernommen, um grundlegende Informationen über die Stichprobe zu gewinnen. Allerdings wurden nur die relevantesten Merkmale erhoben. Diese Merkmale werden am Fraunhofer IAO üblicherweise für eine Zuordnung von Studienteilnehmern verwendet (siehe Beschreibung der Stichprobe). Für die Einteilung nach Unternehmensgröße dienen die Vorgaben des „Benutzerleitfaden zur Definition von KMU“ der Europäische Kommission (2015) als ergänzende Vorlage.[569]

Im ersten Themenblock werden im Detail Einschätzungen zu dem Themenbereich der zukünftigen Produktionssteuerungskennzahlen erhoben. Es existieren mehrere Gründe für die Integration exakt dieser 16 in Kapitel (3.) theoretisch dargelegten Kennzahlen in die quantitative Orientierungsbefragung. Erste Denkanstöße zur Auswahl einiger der Kennzahlen lieferte die Studie „Auf dem Weg zum digitalen Shopfloor Management“ des Karlsruher Instituts für Technologie (KIT) (2018), die somit in diesem Aspekt eine Basisquelle darstellt.[570] Ein weiterer Grund ist, dass die Erhebung der Einstellungen zu einem Großteil der Kennzahlen explizit von dem Fraunhofer IAO und der Betreuerin dieser Arbeit erwartet wurde. Das Grundgerüst an zu untersuchenden Kennzahlen wurde anschließend durch eine Literaturrecherche zu zukünftig wichtigen Steuerungskennzahlen ergänzt. Eine Klassifizierung in organisations- und technologiebezogene sowie mitarbeiterbezogene Kennzahlen ist aufgrund der Vermeidung einer Beeinflussung der Teilnehmerantworten nicht bereits innerhalb des Fragebogens vorgenommen worden.

Die Probanden wurden zunächst aufgefordert anzugeben, welche Kennzahlen für die Steuerung der Produktion aus ihrer Sicht zukünftig besonders wichtig sind. Es sollten bei dieser ersten Teilerhebung acht Kennzahlen ausgewählt und nach deren Relevanz in eine Rangordnung gebracht werden. Bei diesem Rankingverfahren ist es nicht möglich Plätze mehrfach zu vergeben.[571] Es wurde also von jedem Teilnehmer eine eindeutige Reihenfolge erstellt, wobei der gänzliche Ausschluss der übrigen Kennzahlen erfolgte.

Die Teile zwei und drei der quantitativen Orientierungsumfrage beziehen sich auf den zweiten Themenblock, bei welchem es die übergeordnete Zielsetzung ist, Aussagen zu der Relevanz von Anforderungen an die Produktionsarbeit der Zukunft zu erhalten. Die Auswahl explizit die Einstellungen gegenüber diesen 14 Anforderungen zu erforschen begründet sich zum einen darin, dass sich diese durch die auf Organisation und den Menschen gerichtete Literaturrecherche als wichtig ergeben haben. Zum anderen hat das Fraunhofer IAO eine Überprüfung der Wichtigkeit dieser Produktionsanforderungen in der Praxis als Forschungsziel festgelegt. Die Klassifikation der Anforderungen in die Kategorien „Organisation“, „Mensch“ und „Technologie“ wurde in die Gestaltung des Fragebogens übernommen. Innerhalb jeder Kategorie werden die jeweiligen Produktionsanforderungen mithilfe von Erklärungstexten vorgestellt. Diese Anforderungen stellen die Items des Messinstrumentes dar. Laut Definition wird mit Fragebogenitems eine Zustimmung oder Ablehnung zu einer Aussage erhoben.[572] Sie sind ein Bestandteil des hier verwendeten Messinstruments einer Ratingskala. Jedes einzelne Item wird anhand dieser Skala stufenmäßig eingeschätzt. Konkret wurde das explizite Einstellungsmaß der Likertskala angewendet.[573] Als Ordinalskala ist hier eine Rangbildung möglich und die Reihenfolge der Werte ist interpretierbar.[574]

Bei einer solchen Skala sind generell fünf bis sieben Antwortvorgaben denkbar, da das Differenzierungsvermögen der Befragten innerhalb dieses Rahmens nicht überfordert wird.[575] Im Detail erfolgt in dieser Studie eine leichte Anpassung der üblichen Likertskala. Die Skala mit den Zahlenwerten von eins bis sieben wird um Extrempunktbeschriftungen ergänzt.[576] Die Beschriftung der Skala erfolgte verbal und numerisch, da eine rein numerische Beschriftung das Antwortverhalten verzerren kann.[577] Die Einstufung der Wichtigkeit auf der Likertskala erstreckte sich hier somit nach Vorlage einer in Studien üblichen Einteilung von (1) „weniger wichtig bzw. unwichtig“, über (4) „wichtig“, bis hin zu dem Maximalpunkt (7) „sehr wichtig“.[578]

In dem abschließenden dritten Teil des Fragebogens wurden die Probanden dazu aufgefordert, die von ihnen als relevant eingestuften Anforderungen noch einmal als separate Rangfolge aufzulisten. Alle Anforderungen, die im zweiten Teil der Umfrage eine Skalenbewertung von fünf bis sieben erhalten haben, wurden hier als gültige Antworten berücksichtigt. Diese individuelle Auswahl der Umfrageteilnehmer ist in ihrer Länge beliebig, da jedoch die Einstellungen zu 14 Anforderungen erhoben wurden, konnten auch maximal nur so viele Felder ausgefüllt werden. Worauf dieser Umfrageteil genau abzielt folgt im Zusammenhang mit der Umfrageauswertung in Kapitel (5.5).

Ergänzend zu den geschlossenen Fragen bezüglich der Wichtigkeit von Kennzahlen und zukünftigen Produktionsanforderungen, enthält der Fragebogen zwei leere Textfelder, durch welche die Probanden am Ende die Möglichkeit hatten, Kommentare abzugeben und weitere Produktionsanforderungen konkret zu benennen, um individuelle Meinungen und Bedenken zu äußern.

Ein Pretest eines Fragebogens gibt Aufschluss über die praktische Handhabbarkeit und die Verständlichkeit aller Fragen und Formulierungen. Möglicherweise ist es erforderlich Änderungen vorzunehmen, Fragen und Skalen anzupas-

[568] vgl. Schumann (2006), S. 52.
[569] vgl. Europäische Kommission (2015), S. 11.
[570] vgl. Lanza et al. (2018), S. 23ff.
[571] vgl. Schumann (2006), S. 72.
[572] vgl. Mayer (2009), S. 79.
[573] vgl. Haddock und Maio (2014), S. 213.
[574] vgl. Schumann (2006), S. 19ff.
[575] vgl. Nieschlag, Dichtl und Hörschgen (1994), S. 693f.
[576] vgl. Berekhoven, Eckert und Ellenrieder (1999), S. 75.
[577] vgl. Moosbrugger und Kelava (2012), S. 52.
[578] vgl. Spath et al. (2013), S. 50.

sen oder neue Fragen einzubauen.[579] Er gibt Aufschluss über die fachlich Richtigkeit eines Fragebogens.[580] Ein solcher detaillierter Pretest wurde aus zeitlichen Gründen nicht durchgeführt, auch weil davon ausgegangen werden konnte, dass für die hier verwendete Stichprobe keine grundlegenden Verständnisschwierigkeiten bestehen. Allgemein gibt es in der quantitativen Forschung zwei Gütekriterien, die für das Erzielen optimaler Forschungsergebnisse von einem Messinstrument erfüllt werden müssen: Die Reliabilität, die beschreibt, wie zuverlässig eine Messung ist, und die Validität, die sicherstellt, dass mit dem Messinstrument das gemessen wird, was gemessen werden soll.[581] Es wurden hierzu keine konkreten Tests mithilfe von Kennwerten durchgeführt,[582] da sich diese praktischen Nachweise größtenteils ziemlich aufwendig gestalten.[583]

Die durch die quantitative Orientierungsbefragung erhobenen Daten konnten mithilfe von unterschiedlichen Methoden analysiert und ausgewertet werden. Da nicht immer alle Teile des Fragebogens vollständig und komplett richtig ausgefüllt wurden und verwertbar waren, fand je nach Umfrageteil der Einbezug von 29-34 Ergebnissen in die Analyse statt. Die Dateneingabe und Datenaufbereitung zur Auswertung der ausgefüllten Fragebögen erfolgte mit dem Tabellenkalkulationsprogramm EXCEL. Die bei dieser Studie verwendete Methode ist eine häufig angewandte Vorgehensweise.[584] Zur Erläuterung der Ergebnisse werden Grafiken und Tabellen verwendet, welche die erhobenen Daten und gewonnenen Ergebnisse veranschaulichen.[585] Die Daten werden weiterhin mithilfe von statistischen Größen beschrieben. In allen Umfrageteilen erfolgt eine Angabe von Standardabweichungen. Durch diese Maßzahl werden die Verteilung und Streuung der Werte näher angegeben. Diese Größe ist neben dem arithmetischen Mittel eine Standardform der Wertebeschreibung.[586] Bei Umfrageteil zwei wird die durchschnittliche Platzierung auf der Likertskala und somit die Angabe der Wichtigkeit der jeweiligen Anforderung durch die Bildung von Durchschnittswerten (M) abgebildet.

Bei den Teilerhebungen eins und drei wurden Rangfolgen gebildet, weshalb sich die Auswertungen hier etwas komplizierter gestalten. Die Platzierungen in den finalen Rangfolgen werden mithilfe von zwei Ordnungskriterien bestimmt. Das erste Ordnungskriterium ist die Anzahl der Nennungen (N). Das zweite Ordnungskriterium der bewerteten Rangposition (BR) spiegelt die durchschnittliche Rangplatzierung wider. Für eine Vergleichbarkeit der Werte müssen die jeweiligen Werte bei einer Mittelwertbildung immer durch dieselbe Summe geteilt werden.[587] In dem Kennzahlenteil also durch die Summe von 16 und das Anforderungsranking betreffend durch die Summe von 14. Für eine korrekte statistische Auswertung ist hierbei eine Umpolung der Werte erforderlich, da jeder Kennzahl bzw. Anforderung ein Wert zugesprochen werden muss.[588] In Teil eins erhalten also acht Kennzahlen einen Rang von eins bis acht. Alle übrigen, nicht gerankten Kennzahlen erhalten einen Rang von null. Damit die am wichtigsten bewertete Kennzahl den höchsten Wert erhält, wurden die Ränge anschließend umgepolt. Ein Rang von eins entspricht somit einer bewerteten Rangposition (BR) von acht, ein Rang von zwei einer bewerteten Rangposition von sieben. Nur durch diese Vorgehensweise war es möglich, Mittelwerte zur Auswertung heranzuziehen. Dasselbe Vorgehen wurde in Umfrageteil drei bei dem Ranking der 14 Anforderungen angewandt. Die bewertete Rangposition (BR) gibt als Messgröße nachfolgend die umgepolte durchschnittliche Rangposition an.

In dem ersten Erhebungsteil kam es vor, dass Probanden nur Kreuze bei den für sie wichtigsten Kennzahlen gesetzt haben. Diese Daten sind bei der Anzahl der Nennungen (N) berücksichtigt worden, nicht aber bei der bewerteten Rangposition (BR). Dadurch unterscheidet sich hier die Stichprobengröße zwischen beiden Ordnungskriterien. Anschließend an diese Erläuterungen zur Studienmethodik, zu dem Fragebogen und zu den Auswertungsmethoden werden alle Ergebnisse der Studie umfassend betrachtet.

5.3. Studienauswertung Teil 1: Kennzahlen zur Steuerung der Produktion

Dieser Abschnitt widmet sich der ersten Frage der durchgeführten Unternehmensbefragung. Als zentrales Ergebnis dieses Kapitels kristallisieren sich die wichtigsten Kennzahlen für die zukünftige Shopfloorsteuerung heraus. Die gegebenen Antworten der Umfrageteilnehmer sind die Grundlage zur Erstellung einer Kennzahlenrangfolge. Eine Aufgabe der Übersicht ist es, die organisations- und technologiebezogenen sowie die mitarbeiterbezogenen Kennzahlen aus Kapitel (3.) nach deren Wichtigkeit einzuordnen und die „Top 5“ Kennzahlen zur Steuerung der zukünftigen I4.0 Produktion zu determinieren. Grundsätzlich erfolgt in den Auswertungen der Umfrage keine Interpretation von Ergebnissen, sondern nur eine Wertebeschreibung. Die Interpretation der Ergebnisse ist Inhalt der abschließenden Diskussion in Kapitel (7.) und benötigt zunächst die Vorstellung diverser I4.0 Lösungen und Anwendungsfälle zur Umsetzung von organisationalen Produktionsanforderungen in Kapitel (6.) als Basis.

An den Ergebnissen der Umfrage zeigt sich, dass für die Steuerung der zukünftigen Produktion mehrere Kennzahlen infrage kommen. Als Ausschlusskriterium für ungeeignete Kennzahlen wird die Anzahl von fünf oder weniger Nennungen (N) definiert. Wie Tabelle 2 zu entnehmen ist, fallen von den 16 untersuchten Kennzahlen nur die Logistikkosten (n=5; BR=0,62; SD=1,545) und die Instandhaltungskosten (n=5; BR=0,17; SD=0,468) durch dieses Raster. Die restlichen 14 Kennzahlen sind in I4.0 Kennzahlensystemen integrierbar und als relevant zu beurteilen.

[579] vgl. Brosius et al. (2016), S. 131.
[580] vgl. ebenda, S. 96.
[581] vgl. ebenda, S. 51.
[582] vgl. Schumann (2006), S. 40ff.
[583] vgl. Mayer (2009), S. 89.
[584] vgl. ebenda, S. 103.
[585] vgl. Brosius et al. (2016), S. 132.
[586] vgl. Schumann (2006), S. 149f.
[587] vgl. ebenda, S. 143.
[588] vgl. Mayer (2009), S. 110.

Tabelle 2: Studie Teil 1 - Gesamttabelle und Rangfolge der Kennzahlen zur Produktionssteuerung.

Rang	Kennzahl	Primäres Ordnungskriterium: Anzahl Nennungen (N)	Sekundäres Ordnungskriterium: Bewertete Rangposition (BR)	Zugehörige Standardabweichung (SD)
1	Qualität: Right First Time/ Fehlerrate	27	5,00	2,878
2	Produktionskosten	26	4,31	3,106
3	Schnelligkeit/ Durchlaufzeit	24	3,00	2,632
4	Mitarbeitermotivation und Zufriedenheit	22	3,45	2,746
5	Liefertreue	20	3,21	2,944
6	Produktivität/ Produktionsauslastung	19	3,38	3,332
7	Personalproduktivität	17	2,28	2,658
8	Arbeitsschutz und Arbeitssicherheit	14	1,97	2,771
9	Mitarbeiterengagement	13	1,52	1,993
10	Gesamtanlageneffektivität (OEE)	12	1,93	2,975
11	Gesundheitsförderlichkeit und Ergonomie	12	1,17	1,891
12	Wandlungsfähigkeit	11	1,62	2,744
13	Energieverbräuche	8	0,93	1,731
14	Bestandshöhen/Reichweiten	7	0,45	1,021
15	Logistikkosten	5	0,62	1,545
16	Instandhaltungskosten	5	0,17	0,468

Insgesamt ist eine sehr deutliche Abstufung in der Anzahl der Nennungen (N) erkennbar. Die bewertete Rangposition (BR) gibt hier die durchschnittliche Position auf einem Rang von eins bis acht an, während, wie in der Methodik beschrieben, die Ränge aufgrund der statistischen Auswertbarkeit umgepolt werden mussten und eine bewertete Rangposition von acht das beste erzielbare Ergebnis ist. Bei einer identischen Anzahl an Nennungen bestimmt die bewertete Rangposition als sekundäres Ordnungskriterium über die finale Einordnung. Die Abbildungen 11 und 12 werden zur detaillierteren Beschreibung der Teilergebnisse herangezogen.

Die Kennzahl Qualität-Right First Time/Fehlerrate ist mit 27 Nennungen (N=27; BR=5,00; SD=2,878) das wichtigste Produktionssteuerungsinstrument. Die Kennzahlen Produktionskosten (N=26; BR=4,31; SD=3,106), Schnelligkeit/Durchlaufzeit (N=24; BR= 3,00; SD=2,632), Mitarbeitermotivation und Zufriedenheit (N=22; BR=3,45; SD=2,746) und Liefertreue (N=20; BR=3,21; SD=2,944) ergänzen die „Top 5“ der zukünftig essenziellsten Kennzahlen. Auffällig ist, dass die Probanden von den fünf mitarbeiterbezogenen Kennzahlen nur die Mitarbeitermotivation und Zufriedenheit als sehr wichtig einschätzen. Organisations- und technologiebezogene Kennzahlen sind für die Umfrageteilnehmer von leicht höherer Bedeutung. Außerdem ist in den Auswertungsgraphen erkennbar, dass sich alle mitarbeiterbezogenen Kennzahlen auf einer mittleren Wichtigkeitsstufe befinden (N=12-17). Dadurch zeigt sich die Beachtung beider Kennzahlenkategorien in der Praxis. Die technologiebezogenen Kennzahlen Energieverbräuche (N=8; BR=0,93; SD=1,731) und Bestandshöhen/Reichweiten (N=7; BR=0,45; SD=1,021) bilden die Gruppe der am niedrigsten bewerteten, jedoch noch als Steuerungselemente in Betracht zu ziehenden, Kennzahlen.

Dieses Kapitel beantwortet die erste Forschungsfrage und zeigt durch Daten aus der Praxis auf, mit welchen Kennzahlen die Produktion vorrangig gesteuert werden sollte. Es ist, wie bereits in Kapitel (3.2) dargelegt, eine Anforderung an die Produktion der Zukunft, die Organisationsstruktur so auszulegen, dass es möglich ist, wichtige Kennzahlen zu steuern. Im Anschluss wird die Auswertung des Umfrageteils zu den Anforderungen an die zukünftige Produktion geschildert.

5.4. Studienauswertung Teil 2: Organisationale Produktionsanforderungen

Der zweite Teil der quantitativen Orientierungsbefragung setzt sich mit der Fragestellung auseinander, wie die organisationalen Anforderungen an die zukünftige Produktion in der Praxis gesehen werden. Mit den Ergebnissen der Befragung erfolgt an dieser Stelle eine Einstufung der Konzepte aus Kapitel (4.). Es wird weiterhin geklärt, ob diese das Denken der Produktion bereits verändern und worauf Unternehmen ihre Prioritäten setzen.

Mit diesem Teil der Umfrage konnte bewiesen werden, dass alle in dieser Arbeit behandelten Anforderungen an die zukünftige Produktion als wichtig einzuschätzen sind. Des Weiteren zeigt die Untersuchung die grundlegende Bekanntheit der existierenden arbeitsorganisatorischen Konzepte für die Produktion in der Unternehmenspraxis, auch wenn keine Aussage darüber getroffen werden kann, inwiefern diese bereits tatsächlich umgesetzt werden. Die Ergebnisse decken sich somit zu großen Teilen mit den aktuellen Vorstellungen der Forschung. Nachfolgend wird mithilfe von Tabelle 3 eine Erweiterung des Forschungsstandes in Form einer konkreten Wichtigkeitsrangfolge geliefert. Die Tabelle veranschaulicht

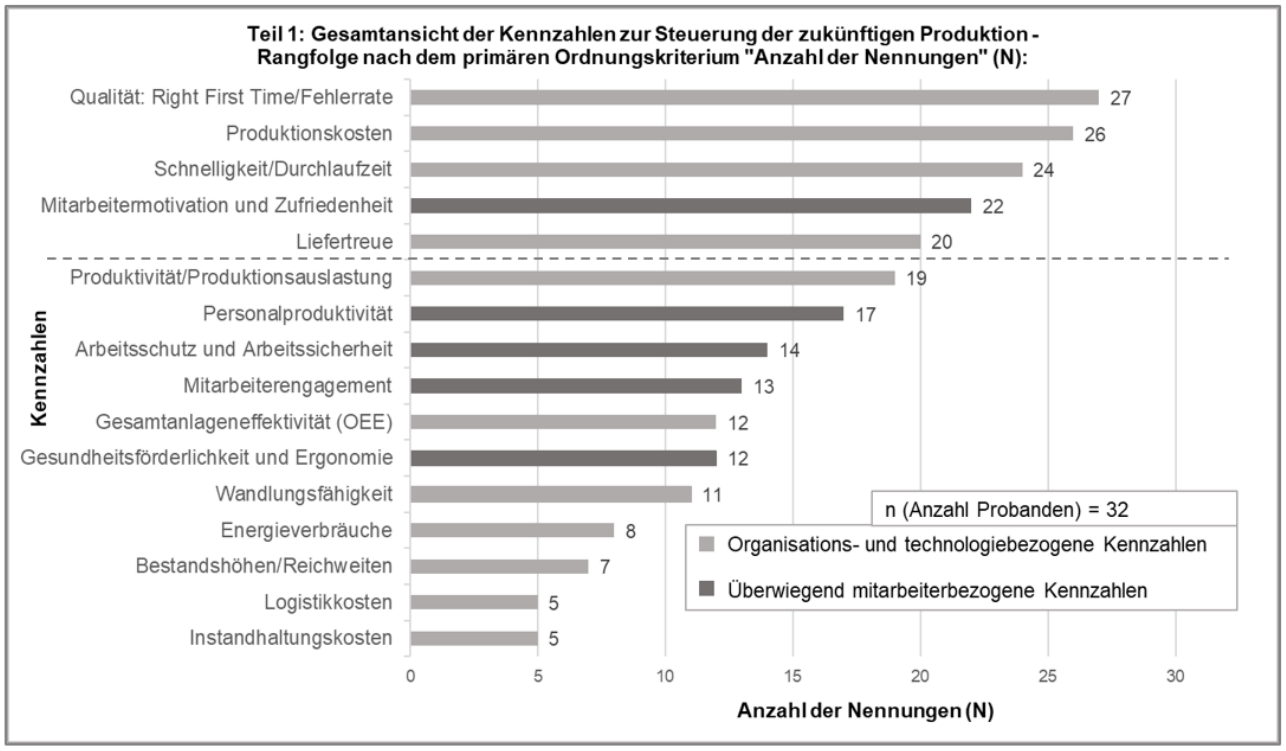


Abbildung 11: Studie Teil 1 - Rangfolge der Steuerungskennzahlen nach dem primären Ordnungskriterium Anzahl der Nennungen (N).

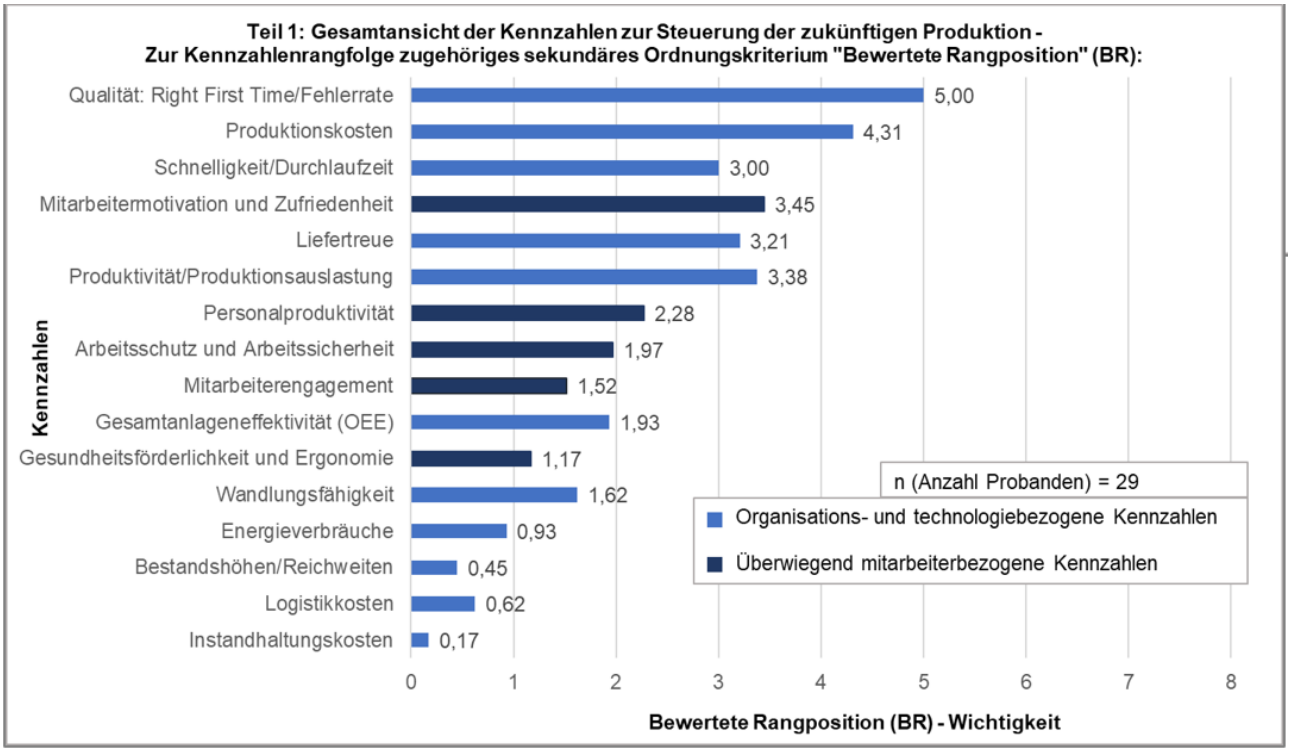


Abbildung 12: Studie Teil 1 - Zur Kennzahlenrangfolge zugehöriges sekundäres Ordnungskriterium bewertete Rangposition (BR).

die Rangpositionen aller Produktionsanforderungen inklusive der zugehörigen Durchschnittswerte (M) und Standardabweichungen (SD).

Im Zuge der Auswertung zu der Bedeutung der Anforderungen, war in diesem Fall keine statistische Umpolung notwendig, weshalb der jeweilige Wert hier tatsächlich den Mittelwert auf der verwendeten Rankingskala von eins bis sieben angibt.

Die Durchschnittswerte der Anforderungen liegen alle zwischen den Bewertungsstufen fünf und sechs. Es sind keine besonders auffälligen Standardabweichungen vorhanden. Die Anforderung einer verbesserten Kommunikation und Abstimmung (M=6,00; SD=1,267) erhält die höchste Wichtigkeitseinstufung, während die Selbstorganisation der Teams (M=5,18; SD=1,547) von den Probanden im Vergleich zu den anderen Anforderungen als eindeutig unwichtiger eingeschätzt wird. Als zweitwichtigste Anforderung ergibt sich die umfassende Vernetzung und Transparenz (M=5,91; SD=0,9) in der Produktion. Eine Ausrichtung der Produktion auf Lean Management 4.0 (M=5,88; SD=0,844) ist ebenso wie die konsequente Mitarbeiterqualifikation (M=5,82; SD=1,058) und die Wandlungsfähigkeit der Produktion (M=5,76; SD=1,393) Bestandteil der fünf wichtigsten Anforderungen an die zukünftige Produktion. Folglich sollte der praktischen Umsetzung dieser Anforderungen zukünftig eine hohe Bedeutung beigemessen werden. Abbildung 13 bietet durch die kategorisierte Zuteilung der Anforderungen eine erweiternde Ansicht der Rangfolge.

In der Gruppe der fünf Anforderungen mit den größten Durchschnittswerten und der höchsten Wichtigkeitseinstufung lassen sich drei der Kategorie „Technologie“ zuordnen, obwohl sich in den insgesamt 14 Anforderungen nur vier technische Produktionsanforderungen finden lassen. Die Mehrzahl der untersuchten Anforderungen (10) bezieht sich vorwiegend auf organisatorische und menschliche Aspekte. Von diesen Anforderungen sind allerdings nur zwei in den „Top 5“ der wichtigsten Anforderungen wieder zu finden. Die Grafik verdeutlicht die Bewertung der Umfrageteilnehmer, dass organisations- und menschenbezogene Anforderungen in der Praxis bisher größtenteils weniger Aufmerksamkeit entgegengebracht wird, als rein technologischen I4.0 Produktionsanforderungen. Der nachfolgende letzte Teil der Unternehmensbefragung zielt darauf ab, diese gewonnenen Erkenntnisse zu spezifizieren.

5.5. Studienauswertung Teil 3: Vergleichendes Ranking der Anforderungen und ergänzende Teilnehmerkommentare

In diesem Kapitel wird der dritte Teil der durchgeführten Studie ausgewertet und diskutiert. In dem abschließenden Teil des Fragebogens wurden die Probanden dazu aufgefordert, die von ihnen als relevant eingestuften Anforderungen noch einmal als separate Rangfolge aufzulisten. Sinn hinter diesem Vorgehen ist es, eine detailliertere Wichtigkeitsabstufung der im zweiten Umfrageteil bewerteten Anforderungen an die zukünftige Produktion zu erhalten. Die Rangpositionen reichen in diesem Umfrageteil bis zu einem Maximum von 14. Wie in der Methodik ersichtlich, ist aufgrund der Auswertbarkeit an dieser Stelle erneut eine Umpolung der Werte erforderlich. Je höher die bewertete Rangposition (BR), desto höher wurde die Anforderung demnach eingestuft (siehe hierzu Messinstrument in Anh. 6).

Die Ergebnisse dieses Umfrageteils sind in Tabelle 4 abgebildet. Der wesentlichste festzustellende Punkt ist die Bestätigung der Wichtigkeit drei der „Top 5“ Anforderungen.

Die Anforderungen einer verbesserten Kommunikation und Abstimmung (N=23; BR=8,35; SD=5,574), einer wandlungsfähigen Produktion (N=19; BR=6,16; SD=5,514) und einer konsequenten Mitarbeiterqualifikation (N=18; BR=6,26; SD=5,597) sind somit definitiv fundamental für die Gestaltung der zukünftigen Produktion.

Die Abbildungen 14 und 15 differenzieren die Anforderungen weiterführend in die drei grundlegenden Kategorien. Anhand der Daten ist zu erkennen, dass die Probanden die Anforderungen der Kategorie „Technologie“ hier insgesamt niedriger bewerten als im zweiten Studienteil. Die extremsten Unterschiede im Vergleich zur vorangestellten Datenerhebung zeigen sich bei der Anforderung einer vernetzten und transparenten Produktion (N=14; BR=4,84; SD=5,854) und einer auf Lean Management 4.0 ausgerichteten Produktion (N=14; BR=4,32; SD=5,281). Die hohe Bedeutung dieser Anforderungen konnte in diesem Umfrageteil nicht bestätigt werden. Eine ergänzende Feststellung ist, dass die Transparenz in der Organisation (N=21; BR=6,39; SD=5,487) genauso wie die Förderung der Mitarbeitermotivation (N=20; BR=6,48; SD=5,353) in dieser Rangfolge einen Platz in der Gruppe der fünf wichtigsten Anforderungen an die zukünftige Produktion einnimmt.

Aufgrund der Gegebenheit, dass dieser Umfrageteil grundsätzlich von weniger Probanden beantwortet wurde (n=31) und nur als Ergänzung für eine weitere Differenzierung dient, wird der zweite Umfrageteil als maßgebliche Quelle zur Bestimmung der Bedeutung der Anforderungen herangezogen. Aufgrund dieser Vorgehensweise ergibt sich durch die Ergebnisse des dritten Umfrageteils eine Erweiterung der Sammlung der fünf wichtigsten Anforderungen an die zukünftige Produktion aus Umfrageteil zwei.

Als konkludierende Erkenntnis der gesamten Studie ergibt sich deshalb neben den fünf wichtigsten Kennzahlen zur Steuerung der zukünftigen Produktion eine Sammlung von sieben Anforderungen, welche die Industrie aktuell und in Zukunft für bedeutsam erachtet. Die Anforderungen sind deshalb in der nachfolgenden Rangfolge und Wichtigkeitseinstufung angeordnet: (1) Verbesserte Kommunikation und Abstimmung, (2) Vernetzung und Transparenz, (3) Ausrichtung auf Lean Management 4.0, (4) Konsequente Mitarbeiterqualifikation, (5) Wandlungsfähigkeit der Produktion, (6) Transparenz in der Organisation und (7) Förderung der Mitarbeitermotivation. Somit wird die „Top 5“ der wichtigsten Anforderungen an die zukünftige Produktion um jeweils eine Anforderung aus der Kategorie „Organisation“ und „Mensch“ ergänzt.

Insgesamt ist festzuhalten, dass durch eine Spanne der Anzahl der Nennungen von (n=12) bis (n=23) alle Anforde-

Tabelle 3: Studie Teil 2 - Gesamttabelle und Rangfolge der Anforderungen an die zukünftige Produktion.

Rang	Anforderung	Durchschnittswert - (M) Wichtigkeit der Anforderung	Zugehörige Standardabweichung (SD)
1	Verbesserte Kommunikation und Abstimmung	6,00	1,267
2	Vernetzung und Transparenz	5,91	0,900
3	Ausrichtung auf Lean Management 4.0	5,88	0,844
4	Konsequente Mitarbeiterqualifikation	5,82	1,058
5	Wandlungsfähigkeit der Produktion	5,76	1,393
6	Transparenz in der Organisation	5,74	1,238
7	Moderne Führung	5,65	1,228
8	Förderung der Mitarbeitermotivation	5,62	1,155
9	Innovationsfähigkeit	5,56	1,501
10	Förderung der Ressourcenschonung und Nachhaltigkeit	5,50	1,330
11	Flexibler Personaleinsatz	5,38	1,074
12	Agilität der Organisation	5,35	1,323
13	Prozessorientierte Arbeitsorganisation	5,32	1,583
14	Selbstorganisation der Teams	5,18	1,547

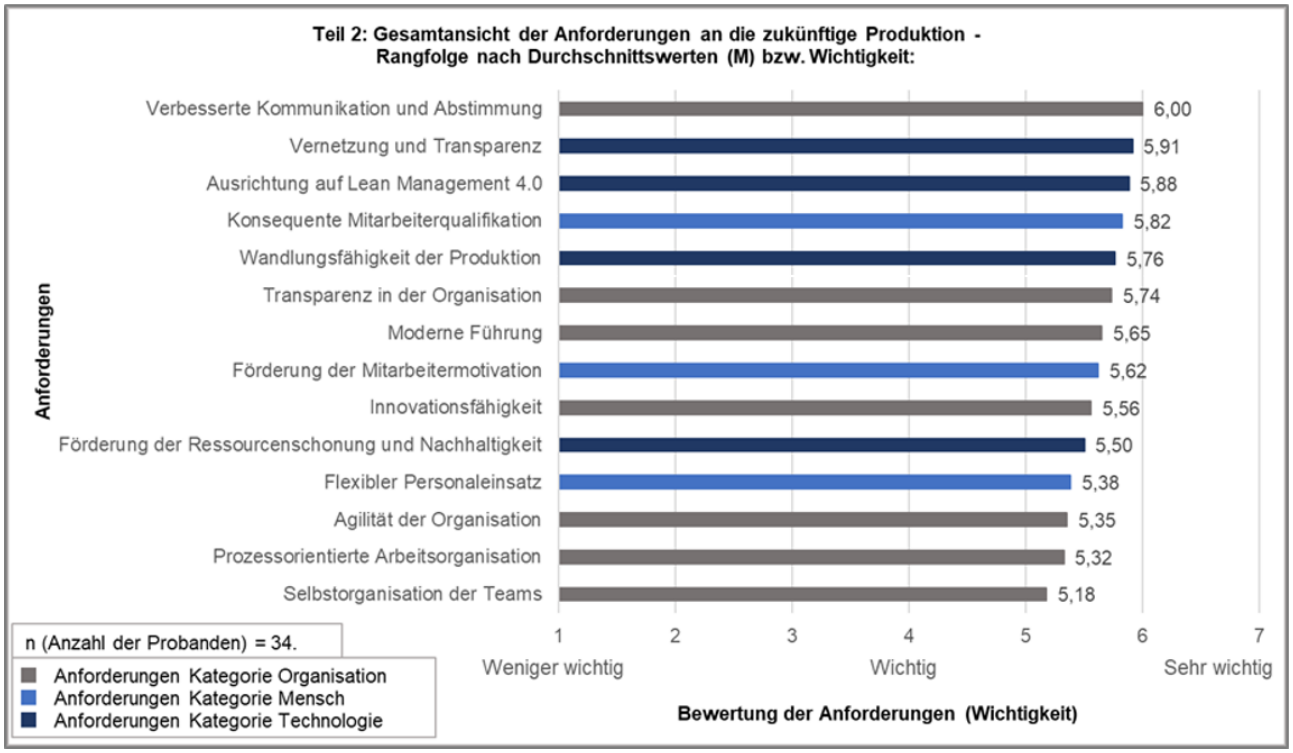


Abbildung 13: Studie Teil 2 - Rangfolge der Anforderungen an die zukünftige Produktion nach Durchschnittswerten (M) bzw. Wichtigkeit.

rungen in diesem Umfrageteil recht nah aneinander liegen. Hierdurch bestätigt sich, dass alle untersuchten arbeitsorganisatorischen Produktionsgestaltungskonzepte grundsätzlich wichtig für die Zukunft sind. Als weitere Schlussfolgerung stellt sich deckend zu Umfrageteil zwei heraus, dass die Selbstorganisation der Teams (n=12; BR=3,71; SD=5,008) insgesamt das für am unwichtigsten gehaltene organisatorische Konzept ist. Zudem bestätigt sich die Einordnung der Anforderung einer modernen Führung als drittwichtigstes Element der zukünftigen Produktion (n=18; BR=6,06; SD=5,591) aus der Kategorie „Organisation“.

Abschließend sind nachfolgender Tabelle ergänzende Kommentare und weitere genannte Anforderungen der Umfrageteilnehmer zu entnehmen (Tab. 5). Diese zusätzlichen Einschätzungen bieten eine Erweiterung der Umfrage und verdeutlichen explizit die Wichtigkeit von Lean Management,

Tabelle 4: Studie Teil 3 - Gesamttabelle zum vergleichenden Ranking der Produktionsanforderungen.

Rang	Kennzahl	Primäres Ordnungskriterium: Anzahl Nennungen (N)	Sekundäres Ordnungskriterium: Bewertete Rangposition (BR)	Zugehörige Standardabweichung (SD)
1	Verbesserte Kommunikation und Abstimmung	23	8,35	5,574
2	Transparenz in der Organisation	21	6,39	5,487
3	Förderung der Mitarbeitermotivation	20	6,48	5,353
4	Wandlungsfähigkeit der Produktion	19	6,16	5,514
5	Konsequente Mitarbeiterqualifikation	18	6,26	5,597
6	Moderne Führung	18	6,06	5,591
7	Prozessorientierte Arbeitsorganisation	16	4,94	5,537
8	Innovationsfähigkeit	15	5,35	5,975
9	Agilität der Organisation	15	4,90	5,688
10	Flexibler Personaleinsatz	15	4,42	5,265
11	Vernetzung und Transparenz	14	4,84	5,854
12	Ausrichtung auf Lean Management 4.0	14	4,32	5,281
13	Förderung der Ressourcenschonung und Nachhaltigkeit	12	3,74	5,234
14	Selbstorganisation der Teams	12	3,71	5,008

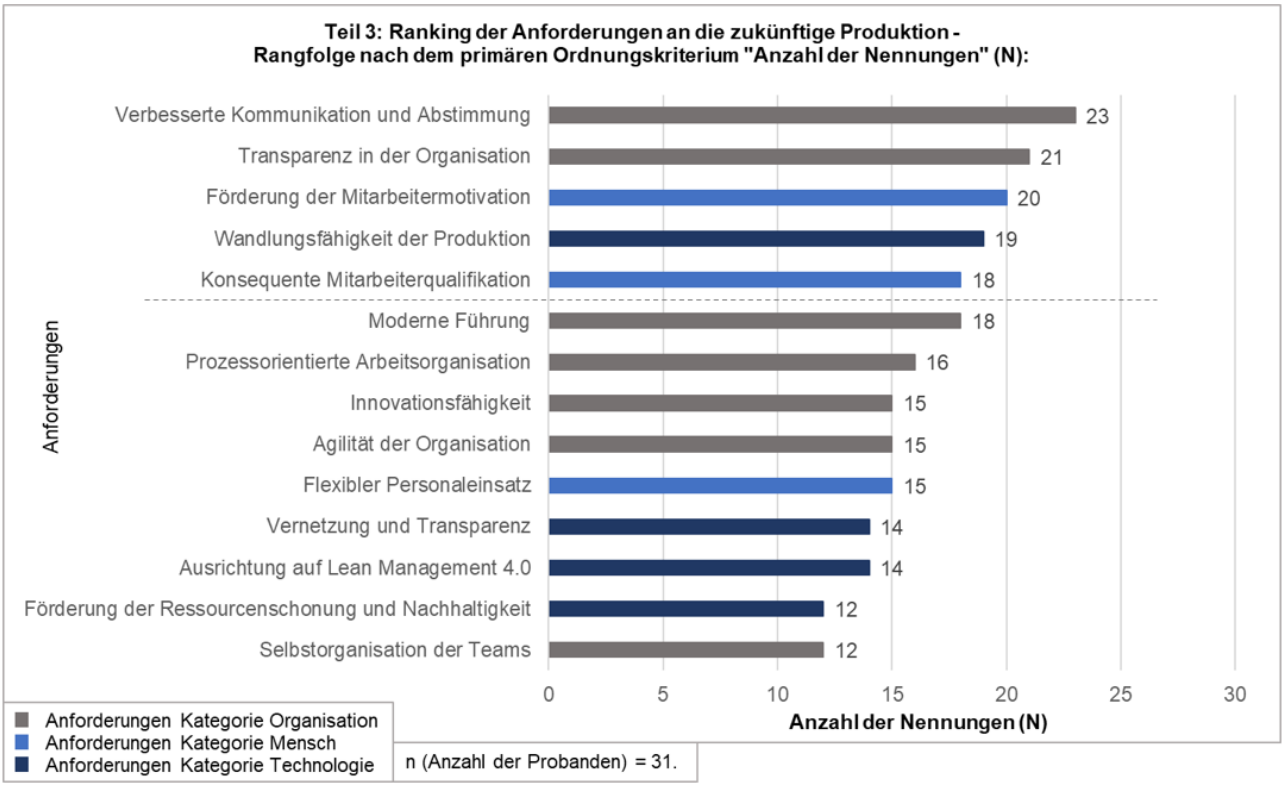


Abbildung 14: Studie Teil 3 - Vergleichendes Ranking der Produktionsanforderungen nach dem primären Ordnungskriterium Anzahl der Nennungen (N).

(digitalen) Prozessoptimierungen, Wandlungsfähigkeit, Agilität und Nachhaltigkeit im Zusammenhang mit I4.0. Der Umgang mit Daten und die Visualisierung von Informationen werden ebenfalls hervorgehoben.

Eine übergeordnete Interpretation der Ergebnisse der Studie und der Teilnehmerkommentare erfolgt in der Diskussion als abschließende Betrachtung zu dieser Arbeit (Kapitel 7.2). Als Erweiterung der quantitativen Orientierungsbefragung stellt das nachfolgende Kapitel I4.0 Anwendungsfälle zur Verfügung, welche die organisationalen Produktionsan-

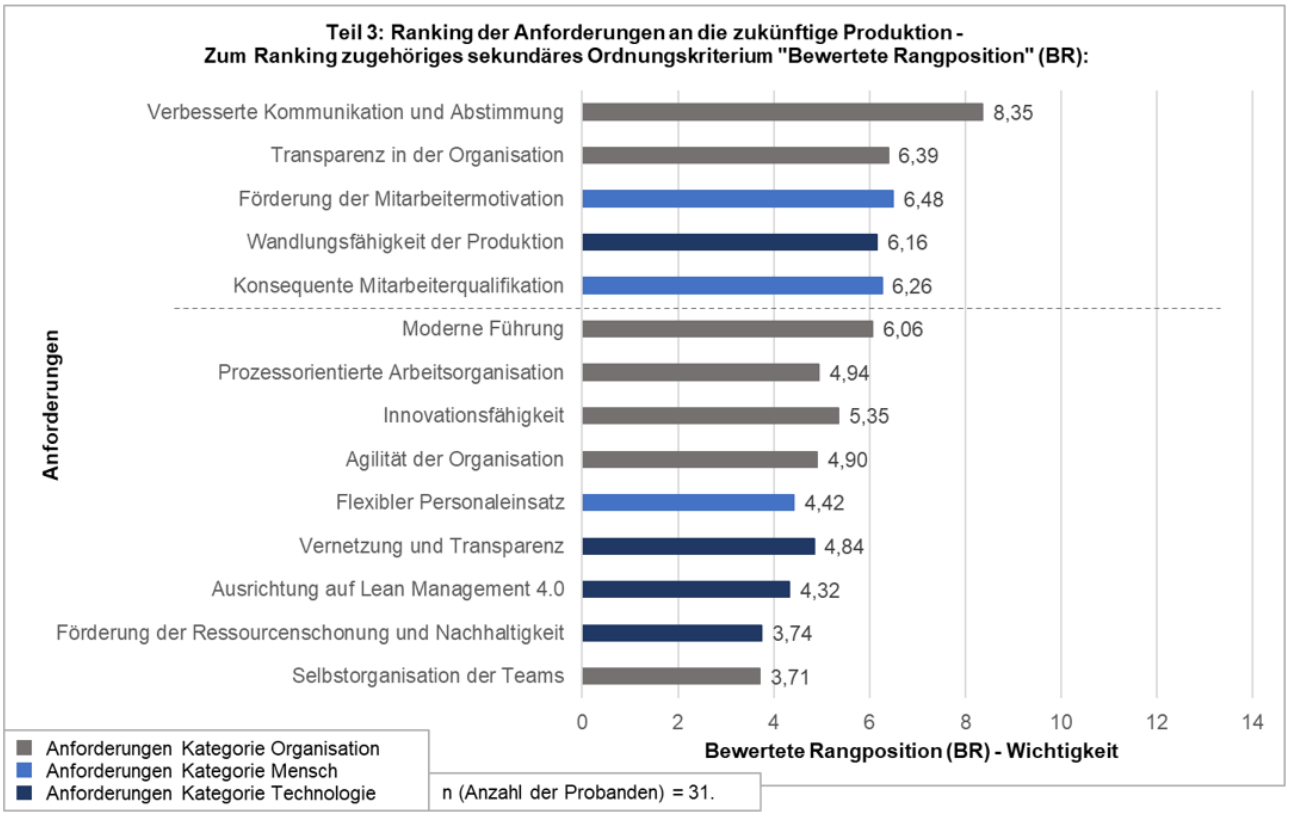


Abbildung 15: Studie Teil 3 - Vergleichendes Ranking der Produktionsanforderungen nach dem sekundären Ordnungskriterium bewertete Rangposition (BR).

Tabelle 5: Studie Teil 3 - Ergänzende Teilnehmerkommentare und weitere Anforderungen.

Unternehmen	Hierarchie/Managementebene/Funktion im Unternehmen	Weitere genannte Anforderungen an die zukünftige Produktion und Kommentare
Daimler AG	Teamleiter	Anforderung: Digitale Wertschöpfungsorientierung: Konsequenter Übertrag des Lean Managements auf Digitale Wertströme. Wo findet Wertschöpfung an Daten statt? Wo existiert Verschwendung in digitalen Prozessen? z.B. Einsatz von Process Mining (Einstufung Saklawert 7). Kommentar: Nicht alles, was mit I4.0 möglich ist, ist auch sinnvoll. Zu Anfang steht immer eine Analyse der Plain Points und eine gezielte Maßnahmenauswahl muss vorgenommen werden
Daimler AG	Teamleiter Digitale Fabrikplanung/ Produktionsplanung	Kommentar: Wandlungsfähigkeit ist für die Automobilindustrie für die nächsten Jahre überlebenswichtig. Agilität ist dafür entscheidend. Eine konsequente Ausrichtung auf schlanke, vollständig digital abgebildete Prozesse in den Administrationsbereichen ist dafür entscheidend. Alle weiteren Anforderungen dienen diesem Ziel. Die zweite Säule ist die Nachhaltigkeit. Nur eine nachhaltige Produktion ermöglicht eine individuelle Mobilität wie wir sie heute kennen.
ELABO GmbH / euromicron AG	Senior Advisor	Anforderung: Umfangreicher Ausbau von EDV und Netzwerktechnik (Einstufung Skalawert 6)
Mann+Hummel GmbH	Strategischer Einkauf - globale Leitung Materialgruppe Metallteile	Anforderung: Kosteneffizienzsteigerungen im Hinblick auf den Gesamtprozess, nicht nur Einzelschrittoptimierungen (Einstufung Skalawert 6).
PTE Inc.	Business Developer	Anforderung: Visualisierung: Darstellung relevanter Informationen, Abläufe via Social Augmented Reality & personalisiertem Dashboard. Moderne, einfach zu verwertende Darstellung von Informationen. (Einstufung Skalawert 7).

forderungen konkret umsetzen.

6. Anwendungsfälle zur Umsetzung von zukünftigen organisationalen Anforderungen an die Produktion (Technologielösungsatlas)

In den vorangestellten Teilen dieser Arbeit werden verschiedene Kennzahlen und Anforderungen der zukünftigen Produktion behandelt. Die quantitative Unternehmensbefragung liefert entsprechende Wichtigkeitsrangfolgen für diese Thematiken. Das Ziel dieses Kapitels ist es, eine Verknüpfung von Theorie und Praxis herzustellen, um festzustellen wie organisatorische I4.0 Konzepte bereits gelebt werden. Der eigens erstellte Technologielösungsatlas für innovative Organisationsformen verdeutlicht die beispielhafte Realisierung von Industrie 4.0 mit unterschiedlichen Anwendungen und technischen Lösungen.

Die Inhalte des Technologielösungsatlases sind aus mehreren Recherchequellen zusammengetragen. Die Plattform Industrie 4.0 stellt auf der stets online aktualisierten, deutschlandweiten „Landkarte Industrie 4.0“ eine umfangreiche Sammlung von Industrie 4.0 Praxisbeispielen zur Verfügung. Die Einträge zeigen Orte von I4.0 Forschungseinrichtungen und Unternehmen, welche unterschiedliche I4.0 Konzepte umgesetzt und genutzt haben.[589] In Ergänzung dazu dient die regionale Landkarte „100 Orte für Industrie 4.0 in Baden-Württemberg“ als weitere Basisquelle.[590] Zudem erwiesen sich einige Veröffentlichungen des Verbandes Deutscher Maschinen- und Anlagenbau e.V. als hilfreich.[591] Wie den Kapiteln (6.2) und (6.3) zu entnehmen ist, dienen mehrere Fraunhofer-Institute als weitere grundlegende Bezugs- und Informationsquellen.

Bei der Lösungskollektion war es problematisch, dass ein Großteil der existierenden I4.0 Lösungen auf technische Aspekte und eher weniger auf die Organisation und den Menschen abzielen. Insgesamt konnten 20 innovative, organisationsbezogene Lösungen zur Unterstützung und Optimierung arbeitsorganisatorischer Abläufe und einer agilen Shopfloor Organisation herausgefiltert werden. Im ersten Teil des Technologielösungsatlas erfolgt die Präsentation der „Top 5“ Lösungen zur Umsetzung der Anforderungen an die zukünftige Produktion. Daraufhin werden ergänzende, prioritätstechnisch leicht niedriger eingeordnete, I4.0 Lösungen vorgestellt. Diese setzen verschiedene Produktionsanforderungen in der Praxis um, jedoch nicht in dem hohen Maße, wie es die „Top 5“ der Lösungen bewerkstelligen. Zunächst werden die verschiedenen I4.0 Lösungen geclustert und eingeordnet.

6.1. Clusterung der Industrie 4.0 Lösungen

In diesem Teilkapitel werden alle selektierten I4.0 Lösungen übersichtlich dargestellt und den Produktionsanforderungen zugeordnet, die sie betreffen und umsetzen. Die Grundlage für die ausgewählte Darstellungsform der Zuordnung als Ergebnismatrix inklusive deren Gestaltung liefern zwei Veröffentlichungen der Plattform Industrie 4.0, in denen ebenfalls I4.0 Anwendungsszenarien diskutiert werden.[592] In der Ergebnismatrix (Abb. 16) wird veranschaulicht, wie stark eine jeweilige organisatorische I4.0 Lösung eines Unternehmens (Zeile), eine Anforderung an die zukünftige Produktion (Spalte) beeinflusst und realisiert. Als Maßstab zur Bewertung der Einflussstärke wird eine farbliche Abstufung (Ampelsystem) genutzt. Enthält ein Feld eine grüne Markierung, so ist ein starker Einfluss vorhanden. Eine gelbe Markierung kennzeichnet einen mittleren Einfluss der I4.0 Lösung auf die Produktionsanforderung. Ist ein Feld unausgefüllt (weiß), zeigt dies, dass kein Einfluss existent ist. Die Produktionssteuerung mithilfe von Kennzahlen wird in diesem Kapitel als fünfzehnte Anforderung an die Produktion gelistet.

Die einzelnen I4.0 Lösungen werden mit eigens erstellten Abbildungen erläutert, um die wesentlichsten Inhalte prägnant zusammenzufassen. Zur Erklärung dienen die einzelnen Inhaltsfelder „Unternehmen“, „Lösung“, „Eingesetzte Technologien“ und „Beschreibung der Lösung“. Die Felder „Nutzen und Verbesserungen“, „Veränderungen in der Organisation“ und „Kategorisierung/Zuordnung zur Tabelle“ sind zur Verdeutlichung der Bedeutung der Lösungen für den Kontext dieser Arbeit maßgeblich. Sie stellen den nötigen verknüpfenden Kontext her. Ist bei der Kategorisierung eine Anforderung kursiv geschrieben, steht dies für einen mittleren Einfluss (gelbes Matrixfeld). Es folgen im ersten Teil der Sammlung die wichtigsten I4.0 Lösungen.

6.2. Technologielösungsatlas - „Top 5“ organisatorische Industrie 4.0 Lösungen

Alle in diesem Teilkapitel präsentierten praktischen I4.0 Anwendungen erhalten eine Platzierung in der Auswahl der fünf geeignetsten Lösungen, da sie entweder generell eine Vielzahl der Produktionsanforderungen umsetzen, oder explizit und detailliert einzelne der dedizierten „Top 7“ Anforderungen der zukünftigen Produktion ansprechen. Diese fünf, zur Umsetzung der organisationalen Produktionsanforderungen am besten geeigneten, I4.0 Lösungen werden nachfolgend durch 15 weitere Anwendungen ergänzt.

[589] vgl. Plattform Industrie 4.0 (2018b), S. 1f, URL siehe Literaturverzeichnis.

[590] Allianz Industrie 4.0 Baden-Württemberg (2020), URL siehe Literaturverzeichnis.

[591] Forum Industrie 4.0 des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA) und Forschungskuratorium Maschinenbau e.V. (FKM) (2017); Forum Industrie 4.0 des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA) (2017); Forum Industrie 4.0 des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA) (2018).

[592] vgl. Arbeitsgruppe Forschung und Innovation der Plattform Industrie 4.0 (2016), S. 36 und Fay, Gausemeier und ten Hompel (2018), S. 10.

Firmen/Lösungen und Anforderungen		Selbstorganisation der Teams	Prozessorientierte Arbeitsorganisation	Agilität der Organisation	Innovationsfähigkeit	Moderne Führung	Transparenz in der Organisation	Verbesserte Kommunikation und Abstimmung	Flexibler Personaleinsatz	Förderung der Mitarbeitermotivation	Konsequente Mitarbeiterqualifikation	Förderung der Resourcenschonung und Nachhaltigkeit (RS und NH)	Wandlungsfähigkeit der Produktion	Ausrichtung auf Lean Management 4.0	Vernetzung und Transparenz	Kennzahlen
Nr.		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Fraunhofer IAO, „KapaflexCy“ - Selbstorganisierte Personalkapazitätsflexibilität für die I4.0															
2	Fraunhofer IPK, „IWEPRO“ - Intelligente selbstorganisierte Werkstattproduktion															
3	Crevis, Personalqualifikation															
4	BLG LOGISTICS/Engelbert Strauß, Industrie 4.0 Hallenlogistiksystem															
5	Technotrans, Visuelle Kommunikation über Ampelsystem															
6	Peakboard, Visualisierungsboard															
7	FESTOOL, Digitales Lean Shopfloor Management															
8	Rota Yokogawa, Digitaler Echtzeit-Produktionsüberblick.															
9	MAN Energy Solutions, Kommissionierung mit Industrie 4.0 Technologien															
10	ESTA Apparatebau, Hallenlüftungssysteme 4.0															
11	KINEMIC, Gestensteuerung mit Wearables															
12	Fraunhofer IAO/aucobo, Mehrmaschinen- Bedienung mit Smartwatches															
13	Picavi, Pick-by-Vision und Kommissionierung mit Datenbrillen															
14	BMW Group, Datenbrille in der Qualitätssicherung															
15	Fraunhofer IAO, KPI Dashboard															
16	Phoenix Contact, Wandlungsfähige Fertigung															
17	.steute, Intelligentes Qualifikations- und Assistenzsystem															
18	TEKA, Intelligentes Raumüberwachungssystem „AIRTRACKER“															
19	Fraunhofer IAO, Qualifizierung am Arbeitsplatz															
20	Schaeffler, Instandhaltung 4.0 „SmartCheck“															

Abbildung 16: Eigene Darstellung - Clusterung und Zuordnung der Industrie 4.0 Anwendungen des Technologielösungsatlases.

Abbildung 17: Top 5: Eigene Darstellung Technologielösungsatlas Nr.1 - Fraunhofer IAO/BorgWarner, Flexibler und selbstorganisierter Personaleinsatz mit „KapaflexCy“.
Quellen: Vgl. Plattform Industrie 4.0 und Fraunhofer-Institut für Arbeitswirtschaft und Organisation (2020), URL siehe Literaturverzeichnis, Fraunhofer-Institut für Arbeitswirtschaft und Organisation (2020c), S. 1f., URL siehe Literaturverzeichnis, Allianz Industrie 4.0 Baden-Württemberg und BorgWarner Ludwigsburg GmbH (2020), URL siehe Literaturverzeichnis und Heller (2015), S. 34f.

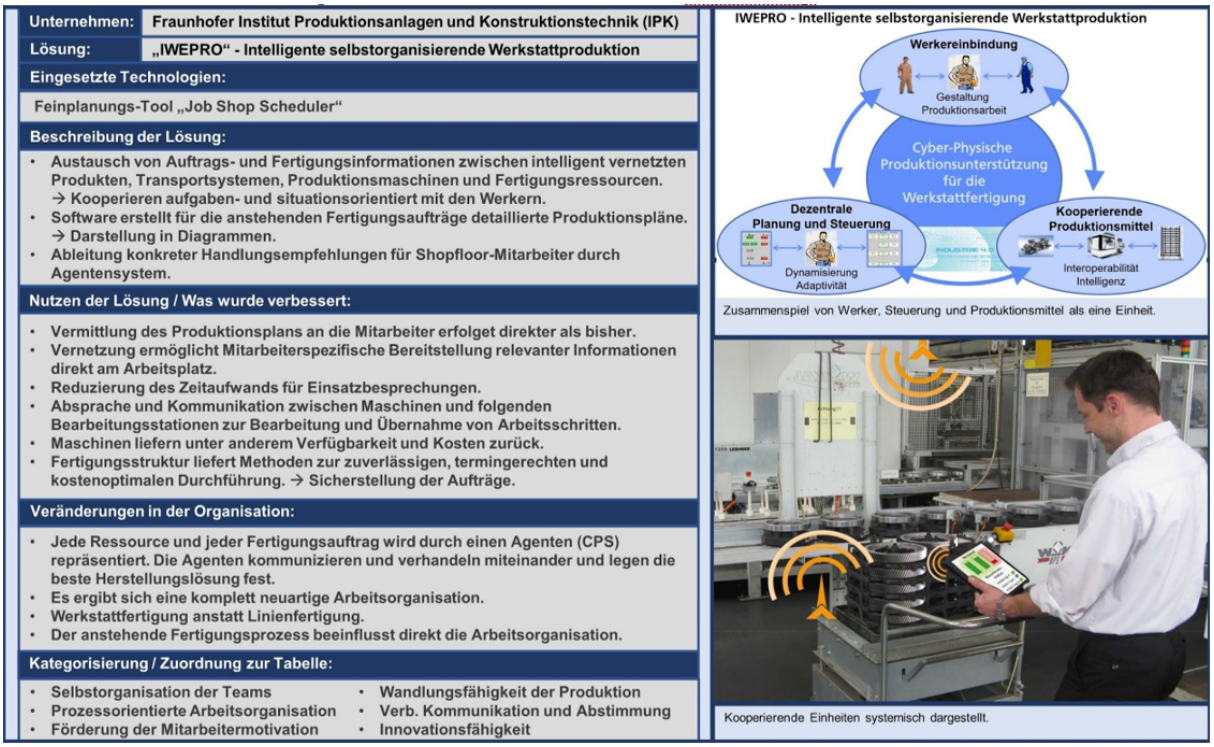


Abbildung 18: Top 5: Eigene Darstellung Technologielösungsatlas Nr.2 - Fraunhofer IPK, Intelligente selbstorganisierte Werkstattproduktion „IWEPRO“.
Quellen: Vgl. Forum Industrie 4.0 des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA) und Forschungskuratorium Maschinenbau e.V. (FKM) (2017), S. 65, Fraunhofer-Institut für Produktionsanlagen und Konstruktionstechnik (2020), S. 1f, URL siehe Literaturverzeichnis und Krüger (2013), S. 10f.

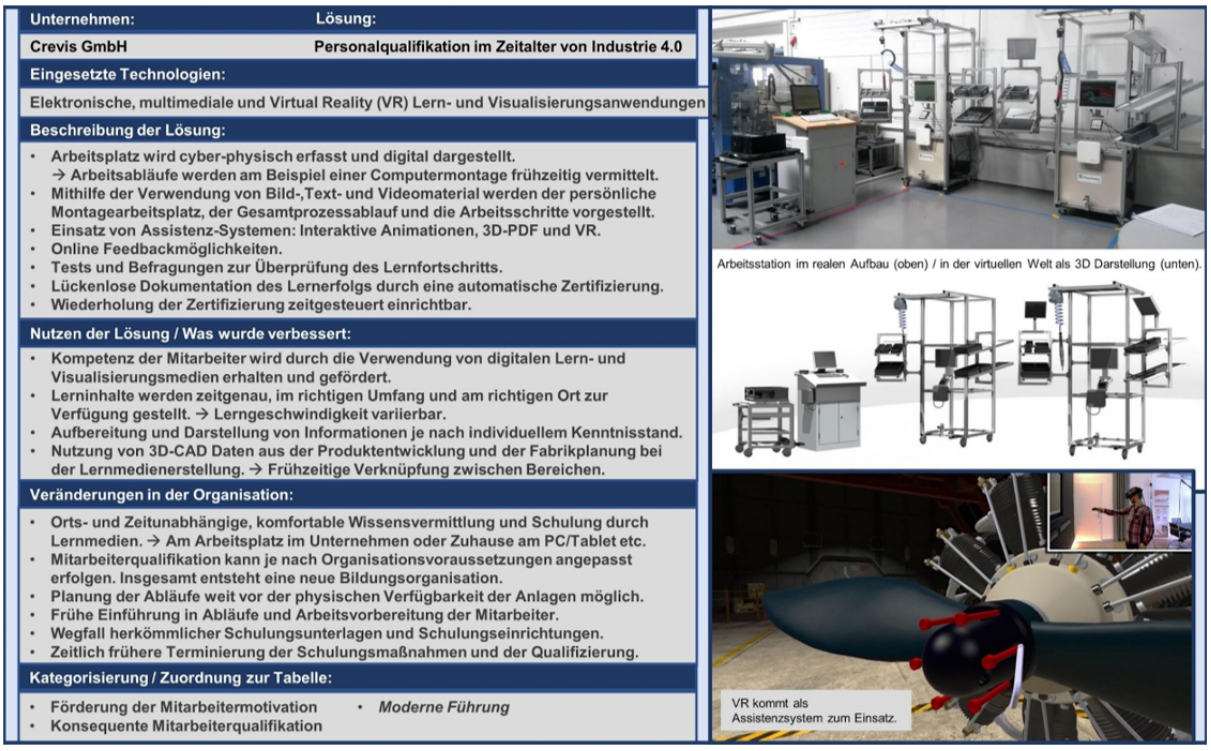


Abbildung 19: Top 5: Eigene Darstellung Technologielösungsatlas Nr.3 - Crevis, Personalqualifikation.
Quelle: Vgl. Allianz Industrie 4.0 Baden-Württemberg (2020), URL siehe Literaturverzeichnis.

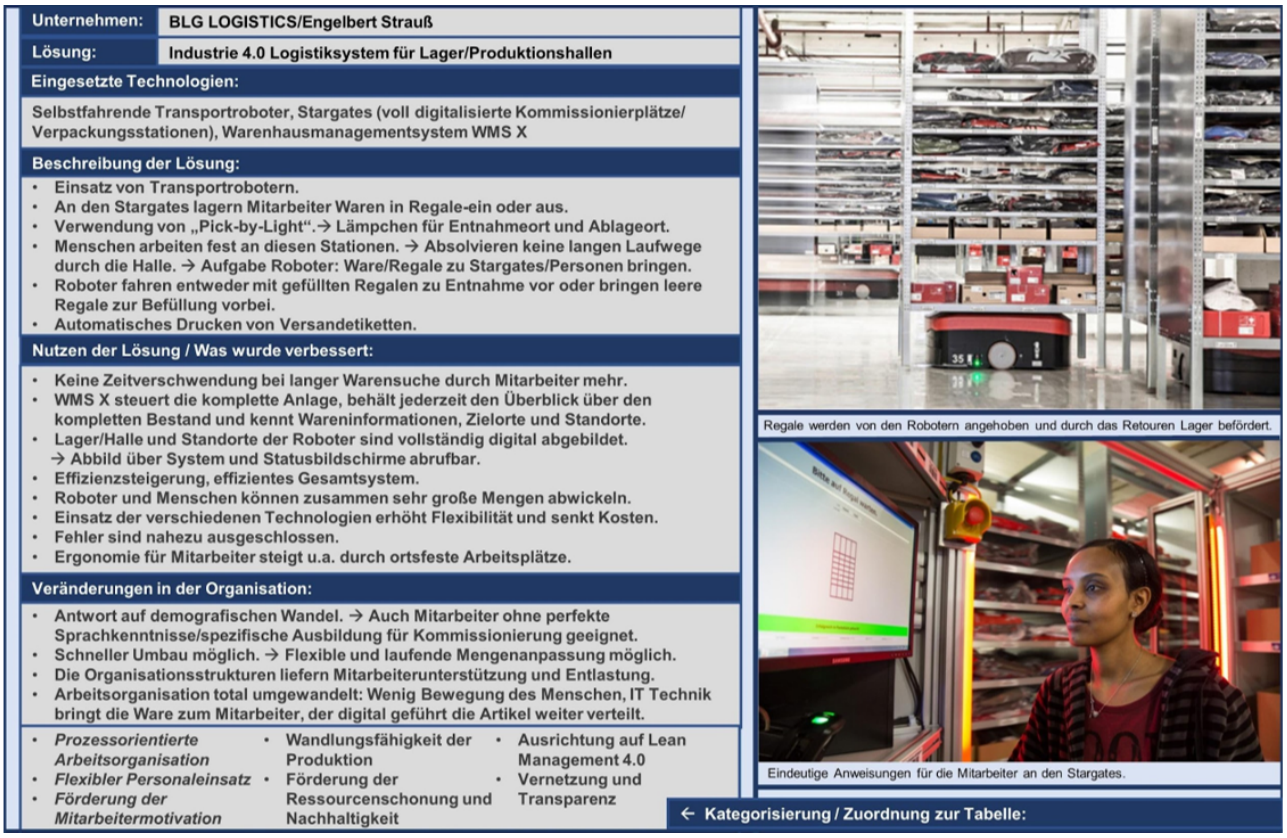


Abbildung 20: Top 5: Eigene Darstellung Technologielösungsatlas Nr.4 - BLG LOGISTICS/Engelbert Strauß, I4.0 Hallenlogistiksystem.
Quelle: Vgl. Raveling (2016), URL siehe Literaturverzeichnis.

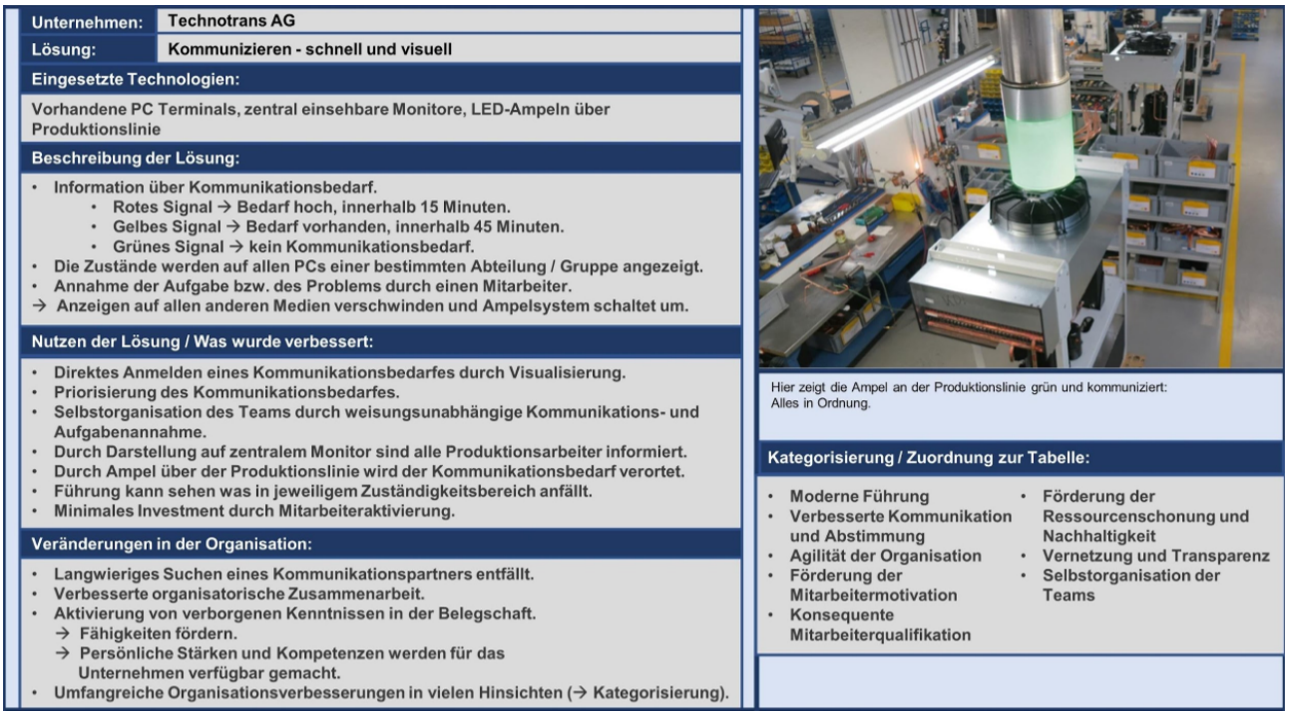


Abbildung 21: Top 5: Eigene Darstellung Technologielösungsatlas Nr.5 - Technotrans, Visuelle Kommunikation über Ampelsystem.
Quelle: Vgl. Meudt et al. (2018), S. 33.

6.3. Technologielösungsatlas – Weitere organisatorische Industrie 4.0 Lösungen

Zusammenfassend stellt der zweiteilige Technologielösungsatlas eine grundlegende Sammlung organisatorische Aspekte verbessernder Industrie 4.0 Lösungen dar. Das Hauptziel der Aufstellung ist die Verdeutlichung, in welche Richtung die Gestaltung der Organisation der Produktion aktuell und in Zukunft verlaufen kann. Einige der vorgestellten Technologien wurden schon vor längerer Zeit als Pilotprojekte durchgeführt und bereits in Unternehmen implementiert, während die Verwendung anderer aufgeführter Lösungen noch eine Zukunftsvision und Leitidee darstellt.

Grundsätzlich ist bei der Implementierung und Nutzung von I4.0 Anwendungen aller Art immer eine konkrete Überprüfung der Ausgangssituation und der individuellen unternehmerischen Rahmenbedingungen zu vollziehen. In jedem Fall sollte ein Sinn zur Einführung und Verwendung einer bestimmten Industrie 4.0 Lösung vorhanden sein, da andernfalls eine Ausnutzung der weitreichenden positiven Wirkungen nicht geschehen kann. Nur wenn der Einsatz von I4.0 Anwendungen analysiert und hinterfragt wird, können Schwachpunkte der bestehenden Produktion tatsächlich ausgemerzt werden und Verbesserungen resultieren.[593] Die wesentlichste zu bewältigende Aufgabe für Unternehmen besteht schlussendlich darin, bei der Auswahl von I4.0 Lösungen überlegt vorzugehen und die optimal passenden Anwendungen zu identifizieren.[594] In der abschließenden Betrachtung werden die in dieser Arbeit behandelten Inhalte diskutiert, erweitert und rekapituliert. Zunächst dient die Vorstellung einer idealtypischen arbeitsorganisatorischen Ablauforganisation der zukünftigen Produktion dazu, Verknüpfungen zwischen den vorangestellten Kapiteln und Erkenntnissen herzustellen.

Hinweis: Aufgrund der besseren Lesbarkeit im Publikationslayout von Junior Management Science, werden zuerst die Abbildungen 22 bis 36 auf den folgenden Seiten dargestellt.

[593] vgl. Dobrowolski (2016), URL siehe Literaturverzeichnis.
[594] vgl. Stief (2018), S. 245.

Abbildung 22: Eigene Darstellung Technologielösungsatlas Nr.6 - Peakboard, Visualisierungsboard.
Quelle: Vgl. Allianz Industrie 4.0 Baden-Württemberg (2020), URL siehe Literaturverzeichnis.

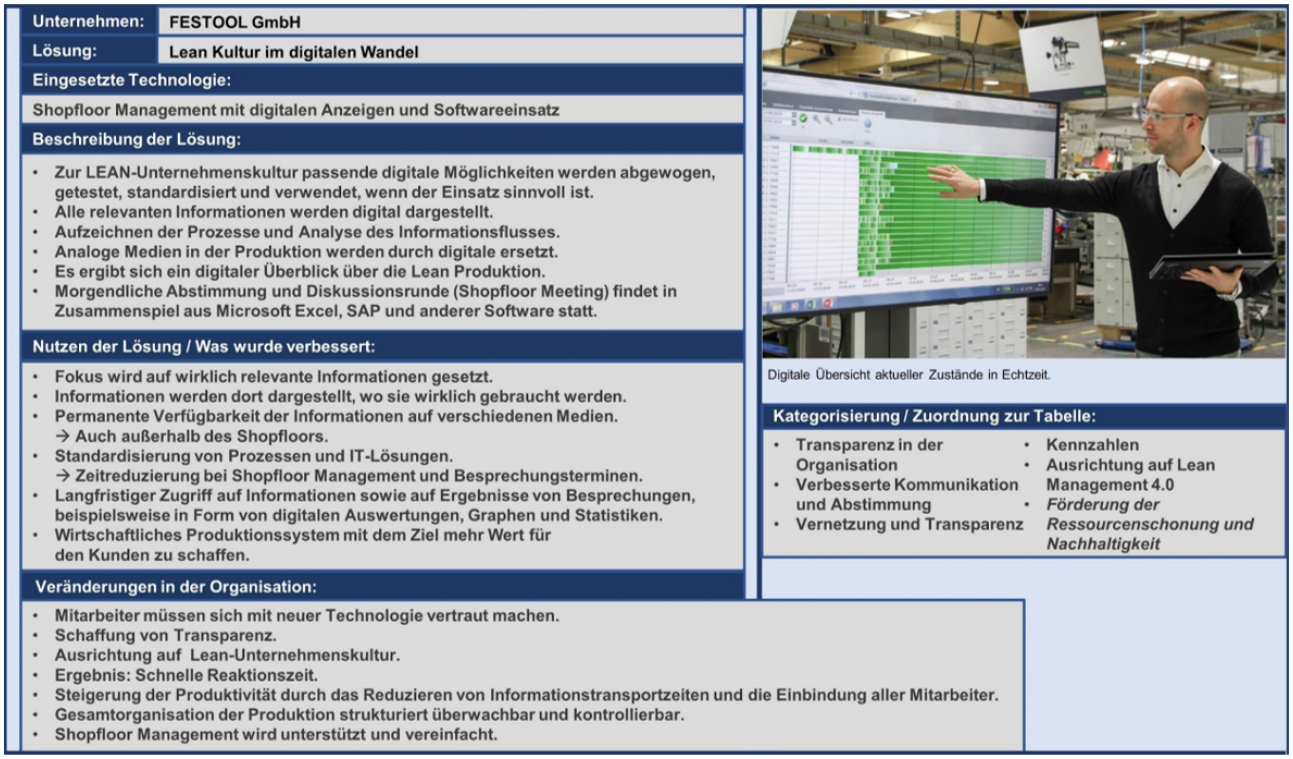


Abbildung 23: Eigene Darstellung Technologielösungsatlas Nr.7 - FESTOOL, Digitales Lean Shopfloor Management.
Quelle: Vgl. Meudt et al. (2018), S. 35.

Unternehmen: Rota Yokogawa GmbH & Co. KG

Lösung:

Digitaler Echtzeit-Produktionsüberblick und termingetreue Versandsteuerung mittels digitalem Pull-System

Eingesetzte Technologien:

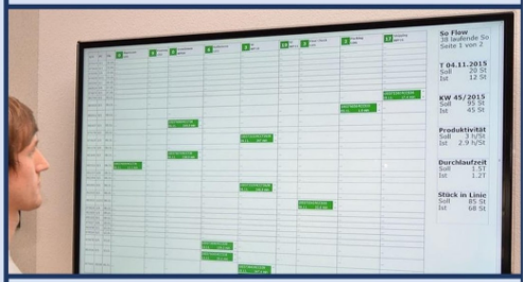
Visualisierung auf Großbildschirm oder Arbeitsplatzbildschirm
Auswertungsprogramm von Echtzeitdaten

Beschreibung der Lösung:

- Verwendung von variantenreichen Durchflussmessgeräten („ROTAMASS“).
- Produktion, Logistik und Versand sorgen unter Verwendung von Echtzeitdaten gemeinsam für Feinsteuerung von Aufträgen und termingerechte Auslieferung.
- Versand: Generell ist je nach Größe des produzierten Gutes eine unterschiedliche Packdauer und ein anderer Platzbedarf für die Verpackungsarbeiten notwendig.
  → Die I4.0 Lösung zeigt alle sich in den letzten 7 Produktionsschritten befindlichen Güter mit deren Größe und aktuellem Standort.
- Produktionssituation kann digital eingesehen und gesteuert werden.

Nutzen der Lösung / Was wurde verbessert:

- Bessere Planung der Mitarbeiterauslastung und der personellen Kapazitäten im Versand, an der Produktionslinie und in der Logistik.
- Flächenbedarf im Versand wird durch optimalere Nutzung der vorhandenen Verpackungs- und Lagerflächen deutlich reduziert.
- Zielgerichteterer Personalressourceneinsatz. → Größere Produkte - mehr Mitarbeiter.
- Fehlerrate von Montagetätigkeiten verringert. → Qualität nachhaltig verbessert.
- Detailplanung der Prozesse (des Versands) möglich.



Digitale Transparenz für die Mitarbeiter:
Alle Aufträge in der Produktion sind zu jedem Zeitpunkt einsehbar.

Kategorisierung / Zuordnung zur Tabelle:

- Selbstorganisation der Teams
- Agilität der Organisation
- Transparenz in der Organisation
- Verbesserte Kommunikation und Abstimmung
- Ausrichtung auf Lean Management 4.0
- Vernetzung und Transparenz

Veränderungen in der Organisation:

- System prüft, ob der Werker berechtigt ist an einem Arbeitsplatz zu arbeiten.
  → Es meldet dem berechtigten Mitarbeiter sofort, wenn sich eine Änderung der Vorschriften bzw. des Arbeitsprozesses am Arbeitsplatz ergeben hat.
  → Einleitung von Kurzeinweisungen über System bei Bedarf (Mitarbeiterschulung).
- (Online) Schulungen der Mitarbeiter sind an ständig ändernde Arbeitsschritte angepasst.
- Umfangreiche, bereichsübergreifende Prozessoptimierung.
- Effizientere Verknüpfung von Logistik und Produktion.
- Mitarbeiterunterstützung in der gesamten Organisation.
- Schaffung von Transparenz und Kontrolle z.B. über Zeitpunkte. → Prognosen möglich.

Abbildung 24: Eigene Darstellung Technologielösungsatlas Nr.8 - Rota Yokogawa, Digitaler Echtzeit-Produktionsüberblick.
Quellen: Vgl. Picavi GmbH (2020c), URL siehe Literaturverzeichnis und Allianz Industrie 4.0 Baden-Württemberg und Rota Yokogawa GmbH & Co. KG (2020), URL siehe Literaturverzeichnis.

Unternehmen: MAN Energy Solutions SE

Lösung:

Einsatz von Datenbrillen im Lager für Kommissionierung und Ansteuerung Maschinen

Eingesetzte Technologien: Datenbrillen, Datenhandschuh mit integriertem Scanner

Beschreibung der Lösung:

- Mit Datenbrillen unterstützte Kommissionierung am automatischen Kleinteilelager.
- Kommunikation der Brille über eine implementierte App mit SAP-System.
- Auf der Datenbrille werden Ein- und Auslagervorgänge und Aufgaben angezeigt.
- Zudem sind Fahrbefehle für die „Lean-Lifte“ zum Warentransport stets sichtbar.
- Die Lagergeräte sind direkt über die Datenbrille ansteuerbar (z.B. Lagerlifte).
- Modernes Bedienkonzept enthält auch Datenhandschuh mit integriertem Scanner.
- Daten des mobilen Scanners werden in Echtzeit an das SAP übertragen.

Nutzen der Lösung / Was wurde verbessert:

- Arbeiter haben immer beide Hände zum Kommissionieren der Teile frei.
- Bei Projektimplementierung wurden Arbeitsprozesse hinterfragt und konsequent nach Lean Prinzipien ausgerichtet.
- Generelle Reduzierung der Anzahl von Eingabegeräten in Produktion und Logistik durch Harmonisierung von Arbeitssystem, Hardware und Software bei dieser Lösung.
- Optimierung intuitiver Arbeitsprozesse mit Wearables.
  → Auch durch die Nutzung der entwickelten SAP-App.

Veränderungen in der Organisation:

- Mehr produktive Arbeit durch Beseitigung zeitraubender Organisationsprozesse und durch Übernahme von Aufgaben durch Datenbrille, Scanner und Software.
- Massive Zeiteinsparungen.
- Organisatorische und arbeitstechnische Unterstützung für die Mitarbeiter.
- Mitarbeiter wurden bei Planung, Prozessgestaltung und technologischer Umsetzung explizit eingebunden. → Förderung von Innovationsfähigkeit und innovativem Denken.
- Arbeitsabläufe werden tiefgreifend verbessert.
  → Grund: Die Wearables für Produktion und Logistik bieten viele Potentiale, um Arbeitsabläufe zu optimieren und digital zu gestalten.

Kategorisierung / Zuordnung zur Tabelle:

- Innovationsfähigkeit
- Ausrichtung auf Lean Management 4.0
- Vernetzung und Transparenz
- *Prozessorientierte Arbeitsorganisation*

Mitarbeiter aus der Logistik mit Datenbrille.

Mitarbeiter aus der Logistik bei der Einlagerung.

Abbildung 25: Eigene Darstellung Technologielösungsatlas Nr.9 - MAN Energy Solutions, Kommissionierung mit I4.0 Technologien.
Quelle: Vgl. Picavi GmbH (2020c), URL siehe Literaturverzeichnis.

Abbildung 26: Eigene Darstellung Technologielösungsatlas Nr.10 - ESTA Apparatebau, Hallenlüftungssysteme 4.0.
Quellen: Vgl. Forum Industrie 4.0 des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA) (2017), S. 20f. und ESTA Apparatebau GmbH & Co. KG (2020), URL siehe Literaturverzeichnis.

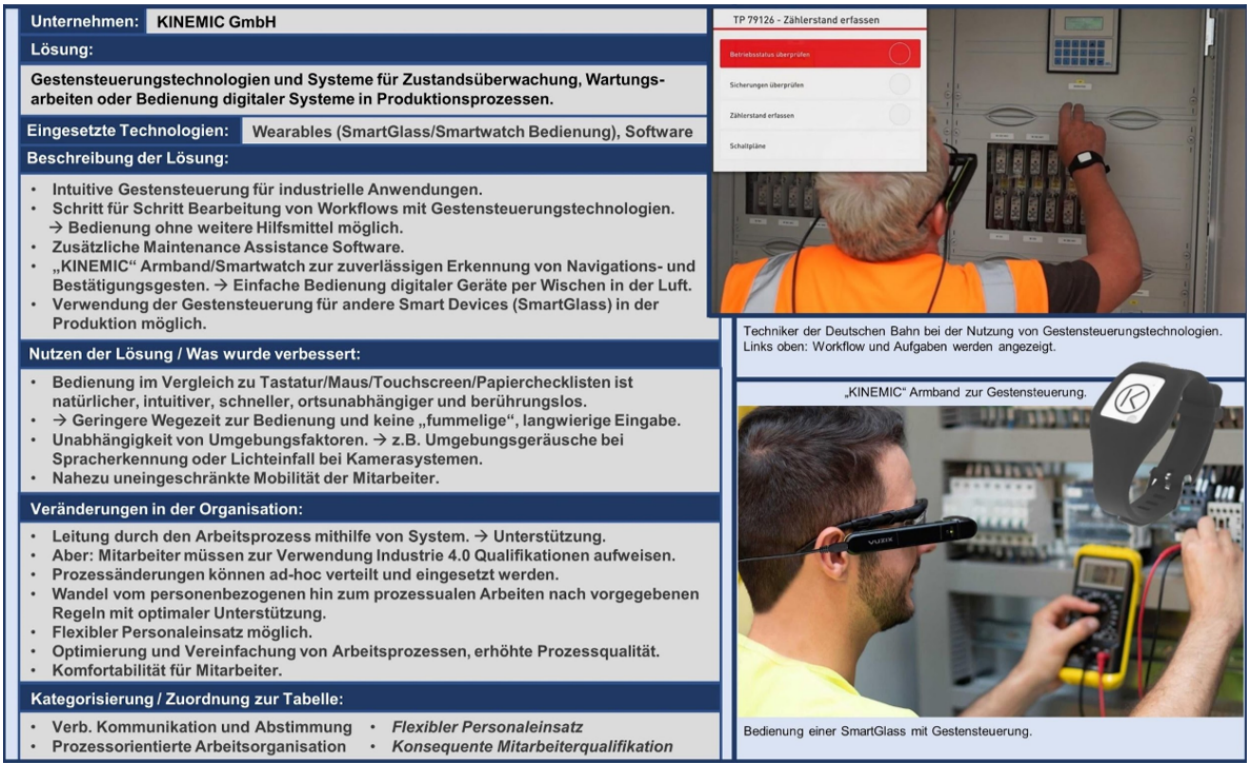


Abbildung 27: Eigene Darstellung Technologielösungsatlas Nr.11 - KINEMIC, Gestensteuerung mit Wearables.
Quellen: Vgl. Picavi GmbH (2020b), URL siehe Literaturverzeichnis, Kinemic GmbH (2020a), URL siehe Literaturverzeichnis und Kinemic GmbH (2020b), URL siehe Literaturverzeichnis.

Unternehmen:

Fraunhofer Institut für Arbeitswirtschaft und Organisation (IAO)/aucobo GmbH

Lösung:	**Eingesetzte Technologien:**
Mobile Mehrmaschinen-Bedienung	Smartwatches

Beschreibung der Lösung:

- Aufgabenmanagement und Austausch über Smartwatches oder andere denkbare mobile Endgeräte.
- Kommunikation zwischen mobilen Mitarbeitern wird ermöglicht.
- Verbindung von Live-Maschinendaten mit dem Arbeiter.
- System meldet Bedarf einer Maschinenbedienung über die Smartwatch.
  → Arbeiter erledigt Bedarf ad-hoc oder gibt den Aufruf an eine Gruppe oder einen Kollegen per Watch weiter.
- Annahme der Aufgabe von Mitarbeiter(n) durch Quittieren.
- In Echtzeit erfasster Kontext aus der Produktion wird zur Anpassung und Optimierung von Produktionsprozessen genutzt.

Nutzen der Lösung / Was wurde verbessert:

- Einfache Smartwatch Konfiguration ermöglicht eigenständige Prozessoptimierung.
- Starke Eliminierung von Mitarbeiterwartezeiten und Stillstandzeiten der Maschinen.
- Fehler können schneller erkannt werden.
- Höhere Mitgestaltungsmöglichkeiten durch Mitarbeiter und einfachere Umsetzung von Erfahrungen.
- Effizienzsteigerungen und geringere Kosten.
- Vereinfachte Kollaboration von Mitarbeitern, da Informationen auf Mitarbeiter angepasst verteilt werden können.

Veränderungen in der Organisation:

- Werker, Planer und Meister können in Absprache selbst Automatisierungen vornehmen und flexibel reagieren.
- Steigerung der Motivation der Mitarbeiter durch hohes Maß an Selbstbestimmung.
- Erhebliche Abstimmungsvereinfachung.
- Bessere Planung, Steuerung und Aufgabenverteilung vorteilhaft für die Organisation.
- Große Auswirkungen auf die Organisation (siehe Kategorisierung).

Kategorisierung / Zuordnung zur Tabelle:

- Selbstorganisation der Teams
- Agilität der Organisation
- Verb. Kommunikation und Abstimmung
- Flexibler Personaleinsatz
- Vernetzung und Transparenz
- *Transparenz in der Organisation*

Die Smartwatch als zentraler Aufgabenmanager.

© Bildausschnitte aus Video.

Auf der Smartwatch erscheinen die Bedienungshinweise zur Maschine mit allen notwendigen Parametern und sonstigen Angaben.

Abbildung 28: Eigene Darstellung Technologielösungsatlas Nr.12 - Fraunhofer IAO/aucobo, Mobile Mehrmaschinen-Bedienung mit Smartwatches.
Quellen: Vgl. Fraunhofer-Institut für Arbeitswirtschaft und Organisation (2020e), S. 1, URL siehe Literaturverzeichnis und Fraunhofer-Institut für Arbeitswirtschaft und Organisation (2019), URL siehe Literaturverzeichnis.

Unternehmen:	**Lösung:**
Picavi GmbH	**Pick-by-Vision, Augmented Reality unterstützte Kommissionierung**

Eingesetzte Technologien:

Verschiedene Datenbrillen mit zugehöriger Software

Beschreibung der Lösung:

- Mit Datenbrillen unterstützte Kommissionierung.
- Scanner, Kameras und Mikrofon in Datenbrille integriert.
- Brille kommuniziert mit Lager- und ERP-System.
- Papiergebundene Pickliste wird ersetzt durch Display-to-go.
- Datenbrillen scannen alle 1D- und 2D-Barcodes.
  → kein händischer Scan mehr nötig.
- Optimierter Workflow wird durch angeschlossene Systeme sichergestellt.
  → Software und Datenbrillen sind aufeinander abgestimmt.
- Produktions-/ Logistikarbeitern wird beim Picking umfangreich assistiert.
- Zusätzliche Serviceberatung und Problemanalyse als Dienstleistung.

Nutzen der Lösung / Was wurde verbessert:

- Schneller und zuverlässiger Materialfluss.
- Vorteile des Kommissioniersystems:
  → Flexibel einsetzbar, intuitive Bedienung, ortsunabhängig, Echtzeitvorgänge.
- Arbeiter haben immer beide Hände zur Teilekommissionierung frei.
- Steigerung der Produktivität und ca. 30% Zeitersparnis.
- Starke Fehlerreduktion durch Gewährleistung eines fehlerfreien Pickings.
- Mitarbeiterschonung: Das schwere Erreichen von Barcodes und die verbundene Anstrengung werden beseitigt.→ Visueller Datenbrillenscan erreicht alle Stellen.

Veränderungen in der Organisation:

- Kürzere Einsatzzeiten pro Aufgabe. → Schnellere Erledigung möglich.
- Wegfall langer Schulungsreihen. → Sehr kurze Eingewöhnung für Mitarbeiter.
- Kommissionieren mit freien Händen → Effektivere Arbeit.
- Schlankere Prozesse.
- Optimierung der gesamten Intralogistik. → Organisation wird vereinfacht.
- Einfaches Handling führt zu hoher Akzeptanz und steigert die Mitarbeiterzufriedenheit.

Der nächste Pick wird angezeigt, immer zwei freie Hände.

Kategorisierung / Zuordnung zur Tabelle:

- Innovationsfähigkeit
- Ausrichtung auf Lean Management 4.0
- Vernetzung und Transparenz
- *Förderung der Mitarbeitermotivation*

Abbildung 29: Eigene Darstellung Technologielösungsatlas Nr.13 - Picavi, Pick-by-Vision und Kommissionierung mit Datenbrillen.
Quelle: Vgl. Picavi GmbH (2020a), URL siehe Literaturverzeichnis, Kinemic GmbH (2020a), URL siehe Literaturverzeichnis und Picavi GmbH (2020c), URL siehe Literaturverzeichnis.

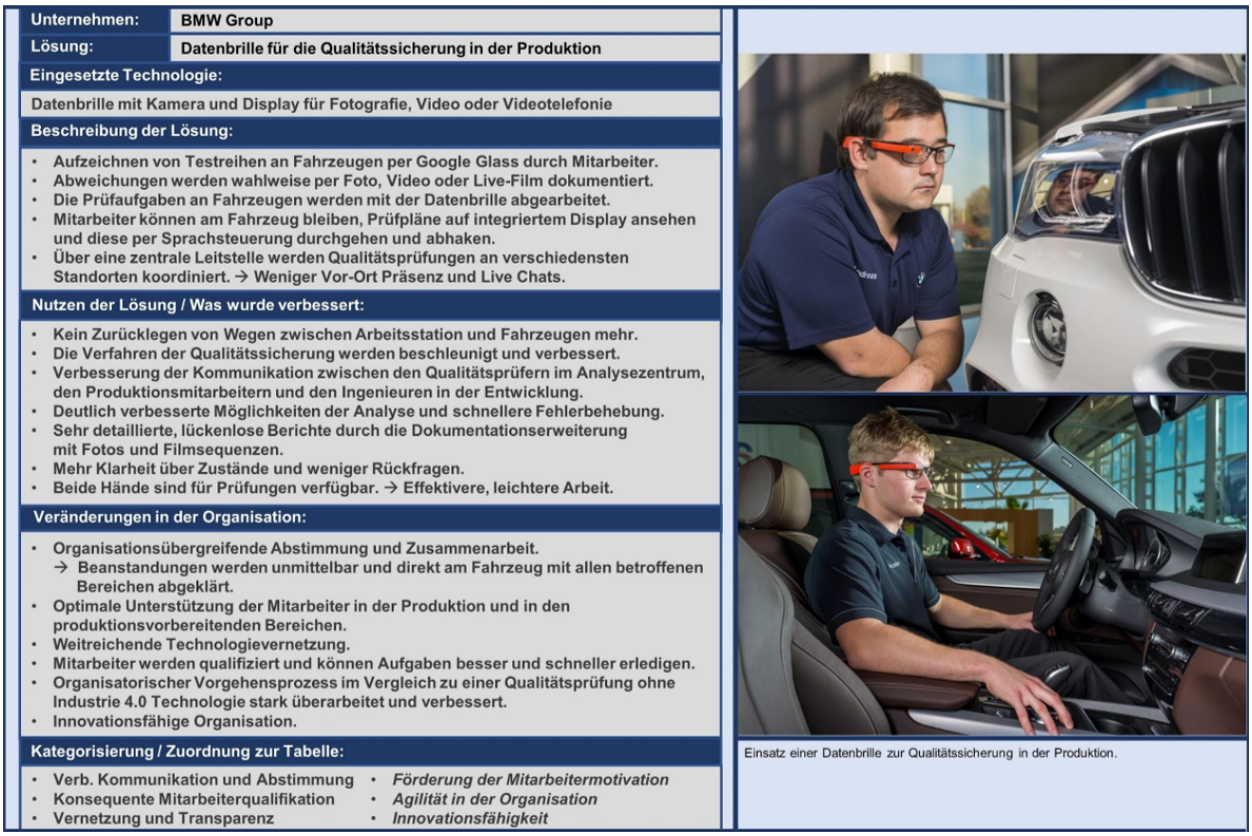


Abbildung 30: Eigene Darstellung Technologielösungsatlas Nr.14 - BMW Group, Datenbrille in der Qualitätssicherung.
Quellen: Vgl. Lucks (2017b), S. 384f. und BMW Group (2020), URL siehe Literaturverzeichnis.

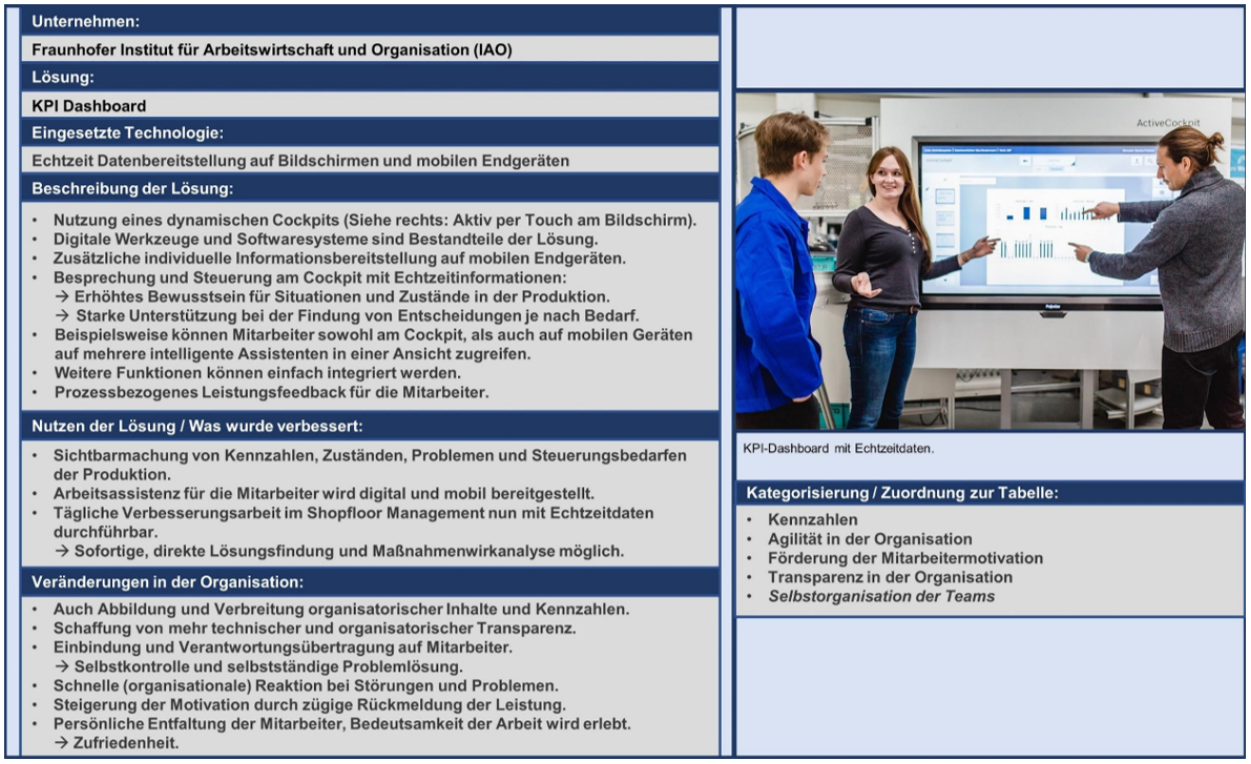


Abbildung 31: Eigene Darstellung Technologielösungsatlas Nr.15 - Fraunhofer IAO, KPI Dashboard.
Quelle: Vgl. Fraunhofer-Institut für Arbeitswirtschaft und Organisation (2020b), S. 1, URL siehe Literaturverzeichnis.

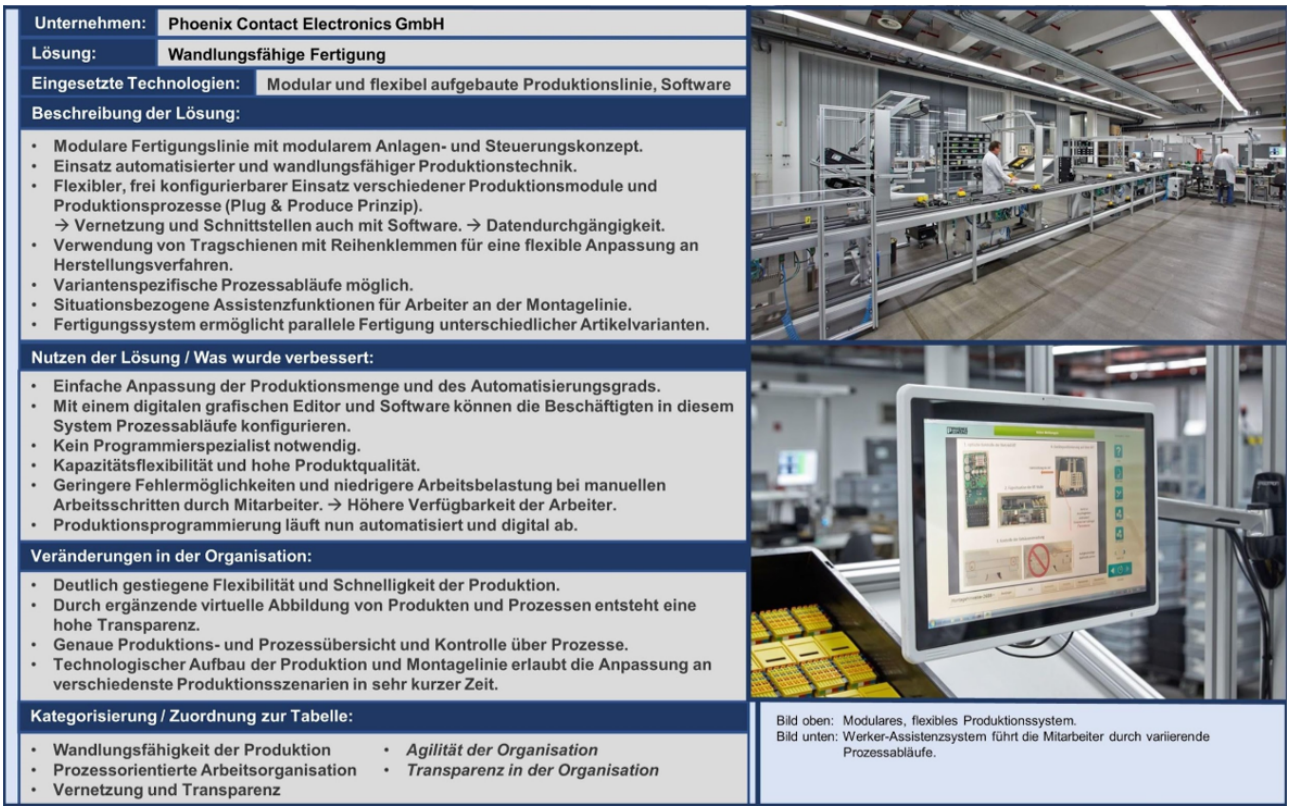


Abbildung 32: Eigene Darstellung Technologielösungsatlas Nr. 16 - Phoenix Contact, Wandlungsfähige Fertigung.
Quellen: Vgl. Plattform Industrie 4.0 und Pheonix Contact Electronics GmbH (2020), URL siehe Literaturverzeichnis, Forum Industrie 4.0 des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA) (2016), S. 39 und Klaes und Schreiber (2015), S. 16ff.

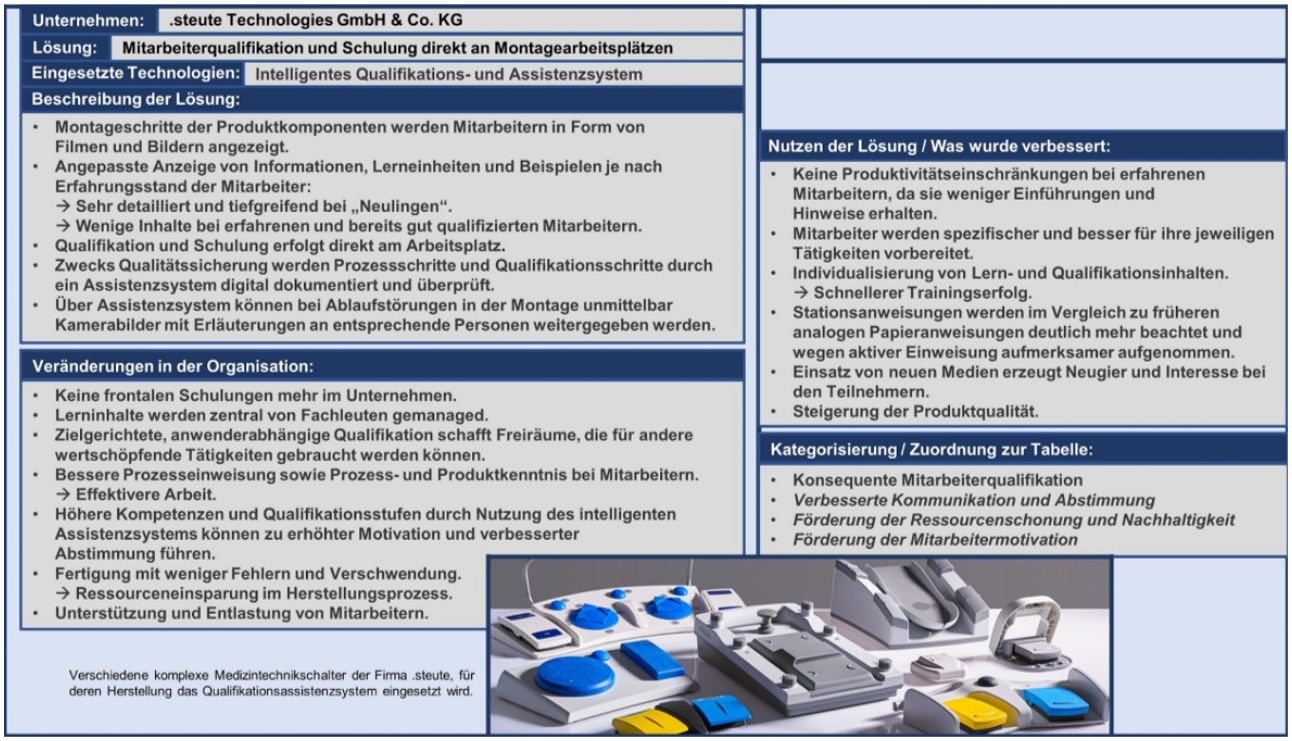


Abbildung 33: Eigene Darstellung Technologielösungsatlas Nr.17 - .steute, Intelligentes Qualifikations- und Assistenzsystem.
Quellen: Vgl. Dumitrescu und Marquardt (2017), S. 648f. und Steute Technologies GmbH & Co. KG (2020), URL siehe Literaturverzeichnis.

Unternehmen:	**TEKA Absaug- und Entsorgungstechnik GmbH**
Lösung:	**Intelligentes Raumüberwachungssystem und digitale Gesundheitsprävention für bessere Arbeitsplatzkonditionen**
Eingesetzte Technologien:	**Netzwerkfähige Luftmessstationen („AIRTRACKER"), IT- Infrastruktur (Software, Schnittstellen, leistungsstarkes Netzwerk)**

Beschreibung der Lösung:

- Das Raummonitoringsystem „Airtracker" wird in Form von smarten Steuerungsboxen unter anderem in Produktionshallen eingesetzt.
- Multisensorsystem überwacht kontinuierlich Raumluftqualität (z.B. Temperatur, Feinstaubbelastung und Luftfeuchtigkeit), Lärmbelastung sowie Anlagenstrom- und Energiebedarf. → Aufdeckung von Wertüberschreitungen und entsprechende automatische Aktioneneinleitung.
- Bedarfsgerechte Steuerung durch Dialog des „Airtrackers" mit Fertigungs-, Filter- und Absauganlagen. → Maschinenbeeinflussung.
- Vernetzte Steuerung und Datenaustausch über die Steuerungsboxen.

→ Beispielsweise Einleitung leistungsstarker Absaugungen von Partikeln jeglicher Art.

- Sammlung, Dokumentation, Auswertung und Verwaltung von Messdaten im Netzwerk mittels zugehöriger „Airtracker" Software.
- → Visualisierung von Informationen auf einem Dashboard.
- Digitales Management. → Abfrage von mobilen Geräten (Smartphone, Tablets etc.).

Nutzen der Lösung / Was wurde verbessert:

- Ortsunabhängige Kontrolle der Bedingungen in der Produktion.
- Digitalisierung des Arbeitsschutzes (Arbeitsschutz und Gesundheitsprävention 4.0).
- Belastungsreduktion in der Fertigung. → Mitarbeiterschutz steht im Vordergrund.
- Vorschrifteneinhaltung leichter kontrollierbar.
- Energieeffizienz und Kostenreduzierung.
- Gefahren- und Unfallminimierung, weniger Störungen.

Veränderungen in der Organisation:

- Ideale, schadstoffarme und klimatechnische Arbeitsumgebung für den Menschen. → Steigerung der Mitarbeiterzufriedenheit durch Mitarbeitergesundheit.
- Hohe Prozesskontrolle und Unterstützung einer optimierten Produktionssteuerung.
- Vergleichbarkeit der Zustände an verschiedenen Standorten und in unterschiedlichen Unternehmensbereichen durch die Vernetzung. → Ganzheitliche Betrachtung möglich.
- Rahmenbedingungen der Produktion können leicht geregelt werden.

Kategorisierung / Zuordnung zur Tabelle:

- Förderung der RS und NH
- Vernetzung und Transparenz
- Kennzahlen
- *Förderung der Mitarbeitermotivation*

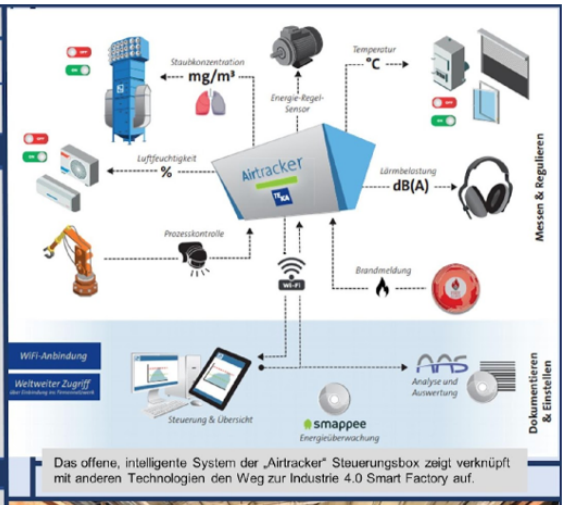


Abbildung 34: Eigene Darstellung Technologielösungsatlas Nr.18 - TEKA, Intelligentes Raumüberwachungssystem „AIRTRACKER".
Quellen: Vgl. Forum Industrie 4.0 des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA) (2018), S. 50, TEKA Absaug- und Entsorgungstechnik GmbH (2018), S. 4 und TEKA Absaug- und Entsorgungstechnik GmbH (2019), URL siehe Literaturverzeichnis.

Unternehmen: Fraunhofer Institut für Arbeitswirtschaft und Organisation (IAO)

Lösung: Qualifizierung und Lernen am Arbeitsplatz

Eingesetzte Technologien:

Lernvideos, Lernsoftware und Lernsystem

Beschreibung der Lösung:

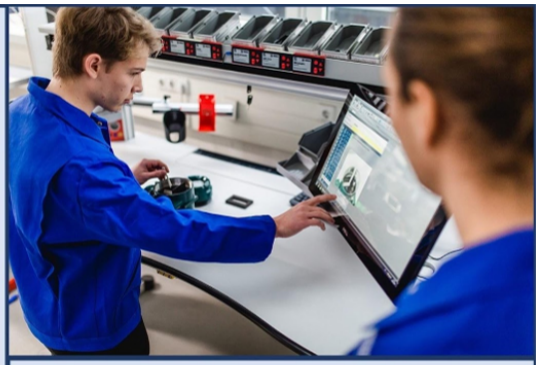
- Nutzung vor allem bei komplexen und schwierige Arbeitsprozessen. → Darstellung von Bewegtbildern und Videos zur Arbeiterführung.
- Wiederholbare Qualifizierungen und kurze Lehreinheiten. → Lernen mit kleinen Wissensteilen.
- Learning on Demand zu Themen die den Arbeits-/Produktionsprozess betreffen.
- Individuelles Lernen und individueller Wissensaufbau.
- Mitarbeiter unterweist sich selbst. Er lernt alleinständig und selbstbestimmt.
- Ergänzende Verwendung von digitalen und mobilen Lernformen im Anlernprozess.

Nutzen der Lösung / Was wurde verbessert:

- Transparente Kontrolle von Lernständen. → Für den Mitarbeiter selbst sowie für Vorgesetzte, die mit Blick in das Softwaresystem die Qualifizierungsmaßnahmen und Fortschritte einsehen können.
- Mitarbeiter kann in Leerlaufzeiten Wissen erweitern. → Effiziente Zeitnutzung.
- Dezentrale und selbstgesteuerte Mitarbeiterschulung.

Veränderungen in der Organisation:

- Selbstständiges Lernen führt zu höherer Motivation.
- Transparente Übersicht der Verteilung von Wissen und der Qualifizierungs- und Kompetenzstände von Mitarbeitern im gesamten Produktionsbereich des Unternehmens.
- Sehr nahes Lernen am tatsächlichen Arbeitsprozess.
- Genaue Abstimmung der Qualifikation auf organisatorische Belange und Anforderungen.



Kategorisierung / Zuordnung zur Tabelle:

- Konsequente Mitarbeiterqualifikation
- Prozessorientierte Arbeitsorganisation
- *Moderne Führung*
- *Selbstorganisation der Teams*

Abbildung 35: Eigene Darstellung Technologielösungsatlas Nr.19 - Fraunhofer IAO, Qualifizierung am Arbeitsplatz.
Quelle: Vgl. Fraunhofer-Institut für Arbeitswirtschaft und Organisation (2020a), URL siehe Literaturverzeichnis.

Abbildung 36: Eigene Darstellung Technologielösungsatlas Nr.20 - Schaeffler, Instandhaltung 4.0 „SmartCheck“.
Quellen: Vgl. Plattform Industrie 4.0, Schaeffler Technologies AG & Co. KG (2020), URL siehe Literaturverzeichnis, Schaeffler Technologies AG & Co. KG (2017), URL siehe Literaturverzeichnis und Schaeffler Technologies AG & Co. KG (2020), URL siehe Literaturverzeichnis.

7. Abschließende Betrachtungen

7.1. Idealtypische arbeitsorganisatorische Ablauforganisation der Produktion

Als Antwort auf die Forschungsfrage, wie ein arbeitsorganisatorischer Ablauf der zukünftigen Produktion aussehen kann, führt die nachfolgende Abbildung 37 alle bisherigen Teilergebnisse zusammen. Dadurch werden in diesem Abschnitt die zentrale Erkenntnis und der Forschungsgewinn dieser Arbeit ausdrücklich hervorgehoben. Dieser Lösungsansatz für die Gestaltung der zukünftigen Produktion basiert auf den für die Industrie 4.0 wichtigsten Kennzahlen, organisationalen Anforderungen und „Best Practices“ der Anforderungsumsetzung. Insgesamt ist die sich ergebende idealtypische Produktionsorganisationsform aus jeweils fünf Steuerungskennzahlen und I4.0 Lösungen sowie aus sieben Anforderungen an die zukünftige Produktion zusammengesetzt.

Mithilfe einer solchen oder ähnlichen Organisationsform, in der Kennzahlen bewusst und zielgerichtet gesteuert werden, sinnvolle und passende I4.0 Technologien und Anwendungen zum Einsatz kommen und zumindest die essenziellsten Produktionsanforderungen etabliert sind, ist es für Unternehmen möglich, den Herausforderungen der neuen Arbeitswelt gerecht zu werden. Die Produktionsarbeit der Zukunft benötigt somit ein geeignetes Layout und einen effizienten Gestaltungsrahmen. Die Ergebnisse aus der durchgeführten Umfrage, der ausführlichen Recherchearbeit und dem Technologielösungsatlas liefern das nachfolgende potenzielle Musterlayout der Produktion (Abb. 37).

Das Resultat ist eine hochflexible Produktion, welche als ein Beispiel für eine weitreichende Ausgestaltung mit I4.0 Komponenten dient. Eine solche Produktion verwendet innovative und effektive Formen der Arbeitsorganisation, wie beispielsweise eine verbesserte Kommunikation und Abstimmung und eine transparente Organisation. In Kombination dazu rücken grundlegende technisch-organisatorische Gesichtspunkte, wie eine starke Vernetzung und Transparenz, die Ausrichtung auf Lean Management 4.0 und die Wandlungsfähigkeit in der Produktion in den Vordergrund. Der Mensch und dessen berufliche Bildung, Qualifikation und Motivation vervollständigen dieses Bild.

Arbeitsorganisationsformen wie diese sind von Interesse, da Unternehmen dadurch die erforderliche Fähigkeit erlangen, sich durch Technik und neue Technologien ergebende Chancen bestmöglich zu nutzen.[595] Im Anschluss an diese Präsentation eines Idealbildes der zukünftigen Produktion erfolgen weitere abschließende Betrachtungen.

7.2. Diskussion und Ausblick

In diesem Kapitel findet eine nähere Diskussion der behandelten Thematiken und der durchgeführten Studie statt. Es wird überdies ein erweiternder Ausblick auf die Zukunft inklusive sich ergebender Herausforderungen im Produktionsbereich gegeben. Eine grundlegende zur Debatte stehende Fragestellung dieser Bachelorarbeit ist es, inwiefern die Organisation und der Mensch mit der Industrie 4.0 verknüpft sind und welche Bedeutung dieser Verbindung beigemessen werden muss. An dieser Stelle ist es die Intention aus einer übergeordneten Perspektive Schlussfolgerungen zu ziehen und diese im Laufe der Arbeit dargestellte Verknüpfung zusammenfassend abzubilden.

Das Vorgehen zur Bearbeitung der Thematik gliedert sich in eine Literaturrecherche, eine Umfragedurchführung und eine Kombination aller Inhalte durch die Erstellung eines organisationsbezogenen I4.0 Technologielösungsatlases. Diese Vorgehensweise war grundsätzlich zielführend und hat nützliche Ergebnisse geliefert. Der Mehrwert des theoretischen Teils besteht darin, auf Grundlage der großen Anzahl an existierenden Quellen zu Industrie 4.0 und arbeitsorganisatorischen Konzepten, explizite, zu beachtende Anforderungen an die zukünftige Produktion zu formulieren. Weiterführend zeigt sich der Mehrwert dieser Arbeit darin, diese Konzepte durch eine Praxisumfrage auf deren Relevanz zu untersuchen und aufzuzeigen, wie eine Umsetzung der Produktionsanforderungen konkret aussehen kann. Der Vorteil der durchgeführten Studie ist, dass sie ein Bild abzeichnet, welche Anforderungen an die Produktion in Zukunft erfüllt sein sollten. Die gebotenen Übersichten über einzelne Wichtigkeitseinschätzungen organisatorischer, menschenbezogener und eher technologiebezogener Gestaltungskonzepte der Organisation der zukünftigen Produktion füllt eine bestehende Forschungslücke, welche bisher nicht in ausreichendem Maße betrachtet wurde.

Der erste und zweite Teil der quantitativen Befragung haben die Erwartungen des Autors erfüllt und Ergebnisse in Form konkreter Abstufungen der Wichtigkeit von Kennzahlen und Anforderungen erbracht. Der dritte Bestandteil der Umfrage kann allerdings kritisch betrachtet werden. Zwar war es möglich mit diesem eine Detaillierung der Rangfolgenbestimmung der Produktionsanforderungen zu erreichen, doch es zeigen sich auch nicht direkt erklärbare Teilergebnisse. Die niedrigen Einstufungen der Anforderungen einer Ausrichtung der Produktion auf Lean Management 4.0 und einer Vernetzung und Transparenz in der Produktion in diesem Umfrageteil auf Rang elf und zwölf, lassen sich schwer erklären, da diese beiden Aspekte im zweiten Umfrageteil in den „Top 5“ der wichtigsten Anforderungen landen. Ein Grund für diese Abweichungen könnte sein, dass die Aufgabenstellung des Fragebogens an dieser Stelle für einige Umfrageprobanden nicht verständlich genug war. Hier besteht somit Potential zur Optimierung des Fragebogens und des Messinstruments.

Die Durchführung und Auswertung der quantitativen Orientierungsbefragung war von vereinzelten Problemen begleitet. Ein erster problematischer Aspekt ist, dass trotz einer eigentlich sehr klaren Beschreibung der jeweiligen Umfrageteile und auszufüllenden Felder einzelne Teile des Fragebogens unvollständig, beziehungsweise nicht korrekt, ausgefüllt wurden, weshalb die Stichprobengröße teilweise variiert. Weiterhin ist ein anzuführender Grund für die Schwierigkeiten eine größere Stichprobe als (n=34) zu erhalten, dass der Fragebogen neben den Erhebungen an den beiden

[595] vgl. Sorge (2003), S. 124f.

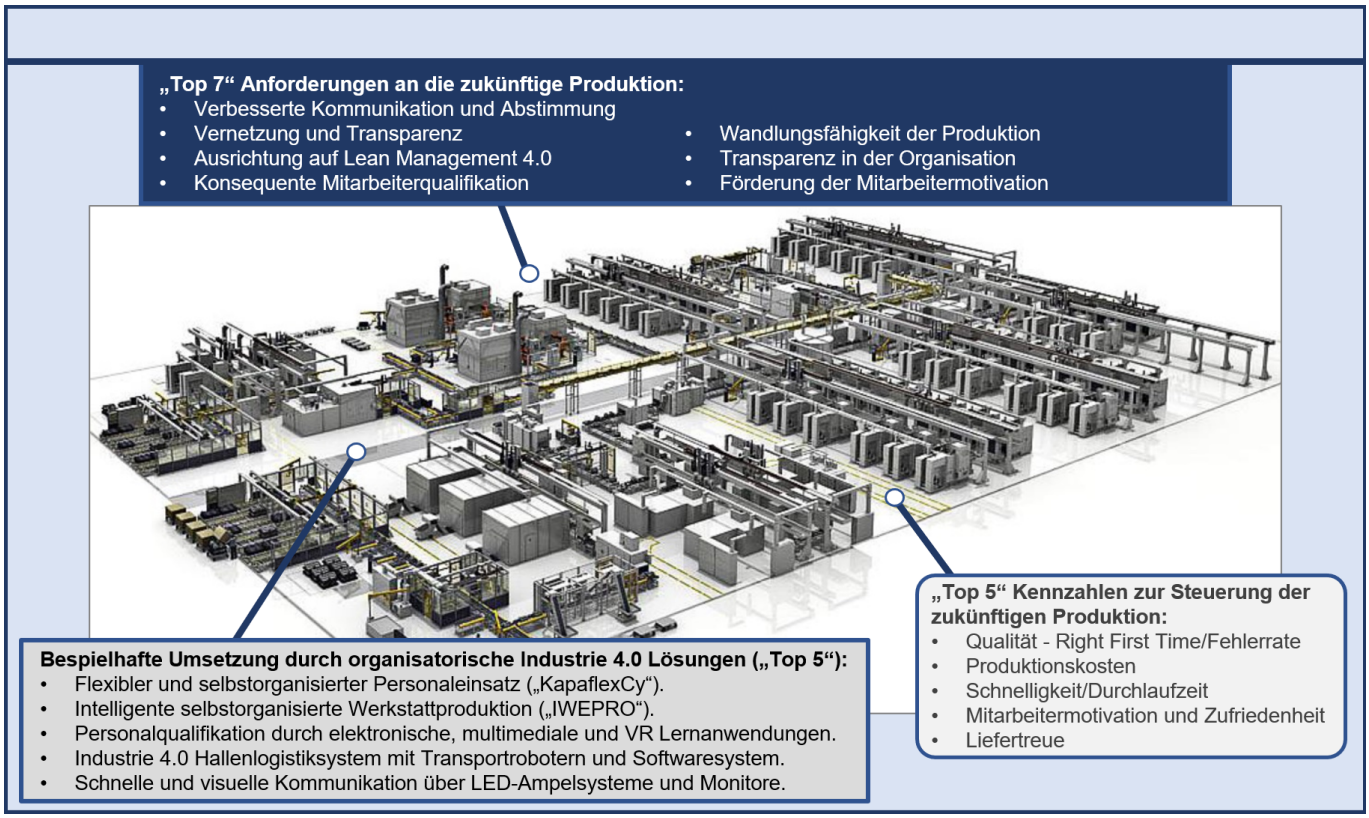


Abbildung 37: Eigene Darstellung - Idealtypische arbeitsorganisatorische Ablauforganisation der zukünftigen Produktion nach Vorlage von Deuse, Lenze, Klenner und Friedrich (2016).
Quelle: Deuse et al. (2016), S. 12.

Eventtagen per E-Mail versandt wurde. Es wurde im Voraus nicht damit gerechnet, dass die Rückmeldungen dabei recht gering ausfallen. Eine niedrige Rücklaufquote von circa 20 Prozent ist bei schriftlichen Befragungen jedoch ein häufiges auftretendes Phänomen.[596]

Trotz allem haben die geschilderten Probleme insgesamt keinen großen Einfluss auf die zentralen Erkenntnisse und Umfrageergebnisse, weshalb die Daten geeignet sind, eine Orientierung zu bieten. Die Studie kann als aussagekräftig angesehen werden, da die Stichprobe zwar kleiner als erhofft, doch ausreichend groß ausfällt. Außerdem trägt die Stichprobe zu einer Aussagekräftigkeit der Studie bei, da sie zu großen Teilen explizit die Einstellungen von Fachleuten im Bereich Produktion und Industrie 4.0 wiedergibt. Zudem ist die Studie grundsätzlich aussagekräftig, weil sich an den Ergebnissen die gute Verständlichkeit aller Fragestellungen, bis auf die des dritten Umfrageteils, erkennen lässt und die gewählten Fragenformate zu der Studienzielsetzung passen.

Obwohl keine Tests zu den Forschungsgütekriterien Reliabilität und Validität durchgeführt wurden, gibt es Anhaltspunkte, weshalb diese bei der vollzogenen Studie vorhanden sind. Alle in dieser Befragung verwendeten Messinstrumente sind in der empirischen Forschung weit verbreitet und weisen bereits grundsätzlich ein hohes Maß an Reliabilität und Validität auf.[597] Es ist beispielsweise ein guter Indikator für eine valide Messung, wenn Erhebungsmethoden verwendet werden, welche bereits erfolgreich in verschiedenen Themenfeldern eingesetzt wurden.[598] Zur Erhebung der Einstellungen zur Wichtigkeit der Produktionsanforderungen wurde in dieser Studie eine bewährte und in vielen verschiedenen Bereichen eingesetzte Skala übernommen und durch einzelne Aspekte aus ebenfalls erfolgreich eingesetzten Messinstrumenten ergänzt. Aus diesen Gründen wird angenommen, dass das Messinstrument valide ist.

Eine Erweiterung der im Rahmen dieser Arbeit durchgeführten Studie ist prinzipiell denkbar und sinnvoll. Eine solche Erweiterung kann zum einen in inhaltlicher Hinsicht und zum anderen im Hinblick auf eine Veränderung des methodischen Vorgehens erfolgen, um ergänzende, tiefergreifendere und weiterführende Erkenntnisse zu gewinnen. Methodisch gesehen ist für die Aufnahme konkreterer Fachmeinungen zu der Wichtigkeit einzelner Steuerungskennzahlen und Anforderungen der Produktion eine Erweiterung durch qualitative Erhebungsmethoden und Experteninterviews vorstellbar. Für die Zeichnung eines umfassenderen Bildes der Meinungen zu Industrie 4.0 und Organisation aus mehreren Teilen des Wirtschaftsstandortes Deutschland wäre es interessant,

[596] vgl. Schumann (2006), S. 133.

[597] vgl. Haddock und Maio (2014), S. 214-218.

[598] vgl. Brosius et al. (2016), S. 58.

Probanden außerhalb von Baden-Württemberg zu befragen.

Zusätzliche adäquate Ausdehnungen der Studie sind eine Wiederholung auf einer höheren Ebene mit einer deutlich größeren Stichprobe und eine Durchführung, die darauf ausgelegt ist, mehr unterschiedliche Zielgruppen zu erreichen. In dieser Hinsicht ist es ein erweiternder Aspekt die Verteilung über die Größe der Unternehmen besser zu streuen, denn in dieser Studie wurden überwiegend Großunternehmen erfasst. Zusätzlich ist hier die Branchentätigkeit der Unternehmen auf produzierende Unternehmen und Unternehmen aus der Metall- und Elektroverarbeitung fokussiert, weshalb eine Umfrageerweiterung mit einer breiteren Branchenstreuung denkbar ist. Ein weiterer Ansatzpunkt für eine Verbesserung der Studie ist die Überarbeitung des Messinstruments im dritten Umfrageteil, beispielsweise dadurch, dass für die Rangfolge weniger auszufüllende Felder zur Verfügung stehen, da es Teilnehmern hier anscheinend schwierig fiel eine mit 14 Feldern weitgefasste konkrete Reihenfolge zu bestimmen. Mithilfe der Abwicklung eines Pretests ist dabei die korrekte Funktionsweise des Messinstruments und das Resultat hochwertiger Ergebnisse abzusichern.

Weiterführend ist es eine Überlegung, andere Verbreitungsmethoden und Wege der Datenerhebung zu verwenden.[599] In Bezug auf die inhaltlichen Studienbestandteile kann eine größere, angepasste Studie darauf abzielen eine höhere Anzahl an Anforderungen der zukünftigen Produktion zu untersuchen. Möglich ist hier zum Beispiel die Integration einiger rein technischer Anforderungen, da in dieser Studie nur technologische Aspekte integriert sind, die hauptsächlich die Organisation betreffen. Diese Vorgehensweise kann einen detaillierteren Vergleich zu den organisations- und menschenbezogenen Anforderungen an die Produktion liefern, um genauer zu bestimmen, welche der zahlreichen vorgestellten Industrie 4.0 Gestaltungskonzepte in Zukunft unbedingt weitflächig umgesetzt werden sollten.

Als Ergebnisse der Literaturrecherche ergeben sich diverse wichtige Kennzahlen und Anforderungen der zukünftigen Produktion. Die Studie grenzt diese Auswahl ein, indem sie fünf Kennzahlen und sieben Anforderungen von ihrer Bedeutung am höchsten einstuft. Die Studienergebnisse sind ausführlich interpretierbar. Von insgesamt 16 Kennzahlen erfassen fünf Umstände, die im Zusammenhang mit dem Menschen in der Produktion stehen. Von diesen Kennzahlen befindet sich nur eine unter den „Top 5“ der wichtigsten Steuerungskennzahlen. Alle übrigen mitarbeiterbezogenen Kennzahlen sind im mittleren Bereich der Rangfolge zu finden. Die Rangfolge zeigt somit, dass Unternehmen bei der Steuerung der I4.0 Produktion mit Kennzahlen technologische Aspekte priorisieren. An dieser Stelle besteht ein Bedarf die Bedeutung und Wirkung mitarbeiterbezogener Kennzahlen weiterführend zu untersuchen. Eine Ergänzung der wichtigsten Steuerungskennzahlen um zusätzliche miterbeiterbezogene Kennzahlen kann für Unternehmen schlussfolgernd eine Lösung für die Schaffung des erforderlichen organisatorischen Rahmens für Industrie 4.0 darstellen.

[599] vgl. Schumann (2006), S. 133.

Anhand der ergänzenden Kommentare der Probanden aus Kapitel (5.5) lassen sich nachfolgende Implikationen ableiten. Komplett digital ablaufende Prozesse und eine umfangreiche Informatisierung sind von Wichtigkeit für die zukünftige Produktion. Auch in digitalen Prozessen kann das Lean Element der Verschwendung auftreten, weshalb in der Industrie 4.0 alle produktionsbezogenen IT-Prozesse mitbetrachtet und Prozessoptimierungstools eingesetzt werden müssen. Generell sollten als organisatorische Rahmenbedingungen in der Industrie 4.0 jegliche Prozesse hinterfragt und optimiert werden. Allerdings zeigt sich hinzukommend, dass eine kritische Hinterfragung der I4.0 Möglichkeiten erforderlich ist. Es sollen nicht unweigerlich alle möglichen Potentiale von Industrie 4.0 ausgeschöpft werden. Vielmehr müssen die für Unternehmen individuell am besten passenden Anforderungen mit optimalen Maßnahmen und exakt zugeschnittenen technischen Anwendungen verwirklicht werden. Ein zusätzlicher Expertenkommentar, dessen Inhalt gleichzeitig mithilfe der I4.0 Lösungen im Technologielösungsatlas bewiesen wurde, betrifft die hohe Wichtigkeit einer Visualisierung von Kennzahlen in der Produktion. Umfangreiche Darstellungen aktueller Kennzahlenstände durch den Einsatz von Dashboards und anderen Medien sind essenziell und müssen in der zukünftigen Produktion Anwendung finden.

Der hohe Einfluss von Organisation und Mensch auf die zukünftige Gestaltung der Produktion hat sich durch die quantitative Orientierungsbefragung bestätigt, da alle arbeitsorganisatorischen Konzepte mit einer Wichtigkeit von mindestens fünf auf der Likertskala eingestuft wurden. Auch wenn sich dadurch die grundlegende Kenntnis der Konzepte in der Praxis zeigt, lassen sich die Ergebnisse insofern interpretieren, als dass organisatorische und menschenbezogene Aspekte im Zusammenhang mit Industrie 4.0 in der Praxis bedeutungsmäßig noch unterbewertet sind. Laut der Literatur ist jedoch eine Schaffung eines arbeitsorganisatorischen Grundrahmens essenziell für das Funktionieren der technologischen I4.0 Seite. Diese Unterbewertung zeigt sich beispielhaft an der Einschätzung der Anforderung selbstorganisierter Teams innerhalb der Studie. Diese Anforderung wurde mit Abstand als unwichtigstes Element eingeschätzt, wobei eigentlich eine enge Verknüpfung zu der hoch bewerteten Anforderung einer modernen Führung besteht. Die Selbstorganisation der Teams ist ein wesentlicher Bestandteil des Einzugs der „New Work“ in die Produktion. Diesem Konzept sollte deshalb mehr Beachtung entgegengebracht werden. Ein möglicher Grund für die niedrige Bewertung kann hierbei sein, dass dieses Konzept ein komplett verändertes Denken der Führung und der Verantwortungsübergabe und -übernahme erfordert, wozu noch weniger Verständnis und Bereitschaft besteht als eigentlich erforderlich.

Generell sollte der Fokus auf die „Top 7“ der organisationalen Anforderungen beziehungsweise die gesamte vorgestellte, idealtypische arbeitsorganisatorische Ablauforganisation gesetzt werden. Auf dieses Grundgerüst aufbauend kann eine Ergänzung durch einzelne, individuell passende arbeitsorganisatorische und menschenorientierte Konzepte

erfolgen. Die generelle Schwierigkeit der Schaffung eines erforderlichen organisatorischen Rahmens zeigt sich anhand der Industrie 4.0 Anwendungsfälle des Technologielösungsatlases in Kapitel (6.). Insgesamt wurden 20 Lösungen gefunden und zusammengestellt, welche die organisationalen Produktionsanforderungen umsetzten. Die Mehrzahl der verwendeten Quellen zu Industrie 4.0 und die zwei online Sammlungen von I4.0 Lösungen in Deutschland „100 Orte für Industrie 4.0 in Baden-Württemberg“ und „Landkarte Industrie 4.0“, setzten eher einen Fokus auf die Technik der Industrie 4.0. Dies bestätigt die Schlussfolgerung, dass die übergreifende Arbeitsorganisation der Industrie 4.0 bisher eine zu kleine Rolle spielt. Dadurch definiert sich gleichzeitig der Auftrag für die Zukunft, hier Veränderungen anzustoßen. Das Verständnis von bereits realisierten organisationsbezogenen Lösungen kann als erster Schritt dienen, um diese zukünftigen Entwicklungen zu gestalten.

Ein konkreter Erklärungsansatz für diese Ergebnisse ist, dass die organisatorischen Anforderungen aus allen drei Kategorien „Organisation“, „Mensch“ und „Technologie“ sehr herausfordernd sind. Die reine technische Installation und Nutzung einzelner I4.0 Lösungen sind vergleichsweise leichtere Punkte als die grundlegenden organisatorischen Rahmenbedingungen und Umgebungen für eine optimale Zusammenarbeit zu schaffen. Eine technische I4.0 Umsetzung fällt Ingenieuren und Produktionsplanern vermutlich leichter, als sich an Organisationsaufgaben zu versuchen oder gar Einstellungsveränderungen der Menschen in der Produktion anzugehen.

Eine zugrundeliegende Forschungsfrage dieser Arbeit ist es, ob das heutige Produktionspersonal dazu geeignet ist, mit den zukünftigen Anforderungen zurecht zu kommen. Hier kann eine klare Verneinung erfolgen. Damit das Produktionspersonal mit den arbeitsorganisatorischen Konzepten, Anforderungen und technologischen Neuerungen zurechtkommt, ist eine konsequente Mitarbeiterqualifikation zwingend erforderlich. Diese Qualifikation ist eine maßgebliche Voraussetzung, um den Wandel hin zu Industrie 4.0 und den unterschiedlichen in dieser Arbeit präsentierten Herausforderungen neuer Produktionsgestaltungskonzepte und Arbeitsumgebungen zu bewerkstelligen. Der Fabrikarbeiter der Zukunft steht weiterhin im Mittelpunkt, ist ein gebildeter Alleskönner mit einem breit aufgestellten Profil und besitzt viele Qualifikationen. Schaut man in die heutige Produktion, so ist zu erkennen, dass dieser Mitarbeitertyp noch nicht annähernd ausreichend verfügbar ist.

Folglich muss hier für eine Umsetzung der vorgestellten organisatorischen Konzepte gezielt und übergreifend im Lehr- und Ausbildungsbereich, in der Hochschulbildung und bezüglich der unternehmensinternen Weiterbildung massiv investiert werden. Die Schaffung dieser Fähigkeiten über Qualifikationsmaßnahmen ist eine der größten zukünftigen Aufgaben für Unternehmen, da die Mitarbeiter mit ihren veränderten Qualifikationsprofilen essenziell für das Funktionieren der zukünftigen Industrie 4.0 Produktion sind.

Die Grenzen der Erkenntnisse einer hohen Bedeutung dieser arbeitsorganisatorischen Produktionskonzepte liegen in der tatsächlichen Durchführbarkeit. Die immensen Veränderungen, welche solche Produktionsgestaltungskonzepte mit sich bringen, verursachen oftmals Veränderungsängste beim Personal. Diesen Ängsten muss zu einem frühen Zeitpunkt entgegengewirkt werden. Im Optimalfall wird ein Umfeld geschaffen, das es leicht macht, eine aktive und intensive Mitarbeit zur Gestaltung der zukünftigen Produktion zu fördern. Weil mit allen vorgestellten Industrie 4.0 Konzepten große Veränderungen in der Arbeitsweise der Belegschaft einhergehen, ist die Rolle des Betriebsrates elementar. In Zusammenarbeit mit dem Betriebsrat müssen Lösungen ermittelt und vereinbart werden, die auch in der veränderten Arbeitswelt den Schutz der Mitarbeiter sicherstellen und unerwünschte beziehungsweise negative Auswirkungen auf die Belegschaft vermeiden.[600] Aus den Ergebnissen dieser Arbeit lässt sich außerdem folgern, dass noch einige Hürden, wie zum Beispiel die durch I4.0 zunehmenden Kontrollmöglichkeiten der Mitarbeiter, zu überwinden sind. Aus einer starken Transparenz folgt eine höhere Nachverfolgbarkeit aller möglichen Aktivitäten des Personals. An dieser Stelle müssen Überlegungen angestellt werden, ob Betriebsräte die Implementierung und Nutzung neuartiger arbeitsorganisatorischer I4.0 Konzepte kritisch sehen und deren Umsetzung zum Schutz der Mitarbeiter eventuell erschweren.

Die Ergebnisse der Literaturrecherche und der Studie zeigen abschließend, dass der unter Anpassungsdruck stehende Produktionsbereich den Herausforderungen der neuen Arbeitswelt nur gerecht werden kann, wenn eine Nutzung der Potentiale der Industrie 4.0 erfolgt und dabei ein Schwerpunkt auf die Organisation und den Menschen gelegt wird. Die Anforderungen in Bezug auf den Menschen und auf die Organisation erwarten ganz andere Typen von Menschen, andere Einstellungen, neue Arbeitsstrukturen und eine andere Organisation des Lebens der Mitarbeiter. Hier muss noch viel Aufklärungsarbeit geleistet werden, damit sich sowohl Einzelpersonen als auch die Unternehmen dessen bewusstwerden.

Weiterhin geht mit den in dieser Arbeit behandelten Veränderungen und neuartigen Organisationskonzepten der Produktion auch ein Wandel der Unternehmenskultur einher. Es ist in den Köpfen der meisten Menschen verankert, dass Fabriken unreine, ölverschmierte, staubige Gegenden sind. Diese Denkweise wird sich in Zukunft weiterentwickeln müssen, denn das Aussehen der Produktion und das Fabrikarbeitsklima nähern sich durch Industrie 4.0 und arbeitsorganisatorische I4.0 Konzepte dem moderner Bürokomplexe mit „smarten“ Mitarbeitern an.[601]

Generell lassen sich wenige Studien finden, die eine Differenzierung einzelner Produktionsanforderungen nach deren Bedeutung anvisieren. An dieser Stelle bietet die im Rahmen dieser Arbeit durchgeführte quantitative Orientierungsumfrage einen ergänzenden Schritt, den thematischen Zusammenhang von Arbeitsorganisation, Mensch und Industrie 4.0 aufzuarbeiten. Allerdings benötigt es hier noch deutlich

[600] vgl. Dobrzanski und Jungkind (2018), S. 687.
[601] vgl. Cole (2017), S. 184.

mehr weiterführende Forschung. Eine entsprechende Untersuchung, welche sich mit diesen Fachbereichen auseinandersetzt, ist die im Jahr 2013 am Fraunhofer IAO durchgeführte Studie zur „Produktionsarbeit der Zukunft - Industrie 4.0“.[602] In dieser sehen es 60 Prozent der befragten Unternehmen bei der Gestaltung von I4.0 als notwendig an, eine systematische Entwicklung von Mitarbeiterkompetenzen anzugehen.[603] Zudem wird die Bedeutung einer Verbesserung der Kommunikation in der Produktion mit mobilen Kommunikationstechniken hervorgehoben.[604] Ein ergänzender Studienhauptbestandteil ist die Betonung der Wichtigkeit des Konzeptes eines flexiblen Personaleinsatzes.[605]

Die Frage, wie sich die Einstellungen zu diesen Themen in der Praxis im Jahr 2019 gestalten, wird mithilfe der quantitativen Orientierungsbefragung dieser Arbeit geklärt. Vergleicht man die Ergebnisse aus beiden Jahren, so ist festzustellen, dass die Wichtigkeit einer konsequenten Mitarbeiterqualifikation und einer verbesserten Kommunikation und Abstimmung in der Praxis weiterhin gesehen wird. Jedoch ist auffällig, dass den Probanden dieser Studie ein flexibler Personaleinsatz als deutlich unwichtiger erscheint, als den Studienteilnehmern von Spath et al. (2013). An dieser Stelle zeigt sich ein Handlungsbedarf, das Konzept des flexiblen Personaleinsatzes zunehmend und verstärkt in die Unternehmenspraxis zu tragen und weitläufig zu verbreiten.

Im Endeffekt ist es für Unternehmen eine Aufgabe, eine ausgewogene Komplettlösung aus den drei Dimensionen „Organisation“, „Mensch“ und „Technologie“ zusammenzustellen.[606] Für die Planung und Umsetzung von Industrie 4.0 gibt es keine Vorlagen oder Methodenbaukästen, es sollte ein individueller Weg gefunden werden.[607] Zusätzlich ist es das korrekte Vorgehen zum Aufbau einer „Smart Factory“ aufgrund der wesentlichen benötigten künftigen Prozesse zuerst eine passende Aufbauorganisation auszuwählen. Auf Basis dieser müssen im Anschluss alle Beteiligten explizit für die neuen Aufgabenbereiche der Arbeit qualifiziert werden. Als weiterer und letzter Schritt werden technologische I4.0 Konzepte ausgewählt und mit Sinnhaftigkeit integriert. Einer der größten Fehler von Unternehmen liegt darin, mit den technischen Lösungen zu starten und die grundlegende Organisation sowie den Menschen zunächst außen vor zu lassen.[608] Schlussendlich gelingt das Arbeiten nach den in dieser Arbeit vorgestellten Prinzipien der „New Work“ nur in einer hochentwickelten Organisation.[609]

Diese Arbeit zeigt auf, wie die zukünftige Organisation der Produktionsarbeit aussehen kann, um eine Wettbewerbsfähigkeit der Industrie zu sichern. Wenn Deutschland die Entwicklung der Industrie 4.0 erfolgreich vollziehen möchte, sind grundlegend einige einschneidende Anpassungen in der Politik und Gesellschaft erforderlich. Zudem muss Digitalisierung und Industrie 4.0 in der kompletten Wirtschaft ankommen und eine Übertragung auf nahezu alle Unternehmen erfolgen. Geschieht dies nicht, kann es passieren, dass Deutschland abgehängt wird und der Wandel der Produktion nicht in dem Maße mitgetragen wird, wie es eigentlich erforderlich wäre.[610]

Hinweis: Die gewonnenen Erkenntnisse und Ergebnisse wurden am Fraunhofer IPA in verschiedenen Projekten und Forschungsbereichen verwertet. Beispielsweise flossen die Erkenntnisse der Bachelorarbeit im Rahmen des Forschungsprojektes „Fluide Fahrzeugproduktion“ in die Formulierung von Wirtschaftlichkeitsindikatoren einer fluiden Fahrzeugproduktion ein. Dabei standen Indikatoren für die Bewertung agiler Produktionsorganisationsstrukturen im Vordergrund. Darüber hinaus werden die gesammelten Praxisbeispiele der Arbeit in einem Organisations-Canvas aufgeführt und dienen somit als Ansatzpunkte für die organisatorische Gestaltung der zukünftigen Produktionsarbeit.

7.3. Zusammenfassung

Die industrielle Produktion ist fortlaufend von einer ansteigenden Komplexität geprägt. Bezüglich einer Beherrschung dieser Komplexität stellt sich die Entwicklung der Industrie 4.0 zunehmend als geeigneter Lösungsansatz heraus. In diesem Kontext veranschaulicht die vorliegende Arbeit im Detail, weshalb das Thema der Industrie 4.0 unweigerlich mit der grundlegenden Produktionsorganisation und dem Menschen zusammenhängt und eine gleichzeitige Betrachtung dieser Aspekte unabdingbar ist.

Die Produktionsarbeit steht aufgrund veränderter Voraussetzungen und Anforderungen zunehmend unter einem Anpassungsdruck. Das klassische Massenfertigungskonzept ist in einem zukünftigen Produktionsumfeld nicht mehr optimal. An dieser Stelle fungiert die Industrie 4.0 als Antwort auf die verlangte Beseitigung der Probleme und Ineffizienzen der Produktion. Die Beherrschung von Herausforderungen der neuen Arbeitswelt wie einem intensiveren Wettbewerb, der digitalen Transformation und dem demografischen Wandel ist eine essenzielle Zielsetzung der Industrie 4.0. Kernbestandteile des Konzeptes sind „Smart Products“ und die „Smart Factory“. Der komplette Bereich der Wertschöpfung wird als digitaler Zwilling auf Grundlage von Sensoren und cyber-physischen Systemen virtuell abgebildet. In diesem hochvernetzten I4.0 System bleibt der Mensch in einer zentralen Rolle, jedoch verändern sich seine Aufgabenbereiche als übergreifender Steuerer immens.

In diesem Zusammenhang sind in der Industrie 4.0 Produktion neue Formen der Arbeit und Organisation erforderlich, für deren Gestaltung in dieser Arbeit diverse Konzepte vorgestellt werden. Mithilfe einer quantitativen Orientierungsbefragung erfolgt eine bedeutungsmäßige Einordnung. Eine grundlegende Anforderung an die zukünftige Produktion ist es, zentrale organisations- und technologiebezogene

[602] Spath et al. (2013).
[603] vgl. Spath et al. (2013), S. 124f.
[604] vgl. ebenda, S. 56-66.
[605] vgl. Spath et al. (2013), S. 67-73.
[606] vgl. Bochum (2015), S. 44.
[607] vgl. Meudt et al. (2018), S. 8.
[608] vgl. Dobrzanski und Jungkind (2018), S. 686.
[609] vgl. Fischer und Köbler (2018), S. 91
[610] vgl. Heyse (2018), S. 10.

sowie mitarbeiterbezogene Kennzahlen zu steuern, da diese eine entscheidende Rolle bei der Überwachung und stetigen Verbesserung der I4.0 Produktion spielen. Im Rahmen der durchgeführten Studie hat sich die Zusammenstellung eines zukünftigen Kennzahlensystems aus nachfolgenden Kennzahlen als ideal herausgestellt. Als Ergebnis dieser Arbeit konnten fünf für die Steuerung der zukünftigen Produktion relevanteste Kennzahlen ermittelt werden. Dies sind die Kennzahlen Qualität - Right First Time/Fehlerrate, Produktionskosten, Schnelligkeit/Durchlaufzeit, Mitarbeitermotivation und Zufriedenheit sowie die Liefertreue.

Als hauptsächliches, zentrales Forschungsresultat dieser Arbeit ergeben sich sieben organisationale Anforderungen an die zukünftige Produktion als wichtigste Konzepte zur Gestaltung der zukünftigen Industrie 4.0 Produktionsorganisation. Die Anforderung einer verbesserten Kommunikation und Abstimmung in der Produktion erfordert eine erhöhte Interaktion zwischen dem Personal über verschiedene Techniken. Als zweitwichtigste Anforderung der zukünftigen Produktion konnte die umfangreiche Vernetzung und Transparenz der gesamten Produktion und ihrer Prozesse, Mitarbeiter und Dinge bestimmt werden. Die gleichfalls wichtige Anforderung einer Ausrichtung der Produktion auf Lean 4.0 kombiniert, formt und unterstützt die schlanken Prozesse der Lean Management Produktion mit I4.0 Methoden und Technologien. Die konsequente Qualifikation, Weiterentwicklung und Schulung des Produktionspersonals für die veränderten Aufgaben der Industrie 4.0 ist eine sehr wichtige Anforderung der zukünftigen Produktion und eine Basis für die Umsetzung vieler arbeitsorganisatorischer Konzepte. Die Wandlungsfähigkeit der Produktion ist das Vermögen einer Fabrik flexibel und angepasst auf verschiedene Auslöser und Bedingungen zu reagieren. Die transparente Darstellung der Produktionsarbeit und Produktionsorganisation ist ebenfalls ein Bestandteil der in dieser Arbeit deduzierten Gruppe der wichtigsten Anforderungen an die Produktion. Die Förderung der Mitarbeitermotivation durch ein optimales Arbeitsumfeld und motivationsfördernde Maßnahmen komplettiert diese Sammlung.

Auf Grundlage der in dieser Arbeit erfolgten Literaturrecherche und quantitativen Orientierungsbefragung erfolgte die Identifikation weiterer signifikanter, zu betrachtender organisationaler Produktionsanforderungen. Eine innovationsfähige Produktion verbessert sich stetig selbst. Es werden Prozesse hinterfragt und kontinuierlich Änderungsideen eingebunden. Zudem ist im Hinblick auf das Konzept der „New Work“ die Produktionsanforderung einer modernen, arbeitsteiligen Führung mit flachen Hierarchien und der Verteilung von Verantwortung zu diskutieren. Bei der Anforderung einer Selbstorganisation der Teams übernehmen Arbeiter flexibel, eigenverantwortlich und mit viel Freiraum verschiedene Fertigungsaufgaben und Rollen. Dies ist genauso wie die prozessorientierte Arbeitsorganisation, bei der eine konsequente Prozessausrichtung und -optimierung angestrebt wird, ein denkbares Organisationskonzept. Die Anforderung einer Agilität in der Produktionsorganisation verfolgt eine zielgerichtete, schnelle Anpassung der Organisation durch agile Organisationsstrukturen und agile Arbeits- und Lernweisen der Mitarbeiter. Ein flexibler Personaleinsatz mit einer selbstständigen Einsatzplanung durch die Mitarbeiter dient dazu, auf Veränderungen und Kapazitätsschwankungen schnell reagieren zu können. Weiterhin ist die Anforderung der Ressourcenschonung und Nachhaltigkeit bei der Gestaltung der Produktionsorganisation der Zukunft eine nicht zu vernachlässigende Thematik.

Der erstellte Technologielösungsatlas für innovative Shopfloor-Organisationformen zeigt beispielhaft auf, wie diese Anforderungen umgesetzt werden können. Zusammenfassend verdeutlicht die idealtypische arbeitsorganisatorische Ablauforganisation der Produktion einen Weg der Gestaltung des organisatorischen Rahmens für die Industrie 4.0. Ein solcher Rahmen muss für eine optimale Implementierung und Nutzung verschiedenster I4.0 Lösung zur Verfügung stehen. Außerdem sichert dieser einen besseren Ablauf der Produktion, ein gutes Arbeitsumfeld und erhöht Produktivität und Effizienz. Zudem ist dieser Rahmen erforderlich, damit sich die Mitarbeiter in der Industrie 4.0 Umgebung bestmöglich einfügen und zurechtfinden.

Letztendlich muss für die Nutzung der Potentiale von Industrie 4.0 die gesamte Organisation von Unternehmen entsprechend ausgerichtet werden. Diese Arbeit zeigt ausführlich, auf welche arbeitsorganisatorischen Konzepte hierbei ein Fokus gelegt werden sollte. Die Umstellung zur Verwendung von Industrie 4.0 Konzepten und vor allem arbeitsorganisatorischer und menschenorientierter Ansätze hat in der breiten Masse der Industrielandschaft jedoch noch einen sehr weiten Weg zu gehen.

Literatur

Abele, E. & Reinhart, G. (2011). *Zukunft der Produktion: - Herausforderungen, Forschungsfelder, Chancen*. München: Hanser.

Adami, W. & Houben, J. (2008). Erfahrungsförderliche Gestaltung von Motivations- und Qualifikationsansätzen in modernen, ganzheitlichen Produktionsorganisationen. In *Montage braucht Erfahrung* (S. 64–95).

Ahrens, D. & Spöttl, G. (2018). Industrie 4.0 und Herausforderungen für die Qualifizierung von Fachkräften. In H. Hirsch-Kreinsen, P. Ittermann & J. Niehaus (Hrsg.), *Digitalisierung industrieller Arbeit* (S. 175–194). Baden-Baden: Nomos.

Allianz Industrie 4.0 Baden-Württemberg (Hrsg.). (2020). *100 Orte für Industrie 4.0 in Baden-Württemberg.* Zugriff am 18.06.2020 auf https://www.i40-bw.de/de/100-orte-fuer-industrie-4-0-in-baden-wuerttemberg/

Allianz Industrie 4.0 Baden-Württemberg & BorgWarner Ludwigsburg GmbH (Hrsg.). (2020). *Digitale Schichtplanung.* Zugriff am 18.06.2020 auf https://www.i40-bw.de/de/100orte/borgwarner-gmbh/

Allianz Industrie 4.0 Baden-Württemberg & Rota Yokogawa GmbH & Co. KG (Hrsg.). (2020). *Termingetreue Versandsteuerung mittels digitalem Pull-System.* Zugriff am 10.06.2020 auf https://www.i40-bw.de/de/100orte/rota-yokogawa-gmbh-co-kg/

Altemeier, K., Bansmann, M., Dietrich, O., Dumitrescu, R. & Nettelstroth, W. (2017). *Auf dem Weg zu Industrie 4.0 - Gestaltung digitalisierter Arbeitswelten, Forschungsbericht der it's OWL Clustermanagement GmbH, Paderborn.* Zugriff auf file:///C:/Users/jan-f/Downloads/2017_Broschuere_Arbeit40.pdf

Antoni, C. H. (2017). Gruppen- und Teamarbeit. In D. Spath, E. Westkämper, H.-J. Bullinger & H.-J. Warnecke (Hrsg.), *Neue Entwicklungen in der Unternehmensorganisation* (S. 161–172).

Arbeitsgruppe Forschung und Innovation der Plattform Industrie 4.0. (2016). *Fortschreibung der Anwendungsszenarien der Plattform Industrie 4.0.* Zugriff auf https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/fortschreibung-anwendungsszenarien.pdf?__blob=publicationFile&v=7

ARENA2036 e.V. (2020). *ARENA2036 - Die flexible Forschungsfabrik: - Die hochflexible Forschungsplattform für die Mobilität und Produktion der Zukunft.* Zugriff am 19.06.2020 auf https://www.arena2036.de/de/

Aßmann, S. (2017). Industrie 4.0 ist bei Bosch bereits Realität. In K. Lucks (Hrsg.), *Praxishandbuch Industrie 4.0* (S. 321–330).

Baethge, M. & Baethge-Kinsky, V. (1998). Jenseits von Beruf und Beruflichkeit? - Neue Formen von Arbeitsorganisation und Beschäftigung und ihre Bedeutung für eine zentrale Kategorie gesellschaftlicher Integration. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung (MittAB)*, *31* (3), S. 461–472.

Banse, G. (2017). Industrie 4.0 aus Sicht der Technikphilosophie und der Technikfolgenabschätzung. In G. Banse, U. Busch & M. Thomas (Hrsg.), *Digitalisierung und Transformation* (S. 117–138). Berlin.

Bauer, W., Schlund, S., Marrenbach, D. & Ganschar, O. (2017). Industrie 4.0 - volkswirtschaftliche Bedeutung für Deutschland: - Auszug aus der Studie des Bitkom. In K. Lucks (Hrsg.), *Praxishandbuch Industrie 4.0* (S. 17–20).

Bauernhansl, T. & Dombrowski, U. (Hrsg.). (2016). *Einfluss von Industrie 4.0 auf unsere Fabriken und die Fabrikplanung, Tagungsband des Fachbeirats Deutscher Fachkongress Fabrikplanung, Ludwigsburg, 20/21.04.2016, Braunschweig.*

Becker, F. (2019). *Mitarbeiter wirksam motivieren: - Mitarbeitermotivation mit der Macht der Psychologie*.

Becker, M. (2005). *Personalentwicklung: - Bildung, Förderung und Organisationsentwicklung in Theorie und Praxis* (4. Aufl.). Stuttgart.

Becker, T. (2017). Digitalisierung und Netzwerkmanagement. In *Rhein-Ruhr-Institut für Sozialforschung und Politikberatung e.V. (RISP) der Universität Duisburg-Essen* (S. 104–108).

Berekhoven, L., Eckert, W. & Ellenrieder, P. (1999). *Marktforschung*. Wiesbaden.

Berg, A. (2018). Industrie 4.0 - Wo steht Deutschland? *Bundesverband Informationswirtschaft Telekommunikation und neue Medien e.V. (Bitkom)*. Zugriff auf https://www.bitkom.org/sites/default/files/pdf/Presse/Anhaenge-an-PIs/2018/Bitkom-Pressekonferenz-Industrie-40-23-04-2018-Praesentation-2.pdf

Berg, A. (2019). Industrie 4.0 - jetzt mit KI. *Bundesverband Informationswirtschaft Telekommunikation und neue Medien e.V. (Bitkom)*.

Berg, M. (2019). Die Führungsrolle 4.0 in agilen Innovationsmodellen. In *Führen in der Arbeitswelt 4.0* (S. 119–136).

Berkholz, D. (2008). Zusammenfassung: Wandlungsfähigkeit: - (k)ein Thema der Zukunft. In P. Nyhuis, G. Reinhart & E. Abele (Hrsg.), *Wandlungsfähige Produktionssysteme* (S. 146–150). Technische Informationsbibliothek u. Universitätsbibliothek and PZH Produktionstechnisches Zentrum.

Bernsteiner, R. & Bauer, F. (2017). IoT-Rahmenmodelle als Grundlage der IT-Integration in mittelständischen Unternehmen. In K. Lucks (Hrsg.), *Praxishandbuch Industrie 4.0* (S. 787–792).

Bieback, K.-J. & Oppolzer, A. (1999). Strukturwandel im Arbeitsschutz. In K.-J. Bieback & A. Oppolzer (Hrsg.), *Strukturwandel des Arbeitsschutzes* (S. 7–42).

Bimos KG. (2020a). *Arbeitssicherheit und Ergonomie: Die schleichenden Gesundheitsrisiken.* Zugriff am 16.06.2020 auf https://www.bimos.com/B/de-de/news/2972/arbeitssicherheit

Bimos KG. (2020b). *Nachhaltigkeit in der modernen Produktion.* Zugriff am 16.06.2020 auf https://www.bimos.com/B/de-de/news/2971/nachhaltigkeit-in-der-modernen-produktion

Binner, H. F. (2005). *Handbuch der prozessorientierten Arbeitsorganisation: - Methoden und Werkzeuge zur Umsetzung*.

Black, J. T. (1991). *The Design of the Factory with a Future*.

BMW Group. (2020). *Sichtprüfung mit Gedächtnis: BMW Group testet Datenbrille zur Qualitätssicherung in der Produktion.* Zugriff am 12.06.2020 auf https://www.press.bmwgroup.com/deutschland/article/detail/T0196242DE/sichtpruefung-mit-gedaechtnis:-bmw-group-testet-datenbrille-zur-qualitaetssicherung-in-der-produktion?language=de

Bochum, U. (2015). Gewerkschaftliche Positionen in Bezug auf „Industrie 4.0". In *Zukunft der Arbeit in Industrie 4.0* (S. 31–45). Springer Berlin Heidelberg.

Böde, E., Hartmann, E. A., Lüdtke, A., Oppenheimer, F., Rötting, M. & Wegerich, A. (2013). *Mensch-Technik-Interaktion Band 3 -: Leitfaden für Hersteller und Anwender.* Berlin Januar. Zugriff auf https://www.digitale-technologien.de/DT/Redaktion/DE/Downloads/Publikation/autonomik-Leitfaden3.pdf?__blob=publicationFile&v=3

Böhle, F. & Bolte, A. (2002). *Die Entdeckung des Informellen: - Der schwierige Umgang mit Kooperation im Arbeitsalltag*.

Borch, H. & Zinke, G. (2008). Aus- und Fortbildung aus einem Guss - Berufsbildung in der Produktionstechnologie. *Berufsbildung in Wissenschaft und Praxis (BWP)*, *37* (4), 43–47.

Börkircher, M., Frank, H., Gärtner, R., Hasse, F., Jeske, T., Lennings, F., … Wüseke, F. (2018). *Digitalisierung & Industrie 4.0 (Teil 1) -: So individuell wie der Bedarf - Produktivitätszuwachs durch Informationen.*

Brosius, H.-B., Haas, A. & Koschel, F. (2016). *Methoden der empirischen Kommunikationsforschung: - Eine Einführung*.

Broy, M. (2010). Cyber-Physical Systems - Wissenschaftliche Herausforderungen bei der Entwicklung. In M. Broy (Hrsg.), *Cyber-Physical Systems* (S. 17–32).

Buhr, D. (2015). *Soziale Innovationspolitik für die Industrie 4.0 - Expertise im Auftrag der Abteilung Wirtschafts- und Sozialpolitik der Friedrich-Ebert-Stiftung (WISO Diskurs)*.

Butollo, F. & Engel, T. (2015). Industrie 4.0 - arbeits- und gesellschaftspolitische Perspektiven: - Zwischen Dystopie und Euphorie. *Zeitschrift Marxistische Erneuerung*, *26* (103), 29–41.

Christian, M. S., Garza, A. S. & Slaughter, J. E. (2011). Work Engagement: A Quantitative Review and Test of its Relations - With Task and Contextual Performance. *Personnel Psychology*, *64* (1), 89–136.

Clement, U. (2006). Arbeit unterhalb der Facharbeiterqualifikation und ihre Herausforderungen für die europäische Berufsbildungspolitik. In U. Clement & M. Lacher (Hrsg.), *Produktionssysteme und kompetenzerwerb* (S. 93–112). Steiner.

Cole, T. (2017). *Digitale Transformation: - Warum die deutsche Wirtschaft gerade die digitale Zukunft verschläft und was jetzt getan werden muss!*

Creusen, U., Gall, B. & Hackl, O. (2017). *Digital Leadership: - Führung in Zeiten des digitalen Wandels*.

Daft, R. L. & Huber, G. P. (1987). How Organizations Learn: - A Communication Framework. In N. Ditomaso & S. Bacharach (Hrsg.), *Research in the Sociology of Organizations* (S. 1–36).

Dale, B. G. (2007). Tools and Techniques: An Overview. In B. G. Dale, T. van der Wiele & J. van Iwaarden (Hrsg.), *Managing quality* (S. 363–381).

Dale, B. G. & Lascelles, D. M. (2007). Levels of TQM Adoption. In B. G. Dale, T. van der Wiele & J. van Iwaarden (Hrsg.), *Managing quality* (S. 111–127).

Dangl, A. (2019). Optimierte Kontrolle und Planung mit Digital Twins für Organisationen: - Das digitale Abbild einer Organisation verschafft Transparenz und Überblick, es erlaubt Simulation, Kontrolle und Planung. *Industrie 4.0 Management (I40M)*, *35* (5), 59–62.

Das, A. (2001). Towards Theory Building in Manufacturing Flexibility. *International Journal of Production Research*, *39* (18), 4153–4177.

Defranceski, M. (2017). Verkürzte Entscheidungsfindung in der Produktion. In B. Vogel-Heuser, T. Bauernhansl & M. ten Hompel (Hrsg.), *Handbuch Industrie 4.0 Bd.1* (S. 139–152).

Denning, S. (2016). *Explaining Agile.* Zugriff am 16.06.2020 auf https://www.forbes.com/sites/stevedenning/2016/09/08/explaining-agile/#78a22954301b

Deuse, J., Achenbach, M. & Lenze, D. (2015). Hybrides Ausbildungskonzept für den Umgang mit Systemkomplexität. In H. Meier (Hrsg.), *Lehren und Lernen für die moderne Arbeitswelt* (S. 149–263).

Deuse, J., Lenze, D., Klenner, F. & Friedrich, T. (2016). Manufacturing Data Analytics zur Identifikation dynamischer Engpässe in Produktionssystemen mit hoher wertschöpfender Variabilität. In C. M. Schlick (Hrsg.), *Megatrend Digitalisierung* (S. 11–26).

Dichtl, S. & Patermann, N. (2017). Shopfloor-Management: - Potenziale durch Transparenz heben. In R. Koether & K.-J. Meier (Hrsg.), *Lean Production für die variantenreiche Einzelfertigung* (S. 199–214). Springer Fachmedien Wiesbaden.

Dobrowolski, P. (2016). *6 Thesen zur wandlungsfähigen Produktion der Zukunft.* Zugriff am 16.06.2020 auf https://industriemagazin.at/a/6-thesen-zur-wandlungsfaehigen-produktion-der-zukunft

Dobrzanski, F. & Jungkind, W. (2018). Mitarbeiter- und Organisationsentwicklung im Rahmen von Industrie 4.0. *Zeitschrift für wirtschaftlichen Fabrikbetrieb (ZWF)*, *113* (10), 686–691.

Dorsch-Schweizer, M. & Schwarz, H. (2007). Beruflichkeit zwischen Arbeitsteilung und Prozessorientierung? *Sozialwissenschaften und Berufspraxis (SuB)*, *30* (2), 300–318.

Drucker, P. F. (1988). The Coming of the New Organization. *Harvard Business Review*, *66* (1), 48–53.

Dumitrescu, R. & Marquardt, R. (2017). Auf dem Weg zu Industrie 4.0 - „it´s OWL" bietet Lösungen für den Mittelstand. In K. Lucks (Hrsg.), *Praxishandbuch Industrie 4.0* (S. 637–652).

Duschlbauer, T., Martin, S. & Saffarnia, P. (2018). Vom Internet der Dinge zum Internet der Menschen. In *Organisationskommunikation im Zeichen der Digitalisierung* (S. 11–12).

Elkins, D. A., Huang, N. & Alden, J. M. (2004). Agile Manufacturing Systems in the Automotive Industry. *International Journal of Production Economics*, *91* (3), 201–214.

ESTA Apparatebau GmbH & Co. KG. (2020). *Filterturm Filtower.* Zugriff am 15.06.2020 auf https://www.esta.com/de/produkte/filterturm-filtower/

Europäische Kommission. (2015). *Benutzerleitfaden zur Definition von KMU*.

Fay, A., Gausemeier, J. & ten Hompel, M. (2018). Einordnung der Beispiele der Industrie 4.0-Landkarte in die Anwendungsszenarien. *Forschungsbericht der Deutschen Akademie der Technikwissenschaften e.V.*.

Fischer, T. & Köbler, J. (2018). Lean-Management und Industrie 4.0: - Warum Lean-Management eine solide Grundlage für die vierte industrielle Revolution darstellt. *Industrie 4.0 Management (I40M)*, *34* (6), 53–55.

Forum Industrie 4.0 des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA). (2016). *Industrie 4.0 konkret - Lösungen für die industrielle Praxis (Auflage 2016), Leitfaden des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA)*.

Forum Industrie 4.0 des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA). (2017). *Industrie 4.0 konkret - Lösungen für die industrielle Praxis (Auflage 2017), Leitfaden des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA)*.

Forum Industrie 4.0 des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA). (2018). *Industrie 4.0 konkret - Lösungen für die industrielle Praxis (Auflage 2018), Leitfaden des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA)*.

Forum Industrie 4.0 des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA) und Forschungskuratorium Maschinenbau e.V. (FKM). (2017). *F&E-Lotse verschafft Überblick - Forschungsprojekte im Themenfeld Industrie 4.0, Forschungsbericht des Verbands Deutscher Maschinen- und Anlagenbau e.V. (VDMA)*.

Franke, S. (2019). Interview: „New Work: Neue Kultur der Zusammenarbeit". In D. Brommer, S. Hockling & A. Leopold (Hrsg.), *Faszination New Work: 50 Impulse für die neue Arbeitswelt* (S. 127–130).

Fraunhofer-Institut für Arbeitswirtschaft und Organisation. (2019). *Innovationsnetzwerk „Produktionsarbeit 4.0" Phase 4.* Zugriff am 20.06.2020 auf https://www.engineering-produktion.iao.fraunhofer.de/content/dam/iao/tim/Dokumente/Innovationsnetzwerk-Prod-Arbeit_4-Phase/InnovationsnetzwerkProduktionsarbeit_Phase4_2020-2022.pdf

Fraunhofer-Institut für Arbeitswirtschaft und Organisation. (2020a). *Beschäftigung flexibilisieren mit der KapaFit-Simulation: - Turbulenzreserven aufbauen!* Zugriff am 13.06.2020 auf https://www.engineering-produktion.iao.fraunhofer.de/content/dam/iao/tim/Dokumente/Kapaflexcy/Kapafit_Dimensionieren.pdf

Fraunhofer-Institut für Arbeitswirtschaft und Organisation. (2020b). *Cognitive Engeneering and Production: - Wir machen produzierende Unternehmen schnell, flexibel und digital!* Zugriff am 20.06.2020 auf https://www.engineering-produktion.iao.fraunhofer.de/content/dam/iao/tim/Dokumente/cognitive-engineering-and-production.pdf

Fraunhofer-Institut für Arbeitswirtschaft und Organisation. (2020c). *Future Work Lab - KPI-Dashboard für Meister.* Zugriff am 05.05.2020 auf https://futureworklab.de/content/dam/iao/digitalisierung/documents/Demonstratoren/kpi-dashboard.pdf

Fraunhofer-Institut für Arbeitswirtschaft und Organisation. (2020d). *Future Work Lab - Qualifizierung mit Lernvideos.* Zugriff am 11.06.2020 auf https://futureworklab.de/content/dam/iao/digitalisierung/documents/Demonstratoren/qualifizierung-lernvideos.pdf

Fraunhofer-Institut für Arbeitswirtschaft und Organisation. (2020e). *Open Lab Day im Future Work Lab.* Zugriff am 16.06.2020 auf https://futureworklab.de/content/dam/iao/digitalisierung/documents/future_work_lab_%20open_lab_days.pdf

Fraunhofer-Institut für Arbeitswirtschaft und Organisation & aucobo GmbH. (2020). *Future Work Lab - Mobile Mehrmaschinenbedienung.* Zugriff am 19.06.2020 auf https://futureworklab.de/content/dam/iao/digitalisierung/documents/Demonstratoren/mehrmaschinen-bedienung.pdf

Fraunhofer-Institut für Produktionsanlagen und Konstruktionstechnik. (2020). *iWePro - Intelligente selbstorganisierende Werkstattproduktion.* Zugriff am 20.06.2020 auf https://www.projekt-iwepro.de/content/dam/ipk/iwepro/de/Dokumente/ps_iWePro_web.pdf

Fraunhofer-Institut für System- und Innovationsforschung. (2020). *Benchmark Wandlungsfähigkeit - Beschreibung der Kennzahlen.* Zugriff am 06.06.2020 auf https://www.industriebenchmarking.eu/download/dywamed/Beschreibung_der_Kennzahlen.pdf

Frey, C. B. & Osborne, M. A. (2013). *The Future of Employment: - How Susceptible are Jobs to Computerisation?*

Gabath, C. (2010). *Risiko- und Krisenmanagement im Einkauf: - Methoden zur aktiven Kostensenkung*.

Gallup Organization. (2017). *State of the Global Workplace*.

Gausemeier, J. (2007). Mehr Interaktion von Wirtschaft und Wissenschaft - Stärken stärken. In J. Gausemeier & H. K. Tönshoff (Hrsg.), *Migration von wertschöpfung* - (S. 43–52).

Geisberger, E. & Broy, M. (2012). *AgendaCPS - Integrierte Forschungsagenda Cyber-Physical Systems.*

Geramanis, O. (2020). Zusammenarbeit 5.0: - Die kooperative Dimension der neuen Arbeitswelt. In *Der Mensch in der Selbstorganisation* (Bd. 78, S. 3–25).

Gibson, I., Yoke, S. W., Stanev, S., Krappe, H., Abul, O. H., Georgoulias, K., … Ovtcharova, J. (2008). Efficient Change Management for the flexible Production of the Future. *Journal of Manufacturing Technology Management*, *19* (6), 712–726.

Glück, M. (2015). Wie entwickelt sich eine Qualität 4.0? *Qualität und Zuverlässigkeit (QZ)*, *60* (8), 39–43.

Golob, N. & Ullrich, M. (2019). New Work und keiner geht hin. In D. Brommer, S. Hockling & A. Leopold (Hrsg.), *Faszination New Work: 50 Impulse für die neue Arbeitswelt* (S. 91–108).

Goschy, W. & Metternich, J. (2016). *25 Jahre Lean Management: - Lean gestern, heute und morgen*.

Gottmann, J. (2016). *Produktionscontrolling: - Wertströme und Kosten optimieren*.

Graf, N., Gramß, D. & Edelkraut, F. (2017). *Agiles Lernen: - Neue Rollen, Kompetenzen und Methoden im Unternehmenskontext*.

Gronau, N., Ullrich, A. & Vladova, G. (2015). Prozessbezogene und visionäre Weiterbildungskonzepte im Kontext Industrie 4.0. In H. Meier (Hrsg.), *Lehren und Lernen für die moderne Arbeitswelt* (S. 125–143).

Gruber, M., Rinner, M., Löscher, T., Almeder, C., Hartl, R. & Katzensteiner, S. (2011). Vorausschauende Produktionsregelung durch simulationsbasierte heuristische Optimierung. In L. März, W. Krug, O. Rose & G. Weigert (Hrsg.), *Simulation und Optimierung in Produktion und Logistik*.

Hacker, W. & Richter, P. (1984). *Spezielle Arbeits- und Ingenieurpsychologie in Einzeldarstellungen (Band 1): - Psychologische Bewertung von Arbeitsgestaltungsmaßnahmen - Ziele und Bewertungsmaßstäbe*.

Hackl, B., Wagner, M., Attmer, L. & Baumann, D. (2017). *New Work: Auf dem Weg zur neuen Arbeitswelt: - Management-Impulse, Praxisbeispiele, Studien*.

Haddock, G. & Maio, G. R. (2014). Einstellungen. In K. Jonas, W. Stroebe & M. Hewstone (Hrsg.), *Sozialpsychologie* (S. 197–229).

Hartmann, E. (2015). Arbeitsgestaltung für Industrie 4.0: Alte Wahrheiten, neue Herausforderungen. In A. Botthof & E. A. Hartmann (Hrsg.), *Zukunft der Arbeit in Industrie 4.0* (S. 9–20).

Hartmann, F. (2017). Zukünftige Anforderungen an Kompetenzen im Zusammenhang mit Industrie 4.0 - Eine Bestandsaufnahme. In *Facharbeit und Digitalisierung* (S. 19–28).

Heidel, R., Hoffmeister, M., Hankel, M. & Döbrich, U. (2017). *Basiswissen RAMI 4.0: Referenzarchitekturmodell mit Industrie 4.0-Komponente Industrie 4.0*.

Heinen, T., Hirsch, B. & Nyhuis, P. (2009). Wandlungsfähigkeit durch Mitarbeiterkompetenz: - Entwicklung eines mitarbeiterorientierten Planungsmodells zur Steigerung der Wandlungsfähigkeit von Fabriken. *Zeitschrift für wirtschaftlichen Fabrikbetrieb (ZWF)*, *104* (10), 836–840.

Heinen, T., Rimpau, C. & Wörn, A. (2008). Wandlungsfähigkeit als Ziel der Produktionssystemgestaltung. In P. Nyhuis, G. Reinhart & E. Abele (Hrsg.), *Wandlungsfähige Produktionssysteme* (S. 19–32).

Heller, T. (2015). *Organisation der Instandhaltung in der Smart Factory: - Dezentral und autonom: Von der Theorie in die Praxis.* Zugriff am 13-06-2020 auf http://publica.fraunhofer.de/eprints/urn_nbn_de_0011-n-3790577.pdf

Hertle, C. (2018). *Shopfloor Management Systeme zur zielgerichteten, systematischen Kompetenzentwicklung in der Produktion: Dissertation*.

Heyse, V. (2018). Einleitung: Mittelstand 4.0 im Spannungsfeld des digitalen Wandels. In V. Heyse, J. Erpenbeck, S. Ortmann & S. Coester (Hrsg.), *Mittelstand 4.0 - eine digitale Herausforderung* (S. 9–16).

Hirsch-Kreinsen, H. (2018). Einleitung: Digitalisierung industrieller Arbeit. In H. Hirsch-Kreinsen, P. Ittermann & J. Niehaus (Hrsg.), *Digitalisierung industrieller Arbeit* (S. 13–32). Nomos.

Hirsch-Kreinsen, H. & Weyer, J. (2014). Wandel von Produktionsarbeit - Industrie 4.0". *Arbeitspapier der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Technischen Universität Dortmund*, *38*.

Hofert, S. (2018). *Das agile Mindset: - Mitarbeiter entwickeln, Zukunft der Arbeit gestalten*.

Hoffmann, F.-J. (2017). iBin - Anthropomatik schafft revolutionäre Logistiklösungen. In B. Vogel-Heuser, T. Bauernhansl & M. ten Hompel (Hrsg.), *Handbuch Industrie 4.0 Bd.1* (S. 47–60).

Hohberger, S. & Damlachi, H. (2017). Performancesteigerung in der Produktion. In S. Hohberger & H. Damlachi (Hrsg.), *Performancesteigerung im Unternehmen* (S. 145–162).

Holtgrewe, U., Riesenecker-Caba, T. & Flecker, J. (2015). *Industrie 4.0 - eine arbeitssoziologische Einschätzung: - Endbericht für die AK Wien.*

Hoppe, S. (2016). Der Schlüssel für Industrie 4.0: Semantische Interoperabilität vom Sensor bis in die Cloud. In C. Manzei, L. Schleupner & R. Heinze (Hrsg.), *Industrie 4.0 im internationalen Kontext* (S. 96–100).

Horzella, A. (2010). *Wertsteigerung im M&A-Prozess: - Erfolgsfaktoren, Instrumente, Kennzahlen: Zugl.: Bayreuth, Univ., Diss., 2009*.

Höpner, A. (2019). *Die Industrie 4.0 stockt - Für einen Durchbruch auf breiter Front gibt es noch viele Hemmnisse.*

Huber, A. S. (2013). Das Ziel Digital Enterprise: Die professionelle digitale Abbildung von Produktentwicklung und Produktion. In U. Sendler (Hrsg.), *Industrie 4.0*.

IHK Darmstadt Service GmbH. (2020). *Mittelstand 4.0 Kompetenzentrum Darmstadt - Anwendungsbeispiele.* Zugriff am 16.06.2020 auf https://kompetenzzentrum-darmstadt.digital/de/Landing/Anwendungsbeispiele/~nm.120/Anwendungsbeispiele.html

International Association of Oil & Gas Producers. (2014). OGP Safety Performance Indicators - 2013 Data. *Forschungsbericht der International Association of Oil & Gas Producers*.

Ittermann, P. & Niehaus, J. (2018). Industrie 4.0 und Wandel von Industriearbeit - revisited. - Forschungsstand und Trendbestimmungen. In H. Hirsch-Kreinsen, P. Ittermann & J. Niehaus (Hrsg.), *Digitalisierung industrieller Arbeit* (S. 33–60).

Ittermann, P., Niehaus, J. & Hirsch-Kreinsen, H. (2015). *Arbeiten in der Industrie 4.0: - Trendbestimmungen und arbeitspolitische Handlungsfelder* (Nr. 308).

Jeske, T., Hille, S. & Börkircher, M. (2015). Produktionsarbeit und Führung in der Industrie 4.0: - Potenziale und Auswirkungen. *Betriebspraxis & Arbeitsforschung - Zeitschrift für angewandte Arbeitswissenschaft*, *o.Jg.* (223), 27–32.

Joos, T. (2012). Plankostenrechnung als Instrument des Produktionscontrollings. In A. Klein & H. Schnell (Hrsg.), *Controlling in der Produktion* (S. 63–84).

Jung, A. (2017). Industrie 4.0 - Die digitale Zumutung. *Der Spiegel* (17/2017), 64–67.

Kagermann, H. (2014). Industrie 4.0 und Smart Services. In Brenner W. & Hess T. (Hrsg.), *Wirtschaftsinformatik in Wissenschaft und Praxis* (S. 243–248). Berlin, Heidelberg.

Kagermann, H., Wahlster, W., Helbig, J., Hellinger, A., Stumpf, V., Kobsda, C. & Treugut, L. (2013). Deutschlands Zukunft als Produktionsstandort sichern - Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0: - Abschlussbericht des Arbeitskreises Industrie 4.0. *Abschlussbericht des Arbeitskreises Industrie 4.0, Positionspapier der Deutschen Akademie der Technikwissenschaften e.V. (acatech) und der Promotorengruppe Kommunikation der Forschungsunion Wirtschaft - Wissenschaft*.

Kagermann, H. & Winter, J. (2017). Industrie 4.0 und plattformbasierte Geschäftsmodellinnovationen. In K. Lucks (Hrsg.), *Praxishandbuch Industrie 4.0* (S. 21–32). Stuttgart: Schäffer-Poeschel Verlag.

Kahn, W. A. (1990). Psychological Conditions of Personal Engagement and Disengagement at Work. *Academy of Management Journal (AMJ)*, *33* (4), 692–724.

Kaltenecker, S. (2018). *Selbstorganisierte Teams führen: - Arbeitsbuch für Lean & Agile Professionals*.

Kaufmann, T. (2015). *Geschäftsmodelle in Industrie 4.0 und dem Internet der Dinge: - Der Weg vom Anspruch in die Wirklichkeit*. Wiesbaden: Springer Vieweg.

Kempkes, J. A., Suprano, F. & Wömpener, A. (2018). Produktion 4.0 mit den richtigen Kennzahlen steuern. *Controlling & Management Review (CMR)*, *62* (4), 56–61.

Kempter, G., Jost, P. & Künz, A. (2020). Gut zu wissen: Technologiegestütztes Lernen während der Arbeit. In S. Wörwag & A. Cloots (Hrsg.), *Zukunft der Arbeit, Perspektive Mensch* (S. 229–239).

Killian, S., Nendel, N., Markert, T. & Riedel, R. (2019). Spielend lernen in der virtuellen Realität: - Weiterbildung durch Gamification. *Industrie 4.0 Management (I40M)*, *35* (2), 53–56.

Kinemic GmbH. (2020a). *Assistenzsysteme besser nutzen - Einfache Bedienung mittels Kinemic Band: - Die Zukunft der Gestensteuerung.* Zugriff am 20.06.2020 auf https://kinemic.com/de/

Kinemic GmbH. (2020b). *Remote-Assistenz.* Zugriff am 22.06.2020 auf https://kinemic.com/de/anwendungen/remote-assistenz/

Klaes, M. & Schreiber, A. (2015). Die intelligente Produktion von morgen: - Anlagen- und Automatisierungskonzept für flexible und wirtschaftliche Fertigung. *Der Konstrukteur*, *47* (4), 16–18.

Klauß, T. & Mierke, A. (2017). *Szenarien einer digitalen Welt - heute und morgen: - Wie die digitale Transformation unser Leben verändert*.

Kletti, J. (2015). *MES - Manufacturing Execution System: - Moderne Informationstechnologie unterstützt die Wertschöpfung*.

Kletti, J. & Schumacher, J. (2014). *Die perfekte Produktion: - Manufacturing Excellence durch Short Interval Technology (SIT)*.

Koch, V., Kuge, S., Geissbauer, R. & Schrauf, S. (2014). Industrie 4.0 - Chancen und Herausforderungen der vierten industriellen Revolution. *Arbeitspapier der Strategy& GmbH und der PricewaterhouseCoopers GmbH*.

Kollmann, T. & Schmidt, H. (2016). *Deutschland 4.0 - Wie die digitale Transformation gelingt*.

Kool, L. & van Est, R. (2015). *Working on the Robot Society: - Visions and Insights from Science concerning the Relationship between Technology and Employment*.

Korge, A. (2017). Konzepte, Methoden, Erfolgsfaktoren in der Produktion. In D. Spath, E. Westkämper, H.-J. Bullinger & H.-J. Warnecke (Hrsg.), *Neue Entwicklungen in der Unternehmensorganisation* (S. 505–512). Berlin, Heidelberg.

Kreggenfeld, N., Prinz, C. & Kuhlenkötter, B. (2016). Mitarbeiterbefähigung in der Industrie 4.0 - Ganzheitlicher Ansatz zur Erfassung sowie zum Management von Mitarbeiter- und Prozesswissen. *Industrie 4.0 Management (I40M)*, *32* (3), 31–34.

Kromrey, H., Roose, J. & Strübing, J. (2016). *Empirische Sozialforschung: - Modelle und Methoden der standardisierten Datenerhebung und Datenauswertung mit Annotationen aus qualitativ-interpretativer Perspektive*.

Krones, M., Strauch, J., Schütze, J. & Müller, E. (2015). PLUG+LEARN - Lehren und Lernen mit wandlungsfähigen Kompetenzmodulen. In H. Meier (Hrsg.), *Lehren und Lernen für die moderne Arbeitswelt* (S. 81–96).

Krüger, J. (2013). *Effektive Produktionsprozesse mit IT - auf dem Weg zu Industrie 4.0: Industrie 4.0 - Zukunftspotentiale für die Produktion.* Zugriff auf https://www.ihk-berlin.de/blueprint/servlet/resource/blob/2272490/4dcdd68cf812f2c51d78a2200bcd4d77/praesentation-va-industrie4-0-krueger-data.pdf

Krügl, S. (2019). Interview: „Wie sieht die Führungskräfteentwicklung der Zukunft aus?“. In D. Brommer, S. Hockling & A. Leopold (Hrsg.), *Faszination New Work: 50 Impulse für die neue Arbeitswelt* (S. 233–242).

Kühl, S. (2015). *Wenn die Affen den Zoo regieren: - Die Tücken der flachen Hierarchien*.

Kühl, S. & Kullmann, G. (1999). *Gruppenarbeit: Einführen, bewerten und weiterentwickeln*.

Laloux, F. (2015). *Reinventing Organizations: - Ein Leitfaden zur Gestaltung sinnstiftender Formen der Zusammenarbeit*.

Lanza, G., Hofmann, C., Stricker, N., Biehl, E. & Braun, Y. (2018). Auf dem Weg zum digitalen Shopfloor Management: - Eine Studie zum Stand der Echtzeitentscheidungsfähigkeit und des Industrie 4.0-Reifegrads. *Forschungsbericht des Karlsruher Instituts für Technologie (KIT)*.

Lanza, G. & Nyhuis, P. (2016). Menschenzentrierte Regelkreise in der Industrie 4.0. In T. Bauernhansl & U. Dombrowski (Hrsg.), *Einfluss von Industrie 4.0 auf unsere Fabriken und die Fabrikplanung, Tagungsband des Fachbeirats Deutscher Fachkongress Fabrikplanung* (S. 24–25).

Laub, M. (2013). Netzwerke gegen die Stromfresser. *Handelsblatt* (67/2013 (08.04.2013)), 6.

Laufer, H. (2013). *Praxis erfolgreicher Mitarbeitermotivation: - Techniken, Instrumente, Arbeitshilfen*.

Leidinger, B. (2017). *Wertorientierte Instandhaltung: - Kosten senken, Verfügbarkeit erhalten*.

Lemmens, V. (2017). Enterprise Operational Intelligence: - Ein neues Instrument zur Strategieumsetzung. *Industrie 4.0 Management (I40M)*, *33* (5), S. 57-60.

Leopold, A. (2019). Growth Hacking HR - Was Unternehmenskultur vom Marketing lernen kann. In D. Brommer, S. Hockling & A. Leopold (Hrsg.), *Faszination New Work: 50 Impulse für die neue Arbeitswelt* (S. S. 243-242).

Leyendecker, B. & Pötters, P. (2018). *Shopfloor Management: - Führen am Ort des Geschehens*.

Liebermeister, B. (2017). *Digital ist egal: Mensch bleibt Mensch - Führung entscheidet*.

Lindner, D. & Amberg, M. (2019). Ist Agilität Voraussetzung oder Folge einer zielgerichteten Digitalisierung? *Industrie 4.0 Management (I40M)*, *35* (4), 30–34.

Lorenz, M., Küpper, D., Rüßmann, M., Heidemann, A. & Bause, A. (2016). *Time to Accelerate in the Race Toward Industry 4.0.*

Losch, R. (2016). *Audi will das Fließband abschaffen.* Zugriff am 20.06.2020 auf https://www.heise.de/newsticker/meldung/Audi-will-das-Fliessband-abschaffen-3504451.html

Lowe, J., Delbridge, R. & Oliver, N. (1997). High-Performance Manufacturing - Evidence from the Automotive Components Industry. *Organization Studies Journal (ORS)*, *18* (5), 783–798.

Lucke, D., Defranceski, M. & Adolf, T. (2017). Cyberphysische Systeme für die prädiktive Instandhaltung. In B. Vogel-Heuser, T. Bauernhansl & M. ten Hompel (Hrsg.), *Handbuch Industrie 4.0 Bd.1* (S. 75–92).

Lucks, K. (2017a). Die Player in der Wetschöpfungskette für Industrie 4.0. In K. Lucks (Hrsg.), *Praxishandbuch Industrie 4.0* (S. 55–72).

Lucks, K. (2017b). Grundlagen und Definitionen einer Industrie 4.0. In K. Lucks (Hrsg.), *Praxishandbuch Industrie 4.0* (S. 3–16).

Lucks, K. (2017c). Interview mit Dr. Stephan Fischer, Head of Software Development, TRUMPF Werkzeugmaschinen GmbH & CoKG. In K. Lucks (Hrsg.), *Praxishandbuch Industrie 4.0* (S. 275–290).

Manzei, C. (2016). Einführung und Überblick. In C. Manzei, L. Schleupner & R. Heinze (Hrsg.), *Industrie 4.0 im internationalen kontext* (S. 9–16). Berlin.

Mättig, B., Jost, J. & Kirks, T. (2018). Erweiterte Horizonte - Ein technischer Blick in die Zukunft der Arbeit. In S. Wischmann & E. A. Hartmann (Hrsg.), *Zukunft der Arbeit* (S. 63–72).

Matuschek, I. (2017). Industrie 4.0 - Transformation der Arbeitsgesellschaft. In *Digitalisierung und Transformation.*

Mayer, H. O. (2009). *Interview und schriftliche Befragung: - Entwicklung, Durchführung und Auswertung*.

Mertens, P., Barbian, D. & Baier, S. (2017). *Digitalisierung und Industrie 4.0 - eine Relativierung*.

Metternich, J., Adolph, S., Hambach, J., Hertle, C., Meudt, T. & Wank, A. (2017). Lean 4.0: Durch Digitalisierung die nächste Stufe der Exzellenz erreichen - Der Darmstädter Ansatz. In K. Lucks (Hrsg.), *Praxishandbuch Industrie 4.0* (S. 193–216).

Metternich, J., Müller, M., Meudt, T. & Schaede, C. (2017). Lean 4.0 - zwischen Widerspruch und Vision. *Zeitschrift für wirtschaftlichen Fabrikbetrieb (ZWF)*, *112* (5), 346–348.

Meudt, T., Hartmann, L. & Metternich, J. (2018). *Leitfaden Industrie 4.0 trifft Lean - Wertschöpfung ganzheitlich steigern*.

MM MaschinenMarkt. (2006). *Innovationsfähigkeit in der Produktion.* Zugriff am 13.06.2020 auf https://www.maschinenmarkt.vogel.de/plattform-unterstuetzt-kmu-bei-der-automatisierung-a-2352/

Möllendorf, J. (2018). Das Wertesystem macht die Unternehmenspersönlichkeit aus. In J. Reinnarth, C. Schuster, J. Möllendorf & A. Lutz (Hrsg.), *Chefsache Digitalisierung 4.0* (S. 155–162). Wiesbaden.

Moosbrugger, H. & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion*.

Müller, E. & Jentsch, D. (2015). Lernort Fabrik - Betriebliche Herausforderungen und aktuelle Lösungsansätze für eine moderne Arbeitswelt. In H. Meier (Hrsg.), *Lehren und Lernen für die moderne Arbeitswelt* (S. 97–109).

Munsch, J.-P. (2020). Standpunkt: Einführung der Selbstorganisation. In S. Wörwag & A. Cloots (Hrsg.), *Zukunft der Arbeit, Perspektive Mensch* (S. 255–261).

Neuhaus, R. (2018). Erst Lean-Management, dann Industrie 4.0! In R. M. Wagner (Hrsg.), *Industrie 4.0 für die Praxis* (S. 53–66).

Nieschlag, R., Dichtl, E. & Hörschgen, H. (1994). *Marketing*.

Niesing, B. (2013). Smart und flexibel fertigen. *weiter.vorn - Das Fraunhofer-*

Magazin, *10* (2), 8–13.

Nowotny, V. (2016). *Agile Unternehmen: - Nur was sich bewegt, kann sich verbessern: Fokussiert, schnell, flexibel*.

Nyhuis, P. (2008). Workshop 2 - Fertigungssteuerung & Logistik - Mensch & Organisation. In P. Nyhuis, G. Reinhart & E. Abele (Hrsg.), *Wandlungsfähige Produktionssysteme* (S. 138–142).

Nyhuis, P., Bellmann, V. K. & Ansari, S. M. (2015). Auswirkungen von globalen Trends auf die Lehr- und Lernkonzepte der Zukunft. In H. Meier (Hrsg.), *Lehren und Lernen für die moderne Arbeitswelt* (S. 163–181).

Ohno, T. (1988). *Toyota Production System: - Beyond Large-Scale Production*.

Orpen, C. (1997). The Effects of Job Enrichment on Employee Satisfaction, Motivation, Involvement and Performance - A Field Experiment. *Journal of Human Relations (JHR)*, *32* (3), 189–217.

o.V. (2019). Digitalisierung und Smart Integrated Factory: - „Wir stehen erst am Anfang“. *Markt & Technik*, *43* (49), 35–36.

Pachow, J., Heins, M., von Bredow, M., Krebs, P. & Wörn, A. (2008). Management der Wandlungsfähigkeit: - Forschungsbedarf für die Produktion von morgen. In P. Nyhuis, G. Reinhart & E. Abele (Hrsg.), *Wandlungsfähige Produktionssysteme* (S. 93–101).

Parsons, T. (1951). *The Social System*.

Pawellek, G. & Schirrmann, A. (2017). Planung der Industrie 4.0 Technologien und Ermittlung von Anwendungsbereichen im Unternehmen. In K. Lucks (Hrsg.), *Praxishandbuch Industrie 4.0* (S. 115–136).

Pfeiffer, S. (2016). Berufliche Bildung 4.0? - Überlegungen zur Arbeitsmarkt- und Innovationsfähigkeit. *Industrielle Beziehungen. Zeitschrift für Arbeit, Organisation und Management*, *23* (1), 25–44.

Pfeiffer, S. & Suphan, A. (2018). Industrie 4.0 und Erfahrung - Das unterschätzte Innovations- und Gestaltungspotenzial der Beschäftigten im Maschinen- und Automobilbau. In H. Hirsch-Kreinsen, P. Ittermann & J. Niehaus (Hrsg.), *Digitalisierung industrieller Arbeit* (S. 275–302).

Pfister, A. & Müller, P. (2019). Psychologische Grundlagen des agilen Arbeitens. In *Führen in der Arbeitswelt 4.0* (S. 33–50).

Picavi GmbH. (2020a). *Picavi Pick-by-Vision - Datenbrillen mit praktischem Barcode-Ringscanner.* Zugriff am 11.06.2020 auf https://picavi.com/technologie/#details

Picavi GmbH. (2020b). *Picavi Pick-by-Vision - Kommissionieren mit Datenbrille.* Zugriff am 11.06.2020 auf https://picavi.com/

Picavi GmbH. (2020c). *Produkte von Picavi - Innovation mit Erfolgsgarantie.* Zugriff am 11.06.2020 auf https://picavi.com/produkte/

Plattform Industrie 4.0. (2015). *Umsetzungsstrategie Industrie 4.0: - Ergebnisbericht der Plattform Industrie 4.0.*

Plattform Industrie 4.0. (2018a). *Factsheet: Online-Landkarte: - Deutschlandweite Anwendungsbeispiele der Industrie 4.0.* Zugriff am 20.06.2020 auf https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/Factsheets/factsheet_online-landkarte.pdf?__blob=publicationFile&v=5

Plattform Industrie 4.0. (2018b). *RAMI 4.0 - Ein Orientierungsrahmen für die Digitalisierung.* Zugriff am 666 auf https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/rami40-einfuehrung-2018.html

Plattform Industrie 4.0. (2019). *Interoperabilität - Unsere Vision für Industrie 4.0: Maschinen sprechen in vernetzten digitalen Ökosystemen interoperabel miteinander.*

Plattform Industrie 4.0 & Fraunhofer-Institut für Arbeitswirtschaft und Organisation. (2020). *KapaflexCy: - Selbstorganisierte Kapazitätsflexibilität für die Industrie 4.0.* Zugriff am 18.06.2020 auf https://www.plattform-i40.de/PI40/Redaktion/DE/Anwendungsbeispiele/096-kapaflexcy-selbstorganisierte-kapazitaetsflexibilitaet-fuer-die-industrie-4-0/beitrag-kapaflexcy-selbstorganisierte-kapazitaetsflexibilitaet-fuer-die-industrie-4-0.html

Plattform Industrie 4.0 & Pheonix Contact Electronics GmbH. (2020). *Wandlungsfähige Fertigung bei Phoenix Contact: - Flexibel durch ein modulares Anlagen- und Steuerungskonzept.* Zugriff am 22.05.2020 auf https://www.plattform-i40.de/PI40/Redaktion/DE/Anwendungsbeispiele/148-wandlungsfaehige-fertigung-bei-phoenix-contact/beitrag-wandlungsfaehige-fertigung-bei-phoenix-contact.html

Plattform Industrie 4.0, Schaeffler Technologies AG & Co. KG. (2020). *Autonome Instandhaltung: - Industrie 4.0: Selbststeuernde Maschinenüberwachung.* Zugriff am 13.06.2020 auf https://www.plattform-i40.de/PI40/Redaktion/DE/Anwendungsbeispiele/141-autonome-instandhaltung-industrie-4-0-selbststeuernde-maschinenueberwachung/beitrag-autonome-instandhaltung-industrie-4-0-selbststeuernde-maschinenueberwachung.html

Posluschny, P. (2007). *Die wichtigsten Kennzahlen*.

Pötter, T. & Steckenreiter, T. (2017). Digital Manufacturing - Prozessindustrie und Life Sciences im Wandel. In K. Lucks (Hrsg.), *Praxishandbuch Industrie 4.0* (S. 455–468).

Probst, G. J. & Büchel, B. (1994). *Organisationales Lernen: - Wettbewerbsvorteil der Zukunft*.

Projektträger Karlsruhe (PTKA) des Bundesministeriums für Bildung und Forschung. (2017). *Industrie 4.0 - Innovationen für die Produktion von morgen.*

Ramezani, H. & Luckow, A. (2019). Big Data, Small Data, and Getting Products Right First Time. In M. Dastbaz & P. Cochrane (Hrsg.), *Industry 4.0 and Engineering for a Sustainable Future* (Bd. 56, S. 77–90).

Ramsauer, C. (2013). Industrie 4.0 - Die Produktion der Zukunft. *WINGbusiness*, *46* (3), 6–12.

Raveling, J. (2016). *BLG LOGISTICS hat einiges auf Lager - zum Beispiel eines der modernsten Kommissioniersysteme Europas: - Roboter in der Lagerhalle erhöhen Flexibilität und senken Kosten.* Zugriff am 11.06.2020 auf https://www.wfb-bremen.de/de/page/stories/digitalisierung-industrie40/blg-logistics-hat-einiges-auf-lager-zum-beispiel-eines-der-modernsten-kommissioniersysteme-europas

Reichmann, T., Kißler, M. & Baumöl, U. (2017). *Controlling mit Kennzahlen: - Die systemgestützte Controlling-Konzeption*.

Reischauer, G. (2015). Industrie 4.0 organisieren: - Ansätze zur Gestaltung der formalen und informalen Organisation. *Zeitschrift Führung und Organisation (ZFO)*, *84* (4), 271–277.

Riffelmacher, P., Kluge, S. & Westkämper, E. (2009). Wandlungsfähige Produktion durch integrierten Einsatz digitaler Werkzeuge. *Industrie 4.0 Management (I40M)*, *25* (3), 29–32.

Rosenstiel, L. v. (2007). *Grundlagen der Organisationspsychologie: - Basiswissen und Anwendungshinweise*.

Rost, M., Stölzel, M. & Kozica, A. (2016). *Neue Arbeitsaufgaben und Anforderungen durch Industrie 4.0: - Eine qualitative Fallstudie in der Automobilindustrie.* Zugriff auf https://wm.baden-wuerttemberg.de/fileadmin/redaktion/m-wm/intern/Dateien_Downloads/Arbeit/Arbeitsmarktpolitik_Arbeitsschutz/A2_Rost_Arbeitswelt-2016_Presentation_Rost_et_al-160919_-_Kopie.pdf

Roth, A. (2018). Industrie 4.0: Die Digitalisierung von Produktion und Logistik. In A. Klein (Hrsg.), *Modernes Produktionscontrolling für die Industrie 4.0* (S. 127–142).

Sanders, F. & Lampe, A. (2011). Gesundheitsmanagement bei Volkswagen Nutzfahrzeuge. In Badura, B., Ducki, A., Schröder, H., Klose, J., Macco, K. (Hrsg.), *Fehlzeiten-Report 2011* (S. 209–219).

Sass, E. (2019). *Mitarbeitermotivation, Mitarbeiterbindung: - Was erwarten Arbeitnehmer?*

Schaeffler Technologies AG & Co. KG. (2017). *Den optimalen Wartungszeitpunkt vorhersagen.* Zugriff am 13.06.2020 auf https://www.schaeffler.com/content.schaeffler.com/de/news_media/stories/digitalization_stories/predictive_maintenance/predictive_maintenance.jsp

Schaeffler Technologies AG & Co. KG. (2020). *Optimierung durch Instandhaltung 4.0 für Intralogistik 4.0.* Zugriff am 13.06.2020 auf https://www.schaeffler.de/content.schaeffler.de/de/produkte-und-loesungen/industrie/industrie-4.0/kundenprojekte-geschaeftsmodelle/instandhaltung-4.0-distributionszentrum-mitte/index.jsp

Schaper, N. (2019a). Arbeitsanalyse und -bewertung. In F. W. Nerdinger, G. Blickle & N. Schaper (Hrsg.), *Arbeits- und Organisationspsychologie* (S. 385–410).

Schaper, N. (2019b). Arbeitsgestaltung in Produktion und Verwaltung. In F. W. Nerdinger, G. Blickle & N. Schaper (Hrsg.), *Arbeits- und Organisationspsychologie* (Bd. 70, S. 411–434).

Schatz, A. & Bauernhansl, T. (2017). Geschäftsmodell-Innovationen. In B. Vogel-Heuser, T. Bauernhansl & M. ten Hompel (Hrsg.), *Handbuch*

Industrie 4.0 Bd.1 (S. 245–260). Berlin, Heidelberg.

Schenk, M., Haase, T., Keller, A. & Berndt, D. (2016). Herausforderungen der Mensch-Technik-Interaktion für die Gestaltung zukünftiger Arbeitssysteme. In C. M. Schlick (Hrsg.), *Megatrend Digitalisierung* (S. 131–139).

Schirmer, D., Blinkert, B. & Buchen, S. (2009). *Empirische Methoden der Sozialforschung: - Grundlagen und Techniken* (Bd. 3175).

Schlick, C., Bruder, R. & Luczak, H. (2018). Arbeitsschutz - Arbeitssicherheit und Gesundheitsschutz. In S. Werther & L. Bruckner (Hrsg.), *Arbeit 4.0 aktiv gestalten* (Bd. 35, S. 385–416).

Schlick, C. M., Heinen, S. & Frenz, M. (2015). Entwicklung eines innovativen Fernlehrgangs einschließlich Mobile Learning zur Förderung reflexiver Handlungsfähigkeit in der Produktionsorganisation. In H. Meier (Hrsg.), *Lehren und Lernen für die moderne Arbeitswelt* (S. 39–56).

Schlund, S. & Pokorni, B. (2017). Industrie 4.0 - Wo steht die Revolution der Arbeitsgestaltung? In G. Banse, U. Busch & M. Thomas (Hrsg.), *Digitalisierung und Transformation* (S. 139–164).

Schmitt, M. (2012). Zwischen Strategie und Produktionsreporting - Produktivitätkennzahlen als Bindeglied. In A. Klein & H. Schnell (Hrsg.), *Controlling in der Produktion* (S. 121–138).

Schneider, M. (2016). *Lean Factory Design: - Gestaltungsprinzipien für die perfekte Produktion und Logistik*.

Schnell, H. (2012a). Effizienzmessung in der Produktion mithilfe von Kennzahlen. In A. Klein & H. Schnell (Hrsg.), *Controlling in der Produktion* (S. 41–62).

Schnell, H. (2012b). Produktionscontrolling: Bedeutung, Selbstverständnis, Aufgaben, Instrumente. In A. Klein & H. Schnell (Hrsg.), *Controlling in der Produktion* (S. 21–40).

Schnell, H. (2018a). Industrie 4.0: Folgen und Empfehlungen für das Produktionscontrolling. In A. Klein (Hrsg.), *Modernes Produktionscontrolling für die Industrie 4.0* (S. 41–54).

Schnell, H. (2018b). Kennzahlen des Produktionscontrollings zur Sicherung der Produktivität. In A. Klein (Hrsg.), *Modernes Produktionscontrolling für die Industrie 4.0* (S. 83–106).

Schocke, K.-O. (2017). Industrie 4.0. In C. H. Glock & E. H. Grosse (Hrsg.), *Warehousing 4.0* (S. 1–16).

Schomburg, F., Sobieraj, A. & Kruse, P. (2016). Paradigmenwechsel in der Führung: - Zukunft ohne Management. In T. Petry (Hrsg.), *Digital Leadership: Freiburg: Haufe* (S. 85–97).

Schröder, W. (2019). *Mitarbeitermotivation - worauf es wirklich ankommt -: Fallbeispiele, psychologische Erkenntnisse, Motivations- und Führungsprinzipien, konkrete Motivationsideen und Kommunikationsregeln zur Motivationssteigerung von Mitarbeitern im Betriebsalltag*.

Schuh, G., Fabry, C., Schmitz-Urban, A. & Siegers, J. (2015). Erfolgreiche Steuerung und Koordination von Business-Communities. In S. Jeschke, A. Richert, F. Hees & C. Jooß (Hrsg.), *Exploring Demographics* (S. 645–656).

Schultz-Wild, L. & Lutz, B. (1997). *Industrie vor dem Quantensprung: - Eine Zukunft für die Produktion in Deutschland*.

Schulze, H. (2000). Erfahrungsgeleitete Arbeit als Leitbild für die Entwicklung und Gestaltung von Produktionssystemen in der industriellen Fertigung. *Dissertation an der Universität Hamburg*.

Schumann, S. (2006). *Repräsentative Umfrage: - Praxisorientierte Einführung in empirische Methoden und statistische Analyseverfahren* (4. Aufl.). München and Wien.

Schwab, K. (2016). *Die Vierte Industrielle Revolution*.

Seel, C. (1995). Gruppenarbeit in der Produktion: Rationalisierungspaket oder kultureller Umbruch? In N. Sokianos (Hrsg.), *Produktion im Wandel* (S. 152–179).

Senderek, R. (2018). Lernförderliche Arbeitssysteme für die Arbeitswelt von morgen. In (S. 87–105).

Sengotta, M. (2003). Controlling-Kennzahlen zur Erfolgsvalidierung bei der Einführung eines Produktionssystems. *Controlling - Zeitschrift für erfolgsorientierte Unternehmenssteuerung*, *15* (5), 245–250.

Sharma, D. & Ali, M. (2010). Framework for Implementing Flexible Automation in Indian Industries. *Global Business and Management Research: An International Journal*, *2* (2), S. 208–223.

Siepmann, D. (2016). Industrie 4.0: - Technologische Komponenten. In A. Roth (Hrsg.), *Einführung und Umsetzung von Industrie 4.0* (S. 47–72).

Sinsel, A. (2020). Wirtschaftlichkeitsbewertung der Smart Factory. In A. Sinsel (Hrsg.), *Das Internet der Dinge in der Produktion* (Bd. 112, S. 131–146).

Soder, J. (2017). Use Case Production. In B. Vogel-Heuser, T. Bauernhansl & M. ten Hompel (Hrsg.), *Handbuch Industrie 4.0 Bd.1* (S. 3–26).

Sorge, A. (2003). Cross-National Differences in Human Resources and Organization. In A.-W. Harzing & J. van Ruysseveldt (Hrsg.), *International Human Resource Management* (S. 117–140).

Spath, D. (2017). Grundlagen der Organisationsgestaltung. In D. Spath, E. Westkämper, H.-J. Bullinger & H.-J. Warnecke (Hrsg.), *Neue Entwicklungen in der Unternehmensorganisation* (S. 3–30).

Spath, D., Dworschak, B. & Zaiser, H. (2016). Kompetenzentwicklung in der Industrie 4.0 - Weiterbildung für die Digitalisierung der produktionsnahen Arbeit. *Industrie 4.0 Management (I40M)*, *32* (3), S. 51–54.

Spath, D., Dworschak, B., Zaiser, H. & Kremer, D. (2015). Kompetenzentwicklung in der Industrie 4.0. In H. Meier (Hrsg.), *Lehren und Lernen für die moderne Arbeitswelt* (S. 113–124).

Spath, D., Ganschar, O., Gerlach, S., Hämmerle, M., Krause, T. & Schlund, S. (2013). *Produktionsarbeit der Zukunft - Industrie 4.0 (Studie)*.

Steegmüller, D. & Zürn, M. (2017). Wandlungsfähige Produktionssysteme für den Automobilbau der Zukunft. In B. Vogel-Heuser, T. Bauernhansl & M. ten Hompel (Hrsg.), *Handbuch Industrie 4.0 Bd.1* (S. 27–44). Berlin, Heidelberg.

Steute Technologies GmbH & Co. KG. (2020). *Meditec Produkte.* Zugriff am 12.06.2020 auf https://www.steute-meditec.com/de/produkte.html

Stich, V., Gudergan, G. & Senderek, R. (2018). Arbeiten und Lernen in der digitalisierten Welt. In H. Hirsch-Kreinsen, P. Ittermann & J. Niehaus (Hrsg.), *Digitalisierung industrieller Arbeit* (S. 143–172).

Stief, S. (2018). *Nutzen und Akzeptanz von Industrie 4.0-Anwendungen*.

Stock, T. (2013). *Ein Verfahren zur Personalplanung und -steuerung und Restrukturierung der Aufbauorganisation für eine bedarfsorientierte und wandlungsfähige Produktion*.

Strauß, R. E. (1996). *Determinanten und Dynamik des Organizational Learning: Zugl.: Freiburg (Breisgau), Univ., Diss., 1996*.

Strauß, R. E. (2019). *Digitale Transformation: - Strategie, Konzeption und Implementierung in der Unternehmenspraxis*.

Strunz, M. (2012). *Instandhaltung: - Grundlagen, Strategien, Werkstätten*.

Tawalbeh, M., Weißflog, L. & Hopf, H. (2019). Wissensmanagement in KMU mithilfe von Assistenzsystemen: - Effizienzsteigernder Einsatz von Assistenzsystemen in KMU. *Industrie 4.0 Management (I40M)*, *3* (35), S. 15-18.

TEKA Absaug- und Entsorgungstechnik GmbH. (2018). Aktuell: - Ampel-Sensor überwacht Raumluft. *Kälte Klima Aktuell (KKA)*, *37* ((Großkälte)), 4.

TEKA Absaug- und Entsorgungstechnik GmbH. (2019). *Teka präsentiert passgenaue Lösungen für die Anlagen-Vernetzung.* Zugriff am 13.06.2020 auf http://www.schweisstechnik.at/detail/%0Bteka-praesentiert-passgenaue-loesungen-fuer-die-anlagen-vernetzung_150873

Tschandl, M., Brunner, U. & Wilfinger, D. (2014). Logistikservice als Kernelement des Logistikcontrollings: - Performance Management des Kundennutzens. In R. Gleich & C. Daxböck (Hrsg.), *Supply Chain- und Logistikcontrolling* (S. 73–96).

Ullrich, A. & Vladova, G. (2015). Qualifizierungsmanagement in der vernetzten Produktion - Ein Ansatz zur Strukturierung relevanter Parameter. In H. Meier (Hrsg.), *Lehren und Lernen für die moderne Arbeitswelt* (S. 57–80).

Ungermann, F., Jacob, A., Verhaelen, B., Itterheim, A., Park, Y.-B., Stricker, N. & Lanza, G. (2019). Die Zukunft der Kennzahlensysteme: - Unternehmenssteuerung durch ein ganzheitliches KPI-Netzwerk auf Basis eines Digital Twins. *Industrie 4.0 Management (I40M)*, *35* (4), 25–29.

Upton, D. M. (1994). The Management of Manufacturing Flexibility. *California Management Review*, *36* (2), 72–89.

von See, B. & Kersten, W. (2018). Arbeiten im Zeitalter des Internets der Dinge: - Wie Qualifikation, Organisation und Führung digital transformiert werden. *Industrie 4.0 Management (I40M)*, *34* (3), 8–12.

Weber, M. (2006). *Schnelleinstieg Kennzahlen*.

Weinreich, U. (2016). *Lean Digitization: - Digitale Transformation durch agiles Management*.

Werner, H. (2014). *Kompakt Edition: Supply Chain Controlling: - Grundlagen,*

Performance-Messung und Handlungsempfehlungen.

Werner, H. C. & Gundrum, E. (2019). Führen in der Arbeitswelt 4.0. In C. Negri (Hrsg.), *Führen in der Arbeitswelt 4.0* (S. 9–16).

Werther, S. & Bruckner, L. (2018). *Arbeit 4.0 aktiv gestalten: - Die Zukunft der Arbeit zwischen Agilität, People Analytics und Digitalisierung*.

Westkämper, E. (1999). Die Wandlungsfähigkeit von Unternehmen. *WT Werkstattstechnik online*, *89* (4), 131–140.

Westkämper, E. & Decker, M. (2006). *Einführung in die Organisation der Produktion*.

Westkämper, E. & Löffler, C. (2016). *Strategien der Produktion: - Technologien, Konzepte und Wege in die Praxis*.

Westkämper, E., Zahn, E., Balve, P. & Tilebein, M. (2000). Ansätze zur Wandlungsfähigkeit von Produktionsunternehmen. *WT Werkstattstechnik online*, *90* (1/2), 22–26.

Wiendahl, H.-P. (2002). Wandlungsfähigkeit: Schlüsselbegriff der zukunftsfähigen Fabrik. *WT Werkstattstechnik online*, *92* (4), 122–127.

Wiendahl, H.-P., Reichardt, J. & Nyhuis, P. (2014). *Handbuch Fabrikplanung: - Konzept, Gestaltung und Umsetzung wandlungsfähiger Produktionsstätten*. Hanser, Wien.

Will-Zocholl, M. (2011). *Wissensarbeit in der Automobilindustrie: -Topologie der Reorganisation von Ingenieursarbeit in der globalen Produktentwicklung*.

Windelband, L. & Dworschak, B. (2018). Arbeit und Kompetenzen in der Industrie 4.0 - Anwendungsszenarien, Instandhaltung und Leichtbaurobotik. In H. Hirsch-Kreinsen, P. Ittermann & J. Niehaus (Hrsg.), *Digitalisierung industrieller Arbeit* (S. 63–80).

Winkler, M., Grau, G. & Tilebein, M. (2015). Lebenslanges Lernen - Neue Ansätze für die Textilwirtschaft. In H. Meier (Hrsg.), *Lehren und Lernen für die moderne Arbeitswelt* (S. 15–38).

Winkler, R. (2019). *Digital Manufacturing: Wie die Wertschöpfung der Produktion in der Zukunft gelingt: - Industrie 4.0 im Spannungsfeld von Kundenanforderungen und Flexibilität*. Zugriff am 16.06.2020 auf https://www.capgemini.com/de-de/2019/03/wertschoepfung-digitale-zukunft/

Wischmann, S. & Hartmann, E. A. (2018a). Prognostizierte Veränderungen der gestaltbaren Arbeitssystemdimensionen. In S. Wischmann & E. A. Hartmann (Hrsg.), *Zukunft der Arbeit* (S. 9–31).

Wischmann, S. & Hartmann, E. A. (2018b). *Zukunft der Arbeit: - Eine praxisnahe Betrachtung*. Berlin, Heidelberg.

Zäh, M. & Reinhart, G. (Hrsg.). (2005). *1st International Conference on Changeable, Agile, Reconfigurable and Virtual Production (CARV 2005)*. München.

Zentralverband Elektrotechnik- und Elektronikindustrie e.V. (2015). *Industrie 4.0: Auf dem Weg zur smarten Fabrik - Die Elektroindustrie geht voran.*

Zimmermann, K. (2017). Digitalisierung der Produktion durch Industrie 4.0 und ihr Einfluss auf das Arbeiten von morgen. In B. Spieß & N. Fabisch (Hrsg.), *CSR und neue Arbeitswelten* (S. 53–72).

Zwirner, H. (2004). Von der Amöbe lernen. *McK Wissen - Das Magazin von McKinsey*, *3* (8), 56–61.

Junior Management Science 7(4) (2022) 1098-1132

Junior Management Science

journal homepage: www.jums.academy

Digital Innovation in Corporations: Deriving a Practical Framework for the Measurement of Success of Digital Innovation Units

Marcel Schemmel

Technische Universität München

Abstract

Confronted with entirely new challenges resulting from digital technologies, established corporations increasingly set up dedicated digital innovation units (DIUs) to foster digital innovation and to explore opportunities for the digital future. Although DIUs recently face criticism with regards to their performance and impact on the core organization, literature lacks in suitable approaches to assess the success of DIUs. Therefore, we derive a practical framework for the measurement of success of DIUs in the course of this research project. We develop this framework by identifying critical success factors (CSFs) and key performance indicators (KPIs). Subsequently, we merge our results with existing literature. To determine these CSFs and KPIs, we designed an explorative, qualitative-empirical case study research approach. The research design is based on a mixed-method approach that combines semi-structured interviews as core component with a supplementary survey. Conducting nine cross-industry case studies, we identified 16 CSFs and 38 objective related KPIs. Thus, the framework derived in this thesis contributes to practice and literature by addressing the existing gap in DIU and performance measurement research.

Keywords: Digital innovation units, performance measurement, critical success factors, key performance indicators, qualitative case studies

1. Introduction

Digital technologies have been a game changer shaking up traditional approaches to innovation and posing entirely new challenges for established corporations (Fichman, Dos Santos, & Zheng, 2014; Yoo, Henfridsson, & Lyytinen, 2010). Numerous industries have been disrupted by digital innovation already (Hund, Drechsler, & Reibenspiess, 2019). Hence, organizations need to maintain existing business while exploring opportunities for the digital future (Fuchs, Barthel, Herberg, Berger, & Hess, 2019). To address these developments and to foster digital innovation, organizations increasingly build up dedicated organizational units (Holotiuk, 2020), which appear in various forms and types (Sindemann & von Buttlar, 2018). Within this study, we define DIUs as "organizational units with the overall goal to foster organizational digital transformation by performing digital innovation activities for existing and novel business areas" (Barthel, Fuchs, Birner, & Hess, 2020, p. 4).

Several studies revealed the high adoption of DIUs among corporations. A large-scale report published by the Boston Consulting Group revealed, that among 570 corporates from Germany, Austria and Switzerland, 19% had already established at least one DIU in 2018 (Brigl, Gross-Selbeck, Dehnert, F., & Steffen, 2019). While as of today, central research and development (R&D) is still perceived as most important source of innovation for corporations, DIUs are expected to replace R&D as most important innovation source in five years from now (Thompson, Bonnet, & Jaballah, 2020), showing the increasing importance of DIUs as corporate innovation vehicle.

Despite the growing interest, DIUs have been criticized recently with regards to their performance and impact on the core organization. The hype around DIUs has flattened as the expectation to transform corporations towards an agile and flexible company was oftentimes not sufficiently met (Meyer, 2020). Frequently, DIUs were established without having a clear concept in place (Kyriasoglou, 2020) and without incorporating clearly defined objectives (Raabe, Horlach, Drews, & Schirmer, 2020). Hence, DIUs might not be able to meet the high expectations of the core organizations, which is one of the reasons why some DIUs were sufficiently downsized or even shut down completely (Kyriasoglou, 2020).

Due to the increasing importance from a practical point of view, DIUs have also gained broad attention by academia

DOI: https://doi.org/10.5282/jums/v7i4pp1098-1132

over the last years. So far, research has covered various DIU related topics, including the innovation process for digital innovation (Fichman et al., 2014), the organizational design (Fuchs et al., 2019; Holotiuk, 2020), the typology of DIUs (Barthel et al., 2020) and objectives (Raabe, Drews, Horlach, & Schirmer, 2021). Although DIUs are facing criticism in regard to their success, related literature is rare. Hence, practitioners and scientific research increasingly ask for approaches to measure the performance of DIUs including success factors and lucid measures (Barthel et al., 2020). So far, literature covered key challenges of measuring design thinking activities at DIUs (Mayer, Haskamp, & De Paula, 2021) and derived several requirements for Performance Measurement Systems (PEMS) in DIUs (Haskamp, Lorson, de Paula, & Uebernickel, 2021).

However, research is still limited when it comes to CSFs for DIUs, which are important areas that should be addressed to be successful and to fulfill respective objectives (Bullen & Rockart, 1981). Moreover, existing literature lacks in regard to objective related KPIs that are or should be used to measure the performance of DIUs. To close this research gap, we aimed at answering the following research questions:

> RQ1: What are CSFs for DIUs to fulfill their objectives?
> RQ2: Which KPIs are or should be used by DIUs to measure their performance?

Due to the novelty of the research field, we designed an explorative, qualitative-empirical case study research approach that utilizes a mixed-method research design with semi-structured interviews as core component. To answer the research questions, we draw data resulting from nine cross-industry case studies as well as additional case related findings from a supplementary survey that is conducted to make the research project more comprehensive.

We merged the identified 16 CSFs and 38 objective related KPIs with existing literature and derived a practical framework for the measurement of success of DIUs that consists of six guiding steps and covers success related factors for DIUs. Thus, we address the existing research gap and contribute to DIU and performance measurement research. In addition, the framework sets the foundation to measure the success of DIUs in a practical way, through guiding DIU employees, responsible for success and performance measurement, in regard to establishing or improving the measurement of success of their DIU.

Initially, Chapter 2 provides a lucid introduction of the evolution of DIUs. Following, Chapter 3 sets the theoretical foundation through closely analyzing performance measurement at DIUs, including different perspective on CSFs, PEMS and the application of KPIs to measure digital innovation. In Chapter 4, an extensive introduction of the applied research design is provided. Subsequently, the results of this study are introduced in Chapter 5. Through merging the findings of this study with existing literature, the new framework for the measurement of success of DIUs is derived and presented in Chapter 6, which additionally covers insights regarding implications, limitations and future research, all being concluded in Chapter 7.

2. The evolution of DIUs

To develop a comprehensive understanding of why and how DIUs are set up by established corporations, it is essential to get an overview of specific DIU features. Therefore, we will now take a closer look at selected underlying concepts, before we subsequently show several approaches to distinguish different types of DIUs. Concluding, we will provide an overview of objectives that are pursued by DIUs.

2.1. Underlying concepts

To better understand the purpose of DIUs, we introduce digital innovation, which is oftentimes fostered by an innovation process particularly for digital innovation, which also enables, in case of its successful execution, organizational ambidexterity. In the following, these concepts are defined consecutively.

Digital innovation: Several definitions for digital innovation have emerged over the last years (Fichman et al., 2014; Nambisan, Lyytinen, Majchrzak, & Song, 2017; Yoo et al., 2010). Compared to Yoo et al. (2010), who rather take a process-based view on digital innovation, the definition provided by Fichman et al. (2014) emphasizes an outcome-based view on digital innovation. Nambisan et al. (2017) combine these views, focusing on digital innovation as process and outcome (Hund et al., 2019). We follow this definition and define digital innovation as "the creation of (and consequent change in) market offerings, business processes, or models that result from the use of digital technology" (Nambisan et al., 2017, p. 224). Moreover, digital innovation can have an internal (i.e. organizational) and external (i.e. market) orientation (Barthel et al., 2020; Kohli & Melville, 2019; Nambisan et al., 2017). Oftentimes, innovation projects involve high degrees of uncertainty (Koen et al., 2001), which can be addressed by applying a multi-stage innovation process (Fichman et al., 2014).

Innovation process for digital innovation: Similar to digital innovation itself, multiple authors have already addressed the innovation process for digital innovation (Fichman et al., 2014; Huesig & Endres, 2019; Kohli & Melville, 2019). Within this paper, we follow Fichman et al. (2014), who split up the process into four consecutive stages: discovery, development, diffusion and impact. The innovation process is applicable for product and business model innovation, as well as organizational, internally-oriented, process innovation (Fichman et al., 2014). Discovery, the first stage, aims at identifying new ideas. The key activities of this stage consist of identifying ideas for digital innovation within an organization and evaluating external ideas for development or adoption. The second stage, development, consists of developing usable digital innovations based on the previously identified and selected ideas. Depending on the type of innovation, this stage includes different activities. For product and business model innovation, the core technology as

well as added complementary services are developed. The core technology and the surrounding service together form the digital innovation. Besides, the selection of technology features and their utilization is applied for process innovation. Diffusion, the third stage, aims at disseminating the developed solution across potential users. Within this stage, the resources, necessary to give potential users access to the solution, are assigned. In addition, diffusion strives for the adoption of the solution by its users. The last stage, impact, strongly focuses on the effects of digital innovation once it is diffused. These can be twofold, improve existing costs or increase revenues. One of the activities is "managing intellectual property and the ecosystem of complementary products and services" (Fichman et al., 2014, p. 336). Moreover, impact aims at improving and adapting the solution in order to take advantage of new opportunities (Fichman et al., 2014). With digital innovation as main objective (Barthel et al., 2020; Raabe et al., 2021), DIUs oftentimes embody particular design features to enable ambidexterity of the core organization (Holotiuk, 2020).

Organizational ambidexterity: Initially introduced by Duncan (1976), organizational ambidexterity has been widely discussed by the scientific community (Birkinshaw & Gibson, 2004; Duncan, 1976; O'Reilly III & Tushman, 2013; Tushman & O'Reilly III, 1996). Organizations oftentimes face the challenge of balancing exploration (e.g. utilizing new opportunities) and exploitation (e.g. leveraging existing resources). Hence, a trade-off in regard to assigning existing resources accordingly arises (March, 1991). Whereas multiple types of ambidexterity exist, structural and contextual ambidexterity are seen as most common (Holotiuk, 2020). Structural ambidexterity describes the need for dual structures within organizations in order to manage innovation appropriately (Duncan, 1976). Contextual ambidexterity is defined as "the capacity to simultaneously achieve alignment and adaptability at a business-unit level" (Gibson & Birkinshaw, 2004, p. 209), allowing employees to individually allocate their resources (e.g. time) to exploration and exploitation operations based on particularly established systems and processes (Duncan, 1976; Gibson & Birkinshaw, 2004; Tushman & O'Reilly III, 1996).

As established organizational designs seem to address the potential and challenges that come along with digital innovation insufficiently (Nambisan et al., 2017; Yoo et al., 2010), organizations increasingly set up dedicated DIUs (Holotiuk et al., 2020; Sindemann & von Buttlar, 2020). DIUs occur in various forms and types and help established organization to foster organizational ambidexterity (Holotiuk, 2020).

2.2. Typology of DIUs

Although DIUs have been widely adopted by established organizations, the research field is still novel from a scientific point of view (Barthel et al., 2020). Hence, taking a practical perspective first, before considering existing literature seems to be an effective approach to fully grasp existing definitions of DIU types. Concluding, several ideal types of DIUs that allow distinguishing DIUs dependent on specific criteria are introduced.

2.2.1. Practical perspective on DIU types

Although various relevant approaches of practitioners to characterize DIUs exist, they lack in consistency and have not been derived in accordance with scientific research guidelines.

Over the last years, several practitioners (e.g. consultancies), have developed multiple approaches to characterize different types of DIUs (etventure, 2021; Ramus & Velten, 2016; Sindemann & von Buttlar, 2017, 2018, 2020). Ramus and Velten (2016) differentiate between four DIU types. Whereas innovation labs and company builder focus on new ideas and startups, accelerators and incubators mainly aim at existing ideas and startups. Similar, Sindemann and von Buttlar (2017) differentiate between innovation labs, accelerators and incubators. Innovation labs utilize internal resources and combine these in interdisciplinary teams with the objective to develop digital innovation. In contrast, accelerators aim at supporting external startups within a given program set up and time frame to develop novel ideas. Characterized through more individual relationships, incubators primarily focus on supporting external startups in regard to developing ideas (Sindemann & von Buttlar, 2017). In 2018, Sindemann and von Buttlar updated this classification and replaced incubators through company builders. Compared to incubators, company builders aim at establishing startups based on novel ideas for digital business models through internal and external resources (Sindemann & von Buttlar, 2018). The latest DIU typology derived by Sindemann and von Buttlar (2020) differentiates DIUs on a more generic level. Whereas startup labs, for example accelerators and incubators, focus on collaborations between the core organization and external startups, innovation labs aim at developing digital innovation mainly in collaboration with employees from the core organization. This does not exclude irregular collaboration with universities, startups and other established companies. Furthermore, DIUs can be clustered, dependent on the market focus of their digital innovation activities, into DIUs that mainly target innovation close to the core business and innovation distant to the core business (Sindemann & von Buttlar, 2020).

Concluding, another DIU characterization was developed by the consulting company etventure (2021). It includes innovation labs, that aim at fostering digital culture and allow employees of the core organization to work on new ideas and business models. Company builders focus on establishing startups based on selected, oftentimes internal, ideas. In contrast, incubators target at supporting existing, external startups. The last type, digital units, enable interdisciplinary teams to develop innovative ideas, business models and novel ways of working (etventure, 2021).

While the DIU types and characterizations derived by practitioners are relevant, "the proposed characterizations are neither consistent nor theoretically sound" (Fuchs et

al., 2019, p. 635), highlighting the necessity for profound, theory-based, DIU classifications (Fuchs et al., 2019).

2.2.2. Theoretical perspective on DIU types

While recently several approaches to classify and characterize DIUs have emerged in scientific research, the development of DIU archetypes, that include relations between organizational dimensions, seems to be a suitable approach to fully grasp the organizational complexity of DIUs.

Similar to the contributions by practitioners, literature addressed DIUs as research field by deriving different ways of characterizing DIUs (Barthel et al., 2020; Fuchs et al., 2019; Holotiuk, 2020; Hund et al., 2019; Raabe et al., 2020). One way of distinguishing DIUs is based on the organizational design, which can be separated into several dimensions, such as motivation & vision, setup & alignment, governance, employees & staffing and operations (Holotiuk, 2020). A more extensive categorization was provided by Fuchs et al. (2019). Likewise, DIUs are distinguished based on their organizational design, however, the organizational dimensions were clustered into categories (i.e. objectives & scope, staffing & collaboration, funding, governance & structure and origins). In addition, Fuchs et al. (2019) derived particular characteristics within each category "providing a theoretically sound and empirically derived taxonomy for the characterization of digital units" (Fuchs et al., 2019, p. 643), which is depicted in Appendix 1. Although this taxonomy could be utilized as a blueprint for designing and establishing DIUs, it does not indicate any relations between specific dimensions and characteristics, for example the impact of objective and scope on staffing and collaboration (Fuchs et al., 2019).

This gap was addressed by Raabe et al. (2020) through identifying two DIU archetypes and linking mechanisms to the core organization. The key criteria to distinguish between these types is their specific focus area: Whereas type one, the coaching & screening type, mainly focuses on the discovery phase of innovation, fostering an innovative, agile mindset within the core organization and novel digital skills, type two, the center of excellence, additionally addresses innovation development, diffusion and the measurement of the innovation impact. An exemplary dependency between different organizational dimensions is the importance of an autonomous budget for being able to cover all innovation process stages. In regard to linking mechanisms, it is important for both types to shift employees between the core organization and the DIU in order to foster collaboration (Raabe et al., 2020).

As many DIUs primarily focus on creating digital innovation (Fuchs et al., 2019; Raabe et al., 2021; Sindemann & von Buttlar, 2020), Barthel et al., 2020 developed a typology of DIUs that emphasizes the "means-ends-relationship between objectives (ends) and design options (means)" (Barthel et al., 2020, p. 2). Organizational design patterns, suitable for specific innovation objectives, were identified. Furthermore, the concept of loose-tight-coupling was applied to identify how tight specific DIU types should be tied to the core organization dependent on the targeted innovation. Following this logic, three ideal types of DIUs were identified. Such ideal types "are theoretically founded abstractions to examine empirically observed real world cases" (Barthel et al., 2020, p. 5). This does not imply, that each real-world case has to follow the specific design pattern of an ideal type. However, similar organizational characteristics were identified (Barthel et al., 2020; Doty & Glick, 1994). The organizational dimensions considered were derived from the taxonomy provided by Fuchs et al. (2019) and are depicted in Appendix 2. Based on the specific DIU characteristics within the categories objective and scope as well as setting and design, the ideal types internal facilitator, external enhancer and external creator were derived by Barthel et al. (2020). Since it is crucial to develop a comprehensive understanding of how different DIUs types can be distinguished based on the specific objective and scope, a profound illustration of each ideal type is provided.

2.2.3. Ideal types: Internal facilitator, external enhancer, external creator

Each type is described based on the specific objective and scope of its digital innovation activities. In addition, suitable, respectively recommended, organizational settings and designs to achieve these objectives are introduced. The combination of both objective and scope and setting and design define the specific ideal type (Barthel et al., 2020).

Internal facilitator

Objective & Scope: The innovation orientation of type one is primarily internal (e.g. business process innovation). Hence, internal facilitators focus on existing business areas and markets. Thus, fostering the digital transformation of the core organization is more important than product- and service-related innovation. With regards to the innovation process, this type usually addresses idea generation, idea selection and innovation development. However, the remaining stages are oftentimes realized by another department within the core organization.

Setting & Design: Ideal type one should be embedded as executive department or separate legal entity, respectively. External cooperation is oftentimes not incorporated. This ideal type is usually integrated into the core organization and therefore tightly tied to it. Projects take place on- and offsite, are ordered by the core organization and autonomously selected.

Due to the internal innovation focus, tight coupling of internal facilitators, especially in regard to embedding and collaboration, seems to be a beneficial setup (Barthel et al., 2020). An overview is provided in Figure 1.

External enhancer

Objective & Scope: Similar to internal facilitators, the second type, external enhancers focus on existing business areas. In contrast to type one, type two is oriented primarily or purely external, respectively. Hence, fostering the market offering including new products, services and business models of the core organization is targeted. Although implementation and commercialization of the innovation is usually taken over by

Category	Dimension		Characteristic				
What? Objective & Scope	Innovation Orientation		Purely Internal	Primarily Internal	Balanced	Primarily External	Purely External
	Market Focus of Innovation		Existing Business Areas			Novel Business Areas	
	Scope of Innovation Process		Idea Generation	Idea Selection	Innovation Development	Innovation Implementation	Innovation Commercialization
How? Setting & Design	Embedding		Executive Department		Line Organization		Separate Legal Entity
	Collaboration	External Cooperation	None		Irregularly		Standard Working Mode
		Internal Ties	Core Organization Integrated		Regular Liaison		Sporadic Liaison
	Projects	Project Selection	Order		Mixed		Autonomously
		Location	Onsite		Mixed		Offsite

Figure 1: Overview of ideal type one: Internal facilitator
Note. Color code: dark blue = frequently identified characteristic, light blue = exceptionally identified characteristic, Figure taken from Barthel et al. (2020, p. 10).

a different department, some DIUs cover all stages of the innovation process.

Setting & Design: Similar to internal facilitators, external enhancers are integrated into the core organization, usually embedded as line organization, sometimes even as executive department. External cooperation takes place in most cases and is sometimes even incorporated as standard mode of cooperation. Projects can be selected autonomously or are ordered by the core organization and are usually conducted offsite.

Compared to type one, type two is usually coupled less tightly to the core organization induced by its embedding as line organization as well as the more frequent external cooperation (Barthel et al., 2020). The graphical overview is depicted in Figure 2.

External creator

Objective & Scope: Similar to type two, external creators are oriented externally. The key difference is the focus on novel business areas (e.g. new customer groups and markets). Therefore, innovation implementation and occasionally commercialization are addressed.

Setting & Design: Type three is usually embedded as separate legal entity and external cooperation is at least incorporated on an irregular basis, oftentimes also as standard working mode. External creators connect with the core organization on a regular basis if not integrated (internal ties). Projects are selected autonomously and conducted offsite.

Especially for creating digital innovation in novel business areas with an external focus, it seems beneficial to have a rather loose coupling to the core organization. Thus, frequent external cooperation and autonomous selection of projects should be realized (Barthel et al., 2020). The overview of ideal type three is provided in Figure 3.

Hybrid type

As the identified ideal types of DIUs are only theoretical abstractions of real-world cases, not all possible forms and types of DIUs are covered. Additionally, there are hybrid DIU types that for example cover both internal and external innovation orientation.

The provided typology by Barthel et al. (2020) is clearly focused on distinguishing different types of DIUs based on the specific innovation objective and scope, including innovation orientation, market focus and scope of innovation process. Thus, DIUs can be classified rather generically, meaning that for example company builders and accelerators can both be categorized as the same type, as long as the innovation objectives overlap. A clear set of objectives is one of the prerequisites for performance measurement (Kaplan & Norton, 1992). Therefore, it is crucial to understand which objectives are targeted by DIUs, that go beyond digital innovation.

2.3. Overview of DIU objectives

Whereas practitioners and scientific research have identified five objectives that are pursued by DIUs, digital innovation is unanimously recognized as primary objective.

So far, practitioners and literature have addressed objectives of DIUs directly (through explicitly identifying DIU objectives) or indirectly (through describing activities of DIUs that pay into specific objectives). An overview of existing contributions is depicted in Table 1.

Within the taxonomy derived by Fuchs et al. (2019, Appendix 1) one dimension specifically addresses the main objective of DIUs, the development of digital innovation. Dependent on the specific DIU type (section 2.2.3), the objective varies in regard to innovation orientation, market focus and stages covered within the innovation process (Barthel et al., 2020). In addition to digital innovation, DIUs oftentimes aim at fostering cultural change and the development of digital expertise within the core organization. These objectives can be realized indirectly, for example through integrating employees into specific projects of the DIU, or directly, for example by conducting workshops (Fuchs et al., 2019).

Building on Fuchs et al. (2019), Raabe et al. (2021) derived two further objectives, organizational design change

Category	Dimension		Characteristic				
What? Objective & Scope	Innovation Orientation		Purely Internal	Primarily Internal	Balanced	Primarily External	Purely External
	Market Focus of Innovation		Existing Business Areas			Novel Business Areas	
	Scope of Innovation Process		Idea Generation	Idea Selection	Innovation Development	Innovation Implementation	Innovation Commercialization
How? Setting & Design	Embedding		Executive Department		Line Organization		Separate Legal Entity
	Collaboration	External Cooperation	None		Irregularly		Standard Working Mode
		Internal Ties	Core Organization Integrated		Regular Liaison		Sporadic Liaison
	Projects	Project Selection	Order		Mixed		Autonomously
		Location	Onsite		Mixed		Offsite

Figure 2: Overview of ideal type two: External enhancer
Note. Color code: dark blue = frequently identified characteristic, light blue = exceptionally identified characteristic, Figure taken from Barthel et al. (2020, p. 11).

Category	Dimension		Characteristic				
What? Objective & Scope	Innovation Orientation		Purely Internal	Primarily Internal	Balanced	Primarily External	Purely External
	Market Focus of Innovation		Existing Business Areas			Novel Business Areas	
	Scope of Innovation Process		Idea Generation	Idea Selection	Innovation Development	Innovation Implementation	Innovation Commercialization
How? Setting & Design	Embedding		Executive Department		Line Organization		Separate Legal Entity
	Collaboration	External Cooperation	None		Irregularly		Standard Working Mode
		Internal Ties	Core Organization Integrated		Regular Liaison		Sporadic Liaison
	Projects	Project Selection	Order		Mixed		Autonomously
		Location	Onsite		Mixed		Offsite

Figure 3: Overview of ideal type three: External creator
Note. Color code: dark blue = frequently identified characteristic, light blue = exceptionally identified characteristic, Figure taken from Barthel et al. (2020, p. 12).

Table 1: DIU objectives covered by practitioners and literature

Author	DIU Objectives				
	Digital innovation	Cultural change	Digital expertise	Organizational design change	Digital innovation ecosystem
Barthel et al. (2020)	x	x	x		
etventure (2021)	x	x	x		x
Fuchs et al. (2019)	x	x	x		
Holotiuk (2020)	x	x	x		
Raabe et al. (2020)	x	x	x		x
Raabe et al. (2021)	x	x	x	x	x
Ramus and Velten (2016)	x				x
Sindemann and von Buttlar (2017)	x	x			x
Sindemann and von Buttlar (2020)	x	x			x

Note. Exemplary contributions.

and participating in, respectively developing, digital innovation ecosystems. The objective organizational design change

incorporates "that DIUs should initiate and enable the realization of various organizational concepts" (Raabe et al., 2021, p. 4), for example structural or conceptual changes. As organizational change might strategically impact the core organization, DIUs should be provided with high levels of authority when addressing this objective. The objective digital innovation ecosystem mainly targets at exploring the underlying ecosystem of digital technologies. Here, DIUs can either establish new digital ecosystems or participate in existing ones (Raabe et al., 2021). Furthermore, Raabe et al., 2021 identified seven areas of activity, directly or indirectly related to DIU objectives. As a detailed illustration would go beyond the scope of this study, an overview is provided in Appendix 3. One exemplary activity is to "develop and leverage digital expertise (and agile methods)" (Raabe et al., 2021, p. 4), which is directly related to the objective digital expertise and indirectly related to achieving cultural change, organizational design change and participating in the digital innovation ecosystem (Raabe et al., 2021).

Whereas the objective digital innovation is addressed by all authors (Table 1), other objectives are covered only partly, highlighting the importance of digital innovation as primary objective of DIUs. Barthel et al. (2020) and Raabe et al. (2021) support this conclusion. A profound, practically oriented study published in 2020, covers various aspects of DIUs, such as outcomes. It reveals, that 90% of the interviewed DIUs have a high or very high focus on digital opportunities that include tested ideas, solutions to problems and application of new technology (Holotiuk et al., 2020), which can be associated with the objective digital innovation.

By analyzing existing literature in regard to selected underlying concepts, approaches to distinguish different types of DIUs and objectives that are pursued by DIUs, we have developed a comprehensive understanding of why and how DIUs are set up by established corporations.

3. Performance measurement at DIUs

Performance measurement is a broad research field with various sub-topics, such as CSFs, PEMS and KPIs (Bourne, Neely, Mills, & Platts, 2003; Neely, Kennerley, & Adams, 2007). There is a strong relation between KPIs and CSFs as "key performance indicators (KPIs) tell management how the organization is performing in their critical success factors and, by monitoring them, management is able to increase performance dramatically" (Parmenter, 2015, p. 4). As we strive to develop a comprehensive understanding of performance measurement at DIUs, it is essential to get an overview of existing literature in regard to CSFs for DIUs. Furthermore, we investigate requirements for PEMS in DIUs. We take a general perspective on KPIs before concluding with ways to use KPIs for measuring digital innovation.

3.1. Critical success factors

CSFs are defined as "the limited number of areas in which satisfactory results will ensure successful competitive performance for the individual, department or organization" (Bullen & Rockart, 1981, p. 7). Furthermore, "CSFs are the few key areas where "things must go right" for the business to flourish" (Bullen & Rockart, 1981, p. 7). In alignment with various other studies (Boynton & Zmud, 1984; Freund, 1988; Holotiuk & Beimborn, 2017; Leidecker & Bruno, 1984), we use this definition as basis for our research project.

3.1.1. Practical perspective on critical success factors

While the contributions by practitioners that cover recommendations and success factors for DIUs lack in theoretical foundation, the common characteristic is the application of clustering approaches that allow to aggregate individual factors into categories.

Similar to the different DIU types that were derived by practitioners over the last years (section 2.2.1), there are several contributions that cover recommendations for DIUs, respectively success factors (Sindemann & von Buttlar, 2017, 2018). Ramus and Velten (2016) for example identified several features that are important for DIUs to consider. The first one is space, meaning that it is important to set up an appropriate location to enable collaboration. Another aspect is infrastructure and technology, which consists of suitable, powerful IT infrastructures. Mentoring and operational support is important as well, as core organizations' employees oftentimes require support to implement agile ways of working (e.g. design thinking), which can be realized through workshops. Furthermore, funding, network and ecosystem are important for DIUs to foster networking and exchange across various players within the digital ecosystem. Two primary factors for the successful establishment of DIUs is teambuilding and talent management, as well as integrating the DIU into the digital strategy of the core organization (Ramus & Velten, 2016).

Sindemann and von Buttlar (2017, 2018) defined success factors of DIUs based on their definition of different types of DIUs (section 2.2.1). In total, six areas of success factors (e.g. governance, topics) were identified and further divided into subfactors (e.g. top management support, degrees of freedom). An overview of exemplary success factors is provided in Table 2.

Further contributions by practitioners addressed "clusters of success" in regard to digital innovation activities of DIUs, divided into innovation discovery, development and scaling (Sindemann & von Buttlar, 2019), as well as important factors with regards to creating digital innovation close and distant to the core business (Sindemann & von Buttlar, 2020).

Similar to the contributions by practitioners regarding DIU types (section 2.2.1), the provided success factors lack in consistency and theoretical foundation, showing the importance for theoretically profound CSFs for DIUs.

3.1.2. Theoretical perspective on critical success factors

Although existing literature provides several important factors and recommendation for DIUs, it lacks in regard to CSFs. However, CSFs were widely identified in other re-

Table 2: Exemplary success factors for DIUs

Governance:	Topics:	Core organization engagement:
• Top management support	• Targeted degree of innovation	• Business unit collaboration
• Degrees of freedom	• Selection criteria & process	• Internal communication
Methodologies:	Scaling:	Network:
• Selection of methodologies	• Integration of innovation	• Collaboration with partners
• Application of methodologies	• Scaling of innovation	• Collaboration with customers

Note. Table adapted from Sindemann and von Buttlar (2017, 2018).

search fields. A commonly observable approach is the aggregation of CSFs into specific categories and dimensions.

Unlike the development of DIU types (section 2.2.2 & 2.2.3), which was already widely addressed by academic research, CSFs of DIUs have been covered only partially so far. Holotiuk (2020) for example, derived "important aspects of the organizational design" (Holotiuk, 2020, p. 13), that allow ambidexterity to develop digital innovation at DIUs. Fuchs et al. (2019) derived a taxonomy for DIUs (section 2.2.2) and defined organizational patterns for DIUs. For example, most DIUs have access to "secured financial funding for their projects" (Fuchs et al., 2019, p. 643) and interdisciplinary teams, consisting of core organization and DIU employees (Fuchs et al., 2019). Similar, Barthel et al., 2020 identified suitable organizational design patterns that "appear to fit the different innovation objectives" (Barthel et al., 2020, p. 12). Furthermore, design features like embedding DIUs into the corporate digital strategy and support of the top management were explicitly described as "prerequisite for any DIU type" (Barthel et al., 2020, p. 13). Moreover, soft aspects like the culture within the core organization, were defined as success factors of DIUs. The focus on digital innovation that addresses real problems of the core organization's business areas is another important factor (Raabe et al., 2021). Similar, Holotiuk (2020) derived several recommendations that support DIUs in accelerating the digital transformation of the core organization, focusing on five dimensions (e.g. setup, strategy). An overview of selected dimensions and exemplary recommendations is depicted in Table 3.

While existing contributions rather focus on general recommendations, CSFs for DIUs to fulfill their objectives have not been derived so far.

In contrast, CSFs have already been widely explored within other areas of research including project management (Belassi & Tukel, 1996; Clarke, 1999), business performance management (Ariyachandra & Frolick, 2008) and agile software development (Abd El Hameed, Latif, & Kholief, 2016; Chow & Cao, 2008). Abd El Hameed et al. (2016) for example, identified CSFs for agile software development and clustered these into the dimensions people, project, product, technical, organization and process. A similar approach of clustering CSFs was applied by Chow and Cao (2008). So far, there are only few scientific contributions in fields related to DIUs, as for example digital transformation (Astri, 2015; Sahu, Deng, & Mollah, 2018; Wu, Huang, & Chen, 2011; Zhou, 2011).

Furthermore, Holotiuk and Beimborn (2017) derived a CSFs framework related to digital business strategy, consisting of 40 factors and clustered into eight dimensions. An overview of the framework is provided in Table 4.

Each dimension contains several CSFs. Sales & customer experience as well as organization were identified as most critical. The other dimensions, are depicted based on the specific importance in a descending order, starting with culture & leadership. Exemplary CSFs within this category are: foster & develop digital mindset, define set of values with focus on digital and allow failure as well as the commitment to transform (Holotiuk & Beimborn, 2017).

By comparing all contributions, including the ones by practitioners (section 3.1.1), a common approach is observable. Identified CSFs are usually aggregated into clusters derived from the underlying concept (e.g. agile software development, digital business strategy).

3.2. Performance measurement systems

Developing a thorough understanding of PEMS, not only from a general point of view, but also in regard to requirements specifically for DIUs, is important for the purpose of this study.

3.2.1. General perspective on performance measurement systems

While PEMS went through several transformations, some PEMS (e.g. balanced scorecard approach) have been widely adopted by corporations. Deriving measures for the (strategic) objectives of an organizations is an approach that can be found in various PEMS. While several approaches for the measurement of innovation exist, literature still lacks in regard to PEMS for DIUs.

PEMS, sometimes also referred to as "management control systems" (Haskamp et al., 2021, p. 3), were widely addressed by academic literate throughout the last decades (Bourne et al., 2003; Neely et al., 2007). PEMS usually include of a set of metrics (Bourne et al., 2003). Although the reasons for measuring performance are manifold (Behn, 2003), PEMS are usually set up by organizations to control their activities by making use of relevant data (Kerssens-van Drongelen & Cooke, 1997). Approaches for performance measurement and corresponding frameworks went through

Table 3: Exemplary recommendations for DIUs

Setup: • Reduce old ties to minimize risk of legacy • Set up small units to reduce financial investments • Ensure flexibility	Strategy: • Define strategic focus to provide guidance • Derive clear objectives to align actions • Look at innovations from start to end
Outcome: • Value all forms of outcomes of the DIU • Make intangible or qualitative outcomes visible and market them	Process: • Cover entire innovation process • Incorporate ideas from all employees • Integrate innovations into main organization

Note. Table adapted from Holotiuk (2020).

Table 4: Categories of CSFs for digital business strategy

	Culture & leadership	
Sales & customer experience	Capabilities & HR competencies Foresight & vision Data & IT Operations Partners	Organization

Note. Table taken from Holotiuk and Beimborn (2017, p. 996).

several transformations over time. Whereas traditional approaches have put a high focus on financial measures and metrics, such as return on investment (ROI), new approaches and frameworks that promote a more balanced PEMS, have emerged (Kaplan & Norton, 1992). Hence, these frameworks address the shortcomings of traditional systems (Bourne et al., 2003). Exemplary PEMS are the performance measurement matrix (Keegan, Eiler, & Jones, 1989), the SMART (strategic measurement and reporting technique) pyramid (Neely et al., 2007), the result – determinants framework (Fitzgerald, Johnston, Brignall, Silvestro, & Voss, 1991) and the balanced scorecard (Kaplan & Norton, 1992; Neely et al., 2007).

The balanced scorecard approach in particular has been widely adopted by organizations (Neely et al., 2007) and was refined in various forms (Gama, Da Silva, & Ataíde, 2007; Kaplan & Norton, 1996; Lawrie & Cobbold, 2004). One of its key elements is the deduction of performance measures (financial and non-financial) based on the strategy and vision of the organization. These performance measures are linked within the framework, including outcome measures as well as related performance drivers. The balanced scorecard consists of four dimensions. First, the financial perspective, which covers the long-term objectives of the organization. Second, the customer perspective, which aims at identifying relevant customer segments and markets for the organization. Third, the internal business process perspective, which includes crucial processes for the organization, such as fulfilling the value proposition for customers. The last perspective, learning & growth, targets at identifying important aspects to ensure long-term growth and future success (Kaplan & Norton, 1992; Kaplan & Norton, 1996). Based on the particular strategic objectives within each perspective, specific measurements are selected and linked to objectives (Kaplan & Norton, 1992; Kaplan & Norton, 1996; Lawrie & Cobbold, 2004).

While existing research highlights the importance and benefits of performance measurement for innovation activities (Davila, Foster, & Oyon, 2009; Ylinen & Gullkvist, 2014), and established approaches like the balanced scorecard have already been adapted to measure the value added by innovation (Gama et al., 2007), there is a gap in regard to measuring the success of digital innovation activities of DIUs.

However, identifying KPIs that are or should be used in DIUs and linking these to the DIU's objectives (section 2.3) seems to be a suitable approach to tap into this field. A similar and widely discussed approach was introduced by Bourne, Mills, Wilcox, Neely, and Platts (2000), who defined four phases for the development of PEMS, starting with the identification of key objectives and related measures (Bourne et al., 2000; Garengo, Biazzo, & Bititci, 2005; Henri, 2006; Neely, 2005).

3.2.2. Requirements for performance measurement systems at DIUs

There are several important requirements for DIUs in regard to PEMS. Due to the novelty of DIUs, these requirements mainly originate from related fields, such as PEMS for innovation.

So far, existing literature exploring success and performance measurement of DIUs is rare, which can be attributed to the novelty of the research field (Barthel et al., 2020). Recently, Haskamp et al. (2021) identified requirements for PEMS particularly important for DIUs. Based on existing literature, different categories of PEMS requirements were identified. These categories cover general requirements, re-

quirements in regard to innovation as well as requirements for digital innovation and agile ways of working. Furthermore, requirements particularly for PEMS in DIUs were derived (Haskamp et al., 2021).

The specific categories are illustrated in the following, starting with general requirements for PEMS (Table 5). General requirements for PEMS (RE 1-4) show a clear focus on data, including its collection, availability and usability. The required data and information for PEMS should be collected in a simple way (RE 1), based on existing organizational activities. In addition, data should be available for specific user groups to utilize and access it in an easy way to make decisions (RE 2 & 3). Furthermore, performance criteria shall be derived from the objectives of the organization and aligned with the corporate strategy (Haskamp et al., 2021).

As the development of digital innovation is the main objective for many DIUs (section 2.3), PEMS requirements in regard to innovation are depicted in Table 6. The requirements related to innovation particularly address relevant aspects that come along with innovation activities (section 2.1). This includes the type of innovation (RE 5, e.g. product) and different stages of the innovation process (RE 6 & 7, e.g. discovery). Furthermore, the requirements emphasize the flexibility of the PEMS, for example in regard to measurement approaches (e.g. input-oriented) and different performance dimensions (e.g. financial) as well as the ability to process different types of data (Haskamp et al., 2021).

An overview of requirements for PEMS systems that are important particularly for digital innovation and agile ways of working is provided in Table 7. PEMS should consider the requirements that result from the innovation process applied for digital innovation (section 2.1). As customer-centricity and market proximity are important aspects of digital innovation, novel metrics are required to address these aspects (RE 8). Furthermore, PEMS should be able to handle the requirements that result from the applied methodologies, such as design thinking (RE 9). Additionally, a certain degree of flexibility is important (Haskamp et al., 2021).

Moreover, Haskamp et al. (2021) identified three additional requirements for PEMS, that are particularly relevant for DIUs (Table 8). RE 10 highlights the importance of incentives in regard to experimenting and the resulting learnings. Here, potential failures that come along with experimenting need to be considered accordingly, meaning, failures should also become visible in selected metrics. However, in a positive way in order to support experimenting. Dependent on the innovation orientation and market focus of the innovation activities pursued within the DIU, the PEMS should be adjusted accordingly and enable data exchange (RE 11). As DIUs usually pursue innovation activities in the early stages of the innovation process (e.g. discovery), standard metrics are hard to apply. Moreover, DIUs are facing the challenge to collect data that is relevant for performance measurement on their own. Therefore, it is crucial to ensure credibility of the data (RE 12) within the PEMS (Haskamp et al., 2021).

After having defined requirements for PEMS at DIUs, it is crucial to analyze ways of measuring digital innovation on a more granular level. More precisely, we seek to explore ways of using KPIs for the measurement of digital innovation.

3.3. Measuring digital innovation

To develop a comprehensive understanding of ways to measure digital innovation, it is essential to take a general perspective on KPIs first, followed by the introduction of challenges, frameworks and exemplary KPIs that can be used to measure digital innovation.

3.3.1. General perspective on key performance indicators

Since KPIs are multifaceted and complex, we exclusively address the key aspects of KPIs from a general perspective. Considering digital innovation activities, deriving KPIs from specific organizational objectives is crucial for the underlying performance measurement.

Similar to PEMS, KPIs have been widely addressed by academic research (Kaplan & Norton, 1996; Kristiansen & Ritala, 2018; Lawrie & Cobbold, 2004). Beyond KPIs, several types of performance measures, such as key result indicators, exist (Parmenter, 2015), stressing the importance to define the term KPI. Among various definitions, one common feature that stands out is the direct relation between organizational objectives and KPIs (Bourne et al., 2003; Macmillan, 2021; Oxford, 2020; Twin & James, 2020). These objectives are oftentimes even described as strategic goals of the organization (Parmenter, 2015). Within this study, we define KPIs as “quantifiable measure used to evaluate the success of an organization, employee, etc. in meeting objectives for performance” (Oxford, 2020, p. 1). As a consequence, KPIs that are or should be used by DIUs to measure their performance should be directly related to the objectives specifically addressed by DIUs, which were introduced in section 2.3. Existing literature recommends limiting the number of KPIs that are used within the PEMS to 15-25. Additionally, the selected KPIs should be linked to specific target values, that are set by the organizations (Kaplan & Norton, 1992; Kaplan & Norton, 1996). Moreover, KPIs can be characterized from several point of views, for example, the separation into financial (e.g. ROI) and non-financial KPIs (e.g. customer satisfaction). Beyond that, KPIs can be of quantitative and qualitative nature and are usually assigned to specific perspectives and areas, as realized in the balanced scorecard approach (Kaplan & Norton, 1992; Kaplan & Norton, 1996; Lawrie & Cobbold, 2004).

Since discussing the general perspective of KPIs in more detail would exceed the scope of this research project, it is crucial to understand the role of KPIs in measuring digital innovation, including associated challenges. Therefore, measurement frameworks for structuring such innovation KPIs as well as suitable, exemplary KPIs need to be discussed. Thereby, these KPIs have the potential to be directly related to main objective of DIUs, digital innovation.

3.3.2. Using key performance indicators to measure digital innovation

Measuring digital innovation appropriately is perceived as challenging endeavor accompanied with several difficul-

Table 5: General PEMS requirements

Number	Description of requirement	Underlying literature
RE 1	The PEMS should allow to collect data in an easy way based on the organizational activities.	(Kerssens-van Drongelen & Cooke, 1997; Micheli & Manzoni, 2010)
RE 2	The PEMS should ensure the availability of data that can be utilized and contextualized. Based on this data, the user of the PEMS should be able to generate insights, derive implications and make decision.	(Globerson, 1985; Lill, Wald, & Munck, 2020; Neely et al., 2000; Wiesche, Bodner, & Schermann, 2012)
RE 3	The PEMS should have a high degree of usability and particularly address the interests of different user groups, such as stakeholders from the core organization but also from the DIU itself.	(Hamilton & Chervany, 1981; Henttonen, Ojanen, & Puumalainen, 2015; Kerssens-van Drongelen & Cooke, 1997)
RE 4	The PEMS should ensure the alignment of the specific performance criteria with the corporate strategy and derive these criteria from the objectives.	(Globerson, 1985; Neely et al., 2000)

Note. Table adapted from Haskamp et al. (2021).

Table 6: PEMS requirements related to innovation

Number	Description of requirement	Underlying literature
RE 5	The PEMS should consider different types of innovation individually (e.g. process, product & business model) and align the required mode of performance measurement and control.	(Barros & Ferreira, 2019; Chiesa, Frattini, Lamberti, & Noci, 2009; Curtis & Sweeney, 2016; Davila et al., 2009; Fichman et al., 2014; Schermann, Wiesche, & Krcmar, 2012; Ylinen & Gullkvist, 2014)
RE 6	The PEMS should include different measurement approaches (e.g. input- & output-oriented). Furthermore, the PEMS should consider different performance dimensions (e.g. financial, learning & knowledge), as well as different stages of the innovation process.	(Chiesa et al., 2009; Henttonen et al., 2015; Micheli & Manzoni, 2010)
RE 7	The PEMS should be able to process quantitative data (e.g. number of interviews) and qualitative data (e.g. customer satisfaction) throughout the applied innovation process.	(R. Adams, Bessant, & Phelps, 2006; Barros & Ferreira, 2019; Said, HassabElnaby, & Wier, 2003)

Note. Table adapted from Haskamp et al. (2021).

Table 7: PEMS requirements related to digital innovation and agile ways of working

Number	Description of requirement	Underlying literature
RE 8	The PEMS should have a close alignment with the innovation process for digital innovation. One reason for this is the new role of the IT department, which has become an important part of the business strategy. Therefore, customer-centricity and market proximity need to be more focused, which requires novel metrics.	(Brynjolfsson & Oh, 2012; Huang, Henfridsson, Liu, & Newell, 2017; Hund et al., 2019; Urbach et al., 2018)
RE 9	Similar to the methodologies applied to create digital innovation, the PEMS should set up accordingly to cope with agile ways of working (e.g. scrum) and exploratory methods (e.g. design thinking), as these approaches require shorter cycle times and a high degree of flexibility in regard to adjustments.	(Basili et al., 2007; Boerman, Lubsen, Tamburri, & Visser, 2015; Lee, Hsu, & Silva, 2020; Mayer et al., 2021)

Note. Table adapted from Haskamp et al. (2021).

ties, that also need to be faced by DIUs. Several approaches that strive to measure innovation performance use KPIs as fundamental component.

Existing literature already addressed difficulties of mea-

Table 8: PEMS requirements related to DIUs

Number	Description of requirement	Underlying literature
RE 10	The PEMS should provide incentives for employees that support experimenting and show resulting learnings.	(Haskamp et al., 2021)
RE 11	The PEMS should enable the exchange of data and information between the DIU and the core organization to support autonomous operations of the DIU.	(Haskamp et al., 2021)
RE 12	The PEMS data should be trustworthy and meaningful.	(Haskamp et al., 2021)

Note. Table adapted from Haskamp et al. (2021).

suring digital innovation, also within the scope of DIUs, and highlights the need for further research in this field (Barthel et al., 2020; Hund et al., 2019; Hund et al., 2019). First of all, difficulties might arise due to differences between traditional and digital innovation (Kohli & Melville, 2019). Beyond that, insufficiently defined objectives of DIUs (Raabe et al., 2020), which might also be caused by the novelty of the research field (Barthel et al., 2020), as well as addressing only selected stages of the innovation process (section 2.1 and 2.2.), might further increase the complexity of measuring digital innovation at DIUs. In addition, measuring specific aspects of digital innovation, for example customer centricity, through metrics is perceived as challenge (Frey, Holotiuk, & Beimborn, 2020; Micheli, Wilner, Bhatti, Mura, & Beverland, 2018). Therefore, qualitative KPIs seem to be more suitable for measuring explorative innovation activities (Bedford, 2015; Kerssens-van Drongelen & Cooke, 1997).

Recently, Mayer et al. (2021) identified several challenges of measuring digital innovation across various DIU areas (e.g. DIU objectives), particularly in the field of design thinking activities: Even though the metrics are derived from exploitive innovation activities, they are frequently used for explorative DIU activities. Additionally, there is a lack of metrics for measuring internal transformation. The expectations of the core organizations towards a diversified innovation portfolio challenges the DIU when it comes to allocating resources to particular projects. Oftentimes, the objectives of the innovation activities are not clearly defined and documented. Therefore, the performance of the DIU might be measured based on conflicting objectives. Beyond that, the time horizon of innovation activities impacts their measurement. Whereas the applied metrics are oftentimes short-term oriented, creating digital innovation and new ventures can be a lengthy process. Thus, applying financial metrics, especially for explorative activities, might be quite challenging. Hence, DIUs prefer qualitative KPIs to report progresses regarding digital innovation initiatives. However, these KPIs oftentimes conflict with the KPIs the core organization usually applies to measure success. On closer examination, one realizes that digital innovation activities might be highly specific, dependent on the project. Since cross-project applicability of KPIs might be challenging, maintaining a flexible way of selecting metrics is crucial. Concluding, softer objectives of DIUs, that go beyond digital innovation (e.g. cultural change and digital expertise) and therefore require more time, might be hard to track and difficult to relate to the DIU (Mayer et al., 2021).

As there is still a lack of existing academic research in regard to performance measurement of digital innovation (Barthel et al., 2020), we also take a closer look at traditional innovation performance measurement, which is, in contrast, widely discussed in literature (Dewangan & Godse, 2014). Several approaches for clustering innovation performance metrics exist (Collins & Smith, 1999; Cruz-Cázares, Bayona-Sáez, & García-Marco, 2013; Gama et al., 2007; Lazzarotti, Manzini, & Mari, 2011; Suomala, 2004). According to Dewangan and Godse (2014), established innovation PEMS lack in adequately addressing early stages of innovation (e.g. discovery). Each stage of the innovation process includes specific activities and outcomes. Instead of using established systems like the balanced scorecard entirely as PEMS, Dewangan and Godse (2014) recommend applying multi-dimensional metrics within each stage of the innovation process, considering selected guiding principles, such as the causal relationship between metrics. While a detailed explanation would exceed the scope of this work an overview of the framework is provided in Table 9, including exemplary KPIs (Dewangan & Godse, 2014).

Assigning selected KPIs to specific stages within the innovation process, as realized in the innovation performance measurement framework by Dewangan and Godse (2014), seems to be an adoptable approach for the innovation process for digital innovation (section 2.1). Complementary, Table 10 provides selected, commonly used, output- and input-related metrics, which were derived through an extensive survey covering top management and senior executives across various regions and industries (Chan, Musso, & Shankar, 2008).

To summarize, we developed a comprehensive understanding of ways to measure digital innovation including related challenges, frameworks and exemplary KPIs that can be used to measure digital innovation.

3.4. Research gap

Although DIUs are perceived as important vehicle for corporate digital innovation and widely established by practitioners, DIUs increasingly face criticism in regard to their success and impact on the core organization (Chapter 1). So

Table 9: Innovation performance measurement framework

	Innovation process stage	
Dimension	Generation & selection of ideas (1)	Incubation of ideas (2)
Non-financial		
Customer	Percentage of ideas generated with customer participation	Number of ideas incubated in collaboration with customers
Internal processes	Ratio of selected ideas to ideas submitted	Percentage of incubated ideas found viable for commercialization
Innovation & learning	Percentage of ideas generated in new domains	Number of patents filed
Financial	Average expenditure per selected idea	Current idea portfolio net present value
Dimension	Commercialization of ideas (3)	Realization of innovation (4)
Non-financial		
Customer	Rate of customer adoption of new offerings	Percentage impact on customer satisfaction index
Internal processes	Rate at which offerings are being launched	Percentage commercial success rate (i.e. innovations that met projections)
Innovation & learning	Number of marketing partners added	Percentage increase in innovation revenues per employee
Financial	Commercialization expenditure for the innovation portfolio	Innovation portfolio ROI realized

Note. Table taken from Dewangan and Godse (2014, p. 542).

Table 10: Selected innovation metrics applied by companies

Output Metrics	Input Metrics
Revenue growth (new products or services)	Number of ideas in the pipeline
Customer satisfaction (new products or services)	R&D spending (as percentage of sales)
Percentage of sales (new products or services)	Number of R&D projects
Number of new products or services launched	Number of employees actively devoted to innovation

Note. Table adapted from Chan et al. (2008).

far, existing literature has addressed various aspects of DIUs, such as underlying concepts (section 2.1), the typology of DIUs (section 2.2) and their objectives (section 2.3). With regard to performance measurement, research has already analyzed CSFs in related fields (section 3.1) and started to lay the foundation for measuring the performance of DIUs (section 3.2 & 3.3). However, existing literature is still limited when it comes to CSFs for DIUs and objective related KPIs to measure the performance of DIUs. To close this gap, we pose the following research questions:

> RQ1: What are CSFs for DIUs to fulfill their objectives?
> RQ2: Which KPIs are or should be used by DIUs to measure their performance?

To answer these research questions and to identify important aspects of performance measurement, such as CSFs and KPIs, relevant to the success of DIUs, we applied a case study approach that utilizes a mixed-method research design.

4. Methodology

In order to develop a comprehensive understanding of the applied methodology, we initially introduce the methodological basis of this study, the case study research approach. As we decided to apply a mixed-method design, the core component (semi-structured interviews) as well as the supplemental component (survey) are explained.

4.1. Case study research approach

Within this study, the case study research approach was applied. In total, a diverse sample that consisted of nine cross-industry cases was investigated through applying a mixed-method research design.

Although, DIUs have already been widely established by organizations, they still represent a novel research field from a scientific point of view (Barthel et al., 2020; Haskamp et al., 2021). An explorative, qualitative-empirical research approach is particularly suitable for such novel research areas (Benbasat, Goldstein, & Mead, 1987; Edmondson & McManus, 2007), as it allows to study recent phenomena in real-world settings (Yin, 2009). To be precise, we applied the

case study research approach according to Eisenhardt (1989) that allows building theory from cases, which is often "novel, testable, and empirically valid" (Eisenhardt, 1989, p. 532).

In total, we investigated nine cases, corresponding to nine DIUs. Hence, we followed a multiple-case study design that allows cross-case analysis to improve the robustness of the study (Benbasat et al., 1987; Yin, 2009). One of the most important steps in case study research is the selection of sample cases. In contrast to research that targets hypothesis-testing, Eisenhardt (1989) recommends theoretical sampling. Therefore, the sample should not be selected randomly, but instead, consist of cases that "are likely to replicate or extend the emergent theory" (Eisenhardt, 1989, p. 537). For the purpose of improving the external validity of the study, we aimed at selecting a diverse set of cases (Benbasat et al., 1987; Yin, 2009). Furthermore, we particularly targeted at a sample that includes multiple DIU types. Hence, we utilized a classification of DIUs provided by Sindemann and von Buttlar (2020) as a starting point for the case selection. DIUs were categorized based on their market focus of digital innovation activities into DIUs that mainly target innovation close to the core business and innovation distant to the core business. This feature is partly aligned with one of the key distinguishing characteristics of the DIU types derived by Barthel et al., 2020, the dimension market focus of innovation within the category objective and scope (section 2.2.3). Moreover, DIUs were classified dependent on the business field of the core organization (Sindemann & von Buttlar, 2020). Concluding, we aimed at selecting cases that differ in regard to industry, size and purpose. An overview of the sample is provided in Table 11.

Case study research design allows to utilize qualitative and quantitative data as "case studies typically combine data collection methods such as archives, interviews, questionnaires, and observations" (Eisenhardt, 1989, p. 534). Due to the growing complexity of recent research endeavors, literature is increasingly derived through the application of mixed-method designs, that allow the combination of qualitative and quantitative data collection and analysis approaches and enable to improve the research scope and quality (Sandelowski, 2000). As identifying important aspects of performance measurement, such as CSFs and KPIs, relevant to the success of DIUs can be classified as complex research endeavor, also caused by the novelty of the research field, we designed a mixed-method research approach.

4.2. Mixed-method research design

The mixed-method research design applied in this study consisted of semi-structured interviews (core component) and a survey (supplemental component). Whereas the data gathered through the interviews was used to answer both research questions, the survey complemented the results.

Combining several methods within one research project can take on different forms (Martha, Sousa, & Mendes, 2007). On the one hand, multiple-method research designs consist of two complete methods, that could both be published independently (Morse, 2003). Mixed-methods research design on the other hand consists of a complete core component and at least one supplementary component that is incomplete and cannot be published independently (Morse & Niehaus, 2009). The core component within the research project aims at answering the key aspects of the research questions and should fulfill the essential requirements to be published independently of the supplementary component. As the dominant and complete method it "must be conducted at a standard of rigor" (Morse & Niehaus, 2016, p. 23). Hence, it has to fulfill the standards of scientific research with regards to the underlying methodology. As the supplemental component consist of at least one incomplete method, it would be better considered as strategy instead of method (Morse & Niehaus, 2016). It "provides explanation or insight within the context of the core component"(Morse, 2010, p. 484) and can be incomplete due to multifaceted reasons, for example due to an insufficient sample size or lack of saturation (Morse, 2010).

We applied a mixed-method design that consisted of a qualitative, inductive core component and a quantitative supplementary component. Whereas both research questions were answered based on the data gathered through the core component, we aimed at making the research project more comprehensive through providing additional, case related information based on survey data (Morse, 2010). Specifically, we used the supplemental survey to identify the respective DIU type of each case. Such research designs are commonly used, but also come along with challenges, for example in regard to aligning data (Harris & Brown, 2010; Schoonenboom & Johnson, 2017). The detailed illustration of the core and supplementary component, including information on data collection and analysis, is provided in section 4.2.1 and 4.2.2.

Another important element of mixed-method research design is the point of interface, which is the point of integrating the supplemental research into the core component (Morse & Niehaus, 2016). Within this research project, the results point of integration was applied, meaning, that the results of the supplemental component were integrated into the results of the core component (Morse & Niehaus, 2016). Furthermore, simultaneous pacing was deployed. Hence, the data collection of both components was conducted at the same time (simultaneously) and independent from each other. From a methodological point of view, the correct notation of this research design is "QUAL + quan" (Morse & Niehaus, 2016, p. 25). Whereas the core component is written in capital letters, the supplemental method is displayed through the use of small letters. The "+" indicates the simultaneous conduction of both components (Morse & Niehaus, 2016). The sample of the core component does not necessarily need to be the different from the sample used within supplemental component (Morse, 2010). Within our research project, the same sample was utilized for both components of the mixed-method approach. Schoonenboom and Johnson (2017) identified further important dimensions that need to be addressed when designing mixed-method

Table 11: Overview of the sample

Case	Industry	Total employees / DIU employees	Revenue*	Interviews**	Role of interview partners
A	Machinery	~12,300 / 26	~2.65 bn. €	1	Managing Director (1)
B	Mechanical Engineering	~10,000 / 150	~2.00 bn. €	1	Head of Digital Innovation (1)
C	Healthcare	~65,600 / 17	~7.4 bn. €	2	Venture Consultant (1) Director Digital Operations (2)
D	Building Materials	~72,500 / 12	~24.05 bn. €	1***	Head of Digital Partnerships (1) Digital Lead Scale-Up and Acceleration (2)
E	Building Materials	~35,000 / 22	~10.00 bn. €	2	Director Digital Hub (1) Director Digital Transformation (2)
F	Multi-Industry	~12,300 / 10	~2.65 bn. €	1	Managing Director (1)
G	Home Appliances	~58,200 / 10	~13.20 bn. €	2	Venture Partner (1) Chief Digital Officer (2)
H	Pharma	~5,300 / 190	~1.53 bn. €	1	Managing Director (1)
I	Consumer Goods	~3,000 / 48	~0.72 bn. €	1	Head of Business Development (1)

Note. *Revenue end of 2019, **Number of interviews, ***One interview with two interview partners conducted.

research. The purpose for applying this methodology was to complement and expand the results of the semi-structured interviews with the results from the survey (Greene, Caracelli, & Graham, 1989; Schoonenboom & Johnson, 2017). The primary theoretical drive within this study was inductive, as semi-structured interviews were used as core component (Schoonenboom & Johnson, 2017).

After having defined the underlying research design, it is important to clearly illustrate the applied methods, including data collection and analysis.

4.2.1. Core component: Semi-structured interviews

Semi-structured interviews were conducted as core component for answering both research questions. The interviews were recorded, transcribed and consequently analyzed based on the content structuring qualitative analysis approach from Kuckartz (2016), utilizing inductive-deductive category formation.

As structured interviews are usually organized around predefined, standardized sets of questions, oftentimes closed-ended, there is little flexibility during the interview (Lune & Berg, 2017). Unstructured interviews on the other hand, are oftentimes conducted without pre-defined interview questions and offer high degrees of flexibility (Gubrium & Holstein, 2002). Semi-structured interviews combine aspects of both structured and unstructured interviews and are usually based on open-ended questions, that enable flexibility and provide guidance during the interviews (W. C. Adams, 2015; Segal, Coolidge, O'Riley, & Heinz, 2006). Within this research project, a semi-structured interview approach was applied.

An important aspect of semi-structured interviews is the development of an interview guideline, that supports the interviewer in organizing and structuring the interview (Creswell & Creswell, 2018). It is crucial to derive the guideline in accordance to the research questions (Rowley, 2012). The interview guideline within this study (Appendix 4) was split up into three parts and consisted of 11 open-ended questions. First, a short introduction including the research purpose, a short notice in regard to the complementary component as well as information concerning the confidentiality of the interviews. Second, two introductory questions that aimed at collecting general, case-related information and creating a pleasant interview atmosphere. Subsequently, the main part, focusing on performance measurement, consisted of questions in regard to CSFs and KPIs (nine questions in total). Due to the complexity of the research purpose, the interview guideline was distributed one day ahead of the actual interview. The interview partners were selected based on their current role within the respective DIU. We aimed at conducting interviews with upper management DIU employees that have a profound understanding of the DIU as a whole and are involved in performance measurement. Dependent on the preferences of the interview partner, the interviews were conducted in English or German. The interviews took place virtually via the online meeting tool "Zoom" and lasted between 25 and 50 minutes. The interview data collection took place from 11/2020 – 01/2021. In total, 12 interviews have been realized. The interviews were recorded, transcribed verbatim and anonymized (Flick, 2013). The primary data gathered through the interviews was triangulated through secondary data retrieved from company websites, management reports and press releases (Yin, 2014). The software program "MaxQDA", which is particularly designed for qualitative data analysis, was utilized to store, analyze and visualize the data.

Various approaches for qualitative data analysis exist (Kuckartz, 2016; Mayring, 2015). An important aspect of

qualitative data analysis is the formation of categories, which can be derived deductive and inductive. While deductive category formation defines the structure of categories based on existing literature before analyzing the data, inductive category formation allows to derive categories through analyzing and abstracting the data iteratively. Therefore, categories are retrieved from the data itself (Mayring, 2015). Within this study, we followed the qualitative content analysis according to Kuckartz (2016). To be precise, the content structuring qualitative analysis that consists of seven phases. We structured the data and applied an inductive-deductive category formation (Kuckartz, 2016; Strike & Rerup, 2016).

Initially, we inductively derived general case insights (e.g. year of foundation, number of employees). A detailed overview including the category, codes, code descriptions and exemplary quotes is depicted in Appendix 5. Furthermore, we enriched this data with the finding generated from the supplemental method (section 4.2.2), following the results point of interface, as described previously. As we aimed at identifying CSFs for DIUs to fulfill their objectives, we analyzed the aggregated data in regard to the objectives that are pursued by each case based on the objectives that were defined within section 2.3. Thus, we applied a deductive approach, which is illustrated in Appendix 6.

To answer RQ1 ("What are CSFs for DIUs to fulfill their objectives?"), the underlying data was coded iteratively while applying an inductive approach. After identifying first order codes, explicit CSFs were derived (depicted as "sub codes" in Appendix 7). Second, we defined categories of CSFs (depicted as "codes" in Appendix 7). Both categories of CSFs and explicit CSFs were derived on a generic level and not directly related to specific DIU objectives. An exemplary overview of the data structure is depicted in Figure 4, including first order codes, second order concepts (explicit CSF) and aggregated themes (categories of CSFs).

The underlying CSF method primarily focuses on identifying information and actions that are crucial to achieve the expected objectives and has been widely applied by practitioners (Bullen & Rockart, 1981; Rockart, 1979). An important benefit of the CSF method is, that implicit CSFs can be made explicit and consequently be addressed appropriately. The CSF method is based on interviews, which target at identifying explicit CSFs (Bullen & Rockart, 1981). Furthermore, the CSF method can be adjusted to the specific research purpose (Bullen & Rockart, 1981; Holotiuk & Beimborn, 2017; Rockart, 1979). Thus, the conducted semi-structured interviews were a suitable approach to identify CSFs for DIUs to fulfill their objectives. In order to derive a more robust set of CSFs, we exclusively considered CSFs that were identified among at least 15% of the cases (Sproull, 2002), meaning at least in two of the sample cases.

To answer RQ2 ("Which KPIs are or should be used by DIUs to measure their performance?"), we applied an inductive approach as well. After identifying first order codes, contextual information for the application of KPIs at DIUs was derived (e.g. challenges and benefits of establishing KPIs at DIUs). A detailed overview that includes information regarding the coding approach is provided in Appendix 8.

Subsequently, specific KPIs were identified (exemplary depicted as "sub codes" in Appendix 9) and aggregated into clusters (exemplary depicted as "codes" in Appendix 9). Both clusters and individual KPIs were directly related to specific DIU objectives as suggested by literature (section 3.3). Similar to the overview of the data structure provided in Figure 4, an exemplary overview of the data structure relevant to KPIs is depicted in Figure 5.

The underlying method for deriving specific KPIs follows the PEMS development approach by Bourne et al. (2000) which is depicted in Appendix 10. We exclusively focused on the first phase, system design, that consists of identifying key objectives and designing measures. Consecutive phases, for example the implementation of measures, would go beyond the scope of this research project.

After having defined the core component of the mixed-method research design comprehensively, it is crucial to introduce the supplementary survey.

4.2.2. Supplemental component: Survey

The core component was complemented with a supplementary survey to identify the DIU type of each case. The supplemental component followed less rigorous scientific guidelines.

Within this study, a survey was conducted as supplementary, quantitative component to identify the DIU type of each case. While Barthel et al., 2020 derived ideal DIU types based on the objective and scope as well as the setting and design of DIUs, we exclusively focus on identifying the DIU type dependent on the specific objective and scope. Information regarding setting and design was primarily used to complement the findings. Hence, we primarily used closed-ended questions that were closely aligned with the typology of DIUs (section 2.2.2 & 2.2.3). The questionnaire consisted of ten questions and was divided into four parts: After providing a short introduction of the purpose of the survey (part one), two general open-ended questions with regards to the background of the participant and the DIU were asked (part two). Consequently, part three consisted of three closed-ended questions that aimed at gathering data in regard to the objective and scope of the DIU. Concluding, part four focused on identifying the specific setting and design of each case based on five closed-ended questions including the response option "other", in case the DIU design differs from the dimensions provided by Barthel et al., 2020. An overview of the questionnaire design is provided in Appendix 11.

The sample of the supplemental component does not necessarily need to be the different from the one used within the core component (Morse, 2010). Therefore, we utilized the same sample for both components. Following the concept of simultaneous mixed-method research design, the survey was distributed individually to the interview partners directly after the interviews. We used a cross-sectional survey design. To conduct the online survey and collect the data, we used the online tool provided by "SurveyMonkey". Survey methodology usually follows rigorous scientific guidelines, for example

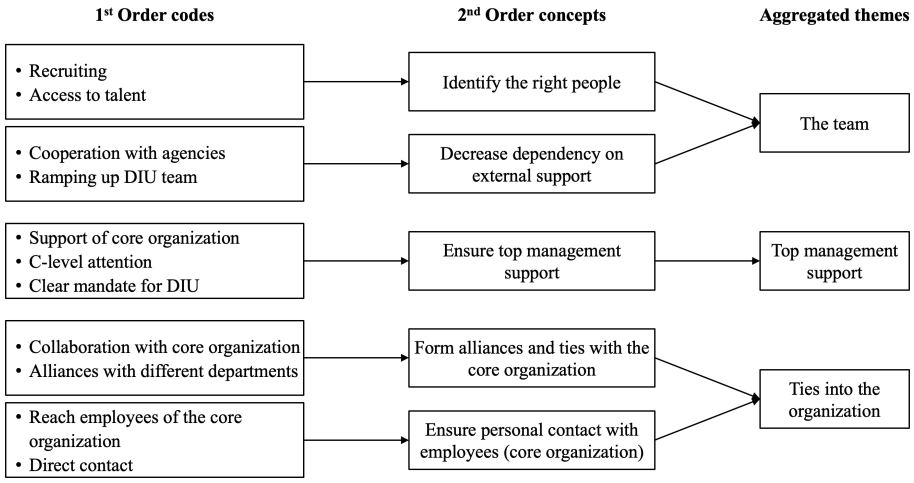


Figure 4: Exemplary data structure CSFs

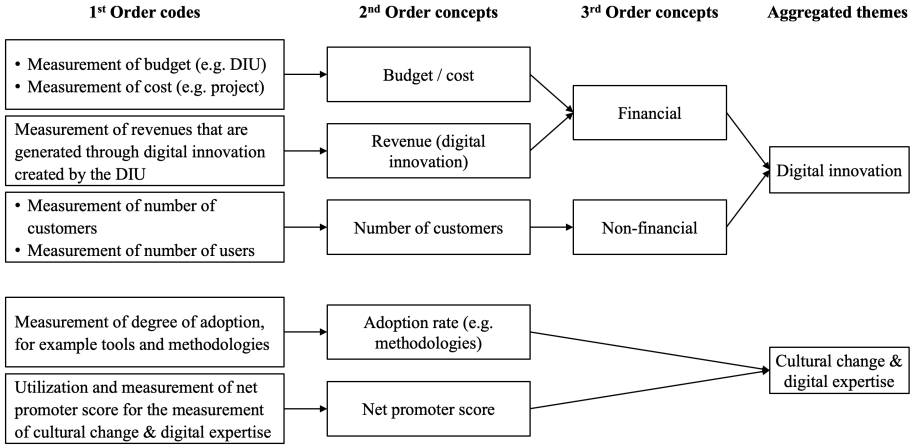


Figure 5: Exemplary data structure KPIs

regard sampling, data collection and analysis (Groves et al., 2009; Laaksonen, 2018; Passmore et al., 2002). However, supplemental components in mixed-method research designs can be incomplete due to multifaceted reasons (Morse & Niehaus, 2016). Within this study, we utilized those less rigorous requirements in regard to sample selection and sample size, as we exclusively focused on the sample of the core component. We analyzed the data individually for each case, instead of applying quantitative descriptive statistics. Thus, the requirements from a scientific research point of view were fulfilled.

To summarize, we developed a comprehensive understanding of the applied methodology. Thus, we set the foundation for presenting the results of this research project.

5. Results

A comprehensive understanding of the results derived within this research project is crucial for its contribution to the existing research gap and its implications for practice. Initially, we merge the findings of the core and supplementary component by introducing the identified DIU types within the sample together with general, case specific information. After defining the objectives that are pursued by the sample cases, we outline the identified CSFs for DIUs to fulfill their objectives. Afterwards, we provide contextual information for applying KPIs at DIUs before concluding with the introduction of KPIs that are or should be used by DIUs to measure their performance.

5.1. DIU types

In total, three DIU types were identified which highlights the diversity of the sample. However, the analyzed DIUs show high degrees of individuality in regard to their development over time, specific activities as well as setting and design.

Based on the results of the conducted survey, the type of each case was derived dependent on the specific objective and scope. Whereas no case was identified as internal facilitator, five cases (Case A-E) were classified as external enhancers and two cases (Case F & G) as external creators. Due to their balanced innovation orientation, Cases H & I were categorized as hybrid types. A summary of the classification is depicted in Table 12.

Table 12: Identified DIU types

	Objective and scope		
Case	Innovation Orientation	Market Focus of Innovation	DIU Type
A	Purely External	Existing Business Areas	External enhancer
B	Primarily External	Existing Business Areas	External enhancer
C	Primarily External	Existing Business Areas	External enhancer
D	Balanced, Primarily External	Existing Business Areas	External enhancer
E	Primarily External	Existing Business Areas, Novel Business Areas	External enhancer
F	Primarily External	Novel Business Areas	External creator
G	Primarily External, Purely External	Existing Business Areas, Novel Business Areas	External creator
H	Balanced	Existing Business Areas	Hybrid
I	Balanced	Novel Business Areas	Hybrid

Note. In case responses of multiple interview partners within one case were inconsistent, the data was triangulated with the results from the semi-structured interview to improve consistence.

Following the logic of the results point of interface within mixed-method research design, the results of the supplementary component were integrated into the results of core component. In the following, general information for each case is provided. This includes data gathered through the interviews (e.g. year of foundation, development over time, exemplary activities) as well as the case specific setup within the organizational dimensions embedding, collaboration (i.e. external cooperation, internal ties) and projects (i.e. project selection, location).

Case A (external enhancer): Throughout its existence, Case A went through several transformations of its activities and strategy. Initially, the DIU focused on transferring deep tech knowledge into the core organization. With 26 employees, Case A realizes venture capital activities and company building for the core organization as well as for external clients (Case A, Interview One). As depicted in Appendix 12, Case A is set up as separate legal entity and its internal ties are best described by regular liaison. External cooperation is seen as standard working mode. Whereas projects are conducted on- and offsite, the projects are mainly ordered and not selected autonomously.

Case B (external enhancer): Established in 2018, Case B has developed towards a DIU with around 150 employees (located in several countries) and is set up as separate business area and legal entity. Initially, Case B was seen as starting point for the digital transformation of the core organization but evolved its focus towards company building (Case B, Interview One). With irregular external cooperation and regular liaison to the core organization, projects are ordered and selected autonomously as well as located on- and offsite (Appendix 13).

Case C (external enhancer): Case C was established as co-innovation lab in 2017 and allows employees from the core organization to become intrapreneurs and work on innovation beyond existing products of the core organization. The responsibilities of its innovation activities are split up between two teams. Whereas the first team (around nine team members) is responsible for ideation, the second team (around nine team members) proceeds with incubation including launch & scale within the established innovation process (Case C, Interview One & Two). As depicted in Appendix 14, Case C is embedded as line organization with regular liaison. Whereas external cooperation is seen as standard working mode for ideation, the incubation and launch & scale team cooperates only irregularly with externals. Projects are ordered and selected autonomously and normally conducted at the facilities of the DIU.

Case D (external enhancer): With around 12 employees, Case D is set up as digital transformation unit that focuses on customer-centric digital initiatives, leveraging innovation created by startups within the core organization and developing towards a more data-driven organization (Case D, Interview One). Embedded as executive department with external cooperation as standard working mode, projects are ordered and selected autonomously as well as conducted on- and offsite. Regarding its internal ties to the core organization, Interviewee One described the relationship as regular liaison, whereas Interviewee Two sees the DIU as core organization integrated (Appendix 15).

Case E (external enhancer): Established in 2018, Case E initially focused on defining strategic priorities in the context of digitalization for the core organization. However, Case E is currently developing towards a more execution-focused DIU. In total, Case E has around 22 employees which are divided into several teams that focus on strategy, project management and creating digital innovation (Case E, Interview One & Two). The graphical overview of the organizational setup depicted in Appendix 16 reveals a discrepancy between the responses of Interviewee One and Two. An example is the embedding of the DIU as executive department and separate legal entity, respectively.

Case F (external creator): Case F (around 10 employees) has a clear focus on supporting its members, which are estab-

lished companies, to validate ideas at an early stage (Case F, Interview One). In contrast to several other cases, Case F only covers idea generation, idea selection and innovation development within the innovation process. Being setup as separate legal entity, Case F utilizes external cooperation as standard working mode. It collaborates regularly with the core organization and project selection as well as location are mixed (Appendix 17).

Case G (external creator): Case G was established in 2018 focusing on scouting and developing new digital business models. Cost reductions within the core organization did not only result in a reduced team size from 22 employees to around 10 employees, but also in the termination of its accelerator activities (Case G, Interview One & Two). Case G mainly focuses on innovation within novel business areas. However, existing business areas are also covered. Similar to Case E, there is a discrepancy in regard to the understanding of the setting and design of Case G between Interviewee One and Two (Appendix 18).

Case H (hybrid type): Established in 2017 as cross functional unit, Case H has become a DIU with around 190 employees that have various backgrounds, such as data science, marketing and e-commerce. With an equal focus on internal and external innovation, Case H was classified as hybrid type. Starting as a pure e-commerce unit, Case H has realized several lighthouse projects that aim at advancing digitalization across the entire value chain of the core organization (Case H, Interview One). As depicted in Appendix 19, Case H is embedded as executive department and core organization integrated. External cooperation is seen as standard working mode. Whereas projects are mainly conducted onsite, the projects are partly ordered by core organization and partly selected by the DIU autonomously.

Case I (hybrid type): Case I was established in 2015 as a corporate startup and focuses almost exclusively on e-commerce activities (Case I, Interview One). Due to its purpose and the balanced innovation orientation, Case I is best classified as hybrid type using external cooperation as standard working mode. Embedded as separate legal entity and core organization integrated, the projects of Case I are selected autonomously and ordered by the core organization as well as mainly conducted onsite (Appendix 20).

After having established a general understanding of the sample, it is important to comprehensively understand which objectives are pursued by each of the sample cases.

5.2. DIU objectives

While most DIUs within the sample define digital innovation as the primary objective, other objectives (e.g. cultural change) are perceived as secondary outcomes. However, all DIU objectives that were introduced in section 2.3 are addressed by the sample (primary or secondary).

Based on the insights generated from the conducted expert interviews, the objectives of each Case (A-I) were identified. A graphical overview is provided in Table 13. We distinguish between primary objectives (marked as x) and secondary objectives (marked as (x)). The table shows that especially the development of digital innovation is pursued as primary objective. Almost all DIUs address the digital innovation ecosystem, cultural change, digital expertise and organizational design change. However, these objectives are oftentimes perceived as secondary outcomes, that occur as by-outcomes when pursuing primary objectives (Case A-I, Interview One & Two).

Each DIU distinguishes itself from others with regards to organizational dimensions. Due to the high degrees of individualism, also in regard to the particular interpretation and implementation of objectives, an overview of each pursued objective is provided. As cultural change and digital expertise are oftentimes seen as secondary objectives and pursued through similar activities (section 2.3, Appendix 3), the findings related to these objectives were aggregated.

Digital innovation: Case A addresses the development of digital innovation through building companies, starting with incubation and targeting at raising external funding as validation and market proof. Whereas Case A primarily focuses on creating digital business value for the core organization, projects with external organizational clients are conducted on a regular basis in order to generate revenue and gain market experience (Case A, Interview One). Similar to Case A, Case B aims at building companies (Case B, Interview One). Case C addresses digital innovation through developing innovations that go beyond physical products (e.g. services, digital tools, business model innovation) in order to contribute to the digital transformation of the core organization (Case C, Interview One & Two). Case D focuses on the digital transformation of the core organization, conducting digital initiatives, that are realized in an agile, customer-centric and business problem-oriented way (Case D, Interview One). By building new services around the core product portfolio, Case E targets at augmenting the existing offering of the core organization as well as increasing its revenue. To realize this objective, Case E applies user-centric methodologies like design thinking, involves various stakeholders from the core organization and utilizes data to create a better understanding of its customers (Case E, Interview One & Two). Case F distinguishes itself from the other cases, as the DIU primarily supports its corporate members in validating early-stage ideas in standardized ways. By doing so, agile methodologies, such as design thinking are applied (Case F, Interview One). Similar to Case A and B, Case G targets at building companies. Case G has a clear focus on creating digital business through scouting, developing, validating and scaling new, standalone digital services (Case G, Interview One & Two). The key objective of Case H (hybrid type) is to digitally transform the different business areas of the core organization. In addition to establishing e-commerce activities, this includes digital innovation projects across the entire value chain (Case H, Interview One). Case I also differs from the other cases in regard to its innovation activities. As depicted in Table 13, the development of digital innovation is not seen as primary objective. Case I focuses on the establishment and development of the e-commerce business for the core organization

Table 13: DIU objectives addressed by sample cases

DIU Objectives	DIU type and case									Count*
	External enhancer					External creator		Hybrid		
	A	B	C	D	E	F	G	H	I	
Digital innovation	x	x	x	x	x	x	x	x	(x)	9
Cultural change	(x)	(x)	(x)	x	(x)	(x)	(x)	x	(x)	9
Digital expertise	(x)	(x)	(x)	x	(x)	(x)	(x)	x	(x)	9
Organizational design change	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	9
Digital innovation ecosystem	x	(x)		x	x	x	(x)	x		7

Note. x = primary objective, (x) = secondary objective, *Count indicates how often each objective was mentioned as primary or secondary objective by cases.

(Case I, Interview One).

Cultural change & digital expertise: At Case A, cultural change and digital expertise are advertised actively towards the core organization. However, these objectives are rather pursued passively through projects that are conducted interdisciplinary with employees from the organization and DIU employees. This results in fostering digital culture and expertise. Hence, these are not primary objectives, but rather beneficial side effects and secondary outcomes (Case A, Interview One). Similar objective structures were identified within the Cases B, C, E-G, I. Cases B, E, F, H, I additionally conduct workshops and trainings for employees from the core organization that address topics like agile methodologies. In contrast, Case D has defined these objectives as primary. Through becoming more data driven, employees should be empowered to make better business decisions. Therefore, a change of mindset is fostered that also impacts the way of tackling specific business problems (Case D, Interview One). This cultural shift and mindset change is also pursued by Case H and fostered through innovation events with external speakers (Case H, Interview One).

Organizational design change: Similar to cultural change and digital expertise, organizational design change is seen as secondary objective by most DIUs. Although, organizational design change was not named explicitly by any case, the explanations in regard to main outcomes and objectives of the DIUs indicate, that organizational design change is pursued implicitly. Case C for example, has split up the responsibilities of the innovation process between two teams. Hence, digital innovation can be handed over in a standardized way for further development, fast integration and growth (Case C, Interview One & Two).

Digital innovation ecosystem: Participating in digital innovation ecosystems or establishing new ones, is seen as one primary objective by the majority of the cases. For example, Case A utilized the digital ecosystem to raise external funding for the ventures that were built by the DIU (Case A, Interview One). Case D leverages the startup ecosystem based on a strong open innovation and platform approach, that allows to integrate innovative startup solutions into the core organization (Case D, Interview One). At Case E, the digital ecosystem (e.g. startups, academic player, innovation networks) is utilized for identifying new technologies and collaboration opportunities (Case E, Interview One & Two). For Case F, which particularly aims at establishing a digital ecosystem for its members, ecosystem activities include knowledge exchange, the utilization of local ecosystems for identifying digital talent and the education of its members in regard to digital expertise (Case F, Interview One). Similar, Case H participates in the digital ecosystem to get access to digital experts that support the DIU in establishing digital expertise within the core organization (e.g. keynote speeches, Case H, Interview One). Case B and G see ecosystem activities as secondary outcome (Case B, G, Interview One & Two). Only Case C and I, do not see the digital ecosystem as part of their objectives at all (Case C, I, Interview One & Two).

Following the case study research approach (section 4.1), we aimed at a diverse sample of DIUs, that addresses all objectives introduced in section 2.3. As we aimed at deriving CSFs for DIUs to fulfill their objectives and at identifying KPIs that are or should be used by DIUs to measure their performance, creating a comprehensive understanding of the objectives pursued within the sample is an important and necessary step to in regard to the robustness and generalization of this study. The results show that all objectives are covered by the sample, either directly or indirectly as primary and secondary objectives. Thus, we have set the basis for presenting the results in regard to CSFs and KPIs.

5.3. Critical success factors

In the course of this study, we derived nine important categories of CSFs that consist of 16 explicit CSFs in total.

To answer RQ1 ("What are CSFs for DIUs to fulfill their objectives?"), several categories of CSFs have been identified. Table 14 shows the aggregated results. Different categories of CSFs are depicted in a descending order (dependent on the number of expert mentions).

In total, ten categories have been derived, consisting of 16 CSFs. However, "incentives" was addressed by one case only. Therefore, it was not identified in at least 15% of the cases

Table 14: Categories of CSFs

Categories of CSFs	DIU type and case									Count*
	External enhancer					External creator		Hybrid		
	A	B	C	D	E	F	G	H	I	
The team	x	x	x		x	x	x	x	x	8
Top management support	x	x		x	x		x	x	x	7
Ties into the core organization	x	x	x	x	x		x			6
Acceptance & trust	x		x	x	x	x	x			6
Methodology		x	x		x	x	x	x		6
Distance & playground	x	x	x	x	x		x			6
Impact	x	x	x		x		x		x	6
Unfair advantage	x	x					x			3
Communication					x			x		2
Incentives		x								1

Note. *Count indicates number of cases that mentioned specific category of CSFs.

(section 4.2.1) and will not be explained further. The same logic was applied to explicit CSFs. As no clear differences between different types of DIUs were observable, CSFs were derived across DIU types consistently.

To summarize the results with regards to RQ1 ("What are CSFs for DIUs to fulfill their objectives?"), an overview of categories of CSFs including explicit CSFs is provided in Table 15.

Based on the categories identified, the underlying explicit CSFs are explained in the following, starting with the team, as mentioned most frequently.

The team: In regard to the team, three CSFs have been identified. First, it is important to identify the right people with relevant experience that are willing to join a corporate DIU (1). This also includes identifying the right team members from the core organization. Relevant experience may include a previous employment in a startup or comparable entrepreneurial experience. This can also be a challenge for DIUs, since aforementioned talent often strives towards founding own startups rather than working for a corporate (Case A-C, E, Interview One & Two). Another CSF is to hire the right digital talent in order to decrease the dependency on external support (2). Being dependent on external support would be too expensive and not sustainable long-term (Case E, H, Interview One & Two). Hence, it is crucial to set up a diverse team with the right mix of skills (3). This includes areas like technology (e.g. data scientists, developers), design (e.g. UX designers), marketing & sales (e.g. e-commerce and marketing experts). Thus, DIUs become more independent (Case A, E, H, Interview One & Two).

Top management support: Almost all cases referred to top management support (CEO, board & senior management) (4) as CSF. The reasons for this are multifaceted: First, the risk of a CEO change, which could result in the termination of initiatives set up by the former CEO. This is oftentimes the case for special, cost intensive initiatives like DIUs (Case A, Interview One). Furthermore, it is important to have a clear mandate from the board that empowers the DIU to work towards its objectives, which is oftentimes associated with (monetary) resources (Case E, Interview One & Two). The top management support also impacts the acceptance and understanding for the DIU's activities across the entire core organization (Case D, Interview One & Two).

Ties into the core organization: Another important category are the ties into the core organization which is closely related to top management support. Although top management is crucial for the success of DIUs, close ties into the core organization are seen as almost as important. Therefore, forming alliances & ties with the core organization (e.g. IT, HR, legal, finance) (5) is seen as CSF for DIUs to achieve their objectives. For example, an alliance with the corporate IT is beneficial for handing over digital innovations created by the DIU during the late phases of the innovation process and during early phases to understand the tech stack of the core organization. Alliances with other departments, such as HR and legal, help the DIU to speed up its operations and to become more efficient. Furthermore, it is important to ensure the personal contact to employees across the entire organization (6). Similar to forming alliances, the DIU benefits from such relationships with regards to operations and innovation activities through leveraging core organization resources. For example, the reach of the core organization into different regions can help to scale digital innovations (Case A-E, G, Interview One & Two).

Acceptance & trust: Closely related to ties into the core organization, gaining acceptance & trust from the core organization (7) is another CSF for DIUs. Some cases recommend working on innovation close to the core first, in order to prove the benefits of digital innovation. Thus, the DIU can increase its acceptance & trust and use resulting autonomy to shift further towards novel business areas (Case A, C-G, Interview One & Two).

Table 15: Categories of CSFs including explicit CSFs

Categories of CSF	Explicit CSFs
The team	Identify the right people with relevant experience that are willing to join a corporate DIU (1)
	Decrease the dependency on external support (2)
	Set up diverse team with the right mix of skills (3)
Top management support	Ensure top management support (CEO, board & senior management) (4)
Ties into the core organization	Form alliances and ties with the core organization (e.g. IT, HR, legal, finance) (5)
	Ensure the personal contact to employees across the entire organization (6)
Acceptance & trust	Gain acceptance & trust from the core organization (7)
Methodology	Select appropriate methodologies according to objectives (8)
	Define responsibilities within the innovation process (9)
	Be fast (10)
Distance & playground	Define suitable distance to the core organization (11)
	Create playground for the DIU and its employees (12)
Impact	Create impact on the core organization (13)
	Select the right topics (14)
Unfair advantage	Utilize the “unfair advantage” (15)
Communication	Establish a clear and effective communication concept (16)

Methodology: Due to the strong focus on digital innovation, it is critical to select appropriate methodologies according to the objectives (8), for example, design thinking. It is important to validate each idea based on specific criteria. Similar to traditional R&D, the cost of pivoting increase throughout the innovation process. This CSF is also closely related to the team, which applies these methodologies in their daily work life. Another CSF within methodology is to define the responsibilities within the innovation process (9). It is fundamental to coordinate the transfer of digital innovation from the DIU to the core organization, including the stage within the innovation process, the department that takes over the project and what exactly will be delivered to the core organization (Case B-D, E-H, Interview One & Two). Furthermore, it is crucial to be fast (10). Otherwise, there is the risk that the management loses interest in specific projects. This CSF might be highly dependent on CSF (3), setting up a diverse team with the right mix of skills (Case A, B, H, Interview One).

Distance & playground: Closely related to methodology, the category distance & playground consists of two CSFs. First, define a suitable distance to the core organization (11), which is important to lose the boundaries of the core organization and to speed up operations. With an increasing distance, DIUs are able to operate more autonomously and agile which might have a positive impact on the outcome (Case A, B, E, G, Interview One & Two). Distance to the core organization is related to another CSF, creating a playground for the DIU and its employees (12). This includes the empowerment of the team, allowing to experiment and to fail, the willingness to take risks and agility (Case A-E, Interview One & Two).

Impact: This category consists of the CSFs create impact on the core organization (13) and select the right topics (14). Creating impact on the core organization includes aligning the activities of the DIU with the strategy of the core organization and creating digital innovation that creates business impact with a relevant size (e.g. revenue increase, cost savings). This is not only critical for the cases that focus on existing business (i.e. external enhancer), but also for the external creators. Furthermore, it is important to focus on scalability and a well-diversified digital innovation portfolio that is self-sustaining (Case A, E, G, I, Interview One and Two). The second CSF within impact, select the right topics (14) is closely related to CSF (13). It includes the identification of relevant opportunity fields and problems. Furthermore, it is important to not only focus on “moonshot” topics, that are of high risk, but also on topics that are close to the core business. Thus, the DIU is able to gain visibility which might result in an improved acceptance within the core organization (Case A-C, G, Interview One & Two).

Unfair advantage: Another CSF for DIUs is to utilize the “unfair advantage” (15). Compared to startups, DIUs are usually slower with regards to operations and innovation activities. However, DIUs can use their unfair advantage (i.e. being part of the core organization) over startups. This includes, for example, access to existing sales structures and resources like knowledge and customers (Case A, B, G, Interview One & Two).

Communication: The last category of CSFs is related to communication. It is critical to establish a clear and effective communication concept (16) to communicate activities and success stories to relevant stakeholders, for example through hosting events and publishing intranet articles. Oftentimes, this also has an educational effect which results in an increased awareness for the DIU within the core organization (Case E, H, Interview One & Two).

The majority of CSFs seems to address the key objective of

DIUs, digital innovation. More detailed analysis with regards to the CSFs that focus on other objectives would go beyond the scope of this research project. As a next step, we take a closer look at the results of the semi-structured interviews in order to address RQ2 ("Which KPIs are or should be used by DIUs to measure their performance?").

5.4. Performance measurement of DIUs

Initially, we provide contextual information about the use of KPIs at DIUs (e.g. challenges and benefits of establishing KPIs at DIUs). Subsequently we introduce different sets of KPIs that are or should be applied at DIUs to measure their performance. In addition, we highlight KPIs that are reported to the core organization and conclude with a short set of KPIs that could be applied to measure the impact of the DIU on the digital transformation of the core organization.

5.4.1. Contextual information: Key performance indicators

So far, none of the cases has utilized an established PEMS (e.g. balanced scorecard approach) to measure its performance. While there are several challenges in regard to establishing KPIs at DIUs (e.g. identification of suitable KPIs), the key benefit is to improve transparency. Furthermore, the KPIs applied with regards to digital innovation usually change and develop alongside the innovation process. The majority of the cases recommends the use of a specific set of KPIs across multiple projects. However, each project usually includes individual KPIs as well.

A detailed overview of the identified contextual information that is relevant for the purpose of this study is provided in the following.

PEMS

Interestingly none of the cases has established an existing PEMS, such as the balanced scorecard approach. While there are first approaches of setting up PEMS, they mainly consist of individual KPIs that do not follow the structure of any established PEMS. The reasons for this are manifold: Most DIUs were founded only a few years ago. Hence, performance measurement was oftentimes addressed only recently. In this context, the DIUs mainly focused on identifying suitable KPIs that can be used to measure their activities and performance (Case A-I, Interview One & Two). Establishing a PEMS like the balanced scorecard is related to high efforts, that would go beyond the requirements for performance measurement as of today (Case C, D, G, Interview One & Two). Furthermore, several cases do not see a good fit between the highly individual activities of their DIUs and the applicability of a standardized PEMS (Case D, G, H, Interview One & Two).

Challenges and benefits of establishing KPIs

In total, five challenges and one major benefit of establishing KPIs at DIUs were identified among the sample. A short overview is provided in Table 16.

Challenges

Identification of suitable KPIs: The key challenge of establishing KPIs mentioned by most cases is the identification of suitable KPIs (Case A-C, E-H, Interview One & Two). As employees might work towards achieving specific KPIs that are related to their compensation, DIUs need to ensure that KPIs actually represent their objectives. Furthermore, projects usually include a high degree of individuality. Hence, identifying KPIs that are applicable for various projects poses a challenge (Case B, Interview One). The complexity and nature of digital innovation further complicates the identification of suitable KPIs (Case A-C, E. Interview One & Two). As DIUs also aim at fostering cultural change within the core organization, purely financial measures would not sufficiently address this objective. However, the identification of KPIs that measure cultural change and digital expertise is challenging (Case E, F, Interview One & Two).

Applicability of traditional KPIs: Another important challenge many DIUs are facing is the applicability of traditional KPIs, which are used for performance measurement within the core organization. Examples for these traditional, oftentimes financial KPIs are ROI and earnings before interest and taxes (EBIT). Core organizations incorporated such KPIs to measure their success. Consequently, DIUs are oftentimes expected to use similar KPIs. However, these KPIs are hardly applicable for the objectives and areas of activities pursued within DIUs (Case C-E, G, Interview One & Two).

Incentives related to KPIs: The establishment of incentives that are attached to KPIs is another challenge DIUs are facing as it is complex to identify the right KPIs which consider that digital innovation might take several years before business impacts are actually realized. Hence, DIUs need to appropriately establish incentives that motive employees to work towards the right KPIs (Case B, D, E, G, Interview One & Two).

Focus on reporting: It is important to keep the tracking and reporting of KPIs rather light, in order for DIUs to be able to focus on their actual objectives (Case C, D, F, Interview One & Two).

Data collection: As many of the KPIs that are applied at DIUs differ from the ones at the core organization, there is oftentimes no system for automated tracking in place. Therefore, collecting related and relevant high-quality data is oftentimes demanding for DIUs (Case G-H, Interview One & Two).

While multiple challenges of establishing KPIs at DIUs were identified, one multifaced key benefit was derived among the sample.

Benefits

Transparency: The key benefit of establishing KPIs identified among all cases is transparency. However, transparency is multifaced and includes several elements: First, KPIs allow showing the value of the DIU to the core organization to justify its existence. Second, KPIs help making the outcomes of the DIU more tangible, observable and comparable over time. Third, KPIs are absolutely required for managing and steering the DIU not only as a whole, but on team and individual level (Case A-I, Interview One & Two).

Journey of KPIs

Another contextual information is the "journey of KPIs"

Table 16: Challenges and benefits of establishing KPIs at DIUs

Challenges & benefits	DIU type and case									Count*
	External enhancer					External creator		Hybrid		
	A	B	C	D	E	F	G	H	I	
Challenges										
Identification of suitable KPIs	x	x	x		x	x	x	x		7
Applicability of traditional KPIs			x	x	x	x	x			5
Establishment of incentives		x		x	x	x	x			5
Focus on reporting			x	x		x				3
Data collection							x	x		2
Benefits										
Transparency	x	x	x	x	x	x	x	x	x	9

Note. x = challenge or benefit identified.

which was identified among the majority of the cases. Journey of KPIs means, that the KPIs applied to measure the objective digital innovation might change and develop alongside the innovation process. While non-financial and qualitative KPIs should be prioritized during early phases of the process (i.e. discovery, development), the applied KPIs should partly develop towards traditional, financial KPIs when moving towards later phases of the process (i.e. diffusion, impact). Some cases even recommend having specific KPIs for each process stage (Case A, C, D, F-H, Interview One & Two).

Cross-project applicability of KPIs

Concluding, the interviews revealed another interesting contextual information concerning the cross-project applicability of KPIs. Due to the high degree of individuality across the conducted projects, it is important to apply customized KPIs for each initiative. However, having a set of common KPIs to foster comparability and to improve the efficiency of performance measurement within the DIU is key (Case A, C, D, F-H, Interview One & Two).

After having introduced relevant contextual information in regard to key performance indicators, we have laid the foundation for answering RQ2 ("Which KPIs are or should be used by DIUs to measure their performance?").

5.4.2. Key performance indicators

In total, 38 KPIs were identified that are or should be used by DIUs to measure their performance. The set was split up into objective related subsets, KPIs that are reported to the core organization and complemented by KPIs that are or should be applied to measure the impact of the DIU on the digital transformation of the core organization.

For measuring the objective digital innovation, 26 KPIs were identified. These can be further clustered into financial (7 KPIs), customer-related (9 KPIs) and other KPIs (10 KPIs). In accordance with section 5.2, KPIs related to cultural change and digital expertise were aggregated (10 KPIs). While no KPI could be identified for the objective organizational design change, two KPIs were derived in regard to digital innovation ecosystem. An overview including all subclusters and the count of each KPI, meaning the number of cases that mentioned a specific KPI, is depicted in Table 17. The high share of KPIs related to digital innovation indicates the importance of the objective within the sample (section 5.2). The KPIs are listed in descending order dependent on their count. In case of the same count, alphabetical sorting was applied.

While an overview of the most important findings is given in the following, Appendix 21 provides a short explanation of each of the identified KPIs.

Digital innovation

Financial: Each of the seven KPIs identified in this cluster is applied by at least two cases within the sample. There are three KPIs included that measure different types of revenue. To be precise, revenues generated through digital innovations created by the DIU, revenues created through services the DIU offers to the core organization and occasionally to external clients and revenues that result from establishing the e-commerce business of the core organization. Another widely applied KPI is the input-related measure budget / cost of the DIU. As many DIUs also create digital innovation that aim at improving processes, the achieved efficiency gains get measured through cost savings. Concluding, several DIUs use ROI to measure their digital innovation activities from a financial point of view (Case A-I, Interview One & Two).

Non-financial – customer: A sub-cluster of non-financial KPIs that are related to customers was identified. While each KPI is explained in Appendix 21, we want to highlight three KPIs, that are widely used among the sample cases: Number of customers, customer / user satisfaction and net promoter score which is closely related to customer satisfaction. Compared to the majority of financial KPIs, customer-related measures are suitable for earlier phases of the innovation process (Case A-I, Interview One & Two).

Non-financial – other: Complementary, the remaining KPIs that were derived to measure the objective digital in-

Table 17: Objective related KPIs of DIUs

KPIs to measure specific objectives	DIU type and case									Count*
	External enhancer					External creator		Hybrid		
	A	B	C	D	E	F	G	H	I	
Objective: Digital innovation										
Financial										
Budget / cost		x	x	x	x	x	x	x	x	8
Revenue (digital innovation)	x	x	x		x	x	x	x		7
Cost savings	x			x	x	x	x	x		6
Revenue (offered services)	x					x	x	x		4
ROI			x		x	x				3
Ratio digital / traditional business				x				x		2
Revenue (e-commerce)								x	x	2
Non-financial: Customer										
Number of customers	x	x	x	x	x		x	x		7
Customer / user satisfaction	x		x		x	x		x	x	6
Net promoter score	x	x	x		x				x	5
Churn rate		x					x			2
Customer engagement				x	x					2
Customer acquisition cost							x			1
Customer centricity			x							1
Customer lifetime value							x			1
Number of interviews						x				1
Non-financial: Other										
Number of projects per stage		x		x	x	x	x	x		6
Total number of projects / tests				x	x		x	x		4
Time to market					x	x		x	x	4
Number of ideas evaluated	x						x	x		3
Share of digital transactions			x					x		2
Number of successful projects		x						x		2
Business model readiness			x							1
Interaction rate			x							1
Learning velocity			x							1
Stability			x							1
Objective: Cultural change & digital expertise										
Adoption rate (e.g. methodologies)				x	x					2
Net promoter score						x	x			2
Number of collaborations					x	x				2
Number of workshops		x				x				2
Employee satisfaction						x				1
Hours of training		x								1
Number of digital trends scouted					x					1
Number of exchanges						x				1
Number of workshop participants		x								1
Social media traction				x						1
Objective: Digital innovation ecosystem										
Number of collaborations					x					1
Number of exchanges						x				1

novation were aggregated within the sub-cluster “other”. The KPIs most frequently used are the number of projects per stage, total number of projects / tests, time to market and number of ideas evaluated. Similar to customer-related measures, these KPIs are suitable for early phases of the innovation process and could even be applied within the first

stage (i.e. discovery, Case A-I, Interview One & Two).

Cultural change & digital expertise

Compared to the KPIs identified for the objective digital innovation, the KPIs applied in regard to cultural change and digital expertise are utilized by fewer DIUs. This might be related to the findings of section 5.2, which shows that most DIUs within the sample define cultural change and digital expertise as secondary objectives. The KPIs used by at least two DIUs are adoption rate (e.g. methodologies), net promoter score, number of collaborations as well as number of workshops (Case B, D-G, Interview One & Two).

Digital innovation ecosystem

With regards to digital innovation ecosystem, the KPIs number of collaborations (e.g. collaborations with startups) and number of exchanges (e.g. exchanges with partners within the digital innovation ecosystem) were identified (Case E-F, Interview One & Two).

In addition, we identified objective related KPIs that are or should be reported by the DIU to the main organization. An overview that follows the same order logic as Table 17 is depicted in Table 18.

The KPIs used by DIUs to measure their objectives (Table 17), also set the basis for the reporting to the core organization. However, not all KPIs that are measured are consequently reported. Furthermore, the count of KPIs, meaning the number of cases that report a specific KPI, is lower for several KPIs compared to the count depicted in Table 17 (e.g. cost savings, number of customers, number of projects per stage). Especially financial and customer related KPIs (e.g. revenue (digital innovation), customer / user satisfaction) seem to be of high importance in regard to reporting. Interestingly the average number of KPIs reported is 6.3 KPIs per case while the average number of measured KPIs per case is 10.4. With 14 KPIs in total, Case E established the most comprehensive set of KPIs that are or should be reported among the sample cases (Case A-I, Interview One & Two).

Concluding, we derived a short set of KPIs that are or should be applied by DIUs to measure their impact on the digital transformation of the core organization (Table 19). While most of these KPIs are only applied by individual DIUs, the KPIs ratio digital / traditional business and net promoter score are applied by multiple DIUs. Similar to the explanations provided for the KPIs that are used to measure specific objectives, a short overview including explanations for this set of KPIs is provided in Appendix 22 (Case A-F, H, Interview One & Two).

Thus, we answered RQ2 ("Which KPIs are or should be used by DIUs to measure their performance?") through deriving several sets of KPIs, that enable DIUs to measure specific objectives, including reported KPIs and metrics to measure the impact of the DIU on the digital transformation of the core organization.

6. Discussion

Following RQ1 ("What are CSFs for DIUs to fulfill their objectives?"), we identified nine re-levant categories of CSFs, which consist of 16 explicit CSFs. We answered RQ2 ("Which KPIs are or should be used by DIUs to measure their performance?") by deriving 38 KPIs that were clustered into objective related subsets (e.g. KPIs to measure digital innovation). Additionally, we identified KPIs that are or should reported to the core organization as well as KPIs to measure the impact of the DIU on the digital transformation of the core organization. Thus, our results address the existing research gap in the field of DIUs and performance measurement.

By answering both research questions and merging our results with existing literature, we aim at deriving a framework for the measurement of success of DIUs. Therefore, we initially discuss our results in association with existing research, following the same content structure as section 5. At the end of each section (section 6.1.1 - 6.1.5), we derive relevant steps for the framework. Subsequently, the framework for the measurement of success of DIUs, which is the key outcome of this study, is presented. Concluding, implications for practice and theory, limitations as well as ideas for future research are introduced.

6.1. Framework development

By discussing and merging our findings with existing literature, six guiding steps that set the foundation for the framework were derived. These steps address various aspects of DIUs, such as different DIU types and objectives, CSFs and relevant KPIs.

6.1.1. DIU types

Using the typology of DIUs derived by Barthel et al. (2020) to identify the type of each sample case, we identified five external enhancers, two external creators and two hybrid types. However, no case was classified as internal facilitator. This indicates that DIUs with an external innovation orientation might be prevalent. We followed Barthel et al. (2020) and exclusively focused on the objective and scope for identifying the DIU type. In addition, we gathered relevant information in regard to the setting and design of each case (Appendix 12 – 20) through conducting the supplemental survey.

The results revealed two major discrepancies: First, differences in regard to the responses of multiple survey participants within one case were identified (e.g. Case D, Appendix 15), for example in regard to the dimensions internal ties and project selection, showing that the setting and design of DIUs can be perceived differently among its employees. Second, the specific setting and design of several cases does not correspond to the setup of ideal types recommended by Barthel et al. (2020) for achieving the specific objectives. However, these design patterns of ideal types "appear to fit the different innovation objectives" (Barthel et al., 2020, p. 12). Thus, we derived two initial steps for the framework based on literature and our findings:

Table 18: Objective related KPIs of DIUs reported to the core organization

KPIs to measure specific objectives	DIU type and case									Count*
	External enhancer					External creator		Hybrid		
	A	B	C	D	E	F	G	H	I	
Objective: Digital innovation										
Financial										
Revenue (digital innovation)	x	x	x		x		x	x		6
Budget / cost		x	x		x				x	4
Revenue (offered services)	x					x	x			3
Cost savings	x				x					2
Ratio digital / traditional business				x				x		2
Revenue (e-commerce)								x	x	2
ROI			x		x					2
Non-financial: Customer										
Customer / user satisfaction			x	x	x	x			x	5
Number of customers			x		x		x	x		4
Net promoter score			x		x				x	3
Customer centricity			x							1
Customer engagement					x					1
Non-financial: Other										
Non-financial: Other										
Time to market					x			x	x	3
Total number of projects / tests				x	x			x		3
Number of projects per stage					x			x		2
Share of digital transactions			x					x		2
Business model readiness			x							1
Interaction rate			x							1
Learning velocity			x							1
Number of successful projects								x		1
Stability			x							1
Objective: Cultural change & digital expertise										
Adoption rate (e.g. methodologies)					x					1
Hours of training		x								1
Number of collaborations					x					1
Number of digital trends scouted					x					1
Number of workshop participants		x								1
Number of workshops		x								1
Objective: Digital innovation ecosystem										
Number of exchanges						x				1

Step 1: Define the objective & scope of the DIU in regard to digital innovation.

Step 2: Dependent on the objective & scope, consider the setting & design of ideal DIU types for establishing a new DIU and for comparison in case the DIU was already established.

6.1.2. DIU objectives

By analyzing existing literature, five objectives pursued by DIUs were identified (i.e. digital innovation (primary objective), cultural change, digital expertise, organizational design change, digital innovation ecosystem). As we aim at considering objective related KPIs within the framework, it was important to understand which DIU objectives are actually addressed by the sample cases, to improve the applicability and generalizability of the framework.

In accordance with existing literature, digital innovation is seen as primary objective among the majority of the sample cases. Cultural change, digital expertise and organizational design change are oftentimes perceived as secondary objectives, that occur as by-outcomes when pursuing main objectives. In addition, several cases defined participating in or creating digital innovation ecosystems a primary objective as well.

To sum up, all objectives identified within literature have

Table 19: KPIs to measure impact on digital transformation

KPIs to measure impact on digital transformation	DIU type and case									Count*
	External enhancer					External creator		Hybrid		
	A	B	C	D	E	F	G	H	I	
Ratio digital / traditional business		x	x	x	x			x		5
Net promoter score		x			x	x				3
Acceptance rate					x					1
Cost savings	x									1
Employee retention					x					1
Employee satisfaction						x				1
Number of successful projects			x							1
Revenue (digital innovation)	x									1
Revenue safeguarded					x					1

been confirmed as primary or secondary objective by the sample cases. Therefore, the sample represents a sufficient basis for answering RQ1 and RQ2. As the objectives pursued within the DIU are crucial in regard to defining suitable KPIs (Chapter 3), we included an objective based step to the framework:

Step 3: Define which objectives are pursued at the DIU that go beyond digital innovation.

6.1.3. Critical success factors

We identified 16 CSFs for DIUs to fulfill their objectives and clustered these into categories of CSFs answering RQ1. While existing contributions by practitioners mainly cover recommendations for DIUs that lack in theoretical foundation, literature only provides important factors for DIUs. However, CSFs have been widely addressed in other research areas (e.g. project management) and occasionally also in related fields to DIUs (e.g. digital business strategy). The common approach observed within literature, is the aggregation of CSFs into clusters derived from the underling concept. Hence, we followed this approach and adapted the overview of CSFs provided in Table 15 through utilizing the structure of the typology provided by Barthel et al. (2020). Dependent on the respective issue addressed within the individual CSFs, each explicit CSFs was assigned to the categories objective and scope and setting and design. An overview is provided in Table 20.

The identified CSFs partly overlap with the factors identified by practitioners and academics (e.g. top management support, Table 2). However, the majority of CSFs have not been derived previously. Thus, we addressed the existing research gap and identified CSFs for DIUs to fulfill their objectives. CSFs were included into the framework through the following step:

Step 4: Consider the explicit CSFs for DIUs to fulfill their objectives within the categories objective & scope and setting & design.

6.1.4. Contextual information: Key performance indicators

Existing literature already highlighted the importance of PEMS for innovation activities. Several attempts of adopting established approaches, such as the balanced scorecard, to measure innovation exist. One key takeaway is the importance of linking KPIs to specific objectives (section 3.2). Furthermore, several important requirements for PEMS in DIUs (section 3.2.2) as well as challenges of measuring digital innovation have been discussed in detail (section 3.3.2). We contribute to the existing discussion by deriving contextual information in regard to the application of KPIs at DIUs. We outlined, that none of the cases has utilized an established PEMS (e.g. balanced scorecard) to measure its performance so far. Furthermore, we identified challenges (e.g. identification of suitable KPIs) and a multifold key benefit (transparency) of establishing KPIs at DIUs. In addition, we introduced the "journey of KPIs", which entails the development of digital innovation KPIs alongside the innovation process. These findings are in accordance with the innovation performance measurement framework provided by Dewangan and Godse (2014) (Table 9) and selected requirements for PEMS in DIUs (Table 5, Table 6, Table 7, Tabl 8). As the requirements derived by Haskamp et al. (2021) set a profound basis for establishing PEMS at DIUs and can be applied as guidelines for selecting specific KPIs to measure the success of DIUs, we added the following step to the framework:

Step 5: Consider the requirements for PEMS in DIUs.

6.1.5. Key performance indicators

By deriving a profound set of 38 KPIs that were clustered into objective related subsets, we answered RQ2. An aggregated overview of the identified KPIs is provided in Table 21.

In addition, we identified KPIs that are reported to the core organization and KPIs that are or should be used to measure the impact of the DIU on the digital transformation of the core organization (Table 22). The derived KPIs partly overlap with the selected innovation metrics identified by Chan et al. (2008) (Table 10) but are more comprehensive

Table 20: Adapted overview of explicit CSFs: Assignment to categories of DIU typology

Category	Explicit CSFs
What? Objective & Scope	Select appropriate methodologies according to objectives. Define responsibilities within the innovation process. Be fast. Create playground for the DIU and its employees. Create impact on the core organization. Select the right topics. Utilize the “unfair advantage”.
How? Setting & Design	Identify the right people with relevant experience that are willing to join a corporate DIU. Decrease the dependency on external support. Set up diverse team with the right mix of skills. Ensure top management support (CEO, board & senior management). Form alliances and ties with the core organization (e.g. IT, HR, legal, finance). Ensure the personal contact to employees across the entire organization. Gain acceptance & trust from the core organization. Define suitable distance to the core organization. Establish a clear and effective communication concept.

Table 21: Objective related KPIs

Objective: Digital Innovation		
Financial		
Budget / cost	Revenue (offered services)	
Revenue (digital innovation)	ROI	Revenue (e-commerce)
Cost savings	Ratio digital / traditional business	
Non-financial: Customer		
Number of customers	Churn rate	Customer centricity
Customer / user satisfaction	Customer engagement	Customer lifetime value
Net promoter score	Customer acquisition cost	Number of interviews
Non-financial: Other		
Number of projects per stage	Share of digital transactions	
Total number of projects / tests	Number of successful projects	Learning velocity
Time to market	Business model readiness	Stability
Number of ideas evaluated	Interaction rate	
Objective: Cultural Change & Digital Expertise		
Adoption rate (e.g. methodologies)	Number of workshops	Number of exchanges
Net promoter score	Employee satisfaction	Number of workshop participants
Number of collaborations	Hours of training	Social media traction
	Number of digital trends scouted	
Objective: Digital Innovation Ecosystem		
Number of collaborations		
Number of exchanges		

(e.g. subsets of KPIs) and particularly applicable to DIUs. In accordance with literature, the KPIs were directly related to objectives (section 3.3.1) and clustered into different subsets (section 3.3.2). With regards to the objective organizational design change, no specific KPI was identified. The reason might be, that this objective was classified secondary by all sample cases.

The KPIs derived in this study contribute to the existing research gap of performance measurement at DIUs and can be utilized as a starting point for establishing PEMS in DIUs. Literature suggests assigning KPIs to specific stages within the innovation process (section 3.3.2, Table 9), which might be considered when selecting and applying KPIs. Based on literature and our findings, we derived the last step of the framework:

Step 6: Select an appropriate number of objective related KPIs dependent on the specific objectives of the DIU (under consideration of the requirements for PEMS in DIUs intro-

Table 22: KPIs to measure impact of DIU on digital transformation

KPIs		
Ratio digital / traditional business	Cost savings	Number of successful projects
Net promoter score	Employee retention	Revenue (digital innovation)
Acceptance rate	Employee satisfaction	Revenue safeguarded

duced in Step 5). Define which of the selected KPIs are reported to the main organization. If necessary, complement the selected KPIs with an additional set of KPIs to measure the impact of the DIU on the digital transformation of the core organization.

6.2. Framework for the measurement of success of DIUs

By aggregating and enriching the derived steps with relevant insights, the practical framework for the measurement of success of DIUs was derived. The six guiding steps within the framework address various aspects of DIUs, including the DIU type (step 1 & 2), DIU objectives (step 3), CSFs (step 4), requirements for PEMS in DIUs (step 5) and KPIs to measure the performance of DIUs (step 6). The main outcome of this research project, the practical framework for the measurement of success of DIUs including all steps, is depicted in Figure 6. In addition, we provide a more comprehensive version of the framework, which was particularly designed for practical application, in Appendix 23.

To make sure that all success related aspects of DIUs derived and aggregated within this study are taken into account, each of the six guiding steps should be considered consecutively, starting with step one. When applying the framework, the different steps can be prioritized to different degrees, depending on specific needs and inefficiencies of the existing performance measurement. Thus, the framework sets the foundation for establishing and improving performance and success measurement at DIUs in a practical and flexible way.

6.3. Implications, limitations and future research

The comprehensive framework derived within this study comes along with several implications for practice: First, the framework can be applied for various DIU types. Second, it can be utilized when establishing new DIUs and adjusting existing ones, as it includes important information in regard to the setting and design of ideal types. Third, the framework helps to assess which CSFs have been realized within the DIU already and which further CSFs should be addressed. Fourth, the framework allows to align existing performance measurement with the recommendations for PEMS in DIUs. Last, the identified KPIs set a profound basis for selecting an appropriate set of KPIs to measure the performance of DIUs. To sum up, the framework sets the foundation to measure the success of DIUs in a practical way, through guiding DIU employees, responsible for success and performance measurement, in regard to establishing or improving the success measurement of their DIU. It helps to address existing challenges DIUs are facing and to increase transparency, for the DIU itself and for the core organization.

Moreover, we contribute to the existing discussion around performance measurement of DIUs by answering both research questions and deriving the framework. Hence, this study has multiple implications for theory. By using the typology of DIUs provided by Barthel et al. (2020), we were able to identify the type of each case within our sample. The objectives of DIUs derived by Raabe et al. (2021) set the foundation to identify the objectives of the sample cases. No further objectives were identified, showing the comprehensiveness of the set of objectives. The theoretically sound identification and clustering of explicit CSFs for DIUs enriched the existing contributions of practitioners and academia. By outlining contextual information in regard to the application of KPIs at DIUs, we contributed further interesting insights with regards to performance measurement of DIUs, which can be considered along the requirements of PEMS in DIUs derived by Haskamp et al. (2021). Existing literature in regard to innovation performance measurement was utilized to identify KPIs that are or should be applied by DIUs to measure their performance. Therefore, we contribute to the theory surrounding performance measurement of digital innovation.

Although we followed the guidelines of the underlying research methodologies rigorously, our findings are not without limitations. We investigated a small sample that consisted of nine cases that differ in company size and industry. However, the sample is geographically limited to Germany and Switzerland. In addition, we were only able to conduct multiple interviews for three cases. The data was gathered and analyzed by one researcher. Future research could investigate a larger sample with cases from several regions and industries, through conducting in depth case studies with multiple researchers and interviews per case to improve generalizability, reliability and reduce interviewer bias.

We exclusively conducted interviews with upper management DIU employees, who have a profound understanding of the success measurement within their DIU. Interview partners from core organizations that are directly involved in the reporting process of DIUs might provide further interesting insights in regard to success measurement of DIUs. As no case was identified as type internal facilitator within the sample, future research could particularly address success measurement of this DIU type. In addition, the derived framework could be applied in future studies to investigate its applicability in real world settings and to identify potential shortcomings. It would be interesting to see how the identified CSFs are actually realized within DIUs and how

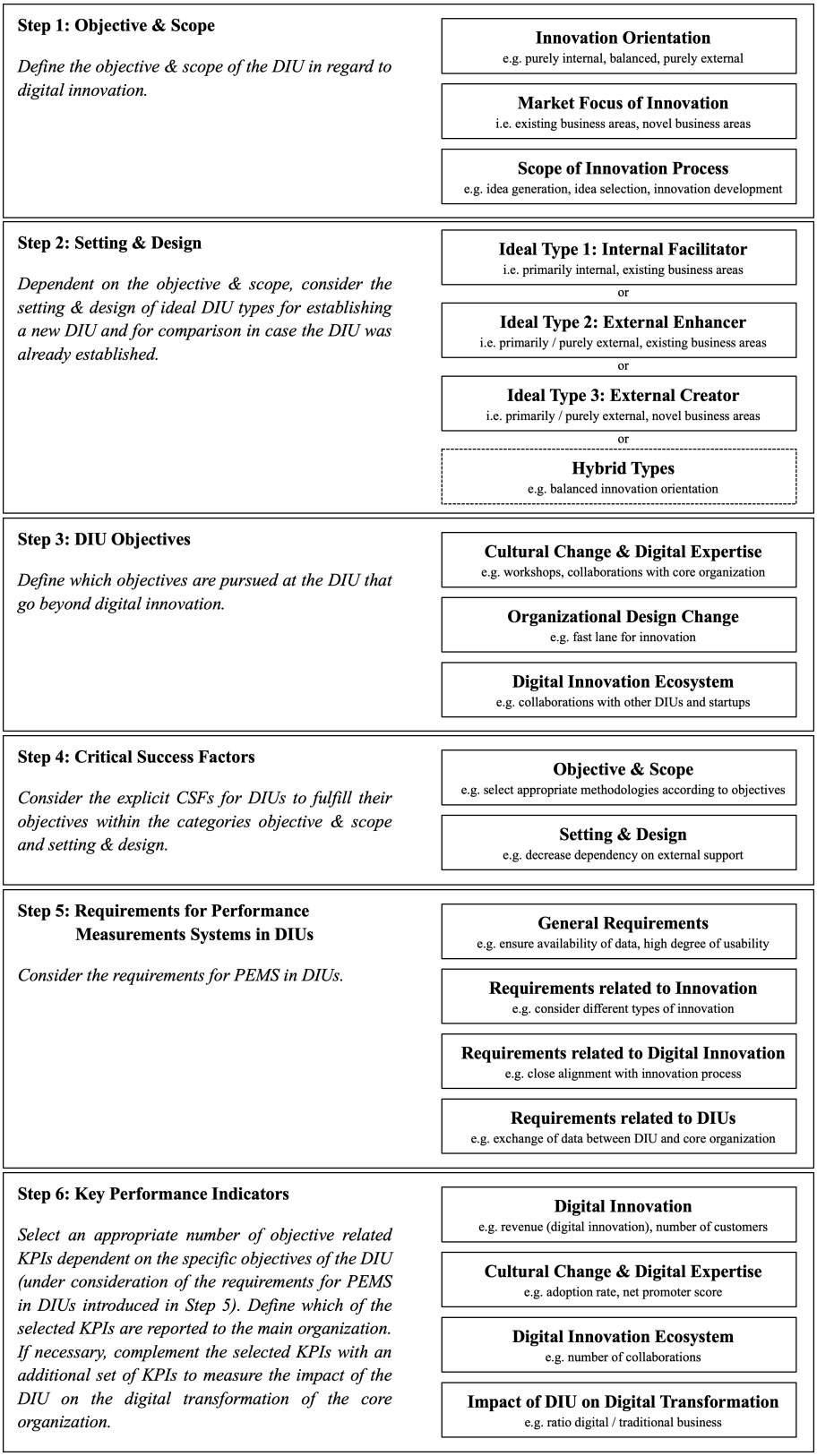


Figure 6: Framework for the measurement of success of DIUs

their implementation impacts the success of DIUs. Similar, future research could analyze how the derived KPIs are implemented within DIUs. Moreover, a subset of KPIs could be applied to measure the success of DIUs, which would enable

performance comparison across multiple DIUs.

Within this study, we only captured the current status of the cases. However, the results from the interviews revealed, that DIUs constantly develop over time, which might also impact performance measurement of DIUs. Future research could investigate the same sample at a later point in time to investigate, if and how CSFs and KPIs have developed over time.

7. Conclusion

This research project deals with digital innovation in corporations. It contributes to DIU and performance measurement research through deriving a practical framework for the measurement of success of DIUs.

Based on the applied case study research design and the mixed-method approach consisting of semi-structured interviews as core component, we answered both research questions in a profound way. We derived 16 CSFs for DIUs to fulfill their objectives. These were assigned to the dimensions of the typology of DIUs (RQ1). Furthermore, we identified 38 KPIs that are or should be used by DIUs to measure their performance and clustered these into objective related subsets. We identified KPIs that are or should reported to the core organization and supplemented the objective related metrics by KPIs that can be used to measure the impact of the DIU on the digital transformation of the core organization (RQ2).

Moreover, we retrieved contextual information in regard to the use of KPIs at DIUs. The results generated through the core component were complemented with data retrieved from the supplementary survey which was utilized to identify the type of each case. Hence, we contributed to the ongoing scientific discussions around different types of DIUs.

Through merging existing literature with the results of this study, a comprehensive framework consisting of six guiding steps was derived that addresses various aspects of DIUs, such as different DIU types, objectives, CSFs and relevant KPIs. Among multiple implications for practice and theory, the framework sets the foundation for establishing and improving performance and success measurement at DIUs in a practical way.

In conclusion, this study is among the first scientific contributions that address the topic success measurement of DIUs. The identified CSFs and KPIs as well as the derived framework are important contributions to this research field, which is facing an emerging interest by practitioners and academia. It does not only provide multiple implications for practice, but also builds the basis for various future research projects that can validate and extend the findings provided within this thesis.

References

Abd El Hameed, T., Latif, M. A. E., & Kholief, S. (2016). Identify and classify critical success factor of agile software development methodology using mind map. *International Journal of Advanced Computer Science and Applications (IJACSA)*, *7*(5), 83-92.

Adams, R., Bessant, J., & Phelps, R. (2006). Innovation management measurement: A review. *International Journal of Management Reviews*, *8*(1), 21-47.

Adams, W. C. (2015). Conducting semi-structured interviews. In *Handbook of practical program evaluation* (4th ed., p. 492-505). New Jersey, USA: John Wiley & Sons. doi: 10.1002/9781119171386

Ariyachandra, T. R., & Frolick, M. N. (2008). Critical success factors in business performance management—striving for success. *Information Systems Management*, *25*(2), 113-120. doi: 10.1080/10580530801941504

Astri, L. Y. (2015). A study literature of critical success factors of cloud computing in organizations. *Procedia Computer Science*, *59*, 188-194. doi: 10.1016/j.procs.2015.07.548

Barros, R. S., & Ferreira, A. M. D. S. d. C. (2019). Bridging management control systems and innovation: The evolution of research and possible research directions. *Qualitative Research in Accounting & Management*, *16*(3), 342-372.

Barthel, P., Fuchs, C., Birner, B., & Hess, T. (2020). Embedding digital innovations in organizations: A typology for digital innovation units. In *Proceedings of the 15th International Conference on Wirtschaftsinformatik*. doi: 10.30844/wi_2020_h1-barthel

Basili, V., Heidrich, J., Lindvall, M., Münch, J., Regardie, M., Rombach, D., ... Trendowicz, A. (2007). Bridging the gap between business strategy and software development. In *Proceedings of the ICIS 2007*.

Bedford, D. S. (2015). Management control systems across different modes of innovation: Implications for firm performance. *Management Accounting Research*, *28*, 12-30. doi: 10.1016/j.mar.2015.04.003

Belassi, W., & Tukel, O. I. (1996). A new framework for determining critical success/failure factors in projects. *International Journal of Project Management*, *14*(3), 141-151. doi: 10.1016/0263-7863(95)00064-X

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The case research strategy in studies of information systems. *MIS quarterly*, *11*(3), 369-386. doi: 10.2307/248684

Birkinshaw, J., & Gibson, C. (2004). Building ambidexterity into an organization. *MIT Sloan Management Review*, *45*(4).

Boerman, M. P., Lubsen, Z., Tamburri, D. A., & Visser, J. (2015). Measuring and monitoring agile development status. In *Proceedings of the 2015 IEEE/ACM 6th International Workshop on Emerging Trends in Software Metrics* (p. 54-62). IEEE. doi: 10.1109/WETSoM.2015.15

Bourne, M., Mills, J., Wilcox, M., Neely, A., & Platts, K. (2000). Designing, implementing and updating performance measurement systems. *International Journal of Operations & Production Management*, *20,7*, 754-771.

Bourne, M., Neely, A., Mills, J., & Platts, K. (2003). Implementing performance measurement systems: A literature review. *International Journal of Business Performance Management*, *5*(1), 1-24. doi: 10.1504/IJBPM.2003.002097

Boynton, A. C., & Zmud, R. W. (1984). An assessment of critical success factors. *Sloan Management Review*, *25*(4), 17-27.

Brigl, M., Gross-Selbeck, S., Dehnert, N., F., S., & Steffen, S. (2019). *After the Honeymoon Ends: Making Corporate-Startup Relationships Work (Boston Consulting Group)*. https://www.bcg.com/publications/2019/corporate-startup-relationships-work-after-honeymoon-ends. (Retrieved 27.02.2021)

Brynjolfsson, E., & Oh, J. (2012). The attention economy: Measuring the value of free digital services on the internet. In *Proceedings of the thirty third international conference on information systems*.

Bullen, C. V., & Rockart, J. F. (1981). *A primer on critical success factors*. Massachusetts.

Chan, V., Musso, C., & Shankar, V. (2008). Mckinsey global survey results: Assessing innovation metrics. *McKinsey Quarterly*, *45*(4), 1–11.

Chiesa, V., Frattini, F., Lamberti, L., & Noci, G. (2009). Exploring management control in radical innovation projects. *European Journal of Innovation Management*, *12*(4), 416-443. doi: 10.1108/14601060910996909

Chow, T., & Cao, D.-B. (2008). A survey study of critical success factors in agile software projects. *Journal of Systems and Software*, *81*(6), 961-971. doi: 10.1016/j.jss.2007.08.020

Clarke, A. (1999). A practical use of key success factors to improve the effectiveness of project management. *International Journal of Project Management*, *17*(3), 139-145. doi: 10.1016/S0263-7863(98)00031-3

Collins, J., & Smith, D. (1999). Innovation metrics: A framework to accelerate growth. *PRISM-CAMBRIDGE MASSACHUSETTS-*, 33–48.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.

Cruz-Cázares, C., Bayona-Sáez, C., & García-Marco, T. (2013). You can't manage right what you can't measure well: Technological innovation efficiency. *Research Policy*, *42*(6-7), 1239-1250. doi: 10.1016/j.respol.2013.03.012

Curtis, E., & Sweeney, B. (2016). Managing different types of innovation: mutually reinforcing management control systems and the generation of dynamic tension. *Accounting and Business Research*, *47*(3), 313-343. doi: 10.1080/00014788.2016.1255585

Davila, A., Foster, G., & Oyon, D. (2009). Accounting and control, entrepreneurship and innovation: Venturing into new research opportunities. *European Accounting Review*, *18*(2), 281-311. doi: 10.1080/09638180902731455

Dewangan, V., & Godse, M. (2014). Towards a holistic enterprise innovation performance measurement system. *Technovation*, *34*(9), 536-545. doi: 10.1016/j.technovation.2014.04.002

Doty, D. H., & Glick, W. H. (1994). Typologies as a unique form of theory building: Toward improved understanding and modeling. *Academy of Management Review*, *19*(2), 230-251. doi: 10.5465/AMR.1994.9410210748

Duncan, R. B. (1976). The ambidextrous organization: Designing dual structures for innovation. *The Management of Organization*, *1*(1), 167-188.

Edmondson, A. C., & McManus, S. E. (2007). Methodological fit in management field research. *Academy of Management Review*, *32*(4), 1155-1179. doi: 10.5465/AMR.2007.26586086

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, *14*(4), 532-550. doi: 10.2307/258557

etventure. (2021). *Was ist ein Digital Lab? Einblicke in das Thema Digital Lab - und wie Unternehmen einen kreativen Raum für neue digitale Geschäftsideen schaffen können.* Retrieved from https://www.etventure.de/digital-lab/

Fichman, R. G., Dos Santos, B. L., & Zheng, Z. (2014). Digital innovation as a fundamental and powerful concept in the information systems curriculum. *MIS quarterly*, *38*(2), 329-354. doi: 10.25300/MISQ/2014/38.2.01

Fitzgerald, L., Johnston, R., Brignall, T., Silvestro, R., & Voss, C. (1991). *Performance measurement in service businesses* (Vol. 69). London, United Kingdom: Chartered Institute of Management Accountants.

Flick, U. (2013). *The sage handbook of qualitative data analysis*. Los Angeles, USA: Sage Publications.

Freund, Y. P. (1988). Critical success factors. *Planning Review*, *16*(4), 20-23. doi: 10.1108/eb054225

Frey, J., Holotiuk, F., & Beimborn, D. (2020). Debating digital innovation: A literature review on realizing value from digital innovation. In *Proceedings of the 15th International Conference on Wirtschaftsinformatik* (p. 1-16). doi: 10.30844/wi_2020_h3-frey

Fuchs, C., Barthel, P., Herberg, I., Berger, M., & Hess, T. (2019). Characterizing approaches to digital transformation: Development of a taxonomy of digital units. In *Proceedings of the 14th International Conference on Wirtschaftsinformatik* (p. 632-646).

Gama, N., Da Silva, M. M., & Ataíde, J. (2007). Innovation scorecard: A balanced scorecard for measuring the value added by innovation. In *Digital Enterprise Technology* (p. 417-424). Boston, MA, USA: Springer. doi: 10.1007/978-0-387-49864-5_49

Garengo, P., Biazzo, S., & Bititci, U. S. (2005). Performance measurement systems in smes: A review for a research agenda. *International Journal of Management Reviews*, *7*(1), 25-47. doi: 10.1111/j.1468-2370.2005.00105.x

Gibson, C. B., & Birkinshaw, J. (2004). The antecedents, consequences, and

mediating role of organizational ambidexterity. *Academy of Management Journal*, *47*(2), 209-226. doi: 10.5465/20159573

Globerson, S. (1985). Issues in developing a performance criteria system for an organization. *International Journal of Production Research*, *23*(4), 639-646. doi: 10.1080/00207548508904734

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, *11*(3), 255-274. doi: 10.2307/1163620

Gubrium, J. F., & Holstein, J. A. (2002). From the individual interview to the interview society. In *Handbook of interview research: Context and method* (p. 3-32). Sage Publications. doi: 10.4135/9781412973588.n3

Hamilton, S., & Chervany, N. L. (1981). Evaluating information system effectiveness-part ii: Comparing evaluator viewpoints. *MIS Quarterly*, *5*(4), 79-86. doi: 10.2307/249329

Harris, L. R., & Brown, G. T. (2010). Mixing interview and questionnaire methods: Practical problems in aligning data. *Practical Assessment, Research, and Evaluation*, *15*. doi: 10.7275/959j-ky83

Haskamp, T., Lorson, A., de Paula, D., & Uebernickel, F. (2021). Bridging the gap - an analysis of requirements for performance measurement systems in digital innovation units. In *Proceedings of the 16th International Conference on Wirtschaftsinformatik*.

Henri, J.-F. (2006). Organizational culture and performance measurement systems. *Accounting, Organizations and Society*, *31*(1), 77-103. doi: 10.1016/j.aos.2004.10.003

Henttonen, K., Ojanen, V., & Puumalainen, K. (2015). Searching for appropriate performance measures for innovation and development projects. *R&D Management*, *46*(5), 914-927. doi: 10.1111/radm.12178

Holotiuk, F. (2020). The organizational design of digital innovation labs: Enabling ambidexterity to develop digital innovation. In *Proceedings of the 15th International Conference on Wirtschaftsinformatik*. doi: 10.30844/wi_2020_j6-holotiuk

Holotiuk, F., & Beimborn, D. (2017). Critical success factors of digital business strategy. In *Proceedings of the 13th International Conference on Wirtschaftsinformatik* (p. 991-1005).

Holotiuk, F., Beimborn, D., Hund, A., Wagner, H.-T., Frey, J., & Moormann, J. (2020). Digital Innovation Labs Report 2020. Frankfurt am Main, Germany.

Huang, J., Henfridsson, O., Liu, M. J., & Newell, S. (2017). Growing on steroids: Rapidly scaling the user base of digital ventures through digital innovaton. *MIS Quarterly*, *41*(1), 301-3014.

Huesig, S., & Endres, H. (2019). Exploring the digital innovation process: The role of functionality for the adoption of innovation management software by innovation managers. *European Journal of Innovation Management*, *22*(2), 302-314. doi: 10.1108/EJIM-02-2018-0051

Hund, A., Drechsler, K., & Reibenspiess, V. A. (2019). The current state and future opportunities of digital innovation: A literature review. In *Proceedings of the 27th European Conference on Information Systems (ECIS)*.

Kaplan, R. S., & Norton, D. (1992). The balanced scorecard - measures that drive performance. *Harvard Business Review*, *January-February*, 71-79.

Kaplan, R. S., & Norton, D. P. (1996). Linking the balanced scorecard to strategy. *California Management Review*, *39*(1), 53-79. doi: 10.2307/41165876

Keegan, D. P., Eiler, R. G., & Jones, C. R. (1989). Are your performance measures obsolete? *Strategic Finance*, *70*(12), 45.

Kerssens-van Drongelen, I. C., & Cooke, A. (1997). Design principles for the development of measurement systems for research and development processes. *R&D Management*, *27*(4), 345-357. doi: 10.1111/1467-9310.00070

Koen, P., Ajamian, G., Burkart, R., Clamen, A., Davidson, J., D'Amore, R., ... Wagner, K. (2001). Providing clarity and a common language to the "fuzzy front end". *Research-Technology Management*, *44*(2), 46-55. doi: 10.1080/08956308.2001.11671418

Kohli, R., & Melville, N. P. (2019). Digital innovation: A review and synthesis. *Information Systems Journal*, *29*(1), 200-223. doi: 10.1111/isj.12193

Kristiansen, J. N., & Ritala, P. (2018). Measuring radical innovation project success: typical metrics don't work. *Journal of Business Strategy*, *39*(4), 34-41. doi: 10.1108/JBS-09-2017-0137

Kuckartz, U. (2016). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (3rd ed.). Weinheim and Basel: Beltz Juventa.

Kyriasoglou, C. (2020). Ausgebremst - mit digitalen labs und viel risikokapital versuchen deutsche konzerne und mittelständler den weg in die zukunft zu finden. die zwischenbilanz ist niederschmetternd. *Manager Magazin*, *02*, 62-66.

Lawrie, G., & Cobbold, I. (2004). Third-generation balanced scorecard: evolution of an effective strategic control tool. *International Journal of Productivity and Performance Management*, *53*(7), 611-623. doi: 10.1108/17410400410561231

Lazzarotti, V., Manzini, R., & Mari, L. (2011). A model for r&d performance measurement. *International Journal of Production Economics*, *134*(1), 212-223. doi: 10.1016/j.ijpe.2011.06.018

Lee, J. H., Hsu, C., & Silva, L. (2020). What lies beneath: Unraveling the generative mechanisms of smart technology and service design. *Journal of the Association for Information Systems*, *21*(6). doi: 10.17705/1jais.00648

Leidecker, J. K., & Bruno, A. V. (1984). Identifying and using critical success factors. *Long Range Planning*, *17*(1), 23-32. doi: 10.1016/0024-6301(84)90163-8

Lill, P., Wald, A., & Munck, J. C. (2020). In the field of tension between creativity and efficiency: a systematic literature review of management control systems for innovation activities. *European Journal of Innovation Management*, *Vol. ahead-of-print*(No. ahead-of-print). doi: 10.1108/EJIM-11-2019-0329

Lune, H., & Berg, B. L. (2017). *Qualitative research methods for the social sciences* (9th ed.). Harlow, Essex, United Kingdom: Pearson.

Macmillan. (2021). *Key performance indicator.* Retrieved from https://www.macmillandictionary.com/dictionary/british/key-performance-indicator

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*(1), 71-87. doi: 10.1287/orsc.2.1.71

Martha, D., Sousa, V. D., & Mendes, I. A. C. (2007). An overview of research designs relevant to nursing: Part 3: Mixed and multiple methods. *Revista Latino-americana de Enfermagem*, *15*(5), 1046-1049. doi: 10.1590/S0104-11692007000500025

Mayer, S., Haskamp, T., & De Paula, D. (2021). Measuring what counts: An exploratory study about the key challenges of measuring design thinking activities in digital innovation units. In *Proceedings of the 54th hawaii international conference on system sciences* (p. 4951-4960). doi: 10.24251/HICSS.2021.602

Mayring, P. (2015). *Qualitative inhaltsanalyse: Grundlagen und techniken* (12. ed.). Weinheim.

Meyer, J.-U. (2020). *Warum Digital Labs überschätzt werden - Innovation in homöopathischer Dosis.* Retrieved from https://www.manager-magazin.de/unternehmen/artikel/digitallabore-werden-in-ihrer-wirkung-schlicht-ueberschaetzt-a-1304512.html (Retrieved 27.02.2021)

Micheli, P., & Manzoni, J.-F. (2010). Strategic performance measurement: Benefits, limitations and paradoxes. *Long Range Planning*, *43*(4), 465-476.

Micheli, P., Wilner, S. J., Bhatti, S. H., Mura, M., & Beverland, M. B. (2018). Doing design thinking: Conceptual review, synthesis, and research agenda. *Journal of Product Innovation Management*, *36*(2), 124-148. doi: 10.1111/jpim.12466

Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In *Handbook of mixed methods in social & behavioral research* (p. 189-208). Thousand Oaks, USA: Sage Publications.

Morse, J. M. (2010). Simultaneous and sequential qualitative mixed method designs. *Qualitative Inquiry*, *16*(6), 483-491. doi: 10.1177/1077800410364741

Morse, J. M., & Niehaus, L. (2009). *Mixed method design: Principles and procedures*. Walnut Creek, CA, USA: Left Coast Press Inc.

Morse, J. M., & Niehaus, L. (2016). *Mixed method design: Principles and procedures*. New York, USA: Routledge.

Nambisan, S., Lyytinen, K., Majchrzak, A., & Song, M. (2017). Digital innovation management: Reinventing innovation management research in a digital world. *Mis Quarterly*, *41*(1), 223-238. doi: 10.25300/MISQ/2017/41:1.03

Neely, A. (2005). The evolution of performance measurement research: De-

velopments in the last decade and a research agenda for the next. *International Journal of Operations & Production Management*, *25*(12), 1264-1277. doi: 10.1108/01443570510633648

Neely, A., Kennerley, M., & Adams, C. (2007). Performance measurement frameworks: a review. In A. Neely (Ed.), *Business performance measurement: Unifying theories and integrating practice* (2nd ed., p. 143-162). Cambridge, United Kingdom: Cambridge University Press.

Neely, A., Mills, J., Platts, K., Richards, H., Gregory, M., Bourne, M., & Kennerley, M. (2000). Performance measurement system design: developing and testing a process-based approach. *International Journal of Operations & Production Management*, *20*(10), 1119-1145. doi: 10.1108/01443570010343708

O'Reilly III, C. A., & Tushman, M. L. (2013). Organizational ambidexterity: Past, present, and future. *Academy of Management Perspectives*, *27*(4), 324-338. doi: 10.5465/amp.2013.0025

Oxford. (2020). *Key performance indicator.* Retrieved from https://www.lexico.com/definition/key_performance_indicator (Retrieved 03.03.2021)

Parmenter, D. (2015). *Key performance indicators: Developing, implementing, and using winning kpis*. New Jersey, USA: John Wiley & Sons.

Raabe, J.-P., Drews, P., Horlach, B., & Schirmer, I. (2021). Towards an intra- and interorganizational perspective: Objectives and areas of activity of digital innovation units. In *Proceedings of the 54th hawaii international conference on system sciences* (p. 5902). doi: 10.24251/HICSS.2021.714

Raabe, J.-P., Horlach, B., Drews, P., & Schirmer, I. (2020). Digital innovation units: Exploring types, linking mechanisms and evolution strategies in bimodal it setups. In *Proceedings of the 15th International Conference on Wirtschaftsinformatik (WI)*. doi: 10.30844/wi_2020_h5-raabe

Ramus, T., & Velten, C. (2016). *Digital Labs - Ideenturbo der digitalen Transformation oder Statussymbol für Vorstände? Managerkreis der Friedrich-Ebert-Stiftung.* Retrieved from https://library.fes.de/pdf-files/managerkreis/12913.pdf (Retrieved 26.03.2021)

Rockart, J. F. (1979). Chief executives define their own data needs. *Harvard Business Review*, *57*(2), 81-93.

Rowley, J. (2012). Conducting research interviews. *Management Research Review*, *35*(3/4), 260-271. doi: 10.1108/01409171211210154

Sahu, N., Deng, H., & Mollah, A. (2018). Investigating the critical success factors of digital transformation for improving customer experience. In *Proceedings of the international conference on information resources management (conf-irm)*. Association For Information Systems.

Said, A. A., HassabElnaby, H. R., & Wier, B. (2003). An empirical investigation of the performance consequences of nonfinancial measures. *Journal of Management Accounting Research*, *15*(1), 193-223. doi: 10.2308/jmar.2003.15.1.193

Sandelowski, M. (2000). Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in Nursing & Health*, *23*(3), 246-255.

Schermann, M., Wiesche, M., & Krcmar, H. (2012). The role of information systems in supporting exploitative and exploratory management control activities. *Journal of Management Accounting Research*, *24*(1), 31-59. doi: 10.2308/jmar-50240

Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *69*(7), 107-131. doi: 10.1007/s11577-017-0454-1

Segal, D. L., Coolidge, F. L., O'Riley, A., & Heinz, B. A. (2006). Structured and semistructured interviews. In *Clinician's handbook of adult behavioral assessment* (p. 121-144). Elsevier. doi: 10.1016/B978-012343013-7/50007-0

Sindemann, T., & von Buttlar, H. (2017). *Konzerne auf den Spuren von Startups - Wie etablierte Unternehmen Innovation Labs, Acceleratoren und Inkubatoren als Instrument der digitalen Transformation nutzen. Infront Consulting & Management GmbH and Capital.* Retrieved from https://www.infront-consulting.com/relaunch20/wp-content/uploads/2019/11/20170622-Infront-Capital-Studie_Digital-Innovation-Units_web.pdf (Retrieved 26.03.2021)

Sindemann, T., & von Buttlar, H. (2018). *Konzerne auf den Spuren von Startups 2018 - Wie etablierte Unternehmen Acceleratoren, Company Builder, und Innovation Labs als Instrumente der digitalen Transformation nutzen. Infront Consulting & Management GmbH and Capital.* Retrieved from https://www.infront-consulting.com/relaunch20/wp-content/uploads/2019/11/20180620_Online_final_komp.pdf (Retrieved 26.03.2021)

Sindemann, T., & von Buttlar, H. (2019). *Konzerne auf den Spuren von Startups 2019 - Der steinige Weg zu substantieller digitaler Innovation. Infront Consulting & Management GmbH and Capital.* Retrieved from https://www.infront-consulting.com/relaunch20/wp-content/uploads/2019/12/Infront_Consulting_Studie-Digital-Innovation-Units_2019.pdf (Retrieved 26.03.2021)

Sindemann, T., & von Buttlar, H. (2020). *Konzerne auf den Spuren von Startups 2020 - Vom Ideation Sprint zum Scaling Marathon. Infront Consulting & Management GmbH and Capital.* Retrieved from https://www.infront-consulting.com/relaunch20/wp-content/uploads/2020/06/Konzerne-auf-den-Spuren-von-Startups-2020.pdf (Retrieved 26.03.2021)

Sproull, N. L. (2002). *Handbook of research methods: A guide for practitioners and students in the social sciences*. Scarecrow press. doi: 10.5860/choice.26-1924

Strike, V. M., & Rerup, C. (2016). Mediated sensemaking. *Academy of Management Journal*, *59*(3), 880-905. doi: 10.5465/amj.2012.0665

Suomala, P. (2004). The life cycle dimension of new product development performance measurement. *International Journal of Innovation Management*, *8*(02), 193-221. doi: 10.1142/S1363919604001039

Thompson, N., Bonnet, D., & Jaballah, S. (2020). *Lifting the Lid on Corporate Innovation in the Digital Age. Capgemini Invent.* Retrieved from https://www.capgemini.com/wp-content/uploads/2020/05/MIT-INVENT-Report_NEW-2020.pdf (Retrieved 26.03.2021)

Tushman, M. L., & O'Reilly III, C. A. (1996). Ambidextrous organizations: Managing evolutionary and revolutionary change. *California Management Review*, *38*(4), 8-29. doi: 10.2307/41165852

Twin, A., & James, M. (2020). *Key Performance Indicators (KPIs). Investopedia.* Retrieved from https://www.investopedia.com/terms/k/kpi.asp (Retrieved 03.03.2021)

Urbach, N., Ahlemann, F., Böhmann, T., Drews, P., Brenner, W., Schaudel, F., & Schütte, R. (2018). The impact of digitalization on the it department. *Business & Information Systems Engineering*, *61*, 123-131. doi: 10.1007/s12599-018-0570-0

Wiesche, M., Bodner, J., & Schermann, M. (2012). Antecedents of it-enabled organizational control mechanisms. In *Proceedings of the 20th European Conference on Information Systems (ECIS)*.

Wu, M.-S., Huang, S.-J., & Chen, L.-W. (2011). The preparedness of critical success factors of it service management and its effect on performance. *The Service Industries Journal*, *31*(8), 1219-1235. doi: 10.1080/02642060903437014

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Thousand Oaks, USA: Sage Publications.

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Thousand Oaks, USA: Sage Publication.

Ylinen, M., & Gullkvist, B. (2014). The effects of organic and mechanistic control in exploratory and exploitative innovations. *Management Accounting Research*, *25*(1), 93-112. doi: 10.1016/j.mar.2013.05.001

Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). Research commentary—the new organizing logic of digital innovation: an agenda for information systems research. *Information systems research*, *21*(4), 724-735. doi: 10.1287/isre.1100.0322

Zhou, T. (2011). Examining the critical success factors of mobile website adoption. *Online Information Review*, *35*(4). doi: 10.1108/14684521111161972

Junior Management Science 7(4) (2022) 1133-1146

Junior Management Science

journal homepage: www.jums.academy

The role of self-enhancement when using status incentives in sales

Die Rolle der Selbstaufwertung beim Einsatz von Statusanreizen im Vertrieb

Niclas Vogt

Ruhr-Universität Bochum

Abstract

The aim of this study is to apply the social psychological theory of self-enhancement to the widespread use of status incentives in the corporate world. More specifically, the potential effect of self-enhancement triggered by the awarding of status incentives is examined based on a fictitious sales team. Special focus is placed on the resulting motivational effect. The research showed that status changes through status incentives (status manipulation) have a very strong effect on employees' motivational performance, because self-enhancement is influenced. The primary finding was that these status changes of individuals within a team may have ambivalent effects on overall motivational performance, partly because in team dynamics an increase in status of individual members is accompanied by the relative decrease in status of the other members. Among other effects, this changes the structure of social comparisons (self-enhancement theory) within the sales team, which possess strong motivational effects.

Zusammenfassung

Ziel dieser Untersuchung ist es, die sozialpsychologische Theorie der Selbstaufwertung auf die in der Unternehmenswelt verbreitete Nutzung von Statusanreizen anzuwenden. Genauer wird sich anhand eines fiktiven Sales-Teams die potenzielle Wirkung der Selbstaufwertung angeschaut, die durch die Vergabe von Statusanreizen ausgelöst wird. Spezieller Fokus wird auf die resultierende Motivationswirkung gelegt. Die Untersuchung hat gezeigt, dass Statusveränderungen durch Statusanreize (Statusmanipulation) eine sehr starke Wirkung auf die Motivationsleistung von Mitarbeitern haben, weil die Selbstaufwertung beeinflusst wird. Die primäre Erkenntnis war, dass diese Statusveränderungen von Personen innerhalb eines Teams ambivalente Effekte auf die Gesamtmotivationsleistung haben können, unter anderem, da in Teamdynamiken eine Statuserhöhung einzelner Mitglieder mit der relativen Verringerung des Status der anderen Mitglieder einhergeht. Neben anderen Effekten ändert dies die Struktur der sozialen Vergleiche (Selbstaufwertungstheorie) innerhalb des Sales-Teams, die starke Motivationswirkungen besitzen.

Keywords: Selbstaufwertung; Statusanreize; Vertrieb; Motivation.

1. Einleitung

1.1. Erklärung der Forschungsfrage

„Effektive Sales-Team-Anreizsysteme sind starke Erfolgsfaktoren für Firmen und stellen eine große unternehmerische Herausforderung dar“ (Bommaraju & Hohenberg, 2018, S. 106).

Die vorliegende Arbeit leistet einen Beitrag zu der wissenschaftlichen Untersuchung der Themen Statusanreize und Selbstaufwertung. Einen großen Teil von unternehmerischen Anreizmethoden machen die Statusanreize aus. Dass der Status einen hohen Stellenwert für viele Menschen hat, ist unumstritten. Diese Tatsache nutzen Unternehmen aus, indem sie den Status als Motivationsfaktor einsetzen. Die Statusanreize werden nachfolgend vor dem Hintergrund der Selbstaufwertung, einem sozialpsychologischen Konzept, analysiert.

Die Selbstaufwertung behandelt im weiteren Sinne die Motivationsmechanismen von Personen. In der nachfolgenden Untersuchung wird der Einfluss der Statusanreize auf die Motivation und die daraus resultierende Vertriebsleistung der Mitarbeiter dargestellt. Hierzu wird das Konzept der Selbstaufwertung als ein wichtiger Faktor in diesem Prozess hinzugezogen.

DOI: https://doi.org/10.5282/jums/v7i4pp1133-1146

Es geht somit um die Persönlichkeit und Leistungsfähigkeit von Menschen. Genauer werden sich vor allem die Personen aus dem Vertriebsbereich angeschaut. Da der Vertrieb ein stark kompetitives Feld ist, spielt der individuelle Status in der Betrachtung womöglich eine größere Rolle als in anderen Berufsfeldern, was dem Thema die Universalität auch im Hinblick auf andere Arbeitsbereiche absprechen kann.

Aus dem untersuchten Thema ergeben sich folgende Forschungsfragen:

- Welche Rolle spielt die Selbstaufwertung beim Einsatz von Statusanreizen im Vertrieb?
- Wie können die Unternehmen gezielt das Konzept der Selbstaufwertung in Sales-Teams nutzen?

1.2. Wissenschaftliche Relevanz

Die Forschung in den Bereichen der Selbstaufwertung und vor allem im Bereich der Statusanreize sind sehr aktuell. Eine logische Konsequenz, wenn man betrachtet, dass es Unternehmen einen enormen Vorteil gibt, motivierte Mitarbeiter zu haben. Dies ist noch ungeachtet von den ökonomischen Vorteilen, die es mit sich bringt, eine Alternative zu finanziellen Anreizen zu haben. Durch gegebene Statusanreize werden nachweislich finanzielle Anreize im Sinne der Kompensation ausgeglichen (vgl. Besley & Ghatak, 2008, S. 206). Dies bedeutet, dass manche Arbeitnehmer in der gängigen Praxis auf Geld verzichten, um einen erhöhten Status zu erlangen. Dieses Phänomen wird sich in Abschnitt 2.2.2 genauer angeschaut.

Es gibt ganze Branchen, in denen Status als Anreiz stärker gewichtet ist als der finanzielle Anreiz. Als Beispiel kann man sich das Militär mit seinem ausgeprägten Abzeichen und Rangsystemen oder die akademische Welt mit ihren vielen Titeln, Stipendien und Preisen anschauen (vgl. Besley & Ghatak, 2008, S. 206). Somit gibt es berechtigte Gründe für die Annahme, dass Personen auf finanzielle Mittel aufgrund von erhöhtem Status verzichten. Eine mögliche Erklärung für dieses Phänomen liefert die Selbstaufwertungstheorie, zu der in dieser Arbeit ein wissenschaftlicher Bezug hergestellt wird.

Auf den Vertrieb bezogen lässt sich feststellen, dass gerade in diesem Bereich der Status eine große Rolle spielt. Somit lässt sich die wissenschaftliche Relevanz aus der Verbindung der beiden Themen/Theorien in einem vertrieblichen Kontext ableiten. Schließlich sind die Forschungsfelder Psychologie und Management maßgeblich zur Beantwortung vertrieblicher Fragestellungen.

1.3. Ziele der Forschung

Im Grundlegenden ist festzuhalten, dass die Forschung auf dem Gebiet der Statusanreize nicht notwendigerweise einen direkten Bezug zu dem Thema der Selbstaufwertungstheorie aufweist. Auch in der aktuellen Forschung werden die beiden Begriffe selten in einem Zusammenhang dargestellt.

Somit besteht das Grundziel der Arbeit in der Beleuchtung der interdisziplinären Schnittstellen dieser beiden Forschungsfelder, Statusanreize verknüpft mit der Selbstaufwertungstheorie. Des Weiteren gilt es einen Vertriebsbezug herzustellen, der sich häufig nicht aus den Begriffen ergibt.

Als konkretes Ziel wird hier die Beantwortung der Forschungsfrage, „Welche Rolle spielt die Selbstaufwertung beim Einsatz von Statusanreizen im Vertrieb?" angestrebt. Des Weiteren soll beantwortet werden, wie Unternehmen genau dieses psychologische Konzept der Selbstaufwertung nutzen können, um Mitarbeiter durch den Anreiz eines höheren Status zu motivieren.

1.4. Gang der Untersuchung

Die Untersuchung startet mit einem theoretischen Teil, der einen Rahmenbezug und erste Theorien zu der Thematik der Selbstaufwertung und dem Thema der Statusanreize geben soll.

Dieses Verständnis ist grundlegend für die nachfolgende Darstellung der Literatur, in dem der aktuelle Stand der Forschung zusammengefasst wird. Anschließend werden die Forschungslücken herausgearbeitet, indem die Überschneidungen der Themenfelder und auch fehlende Überschneidungen dargestellt werden. Nach dem Literatur-Review folgt das Analysekapitel, in dem die zentralen Forschungsergebnisse dargestellt und diskutiert werden. Der Fokus dieses Kapitels liegt darauf, die Wirkung der Statusanreize vor dem Hintergrund der Selbstaufwertung in Sales-Teams darzustellen. Dieses Kapitel endet in der Beantwortung der Forschungsfragen.

In der abschließenden Betrachtung geht es darum, ein Fazit zu ziehen und die Limitationen der Forschung, des Themas und auch dieser Arbeit darzustellen. Nachfolgend werden noch Implikationen für die Forschung und Praxis gegeben, indem Ideen und Vorschläge an Unternehmen gerichtet werden. Als letzter Schritt wird noch ein Ausblick auf die weitere Forschung gegeben. Es wird dargelegt, welche nächsten Schritte nach dieser Arbeit folgen könnten und welche verwandten Themenbereiche interessante Kenntnisse für die weitere Forschung liefern.

2. Theoretische Grundlagen

2.1. Selbstaufwertung

2.1.1. Definition

Zunächst wird der Begriff der Selbstaufwertung für die nachfolgende Verwendung definiert. Durch die hohe Komplexität des Begriffes muss hier eine zweckdienliche Definition genutzt werden.

Selbstaufwertung ist ein Begriff, der in der Sozialpsychologie zu verorten ist. Er bezeichnet grundsätzlich den Drang das Selbst zu bestätigen (vgl. Steele, 1988, S. 297). Es handelt sich um eine Reihe psychologischer Phänomene, bei denen sich die Person tendenziell positiv einschätzt, ohne es zwangsläufig zu sein (vgl. Sedikides & Gregg, 2008, S. 102). Durch diese Einschätzung will das Individuum nicht nur sich

selbst, sondern auch nahestehende Personen davon überzeugen, verdienstlich und lobenswert zu sein (vgl. Sedikides & Gregg, 2007, S. 111). Somit beschreibt die Selbstaufwertung eine verzerrte Perspektive der Selbsteinschätzung.

Ob die Selbstaufwertung vorteilhaft für das Individuum ist, wird von Wissenschaftlern und Forschern kritisch gesehen. Grund dafür sind die unterschiedlichen Ausprägungen und Manifestationen der Selbstaufwertung. Um das Konzept der Selbstaufwertung noch genauer zu beleuchten, bietet es sich an, grundlegende Theorien der Selbstaufwertung anzuschauen.

2.1.2. Bipolare Dimensionen nach Sedikides & Gregg

Die Selbstaufwertung kann als bipolare Dimensionen beschrieben werden. Diese Theorie stammt von den führenden Forschern Sedikides & Gregg. Sie beschreibt vier fundamentale Felder, in denen die Selbstaufwertung auftritt.

Selbstförderung oder Selbstschutz

Eine selbstaufwertende Person tut dies aus dem Beweggrund der Selbstförderung oder des Selbstschutzes (vgl. Sedikides & Gregg, 2008, S. 103).

Das bedeutet, dass das Individuum entweder ein positives Ereignis verstärkt oder ein negatives Ereignis kleinredet (vgl. Arkin, 1981, zitiert nach Sedikides & Gregg, 2008, S. 103). Die empirische Forschung deutet darauf hin, dass der Selbstschutz eine stärkere Priorität als die Selbstförderung der Individuen besitzt, da negative Ereignisse sich stärker auswirken. Um es mit den Worten Baumeisters zu sagen, „bad is stronger than good“ (vgl. Baumeister, Bratslavsky, Finkenauer & Vohs, 2001, S. 323).

Diese Unterscheidung ist einer der fundamentalsten und wichtigsten der Selbstaufwertung, da es das zentrale Motiv der selbstaufwertenden Person darstellt.

Öffentliche oder private Selbstaufwertung

Individuen können sich entweder privat oder öffentlich selbst aufwerten (vgl. Sedikides & Gregg, 2008, S. 104). Dies beschreibt nicht das jeweilige soziale Umfeld, sondern die Gegenüberstellung von nach außen getragener Selbstpräsentation und einer rein intrapsychischen Angelegenheit (vgl. Leary & Kowalski, 1990, S. 322). Die Relevanz dieser Unterscheidung ergibt sich aus der Tatsache, dass die Anwesenheit anderer Personen nicht unwesentlich für die Auswirkung der Selbstaufwertung ist. Die Anwesenheit anderer kann sowohl die Selbstaufwertung verstärken (vgl. Leary, Tchividijian & Kraxberger, 1994, S. 461) als auch die Selbstaufwertung hemmen (vgl. Sedikides, Campbell, Reeder & Elliot, 1998, S. 378).

Selbstaufwertung in wichtigen oder unwichtigen Gebieten

Selbstaufwertung findet meistens in den Gebieten statt, die für das jeweilige Individuum wichtig sind (vgl. Crocker & Knight, 2005, S. 200; Sedikides, Gaertner & Toguchi, 2003, S. 71). Grundsätzlich unterscheiden sich Personen im Hinblick auf die für sie wichtigen Interessenfelder (vgl. Crocker & Wolfe, 2001, S. 598). Somit lassen sich nicht für alle Menschen universell konkrete Interessenfelder vorgeben, in denen die Selbstaufwertung eine wichtige Rolle spielt.

Jedoch lässt sich vermuten, dass es besonders in Gebieten verstärkt stattfindet, in die Personen viel Zeit investieren, wie beispielsweise die eigene Karriere. Des Weiteren lässt es die Annahme zu, dass dies gerade auch in Sales-Teams ein großes Thema sein könnte, da der Vertrieb gemeinhin ein kompetitives Gebiet darstellt.

Aufrichtige oder taktische Selbstaufwertung

Die Selbstaufwertung kann entweder aufrichtig oder taktisch sein (vgl. Sedikides & Strube, 1997, S. 225). Bei der aufrichtigen Selbstaufwertung nutzt eine Person eine Situation, um sich offen und unmittelbar aufzuwerten.

Bei der taktischen Variante verzichtet die Person zunächst auf die Selbstaufwertung, um eine Selbstverbesserung zu ermöglichen. Beispielsweise könnte die Person versuchen, nach wahren Informationen über sich selbst zu suchen (vgl. Trope, 1986, zitiert nach Sedikides & Gregg, 2008, S. 104) anstatt nach selbstbestätigenden Informationen, um in der Zukunft eine gerechtfertigte und dauerhaftere Selbstaufwertung zu gewährleisten (vgl. Pyszczynski, Greenberg & Goldenberg, 2003, zitiert nach Sedikides & Gregg, 2008, S. 104).

Aufrichtige Selbstaufwertung wird in dem sozialen Umfeld oft verurteilt, da ein nach außen getragenes stark positives Selbstbild oft als Arroganz und Überheblichkeit angesehen wird (vgl. Sedikides & Gregg, 2007, S. 165). Der Begriff Narzissmus kann als überzogene Präferenz für aufrichtige Selbstaufwertung gesehen werden (vgl. Vazire & Funder, 2006, S. 154).

Diese bipolaren Dimensionen stellen eine gute Grundlage für die Betrachtung von Vertriebsthemen dar, weil sie sich natürlicherweise auf vertriebliche Themen anwenden lassen. Für das weitere Verständnis der Selbstaufwertungstheorie werden sich nun klassische Ausprägungen der Selbstaufwertung angeschaut.

2.1.3. Klassische Ausprägungen

Die Selbstaufwertung hat sehr viele Ausprägungen und Manifestationen (vgl. Sedikides & Gregg, 2008, S. 104). Vor diesem Hintergrund ist es wichtig zu erwähnen, dass in dieser Arbeit die jeweiligen Theorien nur kurz dargestellt werden können und auch nicht zweckdienliche Theorien ausgelassen werden.

Zu den klassischen und grundlegenden Ausprägungen der Selbstaufwertung gehört die Selbstaufwertungs-Triade. Zu dieser Triade gehören die Manifestationen:

- Überdurchschnittseffekt (Above-Average-Effekt)
- Illusionen von Kontrolle
- Unrealistischer Optimismus

Überdurchschnittseffekt

Der Überdurchschnittseffekt beschreibt die Neigung von Menschen, die eigenen Fähigkeiten als überdurchschnittlich einzuschätzen (vgl. Sedikides & Gregg, 2007, S. 111). Dieser Effekt hängt stark mit dem bekannten Dunning-Kruger-Effekt zusammen, der inkompetenten Personen ein verzerrtes Selbstbild attestiert, die eigenen Fähigkeiten zu überschätzen (vgl. Kruger & Dunning, 1999, S. 1121). Eine aktuelle Studie bestätigt diesen Effekt und findet Evidenz dafür,

dass leistungsschwache Personen sich schlechter einschätzen können als leistungsstarke Personen (vgl. Jansen, Rafferty & Griffiths, 2021, S. 125).

Illusionen von Kontrolle

Personen überschätzen den Grad an Kontrolle, den sie über Ergebnisse und Eventualitäten haben. Der Begriff Illusion von Kontrolle beschreibt die Überzeugung der Personen, Einfluss auf von Natur aus zufälligen Systeme nehmen zu können. Beispielsweise glauben Personen (meist unterbewusst), dass sie das Ergebnis eines Würfelwurfs oder das Ziehen einer bestimmten Karte vorhersagen können (vgl. Sedikides & Gregg, 2007, S. 111).

Im vertrieblichen Kontext könnte das bedeuten, dass obwohl Vertriebsmitarbeitern oft suggeriert wird, jeden potenziellen Kunden überzeugen zu können, dies nicht immer möglich ist. Ein Vertragsabschluss liegt nicht immer innerhalb der Kontrolle des Vertriebsmitarbeiters, unabhängig von seiner Leistung. Dies könnte in einer fälschlichen Versagenszuweisung des Vorgesetzten oder der eigenen Person enden.

Unrealistischer Optimismus

Viele Personen glauben, dass das Schicksal es gut mit ihnen meint (vgl. Sedikides & Gregg, 2007, S. 111; Smits & Hoorens, 2005, S. 84). Unter anderem meinen sie, dass sie mehr positive Lebenserfahrungen machen werden und weniger negative Erfahrungen als ähnliche andere Personen (vgl. Helweg-Larsen & Shepperd, 2001, S. 74; Weinstein, 1980, S. 807; Weinstein & Klein, 1995, S. 132). Als Beispiel unterschätzen Raucher ihr eigenes Risiko, an Krebs zu erkranken, sowohl gegenüber Nichtrauchern als auch gegenüber anderen Rauchern (vgl. Weinstein, Marcus & Moser, 2005, S. 74).

Der unrealistische Optimismus könnte insofern im Vertriebskontext interessant sein, als das junge Berufseinsteiger (auch in Sales) verhältnismäßig wenig verdienen und sie häufig durch Statusanreize, wie z.B. Beförderungen, ihre primäre Arbeitsmotivation ziehen (vgl. Auriol & Renault, 2008, S. 307). Hier wird das Phänomen des unrealistischen Optimismus ausgenutzt, denn nicht jeder kann beliebig hoch aufsteigen, doch glauben die meisten Personen, dass sie es können.

Es folgt eine Darstellung der Strategien, die Individuen anwenden, um sich persönlich aufzuwerten.

2.1.4. Strategische soziale Vergleiche

In dem Prozess der Eigenbewertung kommt das Individuum nicht an sozialen Vergleichen vorbei. Soziale Vergleiche finden dann statt, wenn sich Individuen unsicher über ihre eigene Meinung oder Fähigkeiten sind. Sie werden sich in diesen Situationen zum Zweck der Selbstevaluation mit anderen Personen vergleichen (vgl. Buunk, Collins, Taylor, Yperen & Dakof, 1990, S. 1238). Diese Strategien sind besonders wichtig in der späteren Betrachtung von Sales-Team-Strukturen, da sie grundlegend die Vergleiche innerhalb des sozialen Umfelds analysieren. Dies lässt sich somit sehr treffend auf Sales-Team-Strukturen anwenden.

Es gibt 3 zentrale Vergleichsmöglichkeiten mit anderen Individuen: aufwärtsgerichtete-, abwärtsgerichtete- und lateralgerichtete Vergleiche.

Diese Vergleichsmöglichkeiten gilt es direkt mit dem Motiv nach Status in Verbindung zu setzen, denn Begriffe wie „aufwärts“ und „abwärts“ beziehen sich sowohl auf ein subjektives Gefühl der Über- bzw. Unterlegenheit als auch auf ein objektives Maß einer ordinalen Rangordnung mit dem Status als Kriterium.

Ein *aufwärtsgerichteter* sozialer Vergleich beschreibt einen Vergleich mit einer anderen Person, die von dem Selbst als überlegen angesehen wird. Ein aufwärtsgerichteter Vergleich mit einer solchen Person, kann durch das *Angleichen* des Selbst und deren Eigenschaften zu einer Selbstaufwertung führen (vgl. Collins, 1996, S. 52). Es gibt jedoch Restriktionen für ein solches Angleichen des Selbst. Demnach sollte bspw. der Unterschied zwischen dem Selbst und der Vergleichsperson nicht zu groß sein (vgl. Wheeler, 1966, S. 30) und die vergleichbare Fähigkeit oder der vergleichbare Erfolg müssten erreichbar sein (vgl. Lockwood & Kunda, 1997, S. 91). Ein erfolgreiches Angleichen schafft ein Gefühl der gemeinsamen Überlegenheit.

Sind die Restriktionen nicht erfüllt, kann der Effekt des Vergleiches unterschiedlich ausfallen. Wenn das zuvor genannte Angleichen somit nicht die Folge des aufwärtsgerichteten Vergleichs ist, kann er entweder in einer Inspiration oder in einer Resignation enden. Die Inspiration wäre motivationsförderlich und die Resignation als Unerreichbarkeitssignal motivationshinderlich (vgl. Collins, 2000, zitiert nach Sedikides & Gregg, 2008, S. 106).

Bei *abwärtsgerichteten* sozialen Vergleichen führen Individuen Vergleiche nach „unten“ durch. Das vergleichende Individuum wird als untergeordnet und weniger qualifiziert als das Selbst angesehen. Diese Vergleiche dienen der Ego-Verteidigung, indem ein Überlegenheitsgefühl kreiert wird, bei der das Ego und das Selbstwertgefühl verstärkt werden (vgl. Biernart & Billings, 2002, S. 262).

Lateralgerichtete Vergleiche sind Vergleiche mit Individuen, die dem Selbst als gleichgestellt wahrgenommen werden. Auch diese können der Selbstaufwertung dienlich sein. Vergleiche mit Mitgliedern derselben Gruppe können schützend das Selbstwertgefühl bewahren, wenn die ganze Gruppe als benachteiligt gilt (vgl. S. E. Taylor, Neter & Wayment, 1995, S. 1282).

Das Motiv der Selbstaufwertung ist oft ausschlaggebend dafür, welche der Vergleiche das Individuum primär anstrebt. Somit kann die subjektive Ausnutzung in der Wahl des Motivs dazu führen, dass in dem Vergleich zwischen dem Individuum und einer anderen Person ein tendenziell positives Ergebnis für das Individuum kreiert wird. (vgl. Buunk et al., 1990, S. 1247). **Das Individuum sucht sich seinen Vergleich.**

Die Motivationswirkung dieser einzelnen Vergleiche muss sich im Späteren genauer angeschaut werden, gerade vor dem Hintergrund der Managemententscheidungen. In dem vertrieblichen Kontext, besonders bei Sales-Teams, sind diese Vergleiche deshalb wichtig, weil gerade in diesem Umfeld der Status eine wichtige Rolle spielt. Dies wird in der späteren Analyse unter 4.2 deutlich.

Durch Statusanreize können Unternehmen künstlich

den Status der einzelnen Individuen regulieren und somit das Phänomen der Selbstaufwertung nutzen. Dieser Vorgang wird auch Statusmanipulation genannt (vgl. Auriol & Renault, 2008, S. 308). Im Nachfolgenden werden die theoretischen Grundlagen der Statusanreize dargestellt.

2.2. Statusanreize

2.2.1. Definition

Statusanreize gehören zu den in der Unternehmenspraxis genutzten Anreizsystemen. Sie stellen ein grundlegend anreizbasiertes Führungsinstrument der Unternehmen dar, welches keinen finanziellen Fokus hat (vgl. Bhattacharya & Dugar, 2012, S. 550). Das grenzt die Statusanreize von den primär finanziellen Anreizen, wie z.B. einer Gehaltserhöhung ab.

Statusanreize setzen viel mehr an psychologischen Faktoren, wie der sozialen Anerkennung und Status an (vgl. Ellingsen & Johannesson, 2007, S. 135). Es ist zum Beispiel üblich in Unternehmen den besten Verkäufern Statuspreise, wie Medaillen, Ringe oder „Bester Verkäufer des Jahres"-Schilder in Aussicht zu stellen, um ihre Motivation zu erhöhen (vgl. Nelson, 1994; Podmoroff, 2005, zitiert nach Bhattacharya & Dugar, 2012, S. 549).

Diese Statusanreize sind Beeinflussungsstrategien der Unternehmen, um Mitarbeiter in ihrer relativen Rangordnung zu erhöhen (vgl. Bhattacharya & Dugar, 2012, S. 549). Die Relativität besteht darin, dass die Erhöhung des Status einer Person zwangsläufig die Erniedrigung des Status einer anderen Person zur Folge hat (vgl. Auriol & Renault, 2008, S. 305).

Des Weiteren sind Statusanreize eine begrenzte Ressource, auch positionelles Gut genannt (vgl. Besley & Ghatak, 2008, S. 206). Das bedeutet, dass der Wert eines solchen Anreizes davon abhängt, wie oft er vergeben wird.

2.2.2. Finanzielle und nicht-finanzielle Anreize

Es gibt im Wesentlichen zwei relevante Anreizsysteme für die Vergütung von Arbeitnehmern. Das sind die finanziellen und nichtfinanziellen Anreize. Finanzielle Anreizsysteme sind im klassischen Sinne Gehälter und jeweilige Boni. Nichtfinanzielle Anreize sind alle Anreize, von denen die Person einen nicht primär finanziellen Nutzen hat (vgl. Bhattacharya & Dugar, 2012, S. 550). Somit wäre eine Medaille oder ein Schild, obgleich der Kosten der Anschaffung für das Unternehmen, den nichtfinanziellen Anreizen zuzuordnen.

Die zuvor definierten Statusanreize sind somit klar den nichtfinanziellen Anreizen zuzuordnen, da ihr primärer Fokus auf dem Status liegt und nicht auf finanziellen Zahlungen. Etwaige Überschneidungen können in der weiteren Betrachtung vernachlässigt werden. Ein Beispiel dafür wäre ein Firmenwagen, der sowohl eine finanzielle Zahlung als auch ein Zeichen und Anreiz für Status darstellt.

Nichtfinanzielle Anreize können neben der Motivationswirkung auch deshalb für Unternehmen interessant sein, weil sie eine Kosteneinsparung, vor allem in Personalkosten, darstellen können (vgl. Besley & Ghatak, 2008, S. 206). Dies lässt sich auf folgende einfache Erkenntnis zurückführen: Das Streben nach Status ist in Menschen genauso vorhanden, wie das Streben nach Geld. Somit kann der finanzielle Anreiz durch Statusanreize kompensiert werden. Dies bedeutet in der Praxis, dass Arbeitnehmer auf Geld verzichten, um einen erhöhten Status zu erlangen (vgl. Besley & Ghatak, 2008, S. 206).

3. Stand der Forschung (Literature Review)

3.1. Literaturüberblick

Nachdem nun die relevanten theoretischen Grundlagen für die Arbeit gelegt wurden, gilt es, die verwendete Literatur für die nachfolgende Analyse kurz darzustellen und vorzubereiten.

Hierzu wird die Literatur in Form von Clustern dargestellt, um die einzelnen Gebiete des Themas darzulegen und somit thematische Überschneidungen, Ergänzungen und fehlende Zusammenhänge, also die Forschungslücken, herauszuarbeiten.

Diese Darstellung stellt nicht alle Felder der jeweiligen Themen dar, sondern die Felder, die für die Beantwortung der Forschungsfrage in der Arbeit genutzt wurden. Im Generellen gilt zu sagen, dass die grundlegenden Bereiche der Arbeit aus der Sozialpsychologie, dem Management bzw. der Business-Forschung und dem Marketing stammen. Vereinzelt werden weitere Bereiche, wie z.B. volkswirtschaftliche Modelle tangiert.

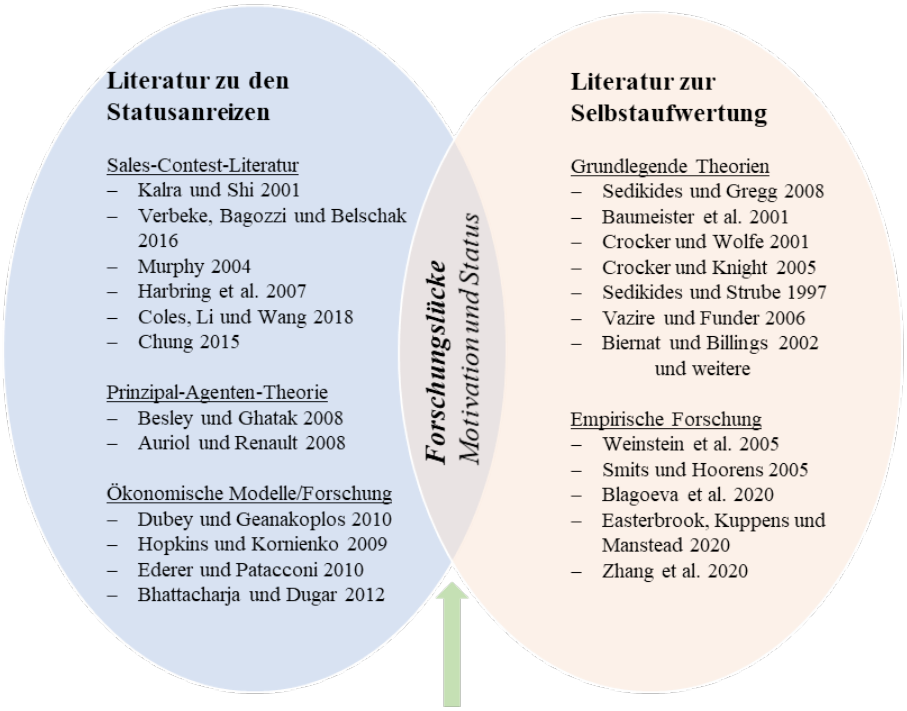
Motivation von Angestellten

Statusanreize werden zur Motivation im vertrieblichen Kontext genutzt. Daher wurde Literatur zur Motivation von Mitarbeitern verwendet. Dies ist ein selbst gewählter Sammelbegriff, der unter anderem Literatur zur Jobzufriedenheit, Literatur zu Motivation, Literatur zu organisatorischem Verhalten und weiterem bündelt. Grund dafür ist, dass diese Literatur sehr viele Schnittstellen und Überschneidungen aufweist. Außerdem verbindet sie die Themen Statusanreize und Selbstaufwertung, weil sie sich mit Motivation und Status beschäftigt, an die beide Themen anknüpfen.

Ein Kernthema dieser Literatur beleuchtet die Möglichkeiten von Personen, auf andere Personen zu wirken, um sie zu motivieren. Der hier verwendete Rahmen bezieht sich auf die Beziehung zwischen Manager und Angestelltem. Zum Beispiel identifizieren Srivastava, Locke & Bartol, 2001 mehrere Faktoren organisatorischer Belohnungssysteme, die die Motivation der Mitarbeiter beeinflussen, das angestrebte Verhalten auszuführen (vgl. Srivastava et al., 2001, S. 959).

Sales Contest-Literatur (Verkaufswettbewerb)

Die Sales-Contest-Literatur stellt eine grundlegende Säule dieser Arbeit dar. Sie ist ein spezifischer Teil der Literatur zur Motivation von Angestellten und deshalb für die Arbeit von Bedeutung, weil sie sich im Kern mit dem Effekt von **Anreizen im Vertrieb** beschäftigt. Hierbei liegt der natürliche Fokus der Literatur häufig auf dem Status beziehungsweise den Statusanreizen.



Forschungslücke: Thematisch verwandte Literatur

Literatur zu subjektivem Wohlbefinden	Literatur zur Motivation von Mitarbeitern
Charles, Hurst und Roussanov 2009	Srivastava, Locke und Bartol 2001
Heffetz 2011	Postlewaite 1988
Huberman und Minns 2007	Knight 1984
Ergänzende Literatur	Hampton 1970
Taylor, Strutton und Thompson 2012	Wildt, Parker und Harris Jr. 1981
Centers und Bugental 1966	Prundeanu, Constantin und Popuşoi 2021
Ellingson und Johanneson 2007	Moncrief, Hart und Robertson 1988

Abbildung 1: Literaturübersicht

Quelle: Eigene Darstellung

Der Sales-Contest ist ein Programm zur Motivation von Vertriebsmitarbeitern, bei dem ihnen Belohnungen angeboten werden, die auf ihren Verkäufen und Ergebnissen basieren. Im Gegensatz zu anderen Kompensationen von Mitarbeitern wird bei Sales-Contests die relative Performance bezüglich des Outputs gegenüber anderen Personen gemessen und belohnt (vgl. Kalra & Shi, 2001, S. 170). Die Literatur beschäftigt sich überwiegend mit empirischen Versuchen zu den Anreizmechanismen von Vertriebsmitarbeitern. Führende Artikel sind von Coles, Li & Wang, 2018 und Verbeke, Bagozzi & Belschak, 2016.

Gegner von Sales-Contests argumentieren, dass ein zu starker interner Wettbewerb zwischen den Mitarbeitern zu einer Sabotage der gegenseitigen Bemühungen, der Leistung

der Mitarbeiter und damit resultierend auch der Leistung des Unternehmens führen könnte (vgl. Harbring, Irlenbusch, Kräkel & Selten, 2007, S. 26f.; Lazear, 1997, S. 195).

Personalökonomik

Weitergehend ist die Personalökonomik, besonders die Prinzipal-Agenten-Theory, sehr verbreitet in der Untersuchung von Statusanreizen. Dies lässt sich darauf zurückführen, dass die Entscheidung der Vergabe von Statusanreizen bei den Prinzipalen liegt und auch mit der Schließung von Verträgen zwischen Prinzipal und Agent einhergeht. Dies liegt zum Beispiel in der zuvor beschriebenen Kompensation von finanziellen und statusbedingten Anreizen begründet. Ein grundlegender Artikel stammt hier von den Forschern Besley & Ghatak, 2008.

Für die Forschungsfrage ist dieses Feld ebenfalls relevant, weil es sich mit dem Nutzen von Prinzipal und Agent im Hinblick auf Vertragsschluss und Motivationswirkung von Statusanreizen beschäftigt.

Eine Schnittstelle zwischen der Sales-Contest-Literatur und der Personalökonomik ist die Turniertheorie. Die Turniertheorie geht davon aus, dass ein Preisgeld oder ein anderer Anreiz, der normalerweise auf eine kleine Anzahl von Gewinnern verteilt wird, im Voraus festgelegt wird (vgl. Verbeke et al., 2016, 4113). Dies hat zur Folge, dass die relative Leistung oder Rangfolge der Ergebnisse anstelle der absoluten Leistung die Gewinner bestimmt. Das bedeutet, dass nur die besten Leistungsträger einen Preis (Positionsgut) gewinnen, egal wie gut oder schlecht diejenigen Personen mit niedrigerem Rang sind. Die Maxime der Turniertheorie liegt darin, dass ein Turnier so gestaltet sein muss, dass die Mitarbeiter die Preise oder Auszeichnungen als erstrebenswert und motivationsstärkend wahrnehmen müssen, damit sie ihre Zeit und ihre Anstrengung investieren, sodass der Wettkampf der Firma förderlich und nicht hinderlich ist (vgl. Verbeke et al., 2016, S. 4113). Kürzlich wurde empirisch festgestellt, dass die Anreize für Industrieturniere in positivem Zusammenhang mit der Unternehmensleistung stehen (vgl. Coles et al., 2018, S. 1451).

Self-Enhancement Literatur (Psychologie)

Die Literatur zu der Selbstaufwertung besteht zunächst aus Theorien aus dem Bereich der Sozialpsychologie. Sie kann als Basis-Literatur verstanden werden und stellt meist keinen aktuellen Forschungsstand dar. Die Grundlage in dieser Arbeit bilden die klassischen Theorien, die bereits im Theoriekapitel dargelegt wurden. Die führenden Forscher auf dem Gebiet sind Constantine Sedikides und Aiden P. Gregg.

Des Weiteren wurden aktuelle empirische Forschungsergebnisse zu dem Thema Selbstaufwertung genutzt, um den gegenwärtigen Stand der Selbstaufwertung einbeziehen zu können und ihn mit der aktuellen Forschung zu den Statusanreizen zu verknüpfen.

Aus Gründen der Übersichtlichkeit wurden hier jeweils nur kurze Einblicke in die verwendete Literatur gegeben. Eine ausführliche Literaturtabelle befindet sich im Anhang der Arbeit. Anschließend an diesen Überblick gilt es die Forschungslücken herauszuarbeiten.

3.2. Forschungslücken

Aus dieser gegebenen Literatur lässt sich ein direkter Schluss auf die Forschungslücke ziehen. In sich geschlossen haben die Themen Statusanreize und Selbstaufwertung eine jeweils hohe Frequenz an Überschneidungen, wodurch sich *intrathematisch* kaum Forschungslücken ergeben. Ausgenommen ist die Sales-Contest-Literatur, weil sie, obwohl in der Praxis spezielle Anreizsysteme häufig genutzt werden, in der Theorie wenig erforscht ist (vgl. Kalra & Shi, 2001, S. 171). Gerade innerhalb der Literatur zu den Statusanreizen, also bei der Sales-Contest-Literatur, Personalökonomik (Prinzipal-Agenten-Theorie) und Motivation von Angestellten (Fokus auf Sales-Force) gibt es so viele Überschneidungen, dass allein die Kategorisierung der einzelnen Themenfelder schon ein Problem darstellt.

Die Forschungslücke liegt in der fehlenden *interthematischen* Konnektivität der Begriffe Statusanreize (im Vertrieb) und der Selbstaufwertung. In keiner Literatur wurden die beiden Begriffe in einen Zusammenhang gestellt, obwohl sie beide durch die großen Themen Status und auch Motivation verbunden werden können. Dies stellt sowohl die Relevanz als auch den explorativen Aspekt der Arbeit dar.

Diese Verbindung der beiden Themenfelder und auch die Diskussion vor dem Hintergrund der Forschungsfragen wird der Fokus des nachfolgenden Analyseteils sein.

4. Statusanreize vor dem Hintergrund der Selbstaufwertung

4.1. Hinführung zur Analyse

Einleitend geht es in dem Analysekapitel darum, die Sinnhaftigkeit von Statusanreizen aus den Blickwinkeln der unterschiedlichen Literatur zu diskutieren. Weitergehend werden die Ergebnisse vor dem Hintergrund der Selbstaufwertung diskutiert, um somit im Abschluss dieses Kapitels eine adäquate Beantwortung der beiden Forschungsfragen zu gewährleisten.

Die Literatur deutet darauf hin, dass die Verwendung von Statusanreizen mit Vor- und Nachteilen verbunden ist. Dies hängt ebenfalls von dem gewählten Ausmaß der Verwendung ab. Viele Vorteile und auch viele Nachteile können sich aus dem Konzept der Selbstaufwertung ableiten lassen, da die Erhöhung des Status einer Person die Vergleichsstrategien aus Kapitel 2.1.4 der Personen innerhalb eines Sales-Teams verändern können. Die enge Verbindung der beiden Themen ist gegeben, da beide Themen sich mit Motivation in Kombination mit dem Status beschäftigen.

Nachfolgend werden Kernthesen zu dem Thema Statusanreize aus dem Blickwinkel der gegebenen Literatur zu Statusanreizen und der Literatur zu der Selbstaufwertung diskutiert.

4.2. Diskussion der Kernthesen

Der Status misst sich immer relativ gesehen zu den anderen Personen der Gruppe. Um den Status eines Agenten innerhalb einer Firma zu erhöhen, wird der Status eines anderen Agenten verringert (vgl. Auriol & Renault, 2008, S. 305). Als Beispiel kann sich in diesem Kapitel ein Sales-Team mit fünf Personen vorgestellt werden. Diese fünf Personen haben im Ausgangszustand ein vergleichbares Statusniveau. Wird nun der Status eines Agenten erhöht, indem er sich durch gute Verkaufszahlen eine Medaille verdient, dann verringert sich gleichzeitig relativ gesehen der Status aller anderen Agenten, die keine Medaille bekommen. Wichtig für die Manager wäre in diesem Fall, dass die positive Motivationswirkung aus der Vergabe des Statusanreizes an das Sales-Team, die negative Motivationswirkung überwiegt (vgl. Besley & Ghatak, 2008, S. 211). Dies ist eines der Hauptargumente gegen die übermäßige Verwendung von Statusanreizen.

Auch in der aktuellen Forschung der Selbstaufwertung ist dieser Effekt bekannt. Es wurde herausgefunden, dass jedes Feedback sowohl negative, als auch positive Auswirkungen auf die Motivation haben kann (vgl. Blagoeva, Mom, Jansen & George, 2020, S. 332).

In diesem Fall gilt es sich an den *aufwärtsgerichteten* sozialen Vergleich der Selbstaufwertung zu erinnern. Dieser hat nur dann eine positive Wirkung für die anderen vier Personen, wenn der Unterschied zwischen dem Selbst und der anderen Person nicht zu groß ist und der Erfolg ebenfalls erreichbar ist. Falls dies nicht gegeben ist, endet die Selbstaufwertung entweder in einer Inspiration, aber wahrscheinlicher Weise in einer Resignation (siehe 2.1.4). Vermutlich hängt dies auch von der Stärke des Anreizes ab. Eine kleine Münze aus Stahl wäre ein schwächerer Anreiz als eine große Münze aus purem Gold.

Neben den Statusanreizen sind jedoch auch finanzielle Anreize innerhalb einer Gruppe von fundamentaler Wichtigkeit. Die Forschung zu subjektivem Wohlbefinden zeigt, dass ein höherer Konsumstatus von Anderen, die Menschen veranlasst, mehr Zeit bezahlter Arbeit zu widmen, um den relativen Konsum aufrechtzuerhalten (vgl. Charles, Hurst & Roussanov, 2009, S. 427f.; Heffetz, 2011, S. 1101; Huberman & Minns, 2007, S. 556). Somit wird insinuiert, dass nicht nur der Status, sondern auch finanzielle Anreize (überwiegend das Gehalt) von den Mitgliedern eines Sales-Teams in einem relativen Verhältnis zu den anderen Mitgliedern gesehen wird.

Statusanreize sind eine knappe Ressource. Dies macht den Status und Statusanreize zu einem sogenannten Positionsgut. Das bedeutet, dass der Wert eines Statusanreizes auch davon abhängt, wie oft er vergeben wird. In dem Bild der fünf Personen innerhalb eines Sales-Teams bedeutet das folgendes: Es macht einen starken Unterschied für die Person, die einen Statusanreiz (hier eine Medaille) bekommt, ob sie die einzige Person ist oder ob beispielsweise drei der fünf Personen eine Medaille bekommen. Grundsätzlich gilt: Je mehr Personen innerhalb eines Sales-Teams einen Statusanreiz bekommen, desto weniger ist er Wert (vgl. Besley & Ghatak, 2008, S. 209). Aus diesem Grundsatz leitet sich der *Crowding-Effekt* ab, der von Besley & Ghatak, 2008 folgendermaßen beschrieben wird: „Wenn jeder ein Positionsgut bekommt, dann geht sein Wert auf Null“ (Besley & Ghatak, 2008, S. 209). Aus diesem Grund versuchen Organisationen in der Praxis darauf hinzuweisen, dass Statusanreize nur an verdienende Personen ausgegeben werden und dass es keinen Markt für solche Belohnungen gibt (vgl. Besley & Ghatak, 2008, S. 209f.).

Der konkrete Zusammenhang zwischen den Statusanreizen und der Selbstaufwertung in einem solchen Team-Konstrukt lässt sich nur vermuten. Grundlegend lassen die Begriffserklärungen, sowie die zuvor dargelegten Theorien vermuten, dass eine Statuserhöhung einer Person ebenfalls mit einer Aufwertung des Selbst beziehungsweise einem positiveren eigenen Weltbild einhergeht. Zusammen mit dem zuvor genannten Aspekt, dass die mehrfache Vergabe von Positionsgütern innerhalb eines Teams zu einer Schwächung des Anreizes führt, kann von der Wirkungskette in Abbildung 2 ausgegangen werden.

Diese Wirkungskette in Abbildung 2 zeigt, dass die Vergabe von Statusanreizen/Positionsgütern einen *verstärkenden* Einfluss auf die Selbstaufwertung der jeweiligen Person hat (H1+). Das Adjektiv *verstärkend* wird hier gewählt, da Begriffe wie *positiv* eine falsche Suggestion der Natur der Selbstaufwertung geben kann. Die Selbstaufwertung ist kein rein positiv und kein rein negativ konnotierter Begriff.

Dieser Effekt wird durch den beschriebenen *Crowding-Effekt* abgeschwächt, also wenn mehrere Personen aus dem Umfeld (hier eines Sales-Teams) ein Positionsgut bekommen (H2-).

Statusanreize bilden eine Alternative zu finanziellen Anreizen. Was zuvor in dem Theoriekapitel differenziert wurde, gilt es nun nochmal argumentativ darzulegen: Ein starkes Argument für die Verwendung von Statusanreizen ist die Kompensation von finanziellen Anreizen (vgl. Besley & Ghatak, 2008, S. 206). Dies bedeutet, dass Agenten, um einen höheren Status zu erlangen, auf monetäre Anreize verzichten und umgekehrt. Aus diesem Grund gibt es Ökonomen, die vorschlagen, sozialen Status direkt in die Nutzenfunktion einfließen zu lassen (vgl. Postlewaite, 1998, S. 779).

Die Literatur weist darauf hin, dass manche Unternehmen dieses Phänomen ausnutzen, indem sie Berufen spezielle, wohlklingende Bezeichnungen geben, um ihn mit einem gewissen Status zu verbinden (vgl. Grant, Berg & Cable, 2014, S. 1201). Als Beispiel wird der Hausmeister zum Facility-Manager oder der Berufseinsteiger zum Junior Consultant. Dies stellt eine direkte Ausnutzung des psychologischen Konzeptes der Selbstaufwertung dar, denn der soziale Status spielt im Rahmen des Selbstkonzeptes und der Selbstaufwertung eine wichtige Rolle, was empirisch nachgewiesen wurde (vgl. Easterbrook, Kuppens & Manstead, 2020, S. 66). Des Weiteren korreliert das Bedürfnis nach Selbstaufwertung mit Materialismus (vgl. Zhang, Hawk, Opree, de Vries & Branje, 2020, S. 386). Somit sind sowohl Geld als auch Status sehr wichtig für Personen. Diese beiden Tatsa-

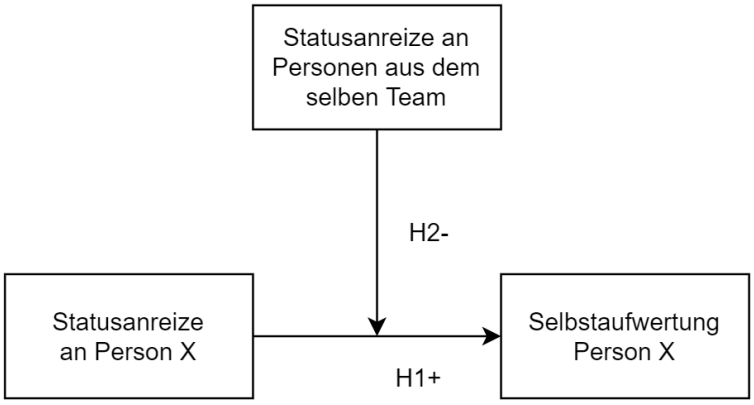


Abbildung 2: Konzeptionelles Modell

Quelle: Eigene Darstellung

chen können das nachfolgende Phänomen erklären:

Es gibt das Phänomen, dass Personen mit einem sehr hohen Einkommen sich mehr um den Status kümmern (vgl. Auriol & Renault, 2008, S. 310), also eine Kompensation zwischen Geld und Status. Um dieses Phänomen zu interpretieren, zogen die Forscher Auriol und Renault die Bedürfnispyramide nach Maslow (1954) hinzu (vgl. Abbildung 3).

Maslow argumentierte, dass es 5-Stufen der Bedürfnisse von Menschen gibt: physiologische Bedürfnisse, Sicherheitsbedürfnisse, soziale Bedürfnisse, Individualbedürfnisse und Selbstverwirklichung. Die wichtigsten Bedürfnisse stehen in der Pyramide unten und nehmen mit der steigenden Stufe an Relevanz ab. Des Weiteren muss eine Stufe zunächst erfüllt sein, bevor die nächste erfüllt werden kann. Beispielsweise sind Individualbedürfnisse unwichtig, wenn die physiologischen Bedürfnisse nicht erfüllt sind, also zum Beispiel, wenn die Person hungert (vgl. Berl, Williamson & Powell, 1984, S. 33).

Zurück auf den Kontext bezogen argumentieren Auriol und Renault, dass je höher die Stufe, desto weniger materialistisch und desto mehr statusbezogen wird das Bedürfnis. Daraus leitet sich ab, dass Personen, deren materialistische Bedürfnisse gesättigt sind, sich nun mehr um ihren Status sorgen. Dies wird bestätigt von Centers & Bugental, 1966, die empirisch nachwiesen, dass die Spitze der Pyramide, also die psychologischen, statusbezogenen Bedürfnisse eine stärkere Rolle bei Personen aus den höheren Einkommensklassen spielen (vgl. Centers & Bugental, 1966, S. 196). Dieses gesamte Phänomen der Verschiebung der Bedürfnisse lässt sich in der Realität sehr gut beobachten. Als Beispiel kann man sich die wohlhabenden Personen vorstellen, die vermehrt durch die Finanzierung von Wohltätigkeitsorganisationen monetäre Mittel aufgeben, um ihren Status in der Gesellschaft zu erhöhen (vgl. Auriol & Renault, 2008, S. 310). Dies zeigt noch einmal anschaulich, dass sowohl Status als auch Geld eine wichtige Rolle im Leben von Menschen spielen.

Mit der Pyramide von Maslow wird argumentiert, dass der monetäre Anreiz grundlegender ist als der Statusanreiz. Streng interpretiert, müsste argumentiert werden, dass er den Forschern widerspricht, die Status und monetäre Anreize als beidseitige Kompensationsgüter darstellen. Nach Maslow würden primär monetäre Anreize gegen Status getauscht und nicht andersherum. Jedoch sehen sich im Fall von Maslow die materialistischen Bedürfnisse zunächst als Grundsicherung (vgl. Berl et al., 1984, S. 33). Die klare Trennung von monetären Anreizen und Statusanreizen ist deshalb schwierig, da man sich auch von monetären Anreizen einen Status erkaufen kann, beispielsweise ein teures Auto als Statussymbol.

Trotz dieses Kompensationseffektes muss festgehalten werden, dass finanzielle Anreize und Statusanreize sich auch oft ergänzen. Das lässt sich daher ableiten, dass hoher Status und hohes Einkommen korrelieren und höherer Status ebenfalls die Basis für höheres Einkommen sein kann (vgl. Auriol & Renault, 2008, S. 306).

Das Streben nach Status der (Vertriebs-)Mitarbeiter hat ambivalente Effekte. Die aktuelle Forschung sieht das Streben nach Status und auch die Statusmanipulation als Managemententscheidung unklar. Es gibt vor allem in der Literatur zu den Sales-Contests viele Befürworter und viele Gegner von übermäßigem Gebrauch von Statusanreizen und der Statusincentivierung. Viele Forscher versuchen durch empirische Studien herauszufinden, ob der positive Effekt oder der negative Effekt von Statusmanipulation, also der Vergabe von Statusanreizen und eines firmeninternen Wettbewerbs überwiegt. Die empirische Forschung deutet darauf hin, dass identische Anreize nicht alle Mitglieder eines Sales-Teams gleich beeinflussen (vgl. Chung, 2015, S. 61).

Das Streben nach Status definiert den Wunsch, dominant und führend zu sein, mit dem entsprechenden Wunsch, die Hierarchie des sozialen Status zu erklimmen (vgl. Cassidy & Lynn, 1989, S. 301). Dies ist mit einem hohen Stellenwert für die Anerkennung und dem Wunsch nach Zustimmung der Vorgesetzten verbunden. Ein positiver Aspekt des Statusstrebens ist die gesteigerte Motivation der Person (vgl. Prundeanu, Constantin & Popuoi, 2021, S. 2), ein negativer Aspekt des Statusstrebens ist, dass der Wunsch nach Ruhm

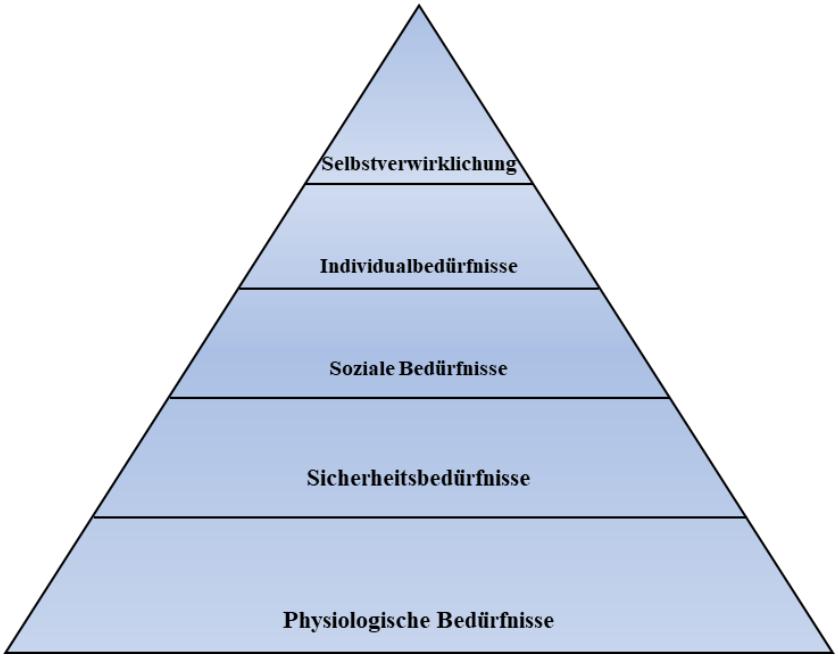


Abbildung 3: Bedürfnispyramide nach Maslow

Quelle: In Anlehnung an Maslow (1954)

unter Umständen mit Betrug in Verbindung gebracht wird (vgl. Knight, 1984, S. 433).

Es hat sich herausgestellt, dass die Verknüpfung attraktiver Belohnungen mit kurzfristigen Zielen (Sales Contest), beispielsweise Statusbelohnungen für gute Verkaufszahlen, zu problematischen Verhaltensweisen führen kann, einschließlich Umsatzverlusten aus anderen Zeiträumen, Verringerung der Zusammenarbeit, Überbewertung von Produkten, Übertreibung von Ansprüchen, Akzeptieren fragwürdiger Kreditrisiken und Aufhebung anderer Verantwortlichkeiten (vgl. Hampton, 1970, S. 87f.; Moncrief, Hart & Robertson, 1988, S. 60; Wildt, Parker & Harris Jr, 1981, S. 63; Wotruba & Schoel, 1983, S. 8). Es handelt sich hier um firmenschädliche Verhaltensweisen der Verkäufer, um in dem relativen Verhältnis zu den anderen Verkäufern gut dar zustehen.

Als ein Grund dafür kann der Faktor Leistungsdruck betrachtet werden. Personen die sich einen höheren Status und/oder Rang erarbeitet haben und womöglich ihre Position innerhalb einer Firma gesichert haben, fühlen meist einen geringeren Leistungsdruck als zuvor in ihrer Karriere (vgl. Murphy, 2004, S. 1267). Währenddessen können untergeordnete Mitarbeiter das Gefühl haben, dass ein Versagen (z.B. Nichterhalten eines Positionsgutes) ihre Chance gefährdet, bei der Firma zu bleiben oder in der Firma aufzusteigen (vgl. Badaracco & Webb, 1995, S. 11). Aus diesem Grund könnten statusniedrige Verkäufer versucht sein, problematische Verhaltensweisen zu adaptieren, um Vorteile zu generieren (vgl. Murphy, 2004, S. 1267). Gleichzeitig könnte die Beteiligung an weniger problematischen Verhaltensweisen generell als Vorteil für die Erlangung des Status angesehen werden, weil das Risiko schwerwiegender Konsequenzen ausgeschlossen wäre (vgl. Murphy, 2004, S. 1269).

Es gibt jedoch auch positive Aspekte des Strebens nach Status. Dass der Status eine wichtige Rolle in der Motivationswirkung von Mitarbeitern spielt, wurde durch die Literatur zu Statusanreizen und Selbstaufwertung bereits dargelegt.

Wirtschaftliche Modelle zeigen formal, wie optimale Statusanreize die Leistung von Personen verbessern können, wenn sie sich um den Status kümmern. Bei dem Konsumieren eines positionellen Gutes üben sie höhere Anstrengungen aus, wenn der Wert der Ränge steigt (vgl. Besley & Ghatak, 2008, S. 210; Dubey & Geanakoplos, 2010, S. 72; Ederer & Patacconi, 2010, S. 348; Hopkins & Kornienko, 2009, S. 552). Auch das empirische Ergebnis von Verbeke, Bagozzi und Belschak deutet darauf hin, dass Statusanreize einen signifikanten Effekt auf das Verkaufsvolumen haben (vgl. Verbeke et al., 2016, S. 4119). Coles, Li und Wang kommen zu dem Ergebnis, dass Anreize im positiven Zusammenhang mit der Unternehmensleistung stehen (vgl. Coles et al., 2018, S. 1451).

Aus dieser Diskussion lässt sich schließen, dass die Vergabe von Statusanreizen sowohl negative als auch positive Effekte haben kann, die es zu beachten gilt.

4.3. Beantwortung der Forschungsfragen

Anhand der durchgeführten Diskussion der Literatur mithilfe der Kernthesen lassen sich die Forschungsfragen folgendermaßen beantworten:

Welche Rolle spielt die Selbstaufwertung beim Einsatz von Statusanreizen im Vertrieb?

Die Bedeutung Selbstaufwertung im vertrieblichen Kontext liegt in der Motivation der Mitarbeiter begründet. Innerhalb eines Sales-Teams spielt der Status der Mitglieder eine wichtige Rolle. Der Status kann durch Statusanreize beeinflusst werden, was sich wiederum auf die Motivation der Mitarbeiter auswirken kann. Diese Wirkung auf die Motivation kann durch das Konzept der Selbstaufwertung erklärt werden. Statusanreize besitzen deshalb eine sehr starke Wirkung, da sie die Selbstaufwertung beeinflussen.

Wie zuvor diskutiert, kann diese Wirkung ambivalent sein und somit sowohl negative als auch positive Auswirkungen haben. Es lässt sich nun vermuten, dass die Selbstaufwertung genau an diesen Auswirkungen einen maßgeblichen Anteil hat. Wie bereits in 2.1.2 (bipolare Dimension) dargelegt, spielt die Selbstaufwertung eine besonders starke Rolle in Gebieten, die für das Individuum wichtig sind. Es kann davon ausgegangen werden, dass die Karriere bei Vertriebsmitarbeitern einen hohen Stellenwert hat, wodurch es ein solches Gebiet darstellt.

Des Weiteren sind strategische Vergleiche für das Individuum wichtig, wie in 2.1.4 beschrieben. Durch die Statusanreize innerhalb eines zuvor statusgleichen Teams, werden Statusunterschiede kreiert (Statusmanipulation). Dies ändert zuvor laterale Vergleiche in abwärts- und aufwärtsgerichtete Vergleiche.

Wie können die Unternehmen gezielt das Konzept der Selbstaufwertung in Sales Teams nutzen?

Die Antwort dieser Frage leitet sich aus der Beantwortung der ersten Forschungsfrage ab. Unternehmen können Statusanreize nutzen, um die Dynamik und die Hierarchie innerhalb einer Gruppe zu verändern.

Indirekt nutzen Unternehmen durch den Einsatz von Statusanreizen immer auch die Wirkung der Selbstaufwertung. Beispielsweise kann die Selbstaufwertung aus Selbstschutz bei den nun statusniedrigeren und die Selbstaufwertung aus Selbstförderung der statushöheren Mitarbeiter ausgelöst werden.

Generell gilt zu sagen, dass durch die Statusveränderung viele Effekte der Selbstaufwertung eintreten können. Wie zuvor erwähnt, können diese ambivalente Effekte haben, was darauf schließen lässt, dass auch die Selbstaufwertung bei Statusveränderung ambivalente Effekte haben kann (vgl. Blagoeva et al., 2020, S. 351).

Die Selbstaufwertung lässt sich als externer Effekt beschreiben. Das bedeutet, dass dieser, wenn auch nicht intendiert, immer präsent ist, wenn eine Statusänderung auftritt. Unternehmen sollten sich über dies im Klaren sein.

5. Abschließende Betrachtung

5.1. Fazit

Zusammenfassend lässt sich sagen, dass das Thema der beiden großen Forschungsfelder der Statusanreize und der Selbstaufwertung sich sehr gut für die Forschung eignet. Es stellt eine interdisziplinäre Schnittstelle zwischen der Psychologie und dem Management dar, die die Forschung im Bereich Marketing und Vertrieb bereichert. Des Weiteren lassen sich nachfolgend Implikationen für die Praxis formulieren. Die Sinnhaftigkeit der Statusanreize im Vertrieb konnte detailliert herausgearbeitet werden und sowohl die klassischen Theorien als auch die aktuelle Forschung zu der Selbstaufwertung ließen sich darauf anwenden.

Eines der früheren Forschungsziele war es nicht herauszufinden, **welche**-, sondern **ob** die Selbstaufwertung eine Rolle beim Einsatz von Statusanreizen im Vertrieb spielt. Nachdem diese Frage trotz der fehlenden Überschneidungen der Literatur schnell verifiziert werden konnte, lohnte es sich, die Forschungsfrage neu zu formulieren.

Es stellte sich schnell heraus, dass die Selbstaufwertung bei Statusveränderungen einer Gruppe **immer** eine Rolle spielt, wenn auch nicht immer im Vordergrund, sondern meistens als eine Art externer Effekt. Des Weiteren scheint offensichtlich, dass Unternehmen und Manager bisher die Statusanreize im Vertrieb nicht bewusst mit der Wirkung auf die Selbstaufwertung der einzelnen Teammitglieder genutzt haben. Doch ließen sich direkte Parallelen zwischen den Ergebnissen der Motivationswirkung der Statusanreize und der Selbstaufwertung in Bezug auf Statusveränderungen erkennen. Status und Motivation verbindet die beiden Begriffe.

5.2. Limitationen

Die literaturbasierte Forschungsarbeit unterliegt einigen Limitationen, die hier dargestellt werden. Die erste Einschränkung liegt darin, dass durch die hohe Komplexität des Themas „Die Rolle der Selbstaufwertung beim Einsatz von Statusanreizen im Vertrieb“ ein starker Fokus gesetzt werden musste. Dieser lag in der Betrachtung eines kleinen Sales-Teams, welches durch Statusanreize motiviert wird. Das Thema in all seinen Ausprägungen und Facetten zu durchleuchten, würde den Rahmen einer solchen Arbeit übersteigen. Gerade der Begriff der Selbstaufwertung musste sehr zweckdienlich definiert werden, um Hypothesen der Auswirkung der Statusanreize und die Rolle der Selbstaufwertung aufstellen zu können.

Dies stellt direkt die zweite wesentliche Einschränkung dar: Durch die fehlende direkte Verbindung von dem Konzept der Selbstaufwertung und der Statusanreize konnte diese nur durch Hypothesen aufgestellt werden. Somit stellt das Thema direkt die eigentliche Forschungslücke dar.

Weiterhin musste hier von einer eindimensionalen Auswirkung der Statusanreize auf die Selbstaufwertung und die resultierende Motivationswirkung ausgegangen werden. Dies macht besonders die Bewertung der Statusanreize **generell** schwierig, da natürlich mehrere Effekte mit der Vergabe von Statusanreizen einher gehen, hier aber oft nur aus der Perspektive der Selbstaufwertung argumentiert wurde. Die Selbstaufwertung ist also nicht der einzige Grund für die veränderte Motivation.

5.3. Implikationen für die Praxis

Trotz der dargelegten Limitationen erlangt die Arbeit einen Erkenntnisgewinn, der es erlaubt, Implikationen für die Praxis zu formulieren.

Die wichtigste Implikation der Arbeit ist, dass Statusanreize mit Bedacht zu vergeben sind. Das Streben nach Status und die Verwendung von Statusanreizen können in einem Sales-Team sowohl positive als auch negative Effekte auf die Motivationswirkung der Personen haben, die einen Statusanreiz bekommen oder auch nicht bekommen. Dies lässt sich durch die Natur der Selbstaufwertung beschreiben, die sehr individuell ausfällt.

Obwohl diese Ambivalenz vorliegt, kann die Vergabe von Statusanreizen nunmehr mit einer starken Motivationswirkung einher gehen und somit ein wichtiges Mittel für Unternehmen und Vorgesetzte darstellen. Nicht zu vergessen ist, dass Statusanreize finanzielle Anreize bis zu einem gewissen Grad kompensieren und somit Kosten für der Motivation der Mitarbeiter gespart werden können.

Als konkrete Idee für die Praxis kann somit eine Einführung eines Anreizsystems in Sales-Teams gedacht werden, beispielsweise eine Rangordnungstafel nach Verkäufen. Wichtig wäre hier, dass die Verkäufer ein „ähnliches Level“ besitzen, sodass sie sich gegenseitig motivieren können und werden (siehe 2.1.4). Die Personen, die am Ende des Jahres oben auf der Liste stehen, bekommen nun einen Statuspreis. Dies kann von einer Medaille bis zu einem Schild alles sein.

In gewisser Hinsicht deckt sich dieses Ergebnis vor dem Hintergrund der Selbstaufwertung mit den Ergebnissen der aktuellen Forschung der Sales-Contest-Literatur.

5.4. Ausblick auf weitere Forschung

Da sowohl die Forschung zu den Statusanreizen als auch zu der Selbstaufwertung eine gewisse Aktualität genießt, bleibt sie in vielerlei Hinsicht stark forschungsrelevant. Dabei ist der konkrete Zusammenhang zwischen der Selbstaufwertung und den Statusanreizen, der in der vorliegenden Arbeit nicht vollständig geklärt werden konnte, von hoher Bedeutung und sollte durch weiterführende Forschungsarbeiten untersucht werden. Konkret würde die empirische Untersuchung der Hypothese in Abbildung 2 einen starken Mehrwert sowohl für die Forschung als auch für die Praxis im Vertrieb bieten.

Des Weiteren sollte die Selbstaufwertung im Vertriebsbezug genauer untersucht werden. Es gibt durchaus Literatur zu der Selbstaufwertung in Bezug auf Kunden im Kaufprozess (vgl. D. G. Taylor, Strutton & Thompson, 2012, S. 13), aber sehr wenig zu den Sales-Teams im Verkaufsprozess. Dies sollte von Forschern adressiert werden.

Literatur

Arkin, R. M. (1981). *Self-presentation styles, in Impression management theory and research*. Tedeschi, Hrsg. New York: James T. Tedeschi, Hrsg. New York: Academic Press.

Auriol, E. & Renault, R. (2008). Status and incentives. *The RAND Journal of Economics*, *39* (1), 305–326.

Badaracco, J. L. & Webb, A. P. (1995). Business Ethics: A view from the trenches. *California Management Review*, *37* (2), 8–28.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C. & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5* (4), 323–370.

Berl, R. L., Williamson, N. C. & Powell, T. (1984). Industrial salesforce motivation: A critique and test of maslow's hierarchy of need. *Journal of Personal Selling & Sales Management*, *4* (1), 32–39.

Besley, T. & Ghatak, M. (2008). Status incentives. *American Economic Review*, *98* (2), 206–211.

Bhattacharya, H. & Dugar, S. (2012). Status incentives and performance. *Managerial and Decision Economics*, 549–563.

Biernart, M. & Billings, L. S. (2002). Standards, expectancies, and social comparison. In *Blackwell handbook of social psychology: Intraindividual processes, abraham tesser and norbert schwarz, hrsg. oxford, UK: Blackwell publishing ltd* (S. 257–283). Oxford, UK: Blackwell Publishing Ltd.

Blagoeva, R. R., Mom, T. J. M., Jansen, J. J. P. & George, G. (2020). Problem-Solving or Self-Enhancement? A power perspective on how CEOs affect R&d search in the face of inconsistent feedback. *Academy of Management Journal*, *63* (2), 332–355.

Bommaraju, R. & Hohenberg, S. (2018). Self-Selected Sales Incentives: Evidence of their Effectiveness, Persistence, Durability, and Underlying Mechanisms. *Journal of Marketing*, *82* (5), 106–124.

Buunk, B. P., Collins, R. L., Taylor, S. E., Yperen, N. W. V. & Dakof, G. A. (1990). The affective consequences of social comparison: Either direction has its ups and downs. *Journal of Personality and Social Psychology*, *59* (6), 1238–1249.

Cassidy, A. & Lynn, R. (1989). A multifactorial approach to achievement motivation: The development of a comprehensive measure. *Journal of Occupational Psychology*, *62*, 301–312.

Centers, R. & Bugental, D. E. (1966). Intrinsic and extrinsic job motivations among different segments of the working population. *The Journal of applied psychology*, *50* (3), 193–197.

Charles, K. K., Hurst, E. & Roussanov, N. (2009). Conspicuous consumption and race *. *Quarterly Journal of Economics*, *124* (2), 425–467.

Chung, D. J. (2015). How to really motivate sales people. *Harvard business review*, *93* (4), 54–61.

Coles, J. L., Li, Z. & Wang, A. Y. (2018). Industry tournament incentives. *The Review of Financial Studies*, *31* (4), 1418–1459.

Collins, R. L. (1996). For better or worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin*, *119* (1), 51–69.

Collins, R. L. (2000). Among the better ones: Upward assimilation in social comparison. In *Handbook of social comparison: Theory and research, Jerry M. Suls and Ladd Wheeler, Hrsg. : Kluwer Academic* (S. 159–171). Suls and Ladd Wheeler, Hrsg.: Kluwer Academic.

Crocker, J. & Knight, K. M. (2005). Contingencies of self-worth. *Current Directions in Psychological Science*, *14* (4), 200–203.

Crocker, J. & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review*, *108* (3), 593–623.

Dubey, P. & Geanakoplos, J. (2010). Grading exams: or? *Games and Economic Behavior*, *69* (1), 72–94.

Easterbrook, M. J., Kuppens, T. & Manstead, A. S. R. (2020). Socioeconomic status and the structure of the self-concept. *The British journal of social psychology*, *59* (1), 66–86.

Ederer, F. & Patacconi, A. (2010). Interpersonal comparison, status and ambition in organizations. *Journal of Economic Behavior & Organization*, *75* (2), 348–363.

Ellingsen, T. & Johannesson, M. (2007). Paying respect. *Journal of Economic Perspectives*, *21* (4), 135–149.

Grant, A. M., Berg, J. M. & Cable, D. M. (2014). Job titles as identity badges: How self-reflective titles can reduce emotional exhaustion. *Academy of Management Journal*, *57* (4), 1201–1225.

Hampton, D. R. (1970). Contests have side effects too. *California Management Review*, *12* (4), 86–94.

Harbring, C., Irlenbusch, B., Kräkel, M. & Selten, R. (2007). Sabotage in corporate contests - An experimental analysis. *International Journal of the Economics of Business*, *14* (3), 367–392.

Heffetz, O. (2011). A test of conspicuous consumption: Visibility and income elasticities. *Review of Economics and Statistics*, *93* (4), 1101–1117.

Helweg-Larsen, M. & Shepperd, J. A. (2001). Do Moderators of the Optimistic Bias Affect Personal or Target Risk Estimates? A Review of the Literature. *Personality and Social Psychology Review*, *5* (1), 74–95.

Hopkins, E. & Kornienko, T. (2009). Status, affluence, and inequality: Rank-based comparisons in games of status. *Games and Economic Behavior*, *67* (2), 552–568.

Huberman, M. & Minns, C. (2007). The times they are not changin': Days and hours of work in Old and New Worlds. *Explorations in Economic History*, *44* (4), 538–567.

Jansen, R. A., Rafferty, A. N. & Griffiths, T. L. (2021). *A rational model of the Dunning-Kruger effect supports insensitivity to evidence in low performers.* Nature human behaviour.

Kalra, A. & Shi, M. (2001). Designing optimal sales contests: A theoretical perspective. *Marketing Science*, *20* (2), 170–193.

Knight, J. A. (1984). Exploring the compromise of ethical principles in science. *Perspectives in biology and medicine*, *27* (3), 432–442.

Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77* (6), 1121–1134.

Lazear, E. (1997). Incentives in basic research. *Journal of Labor Economics*, *15* (1), 167–197.

Leary, M. R. & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, *107* (1), 34–47.

Leary, M. R., Tchividijian, L. R. & Kraxberger, B. E. (1994). Self-presentation can be hazardous to your health: Impression management and health risk. *Health Psychology*, *13* (6), 461–470.

Lockwood, P. & Kunda, Z. (1997). Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology*, *73* (1), 91–103.

Maslow, A. (1954). *Motivation and personality*. New York: Harper & Row.

Moncrief, W. C., Hart, S. H. & Robertson, D. (1988). Selling & sales management in action: Sales Contests: A new look at an old management tool. *Journal of Personal Selling & Sales Management*, *8* (3), 55–61.

Murphy, W. H. (2004). In pursuit of short-term goals: Anticipating the unintended consequences of using special incentives to motivate the sales force. *Journal of Business Research*, *57* (11), 1265–1275.

Nelson, B. (1994). *1001 ways to reward employees*. New York: Workman Publishing.

Podmoroff, D. (2005). *365 ways to motivate & reward your employees every day: With little or no money*. Ocala, FL: Atlantic Publishing.

Postlewaite, A. (1998). The social basis of interdependent preferences. *European Economic Review*, 779–800.

Prundeanu, O., Constantin, T. & Popuoi, S. A. (2021). Climb up your ego! Narcissistic status pursuit and motivation to lead. *Personality and Individual Differences*, *177* (3), 1–7.

Pyszczynski, T., Greenberg, J. & Goldenberg, J. L. (2003). Freedom versus fear: On the defense growth and expansion of the self. In *Handbook of self and identity, Mark R. Leary and June P. Tangney, Hrsg. : The Guilford Press* (S. 314–343). Leary and June P. Tangney, Hrsg.: The Guilford Press.

Sedikides, C., Campbell, W. K., Reeder, G. D. & Elliot, A. J. (1998). The self-serving bias in relational context. *Journal of Personality and Social Psychology*, *74* (2), 378–386.

Sedikides, C., Gaertner, L. & Toguchi, Y. (2003). Pancultural self-enhancement. *Journal of Personality and Social Psychology*, *84* (1), 60–79.

Sedikides, C. & Gregg, A. P. (2007). Portraits of the self. In *The Sage handbook of social psychology. Concise Student Edition, Michael A. Hogg and Joel Cooper, Hrsg. Los Angeles: Sage* (S. 110–138). Concise Student Edition, Michael A. Hogg and Joel Cooper, Hrsg. Los Angeles: Sage.

Sedikides, C. & Gregg, A. P. (2008). Self-enhancement: Food for thought. *Perspectives on Psychological Science*, *3* (2), 102–116.

Sedikides, C. & Strube, M. J. (1997). Self-Evaluation: To thine own self be good, To thine own self be sure, to thine own self be true, and to

thine own self be better. *Advances in experimental social psychology*, *29*, 209–269.

Smits, T. & Hoorens, V. (2005). How probable is probably? It depends on whom you're talking about. *Journal of Behavioral Decision Making*, *18* (2), 83–96.

Srivastava, A., Locke, E. A. & Bartol, K. M. (2001). Money and subjective well-being: It's not the money. *it's the motives. Journal of Personality and Social Psychology*, *80* (6), 959–971.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In *Advances in experimental social psychology* (Bd. 21, S. 261–302). Leonard Berkowitz, Hrsg.: Elsevier.

Taylor, D. G., Strutton, D. & Thompson, K. (2012). Self-enhancement as a motivation for sharing online advertising. *Journal of Interactive Advertising*, *12* (2), 13–28.

Taylor, S. E., Neter, E. & Wayment, H. A. (1995). Self-evaluation processes. *Personality and Social Psychology Bulletin*, *21* (12), 1278–1287.

Trope, Y. (1986). Self-enhancement and self-assessment in achievement behavior. In *Foundations of social behavior, Richard M. Sorrentino and E. T. Higgins, Hrsg. : Guilford Press* (S. 350–378). Sorrentino and E. T. Higgins, Hrsg.: Guilford Press.

Vazire, S. & Funder, D. C. (2006). Impulsivity and the self-defeating behavior of narcissists. *Personality and social psychology review: an official journal of the Society for Personality and Social Psychology, Inc*, *10* (2), 154–165.

Verbeke, W., Bagozzi, R. P. & Belschak, F. D. (2016). The role of status and leadership style in sales contests: A natural field experiment. *Journal of Business Research*, *69* (10), 4112–4120.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, *39* (5), 806–820.

Weinstein, N. D. & Klein, W. M. (1995). Resistance of personal risk perceptions to debiasing interventions. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*, *14* (2), 132–140.

Weinstein, N. D., Marcus, S. E. & Moser, R. P. (2005). Smokers' unrealistic optimism about their risk. *Tobacco control*, *14* (1), 55–59.

Wheeler, L. (1966). Motivation as a determinant of upward comparison. *Journal of Experimental Social Psychology*, *1*, 27–31.

Wildt, A. R., Parker, J. D. & Harris Jr, C. E. (1981). Sales contest: What we know and what we need to know. *Journal of Personal Selling & Sales Management*, *1* (1), 57–64.

Wotruba, T. R. & Schoel, D. J. (1983). Evaluation of salesforce contest performance. *Journal of Personal Selling & Sales Management*, *3* (1), 1–10.

Zhang, Y., Hawk, S. T., Opree, S. J., de Vries, D. A. & Branje, S. (2020). Me, We, and Materialism: Associations between Contingent Self-Worth and Materialistic Values across Cultures. *The Journal of psychology*, *154* (5), 386–410.

Junior Management Science 7(4) (2022) 1147-1163

Junior Management Science

journal homepage: www.jums.academy

Balancing player values in German professional football – A critical analysis of German GAAP

Bilanzierung von Spielerwerten im deutschen Profi-Fußball – Eine kritische Analyse der handelsrechtlichen Vorschriften

Malte Ludwigs

Heinrich-Heine-Universität Düsseldorf

Abstract

Rising transfer fees in professional football are causing an increasing influence of player transfers on the balance sheet of the clubs. However, proper accounting and valuation of player assets entails some ambiguities and problems in the context of accounting. The aim of the present work is to illustrate the accounting treatment of a player transfer in accordance with the provisions of the German Commercial Code. Furthermore, it is considered to what extent the accounting practice in football fulfils the basic purposes of accounting. At the beginning there is a presentation of the legal framework of player transfers. Subsequently, the rules for the recognition and measurement of player assets are explained in more detail in order to finally show which problems arise from the current accounting practice. The results show that clubs have to take into account additional football-specific items in addition to the general regulations of the German Commercial Code (HGB). This leads to a comprehensive regulatory construct that nevertheless leaves the clubs room for manoeuver in terms of balance sheet policy. The balance sheet item player assets therefore has only a limited significance with regard to the actual economic circumstances in the club.

Zusammenfassung

Steigende Ablösesummen im professionellen Fußball sorgen für eine zunehmende Einflussnahme von Spielertransfers auf das Bilanzbild der Vereine. Eine sachgerechte Bilanzierung und Bewertung von Spielervermögen, birgt aber im Rahmen der Rechnungslegung einige Unklarheiten und Probleme. Das Ziel der vorliegenden Arbeit ist es, die bilanzielle Behandlung eines Spielertransfers nach den Vorschriften des Handelsgesetzbuches zu veranschaulichen. Des Weiteren wird betrachtet, inwieweit die Bilanzierungspraxis im Fußball die grundsätzlichen Zwecke der Rechnungslegung erfüllt. Zu Beginn erfolgt eine Darstellung des rechtlichen Ordnungsrahmens von Spielertransfers. Anschließend werden die Vorschriften für den Ansatz und die Bewertung von Spielervermögen näher erläutert, um abschließend aufzuzeigen, welche Problemstellungen sich aus der aktuellen Bilanzierungspraxis ergeben. Die Ergebnisse zeigen, dass Vereine neben den allgemeinen Vorschriften des HGB zusätzliche fußballspezifische Posten zu berücksichtigen haben. Dieses führt zu einem umfassenden Regelungskonstrukt, dass den Vereinen dennoch einen bilanzpolitischen Spielraum überlässt. Die Bilanzposition Spielervermögen hat somit nur eine eingeschränkte Aussagekraft hinsichtlich der tatsächlichen wirtschaftlichen Verhältnisse im Verein.

Keywords: Bilanzierung; Fußball; Spielerwerte; Spielertransfers.

1. Einleitung

Die Vereine der Fußball-Bundesliga haben sich im Laufe der Zeit zu bedeutenden Wirtschaftsunternehmen entwickelt. In der Saison 2019/2020 generierten die 18 Vereine der ersten Bundesliga einen Gesamtumsatz von 3,8 Milliarden Euro. Lediglich in den zwei zurückliegenden Spielzeiten waren die Erträge größer (vgl. DFL, 2021a, S. 18f.). Dieses kann auf die im Frühjahr 2020 eingetretene COVID-19-Pandemie zurückgeführt werden, die zu einer Einschränkung des Spielbetriebes und folglich zu einem Umsatzrückgang ge-

DOI: https://doi.org/10.5282/jums/v7i4pp1147-1163

führt hat. Angesichts des ansonsten starken wirtschaftlichen Wachstums in der Fußballbranche sind auch die im Rahmen eines Spielertransfers gezahlten Ablösesummen in den vergangenen Jahren beachtlich angestiegen. Die Transferausgaben aller 18 Vereine der Bundesliga Saison 2019/2020 beliefen sich auf einen Höchstwert von 910 Millionen Euro, mit einem prozentualen Anteil von 23 % an den gesamten Aufwendungen (vgl. DFL, 2021a, S. 19). Vereine im professionellen Fußball investieren zum Teil hohe zweistellige Millionensummen, um die Wettbewerbsfähigkeit ihres Teams im nationalen und auch im internationalen Vergleich zu gewährleisten. Zudem stellen die transferierten Spieler als Humankapital die entscheidenden Werttreiber im Verein dar. Sie sorgen nicht nur für den sportlichen Erfolg auf dem Spielfeld, sondern mit ihrem Vermarktungspotenzial auch für zentrale Einnahmequellen neben dem Spielfeld (vgl. Weber, 2016, S. 290). Aufgrund der stetig steigenden Ablösesummen und der zentralen Bedeutung der Spieler für einen Verein nehmen die Spielertransfers bei der Beurteilung der wirtschaftlichen Lage der Vereine eine immer größere Stellung ein. Die Bilanzierung und Bewertung von Spielertransfers erhält dadurch eine essenzielle Rolle bei Rechnungslegungs- und Finanzierungsfragen und begründet gleichzeitig die nähere Betrachtung dieser Thematik. Das aktivierte Spielervermögen, das den bilanziellen Wert der Spieler veranschaulicht, betrug in der Saison 2019/2020 kumuliert über alle 18 Vereine 1,47 Milliarden Euro und stellt somit mit knapp 37 % den größten Vermögensposten in der Bilanz dar (vgl. DFL, 2021a, S. 17). Das Spielervermögen hat daher einen erheblichen Einfluss auf das Bilanzbild und folglich auch auf die Vermögensstruktur der Vereine. Aus dieser Erkenntnis ergibt sich zwangsläufig die Frage nach der sachgerechten Bilanzierung des Spielers als wesentlicher Vermögenswert in der Bilanz der Vereine.

Vor diesem Hintergrund soll die vorliegende Arbeit die bilanzielle Behandlung eines Spielers vom Kauf durch einen aufnehmenden Verein bis hin zu einem möglichen Folgetransfer mit ihren verschiedenen Komponenten veranschaulichen. Die Arbeit bezieht sich dabei auf die Rechnungslegungsvorschriften des Handelsgesetzbuches (HGB), die zusammen mit den erlassenen Rechtssprechungen des Bundesfinanzhofes (BFH) die Grundlage für die Bilanzierung von Spielertransfers bilden.

Hinsichtlich des Aufbaus der Arbeit soll zunächst im zweiten Kapitel verdeutlicht werden, welche Organisationsstrukturen im Profi-Fußball bestehen, um im Anschluss daran die rechtlichen Grundlagen zur Bilanzierung von Spielertransfers zu erläutern. Kapitel drei soll zunächst aufzeigen, wie sich das Transfersystem und damit auch die Bilanzierungspraxis im Laufe der Zeit verändert haben, um dann die Bilanzierungsvorschriften für den Ansatz und die Bewertung von Spielervermögen in der Bilanz eines Vereines näher zu erläutern. Daran anknüpfend soll in Kapitel vier veranschaulicht werden, wie sich die Umsetzung der normativen Regelungen aus Kapitel drei in der Praxis widerspiegeln. Abschließend sollen die Probleme beziehungsweise (bzw.) Kritiken, die sich aus der aktuellen Bilanzierungspraxis ergeben, diskutiert werden.

2. Rahmenbedingungen im deutschen Lizenzfußball

2.1. Institutionen und Verbände im Profi-Fußball

Im Profi-Fußball bestehen im Gegensatz zu anderen Branchen vielfältige besondere Aspekte. Dieses können zum einen der kaum prognostizierbare sportliche Erfolg und die damit verbundenen Finanzmittelzuflüsse sein oder auch die Notwendigkeit der Verbandszugehörigkeit eines Vereines, um am Spielbetrieb teilzunehmen (vgl. Weber, 2017, S. 286). Durch die in der Vergangenheit immer stärker werdende wirtschaftliche Bedeutung der Fußballbranche ergibt sich die Notwendigkeit angemessener Strukturen hinsichtlich der rechtlichen und organisatorischen Durchführung des Spielbetriebes. Die zentralen Organe im Fußball sind die Verbände, Vereine und die Fußballspieler selbst. Die Verbandsstrukturen, die sich aus nationalen und internationalen Verbänden zusammensetzen, fungieren als Organisatoren des Spielbetriebes und weisen eine hierarchische Rangordnung auf (vgl. Hierl & Weiß, 2016, S. 7). An der obersten Stelle steht der Fußballweltverband FIFA, darunter folgen sechs Kontinentalverbände mit der UEFA als europäischer Dachverband. Die UEFA mit ihren 55 untergeordneten nationalen Fußballverbänden ist bestrebt, für einen fairen und nachhaltigen Wettbewerb zu sorgen, um die langfristige Existenz des europäischen Fußballs zu gewährleisten (vgl. Küting & Strauß, 2011, S. 65). Der Deutsche Fußball-Bund (DFB) ist ein Mitglied der UEFA und vertritt den deutschen Profi-Fußball als Nationalverband. Infolgedessen das der DFB Mitglied der UEFA und somit auch der FIFA ist, haben sich der DFB, seine Mitglieder sowie Spieler und Vereine an die Vorschriften und Regelungen der beiden Verbände zu halten (vgl. DFL, 2019b, S. 4). Solche verbindlichen Bestimmungen sind insbesondere bei der Betrachtung von Spielertransfers zu berücksichtigen, da sich hierbei Vorgaben bezüglich Reglements und Vorschriften für die bilanzielle Behandlung von Spielern ergeben. Der DFB hat mit Gründung des Ligaverbandes „Die Liga Fußball-Verband e.V." im Jahr 2000 die Durchführung des Spielbetriebs der ersten und zweiten Bundesliga ausgegliedert und somit vom Amateurbereich separiert (vgl. Thyll, 2004, S. 169). Der Ligaverband, seit dem Jahr 2016 firmiert als „Deutsche Fußball Liga e.V.", setzt sich aus den lizenzierten Vereinen und Kapitalgesellschaften der ersten und zweiten Bundesliga zusammen. Gemeinsam mit ihrer Tochtergesellschaft, der „Deutsche Fußball Liga GmbH" (DFL) führt sie das operative Geschäft im deutschen Lizenzfußball (vgl. DFL, 2019b, S. 1). Dabei ist die DFL als ordentliches Mitglied des DFB den Satzungen und Ordnungen des DFB und damit auch den internationalen Verbänden unterworfen (vgl. DFL, 2019b, S. 3f.). Somit kann der DFB seine Interessen und Ziele sicherstellen, ohne einen unmittelbaren Einfluss auf den Spielbetrieb zu nehmen. Der DFL obliegen zentrale Aufgaben wie die Durchführung des Spielbetriebs in den Lizenzligen sowie die Erteilung der Lizenzen an Vereine, Kapitalgesellschaften und Spieler, damit sie zur

Teilnahme am Spielbetrieb berechtigt sind (vgl. DFL, 2019b, S. 5). Die DFL bedient sich bei ihrer Arbeit einem umfassenden Regelungskonstrukt, dem sogenannten (sog.) Ligastatut. Dieses beinhaltet neben allgemeinen Bestimmungen für den ordnungsgemäßen Spielbetrieb auch Vorschriften für die Lizenzerteilung an Vereine und Kapitalgesellschaften in der Lizenzierungsordnung (LO) sowie an Spieler in der Lizenzordnung Spieler (LOS) (vgl. DFL, 2019b, S. 7).

Vereine und Kapitalgesellschaften, die am nationalen Spielbetrieb teilnehmen wollen, sind verpflichtet, sich an die Satzungen und Ordnungen des DFB und der DFL zu halten und für jede Spielzeit eine Lizenz bei der DFL zu beantragen. Dass Kapitalgesellschaften überhaupt am Spielbetrieb teilnehmen dürfen, ist auf eine Abwandlung der Satzung des DFB aus dem Jahr 1998 zurückzuführen. Bis zu diesem Zeitpunkt war es nur eingetragenen Vereinen (e.V.) erlaubt, eine Lizenz zu beantragen (vgl. Teichmann, 2007, S. 116). In Anbetracht dessen, dass sich viele Vereine mittlerweile zu mittelständischen Unternehmen entwickelt haben, wurde die Rechtsform des e.V. nicht mehr als sachgerecht angesehen. Deswegen wurde für die Vereine die Möglichkeit eröffnet, ihre Lizenzspielerabteilung vom Mutterverein auf eine Kapitalgesellschaft auszugliedern, um somit in einer für ihre Größe angemessenen Organisationsform agieren zu können (vgl. Littkemann, Brast & Stübinger, 2003, S. 416f.). Von den 18 Vereinen, die in der Saison 2019/2020 am Spielbetrieb der ersten Bundesliga teilnahmen, hatten 13 Vereine ihre Lizenzspielerabteilung als AG, GmbH oder als GmbH & Co KGaA auf eine Kapitalgesellschaft ausgegliedert. Die übrigen fünf Vereine wurden weiterhin in ihrer ursprünglichen Rechtsform des e.V. geführt (vgl. Kicker, 2020). Von Seiten der DFL bestehen hinsichtlich der am Spielbetrieb teilnehmenden Kapitalgesellschaften aber dahingehend Einschränkungen, dass die Lizenz nur dann erteilt wird, wenn der Mutterverein mehrheitlich an ihr beteiligt ist, also mindestens 50 % der Stimmrechtsanteile zuzüglich eines weiteren Stimmrechts besitzt (vgl. DFL, 2019b, S. 9f.). Die Ausgliederung bedeutet für die Vereine aber nicht nur einen Zugang zu neuen Finanzierungsmöglichkeiten am Kapitalmarkt sowie die Einbindung verschiedener Stakeholder, sondern auch verschärfte Rechnungslegungs- und Publizitätspflichten (vgl. Littkemann et al., 2003, S. 416). Die kodifizierten Rechnungslegungsvorschriften für Kapitalgesellschaften gelten ebenso für Sportmannschaften und werden zusätzlich noch durch Vorgaben seitens der Verbände ergänzt, sodass sich ein umfassendes Rechnungslegungskonstrukt ergibt. Dieses soll im Kapitel 2.2 näher beleuchtet werden.

Nach der Betrachtung der Verbände und der Vereine als wichtige Organe im Fußball soll zum Abschluss dieses Kapitels der Fokus noch einmal auf die Fußballspieler selbst gelegt werden. Wie bereits in der Einleitung erwähnt, stellen die Spieler für die Vereine die wesentlichen Produktionsfaktoren dar. Sie können aber auch als Bestandteil eines Transfergeschäftes dienen, wodurch die Vereine unmittelbar hohe Einnahmen generieren können. Der Spielertransfer zwischen zwei Vereinen umfasst dabei eine Vielzahl von rechtlichen Aspekten und Bestimmungen seitens der Verbände, die von den beteiligten Parteien zu berücksichtigen sind. In Bezug auf eine akkurate Anwendung der Bewertungsvorschriften auf das Spielervermögen ist eine Analyse der Rechte und Pflichten der beteiligten Parteien bei Spielertransaktionen unerlässlich. Bei der Betrachtung der Spielertransfers sind zunächst einmal die verbandsrechtlichen Vorgaben im Hinblick auf die Vertragsgestaltung zu berücksichtigen. So haben die Verträge sowohl den Bestimmungen der Verbände, als auch den zivilrechtlichen Anforderungen zu genügen (vgl. Seip, 2017, S. 22). Das FIFA-Transferreglement beinhaltet die verbindlichen Bestimmungen bezüglich Status und Transfer von Spielern und legt somit die Grundregeln für internationale Transfers fest (vgl. Seip, 2017, S. 23). In Deutschland enthält die LOS die grundlegenden Bestimmungen für Spielertransfers, bezieht sich bei ihren Regelungen aber ebenfalls auf die Anordnungen des FIFA-Transferreglements (vgl. Pellens, Küting & Schmidt, 2019, S. 2474f.). Die LOS reguliert die Kriterien für die Erteilung der Spielerlizenzen, Spielberechtigungen sowie Wechselbestimmungen der Spieler (vgl. DFL, 2021c, S. 3). Um am Spielbetrieb der Lizenzligen teilnehmen zu können, benötigen die Spieler eine Lizenz, eine Spielerlaubnis sowie eine Spielberechtigung (vgl. Seip, 2017, S. 32). Die Lizenz erhält der Spieler nach § 1 LOS durch einen vereinsunabhängigen Lizenzvertrag mit der DFL. Diese wird unbefristet erteilt und bleibt auch bei einem Vereinswechsel bestehen, begründet aber kein Arbeitsverhältnis zwischen Spieler und DFL. Inhaltlich berechtigt sie die Spieler zur Teilnahme am Spielbetrieb der Lizenzligen und regelt die Unterwerfung der Spieler unter die Satzungen und Ordnungen der DFL (vgl. DFL, 2021c, S. 4). In § 2 LOS sind die Voraussetzungen für die Lizenzerteilung geregelt. Der Spieler hat den unterzeichneten Lizenzvertrag mit der DFL, einen Nachweis der Sporttauglichkeit sowie einem unter Vorbehalt der Lizenzerteilung gültigen Arbeitsvertrag mit einem lizenzierten Verein nachzuweisen (vgl. DFL, 2021c, S. 4). In diesem Arbeitsvertrag müssen gemäß § 6 LOS spezifische Vereinbarungen enthalten sein wie zum Beispiel (z.B.), dass der Spieler dem Verein die Rechte zur Nutzung und Verwertung an den vermögenswerten Bestandteilen seiner Persönlichkeit für den Zeitraum des Arbeitsverhältnisses zu Verfügung stellt (vgl. DFL, 2021c, S. 13f.). Der Spieler wird nach der Lizenzerteilung als Lizenzspieler angesehen und wäre berechtigt, mit einem lizenzierten Verein am Spielbetrieb teilzunehmen. Zur Ausübung dieser Spielberechtigung bedarf es aber gemäß § 13 LOS der Erteilung der Spielerlaubnis durch die DFL. Diese ist von dem Verein, der den Spieler verpflichtet zu beantragen und setzt voraus, dass zwischen Spieler und Verein ein wirksamer Vertrag vorliegt sowie das von Seiten des Spielers keine anderweitigen vertraglichen Verbindungen mit einem anderen Verein bestehen (vgl. DFL, 2021c, S. 18f.). Die Spielerlaubnis als exklusive Nutzungsmöglichkeit und damit auch die Spielberechtigung des Lizenzspielers wird dem beantragenden Verein für die Dauer des Arbeitsverhältnisses erteilt. Sie endet mit der Laufzeit des Arbeitsvertrages bzw. bei vorzeitiger Auflösung (vgl. DFL, 2021c, S. 19). Liegen alle notwendigen Unterlagen und Verträge vor, so sind diese in das Transfer-Online-Registrierungssystem des DFL hochzula-

den und werden dort abschließend geprüft. Entsprechen die Unterlagen den verbandsrechtlichen Vorgaben zu Erteilung der Spielerlaubnis, ist der Spielertransfer abgeschlossen (vgl. Seip, 2017, S. 35f.). Der Spieler ist registriert und darf im Spielbetrieb eingesetzt werden.

Im Rahmen eines Spielertransfers bestehen zudem zwischen den beteiligten Parteien verschiedene vertragliche Beziehungen. Der Transfervertrag steht dabei im Mittelpunkt der Betrachtung. Dieser beinhaltet die Modalitäten bezüglich des Spielerwechsels zwischen aufnehmenden und abgebenden Verein und veranschaulicht insbesondere im Hinblick auf die bilanzielle Behandlung, wofür Transferentgelte bezahlt werden bzw. worin die Leistungspflichten der Vereine bestehen (vgl. Pellens et al., 2019, S. 2475). Damit der Spieler einen neuen Arbeitsvertrag mit dem aufnehmenden Verein abschließen kann und somit für ihn spielberechtigt wäre, hat er zuerst den bestehenden Arbeitsvertrag mit dem abgebenden Verein aufzulösen. Zudem kommt aufgrund der regelmäßigen Beteiligung eines Spielervermittlers bei Transfergeschäften auch ein Vertrag zwischen dem Spielervermittler und dem aufnehmenden Verein zustande (vgl. Seip, 2017, S. 37). Die Abbildung 1 soll abschließend noch einmal die wesentlichen Aspekte eines Spielertransfers veranschaulichen.

2.2. Besonderheiten des Lizenzierungsverfahrens

Nachdem im vorherigen Kapitel bereits intensiver auf die Organe und auf die Lizenzerteilung an die Spieler eingegangen wurde, soll nun ein Blick auf die Lizenzerteilung an die Vereine und Kapitalgesellschaften geworfen werden. Die Voraussetzung für die Teilnahme an den Lizenzligen ist neben der sportlichen Qualifikation, wie auch bei den Spielern, an die Erteilung einer Lizenz durch den Verband geknüpft. Auf nationaler Ebene enthält das jährliche Lizenzierungsverfahren der DFL die verbandsrechtlichen Rahmenbedingungen für die Vereinslizenzierung. Maßgeblich für die Lizenzerteilung ist dabei die LO, die als Regelungswerk die Anforderungen an die Rechnungslegung und Prüfung enthält und gleichzeitig als Kontrollsystem dient (vgl. Thyll, 2004, S. 168). Die LO ist somit die Rechtsgrundlage für das Lizenzierungsverfahren und legt die Rechte und Pflichten der beteiligten Parteien fest. Das nationale Lizenzierungsverfahren berechtigt beim erfolgreichen Durchlaufen zur Teilnahme am Spielbetrieb der nationalen Lizenzligen, da die DFL darüber hinaus aber auch die Mindestanforderungen der UEFA in Bezug auf die Vereinslizenzierung vollständig übernommen hat, erwirbt der Lizenzbewerber auch gleichzeitig die Berechtigung zur Teilnahme am internationalen Wettbewerb (vgl. DFL, 2021b, S. 4).

Die Erteilung der Lizenz, das heißt (d.h.) der Abschluss eines Lizenzvertrages zwischen der DFL und den Vereinen, stellt eine vertragliche Vereinbarung dar. Sie bindet die Lizenzinhaber an die Ordnungen und Regelungen der DFL sowie des DFB (vgl. DFL, 2021b, S. 5). Das originäre Ziel des Lizenzierungsverfahrens ist die kurze und langfristige Sicherstellung des Spielbetriebes. Des Weiteren ist die DFL gemäß § Präambel der LO bestrebt, die Integrität des Wettbewerbes sowie die nationale und internationale Wettbewerbsfähigkeit durch das Lizenzierungsverfahren zu unterstützen (vgl. DFL, 2021b, S. 3). Zur Zielerreichung hat die DFL in § 2 LO Kriterien in verschiedenen Fachbereichen definiert, die die Lizenzbewerber für eine erfolgreiche Lizenzerteilung erfüllen müssen. Die Vereine haben somit neben finanziellen Kriterien auch sportliche, rechtliche oder auch infrastrukturelle Mindestkriterien im Lizenzierungsverfahren einzuhalten (vgl. DFL, 2021b, S. 6). Im Mittelpunkt der verschiedenen Anforderungen steht insbesondere die Einhaltung der finanziellen Kriterien. Dieses dient dem Verband zum einen zur Bewertung der wirtschaftlichen Leistungsfähigkeit der Lizenzbewerber und zum anderen als frühzeitige Eingriffsmöglichkeit bei einer finanziellen Schieflage eines teilnehmenden Vereins. Somit soll verhindert werden, dass eine Mannschaft während der Spielzeit aus finanziellen Gründen ausscheidet und somit Nachteile für Mitbewerber entstehen (vgl. Galli, 2012, S. 1633).

Zum Nachweis der wirtschaftlichen Leistungsfähigkeit haben die Vereine und Kapitalgesellschaften der DFL im Sinne der Lizenzierungsordnung einen Jahres- bzw. Konzernabschluss vorzulegen. Bei der Aufstellung dieser Abschlüsse ergeben sich einige branchenspezifische Besonderheiten. Dieses begründet sich daher, dass die Fußballvereine nicht allein die handelsrechtlichen Vorgaben des HGB zu berücksichtigen haben, sondern zusätzlich auch die aufgrund des Lizenzvertrages abgeschlossenen Bestimmungen und Vorgaben der Verbände hinsichtlich des Abschlusses zu beachten haben (vgl. Huwer, 2019, S. 273). Die vom Verband auferlegte Rechnungslegungspflicht gilt gleichermaßen für die Vereine mit der Rechtsform e.V., die ansonsten nur in geringem Umfang Prüfungs- und Publizitätspflichten unterliegen. So fordert die DFL von allen Lizenzbewerbern, unabhängig von ihrer Rechtsform, einen Jahresabschluss nach den handelsrechtlichen Vorschriften für Kapitalgesellschaften gemäß §§ 264 ff. HGB aufzustellen (vgl. Küting & Strauß, 2010, S. 796). Der einzureichende Jahres- bzw. Konzernabschluss erstellt nach den handelsrechtlichen Normen und erweitert um fußballspezifische Aspekte soll dabei Auskunft geben über die Vermögens-, Finanz- und Ertragslage sowie die Planung für die aktuelle und zukünftige Saison. Zu diesem Zweck hat der Lizenzbewerber der DFL vor und während der Spielzeit gemäß § 8 LO vergangenheitsbezogene Jahres- bzw. Zwischenabschlussinformationen sowie zukünftige Finanzinformationen offenzulegen (vgl. DFL, 2021b, S. 24). Die einzureichenden Unterlagen, die den Nachweis der wirtschaftlichen Leistungsfähigkeit erbringen sollen, dienen vorrangig der Überprüfung der Liquiditätssituation, da diese die Finanzierung des Spielbetriebes sicherstellt (vgl. Hierl & Weiß, 2016, S. 68-70).

Innerhalb dieser Abschlüsse haben die Lizenzbewerber die Verhältnisse ihrer Vermögens- und Schuldensituation in Bilanzform sowie eine Gegenüberstellung von Erträgen und Aufwendungen innerhalb einer Gewinn- und Verlustrechnung (GuV) darzustellen. Aufgrund der zu berücksichtigenden Vorgaben seitens der Verbände ergeben sich bei der Aufstellung der Bilanz und GuV einige Unterschiede in Bezug

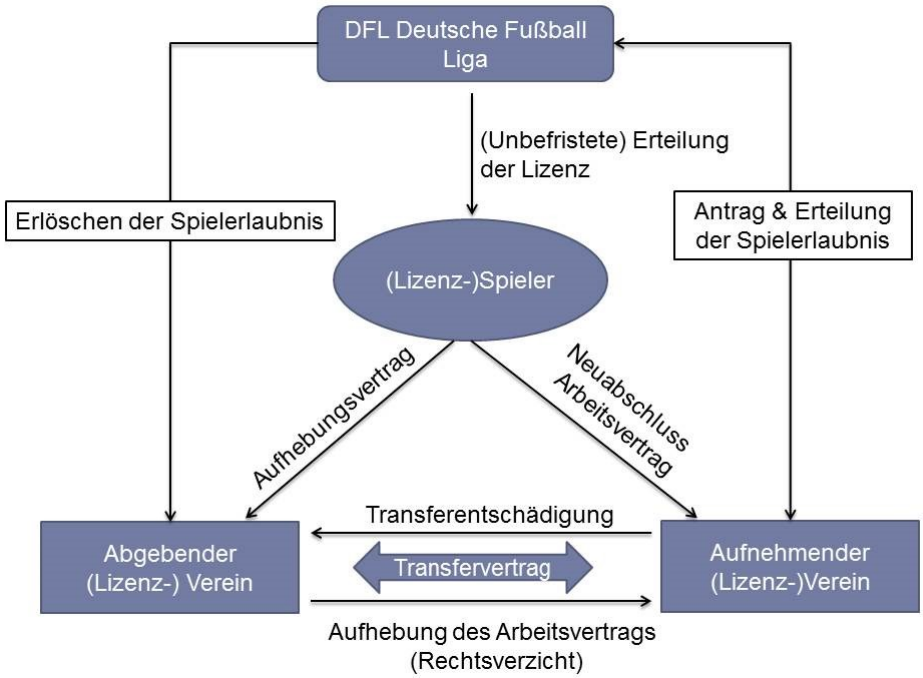


Abbildung 1: Beziehungsgeflechte eines Spielertransfers

Quelle: In Anlehnung an Pellens et al., 2019, S. 2476.

zur handelsrechtlichen Regelung. So ist die einzureichende Bilanz entsprechend der Bilanz für Kapitalgesellschaften nach § 266 HGB gegliedert, beinhaltet darüber hinaus aber mit der Position des Spielervermögens im Anlagevermögen einen fußballspezifischen Posten (vgl. Littkemann, Brast & Stübinger, 2002, S. 1198). Der Bilanzposten des Spielervermögens bzw. der Spielerwerte resultiert also allein aus den Vorschriften der Verbände und stellt im späteren Verlauf dieser Arbeit auch den zentralen Posten dar, wenn es um die bilanzielle Behandlung von Spielertransfers geht. In Bezug auf die GuV der Fußballvereine zeigt sich, dass sie zwar nach dem Standardschema gemäß § 275 Abs. 2 HGB im Gesamtkostenverfahren aufgestellt worden ist, aber eine detailliertere Aufschlüsselung einzelner Posten aufweist (vgl. Littkemann et al., 2002, S. 1999f.). So sind z.B. die Umsatzerlöse gegliedert nach Spielbetrieb, Werbung, mediale Verwertungsrechte oder auch Transfererlöse. Selbst die Abschreibung auf das Spielervermögen erhält eine eigene Position, wodurch deutlich wird, welche hohe Relevanz der Spielerbilanzierung zukommt. In Anbetracht der Erweiterung der handelsrechtlichen Rechnungslegung durch die Vorgaben der Verbände in Bezug auf die Abschlüsse und Prüfung ergibt sich ein umfangreiches Regelungswerk für die Vereine. Dieses wird noch einmal relevant, wenn es um den Ansatz und die Bewertung von Spielervermögen geht, da dort zum Teil unterschiedliche Vorschriften bestehen.

3. Bilanzielle Behandlung von Spielertransfers im Profi-Fußball

3.1. Handelsrechtliche Anforderungen an die Rechnungslegung

Die am Spielbetrieb teilnehmenden Vereine und Kapitalgesellschaften unterliegen wie in Kapitel 2.2 dargestellt, unterschiedlichen Rechnungslegungspflichten. Dazu sollte aber noch aufgeführt werden, dass die verbandsrechtlichen Vorschriften keinen vorrangigen Einfluss auf den handelsrechtlichen Jahresabschluss einnehmen, sondern als explizite Ergänzungen dienen. Dieses zeigt sich darin, dass die bilanziellen Vorgaben innerhalb der LO zumeist auf den handelsrechtlichen Normen des HGB beruhen (vgl. Huwer, 2019, S. 277f.).

Da die Lizenznehmer gemäß der LO unabhängig von ihrer Rechtsform zur externen Rechnungslegung verpflichtet sind, stellt sich die Frage, welchen Zweck der damit verbundene Jahresabschluss erreichen soll. Im Allgemeinen dient der Jahresabschluss mit seinen Bestandteilen der Informationsvermittlung über die Vermögens-, Finanz- und Ertragslage eines Unternehmens (vgl. Schäfer, 2019, S. 43). Insbesondere die unterschiedlichen Stakeholder haben ein Interesse an der aktuellen wirtschaftlichen Lage ihres Vereins. Daher ist eine Anforderung an die Rechnungslegung, die Rechnungslegungsadressaten so zu informieren, dass sie auf Basis dieser Informationen ihre Entscheidungen weitestgehend ohne Risiken treffen können. In der Fußballbranche sind z.B. Gläubiger, Werbepartner sowie die Verbände als externe Adressaten

anzusehen. In Bezug auf externe Adressaten sollen die Informationen der Berichterstattung als Entscheidungsinstrument für mögliche Beteiligungen, aber auch als Rechenschaft und Kontrolle z.B. für die Prüfung im Rahmen des Lizenzierungsverfahrens dienen (vgl. Schüler, 2016, S. 5f.). Die Informationen nutzen ebenfalls den internen Adressaten wie Anteilseignern und Mitgliedern. Sie können als Selbstinformation über die Vermögenslage verwendet werden, um damit zurückliegende Investitionsentscheidungen zu bewerten und zukünftige Investitionen planen zu können (vgl. Huwer, 2019, S. 7).

Zur Erfüllung dieser Zwecksetzungen bedarf es einiger Grundprinzipien, wie z.B. den Grundsätzen ordnungsgemäßer Buchführung (GoB) oder auch dem Gläubigerschutz. Die GoB, verankert in § 264 Abs. 2 HGB, fordern von den Kapitalgesellschaften, dass ihr Abschluss eine den tatsächlichen Umständen entsprechendes Abbild der Vermögens-, Finanz- und Ertragslage darstellt (vgl. Schüler, 2016, S. 7). So sind also zwingend die richtigen Informationen in die Jahresabschlüsse aufzunehmen, da ansonsten die Adressaten kein adäquates Urteil über die wirtschaftliche Lage des Unternehmens fällen können. Gerade im Hinblick auf Spielertransaktionen und die damit verbundenen hohen Transfersummen ist es unabdingbar, dass die Spielerwerte sachgerecht bilanziert werden, damit die aus der Bilanz gewährten Informationen den Adressaten ein wahrheitsgemäßes Bild vermitteln. Dieses ist auch im Sinne des Gläubigerschutzes, da eine unzureichende oder fehlerhafte Bilanzierung zu ungenauen Informationen über die Vermögenslage führt und möglicherweise das Risiko einer Fehlinvestition erhöht. Um den Gläubigern einen sicheren und vergleichbaren Einblick in die Rechnungslegung der Vereine zu gewähren, haben diese zudem die Stetigkeit der Bilanzansätze- und Bewertungen sicherzustellen sowie das nur die Spieler als Vermögensgegenstände in die Bilanz aufgenommen werden, die dem Verein wirtschaftlich zurechenbar sind und den Gläubigern folglich als Schuldendeckungspotenzial dienen können (vgl. Schüler, 2016, S. 11).

Angesichts der immer komplexer werdenden Konzernstrukturen, dadurch das viele Vereine mittlerweile neben ihrer Lizenzspielermannschaft auch andere Teilgebiete ausgliedern, zeigt sich die Notwendigkeit einer konsolidierten Rechnungslegung. Nur ein auf Mutterebene aufgestellter Konzernabschluss mit allen ausgegliederten Tochtergesellschaften würde einen den tatsächlichen Umständen entsprechendes Abbild der gesamten wirtschaftlichen Lage vermitteln und somit den Erfordernissen der GoB gerecht werden. (vgl. Küting & Strauß, 2011, S. 71). Die in diesem Kapitel aufgezählten Rechnungslegungsanforderungen seitens des Handelsrechts, aber auch der Verbände, sind bei der Untersuchung der bilanziellen Behandlung des Spielertransfers besonders zu berücksichtigen.

3.2. Historische Entwicklung der Bilanzierungspraxis von Spielervermögen

Die bilanzielle Behandlung von Transfergeschäften wurde in der Vergangenheit in der deutschen bilanzrechtlichen Literatur umfangreich diskutiert. Der Gegenstand der Diskussion, der auch Bestandteil verschiedener Urteile des BFH war, ist die Frage, ob Ablösezahlungen für Spieler überhaupt als Vermögensgegenstand bilanziert werden dürfen. In diesem Kapitel soll entlang der zentralen Urteile die Entwicklung des Transfersystems und die sich daraus resultierenden Auswirkungen auf die Bilanzierungsvorschriften aufgezeigt werden.

Bis Anfang der 1970er-Jahre existierten in der nationalen Satzung keine rechtlichen Rahmenbedingungen für die Bilanzierung von Spielervermögen. Daher kam es bei den Vereinen zu sehr differenzierten Herangehensweisen bei der Bilanzierung. So wurden Transferentschädigungen beispielsweise (bspw.) sofort als gewinnmindernder Aufwand verbucht oder auch als Posten im Umlaufvermögen ausgewiesen (vgl. Steiner & Gross, 2005, S. 532). Ein erstes Urteil vom BFH zu diesem Problemkreis wurde am 13. Mai 1987 gesprochen. Ausgehend von diesem Urteil waren die aufnehmenden Vereine zwar verpflichtet zur Zahlung einer Transferentschädigung an den abgebenden Verein, die Aktivierung dieser Zahlungen in der Bilanz wurde jedoch abgelehnt, da der BFH Bedenken an der Entgeltlichkeit des Erwerbs hatte (vgl. Euler, 2020, S. 767). Fünf Jahre später revidierte der BFH seine ursprüngliche Rechtshaltung und erließ mit dem Urteil vom 26.08.1992 eine neue Rechtsprechung. In dieser wurde festgehalten, dass die Transferentschädigungen die bei einem Wechsel gezahlt wurden, als aktivierungspflichtige Anschaffungskosten der Spielerlaubnis auszuweisen sind (vgl. BFH, Urteil vom 26.08.1992 – I R 24/91 BStBI. 1992 II, S. 977). Der BFH befand, dass es sich bei der Spielerlaubnis um ein konzessionsähnliches Recht nach § 266 Abs. 2 HGB handelt und dieses somit als immaterieller Vermögensgegenstand anzusetzen ist. Die Transferentschädigung „tritt wirtschaftlich an die Stelle einer Gegenleistung für die Übertragung eines nach allgemeinen Rechtsgrundsätzen nicht bestehenden “Rechts am Spieler““ (BFH, Urteil vom 26.08.1992 – I R 24/91 BStBI. 1992 II, S. 980). Zu berücksichtigen ist bei diesem Urteil, dass es zu einer Zeit getroffen wurde, in der Transferzahlungen grundsätzlich immer dann entrichtet wurden, wenn ein Spieler den Verein wechselte unabhängig davon, ob er einen auslaufenden Arbeitsvertrag besaß oder der Wechsel aus einem bestehenden Arbeitsverhältnis erfolgte (vgl. Galli, 2003, S. 811). Zudem war die Zahlung einer Transfersumme durch den aufnehmenden Verein eine Grundvoraussetzung für die Erlangung der Spielerlaubnis durch den damals zuständigen DFB. Aufgrund dieses engen Zusammenhanges zwischen der Erteilung der Spielerlaubnis und der Transferentschädigung rechtfertigte der BFH den Ansatz der Ablösezahlung als Anschaffungskosten für die Spielerlaubnis in der Bilanz (vgl. Kaiser, 2004, S. 1110). Das der BFH die Spielerlaubnis als vermögenswertes Recht ansah, war in der Literatur umstritten, da es eigentlich als Ordnungsrecht für den Spielbetrieb konzipiert war (vgl. Euler, 2020, S. 759).

Die Kritik an dieser Bilanzierungspraxis wurde mit dem Urteil des Europäischen Gerichtshofs (EuGH) vom 15.12.1995 dem sog. „Bosman-Urteil“ nochmals verstärkt. Durch dieses Urteil wurden die Rahmenbedingungen des Transfersystems

fundamental verändert. Der EuGH erklärte das bisherige Transfersystem, nachdem ein Spieler auch nach Vertragsablauf nur gegen eine Ablösezahlung den Verein verlassen durfte, für unzulässig vereinbar mit dem Recht der Freizügigkeit der Arbeitnehmer in Artikel 48 EG-Vertrag (vgl. Elter, 2011, S. 383). Infolge dieses Urteils und der daraus resultierenden Änderung der verbandsrechtlichen Transferbestimmungen war eine Ablösezahlung allein dann notwendig, wenn ein Spieler vor dem Ende seiner Vertragslaufzeit den Verein verlassen wollte. Der bisherige Verein hat damit bei einem Wechsel eines Spielers nach Vertragsablauf keine Ansprüche mehr gegenüber dem aufnehmenden Verein (vgl. Galli, 2003, S. 811). Die Vereine gingen allerdings dazu über, mit ihren Spielern längerfristige Verträge zu vereinbaren, um einen ablösefreien Wechsel zu verhindern und somit das System der Transferentschädigung aufrechtzuerhalten (vgl. Elter, 2011, S. 384). Der DFB hat auf die neue Rechtslage reagiert und die verpflichtende Voraussetzung der Zahlung einer Transferentschädigung für die Spielerlaubnis aus den Satzungen entfernt (vgl. Elter, 2011, S. 385). Aufgrund der Veränderung der verbandsrechtlichen Rahmenbedingungen und damit des Wegfalls eines maßgeblichen Grundes für die Aktivierung der Spielerlaubnis, nämlich dem entgeltlichen Erwerb, ist in der Literatur und auch bei den Vereinen die Frage aufgekommen, ob es überhaupt noch gerechtfertigt ist, Spielerwerte zu aktivieren oder ob sie direkt als Betriebsausgaben zu erfassen sind (vgl. BFH Urteil vom 14.12.2011 – I R 108/10 BStB 2012 II S. 239).

In einer Rechtsprechung des BFH vom 14.12.2011 wurde im Kern das BFH-Urteil von 1992 bestätigt und fortgeführt. Es beinhaltet allerdings die Abwandlung, dass die Ablösezahlungen die bei einem Vereinswechsel anfallen, „als Anschaffungskosten auf das immaterielle Wirtschaftsgut der exklusiven Nutzungsmöglichkeit “an dem Spieler“ zu aktivieren und auf die Vertragslaufzeit abzuschreiben“ sind (BFH Urteil vom 14.12.2011 – I R 108/10 BStB 2012 II S. 238). Die Ablösezahlungen sind somit weiterhin als Anschaffungskosten zu aktivieren, nun aber nicht mehr für die Spielerlaubnis selbst, sondern für die aus der Spielerlaubnis entstehende exklusive Nutzungsmöglichkeit an dem Spieler. Die Rechtsprechung von 2011 stellt auch die aktuelle Grundlage zur bilanziellen Behandlung des Spielervermögens dar. Dieses soll in den folgenden Kapiteln näher betrachtet werden.

3.3. Bilanzierung und Bewertung von Spielervermögen nach HGB

3.3.1. Bilanzansatz von Spielerwerten als immaterieller Vermögensgegenstand

Hinsichtlich der bilanziellen Behandlung von Spielerwerten ist zunächst die grundsätzliche Bilanzierungsfähigkeit zu überprüfen. Nach nunmehr gefestigter Rechtsprechung des BFH aus dem Jahre 2011, ist der Aktivierungsgegenstand in der Bilanz die aus der Spielerlaubnis resultierende exklusive Nutzungsmöglichkeit an dem Spieler. Die vom aufnehmenden Verein gezahlte Transferentschädigung dient dabei als Gegenleistung für die Übertragung der verbandsrechtlichen Spielerlaubnis und den damit garantierten Einsetzungs- und Vermarktungsfähigkeiten des Spielers (vgl. Galli, 2012, S. 677). Die Nutzungsmöglichkeit am Spieler ist zwar im wesentlichen Sinne auch ein Bestandteil des Arbeitsvertrages zwischen Verein und Spieler, geht aber aus Sicht des BFH aufgrund der verbandsrechtlichen Absicherung durch die Spielerlaubnis in Verknüpfung mit dem Transfersystem darüber hinaus und ist somit von den Leistungspflichten des Arbeitsvertrages gesondert als eigenständiger Vermögenswert anzusehen (vgl. BFH Urteil vom 14.12.2011 – I R 108/10 BStB 2012 II S. 242). Ausgehend von dieser Darstellung wird ebenfalls ersichtlich, dass nicht wie zum Teil vermutet der Spieler selbst als Vermögensgegenstand angesetzt wird, sondern das aus der Spielerlaubnis resultierende Nutzungsrecht bzw. der vermögensmäßige Vorteil. In der Rechtsprechung wird gesondert darauf hingewiesen, dass es nicht zu einer Bilanzierung des Menschen als Humankapital kommt, da dieses gegen die im Grundgesetz geschützte Menschenwürde verstoßen würde (vgl. BFH Urteil vom 14.12.2011 – I R 108/10 BStB 2012 II S. 241).

Nachdem nun herausgestellt wurde, was generell als Bilanzansatz in Frage kommt, ist als Nächstes zu klären, inwieweit die Spielerlaubnis mit ihrem exklusiven Nutzungsrecht die Voraussetzungen eines immateriellen Vermögensgegenstandes erfüllt und folglich auch angesetzt werden kann. Zu berücksichtigen ist dabei, dass es sich aus Sicht des BFH bei der Spielerlaubnis um ein konzessionsähnliches Recht, d.h. um die Übertragung eines Nutzungsrechts und somit um einen immateriellen Vermögensgegenstand nach § 266 Abs. 2 HGB handelt (vgl. BFH, Urteil vom 26.08.1992 – I R 24/91 BStBI. 1992 II, S. 978). Zur Feststellung der Bilanzierungsfähigkeit werden die notwendigen Vermögensgegenstandseigenschaften, die sich aus dem Handelsrecht, aber auch aus der Sicht des BFH ergeben, näher betrachtet. Ein zu aktivierender immaterieller Vermögensgegenstand muss demnach unter der Voraussetzung des wirtschaftlichen Eigentums im Sinne der abstrakten Aktivierungsfähigkeit selbstständig verkehrsfähig und bewertbar sowie in Bezug auf die konkrete Ansatzfähigkeit entgeltlich erworben sein (vgl. Schubert & Huber, 2020, § 247, Rn. 389-390).

Die Tatbestandsvoraussetzung der Verkehrsfähigkeit im Sinne von Einzelveräußerbarkeit bzw. –verwertbarkeit, d.h. die Möglichkeit der einzelnen Übertragung bzw. Veräußerung eines bestimmten wirtschaftlichen Wertes wird durch den BFH bejaht (vgl. Schubert & Huber, 2020, § 247, Rn. 390). Nach Auffassung des BFH ist es ausreichend, dass der abgebende Verein auf die Spielerlaubnis verzichtet, um dem aufnehmenden Verein die Möglichkeit für eine Neuerteilung zu erlauben und somit das exklusive Nutzungsrecht auf den neuen Verein übergeht (vgl. BFH Urteil vom 14.12.2011 – I R 108/10 BStB 2012 II S. 242). Die Vereine können für die Dauer des Arbeitsverhältnisses mitbestimmen, ob der Spieler den Verein frühzeitig gegen eine Ablösezahlung verlassen darf und folglich, ob der Nutzen aus der Spielerlaubnis wirtschaftlich übertragen wird, wodurch das Kriterium der abstrakten Übertragbarkeit begründet wird (vgl. BFH Urteil vom 14.12.2011 – I R 108/10 BStB 2012 II S. 242). Als weitere Voraussetzung für den Ansatz eines Vermögensgegen-

standes ist die selbstständige Bewertbarkeit erforderlich. Sie folgt aus dem Grundsatz der Einzelbewertung, nach dem der einzelne Vermögenswert in Bezug auf das Gesamtvermögen abgrenzbar ist und sich ein gesonderter Wert ermitteln lässt (vgl. Schubert & Huber, 2020, § 247, Rn. 390). Der BFH bejaht auch dieses Kriterium, indem er dazu ausführt, dass die exklusive Nutzungsmöglichkeit am Spieler aufgrund der unterschiedlich hohen Transferzahlungen, die zwischen den Vereinen abgesprochen werden, einen eigenständigen Wert besitzt (vgl. BFH Urteil vom 14.12.2011 – I R 108/10 BStB 2012 II S. 243).

Bei der Überprüfung der konkreten Ansatzfähigkeit für einen immateriellen Vermögensgegenstand sind neben der Bedingung des entgeltlichen Erwerbs auch etwaige Aktivierungsgebote,- wahlrechte und -verbote gesondert zu berücksichtigen. Dieses ist dahingehend relevant, da für entgeltlich erworbene immaterielle Vermögensgegenstände gemäß § 246 Abs. 1 HGB eine Aktivierungspflicht besteht, während für originär erworbene Vermögenswerte gemäß § 248 Abs. 2 HGB ein Ansatzverbot bzw. -wahlrecht besteht (vgl. Schubert & Huber, 2020, § 266, Rn. 59-60). Ein entgeltlicher Erwerb setzt verschiedene Konkretisierungen voraus. So muss es sich zum einen um einen abgeleiteten Erwerb auf dem Markt handeln und zum anderem müssen einmalig festgelegte Anschaffungskosten für den Erwerb eines Vermögenswertes von einem Dritten vorliegen (vgl. Schubert & Huber, 2020, § 247, Rn. 390-391). Daran anknüpfend sollte der immaterielle Vermögenswert zudem als Bestandteil im Erwerbsvorgang beteiligt sein, sodass Leistung und Gegenleistung einen direkten Bezug zueinander aufweisen (vgl. Fross, 2012, S. 196). Von Seiten des BFH wird dazu ausgeführt, dass obwohl die Spielerlaubnis vom Verband erteilt wird, die Transferentschädigungen aber an den abgebenden Verein gezahlt werden, zwischen den beiden Handlungen ein enger Veranlassungszusammenhang besteht, sodass man von einem abgeleiteten Erwerb ausgehen kann (vgl. BFH, Urteil vom 26.08.1992 – I R 24/91 BStBI. 1992 II, S. 980). Dies bestätigte der BFH auch in seiner Rechtsprechung von 2011. Dort bekräftigte er, dass wenn ein Verein im Zuge eines Vereinswechsels ein Entgelt in Form einer Ablösezahlung an den abgebenden Verein zahlt, es sich dabei um Anschaffungskosten für den Erwerb des exklusiven Nutzungsrechts am Spieler handelt (vgl. BFH Urteil vom 14.12.2011 – I R 108/10 BStB 2012 II S. 243).

Daraus lässt sich folgern, dass entgeltlich erworbene Spieler nach den Vorschriften des HGB und des BFH als immaterielle Vermögensgegenstände in Höhe der gezahlten Ablösesumme in der Bilanz zu aktivieren sind. Demgegenüber dürfen ablösefreie Spieler nicht in der Bilanz aktiviert werden. Sie erfüllen zwar die Kriterien der abstrakten Aktivierungsfähigkeit, ihnen fehlt es aber an der Voraussetzung des entgeltlichen Erwerbs der Nutzungsmöglichkeit mittels Anschaffungskosten. Für bestimmte selbsterstellte immaterielle Vermögensgegenstände des Anlagevermögens besteht seit der Umsetzung des Bilanzmodernisierungsgesetzes (BilMoG) im Jahre 2009 gemäß § 248 Abs. 2 HGB eine Möglichkeit zum Ansatz. Dieses stellt eine Besonderheit dar, da bis dahin nur entgeltlich erworbene immaterielle Vermögenswerte aktivierungsfähig waren (vgl. Förschle & Usinger, 2010, § 248, Rn. 11). Im Kontext von Fußballvereinen bedeutet dies, dass für Jugend- und Nachwuchsspieler ein Aktivierungswahlrecht entstanden ist. Dieses setzt jedoch voraus, dass die Kosten für die einzelnen Spieler genau dokumentiert und zugeordnet werden können. Von Seiten der UEFA wird die Aktivierung der Ausbildungskosten für Nachwuchsspieler allerdings untersagt, womit in den meisten Fällen bei den Vereinen von einem Bilanzansatz abgesehen wird (vgl. Handschin, 2012, S. 1284).

Für den handelsrechtlichen Bilanzansatz kommt es daher abschließend auf die Art des Spielerwechsels an, ob ein Spieler ablösefrei oder gegen eine Transferzahlung den Verein wechselt. Nachdem in diesem Kapitel die Frage geklärt wurde, was generell in der Bilanz zum Ansatz kommt, soll im folgenden Kapitel veranschaulicht werden, wie sich die Höhe der Anschaffungskosten für das exklusive Nutzungsrecht zusammensetzt.

3.3.2. Zugangsbewertung von Spielerwerten

Nach Klärung der Frage, ob es zu einem Ansatz eines immateriellen Vermögensgegenstandes in der Bilanz der Vereine kommt, stellt sich nun die Frage, in welcher Höhe die Nutzungsmöglichkeit am Spieler im Zugangszeitpunkt zu aktivieren ist. Gemäß den handelsrechtlichen Vorschriften veranschaulichen die Anschaffungskosten den Ausgangspunkt der Zugangsbewertung. Wird ein Spieler mittels einer Transferzahlung aus seinem laufenden Vertrag bei einem anderen Verein herausgekauft, so stellen die gezahlten Entgelte an den abgebenden Verein die Anschaffungskosten des immateriellen Vermögensgegenstandes dar. Maßgeblich für die Zugangsbewertung von entgeltlich erworbenen Spielerwerten ist, dass sie gemäß § 253 Abs. 1 HGB höchstens mit ihren Anschaffungskosten anzusetzen sind. Die Anschaffungskosten stellen somit die Wertobergrenze im Rahmen der Zugangsbewertung dar. Nach handelsrechtlicher Definition, die auch vom BFH in seiner Rechtsprechung übernommen wurde, handelt es sich bei den Anschaffungskosten gemäß § 255 Abs. 1 Satz. 1 HGB um „die Aufwendungen, die geleistet werden, um einen Vermögensgegenstand zu erwerben und in einen betriebsbereiten Zustand zu versetzen, soweit sie dem Vermögensgegenstand einzeln zugeordnet werden können“ (BFH Urteil vom 14.12.2011 – I R 108/10 BStB 2012 II S. 243). Der Begriff der Anschaffungskosten umfasst dabei neben dem reinen Anschaffungspreis auch Preisminderungen, Nebenkosten sowie nachträglich anfallende Kosten (vgl. Schubert & Gadek, 2014, § 255, Rn. 50). Anschaffungsnebenkosten und nachträglich anfallende Kosten werden aber nur dann zu den gesamten Anschaffungskosten hinzugerechnet, wenn sie in einem direkten wirtschaftlichen Zusammenhang mit der Anschaffung anfallen (vgl. Schubert & Gadek, 2014, § 255, Rn. 70-75). Der Zeitpunkt der Aktivierungspflicht der Aufwendungen beginnt generell dann, wenn Handlungen zum Erwerb eines bestimmten Vermögenswertes vorgenommen werden (vgl. Schubert & Gadek, 2014, § 255, Rn. 33). In der Fußballbranche können der Abschluss

eines Transfervertrages und die Erteilung der Spielerlaubnis durch den Verband zumeist als Aktivierungszeitpunkt angesehen werden, wobei Kosten auch bereits vor oder nach diesem Zeitpunkt anfallen können (vgl. Schiffers & Feldgen, 2015, S. 504). Dass es im Zusammenhang mit der Zugangsbewertung bei Spielerwerten zu anderen Kostenkategorien kommt, als bspw. bei einem Kauf einer Maschine, soll in der Abbildung 2 veranschaulicht werden.

Die Ablösezahlung die geleistet wird, um den Spieler aus einem laufenden Vertrag zu erwerben, stellt die wesentliche Komponente der Anschaffungskosten dar. Zu den aktivierbaren Anschaffungsnebenkosten zählen im Fußballsektor insbesondere die Spielervermittlerhonorare, Handgelder und Förderungsentschädigungen. Da mittlerweile in vielen Vertragsverhandlungen ein Spielervermittler mitwirkt, um eine Einigung zwischen den beteiligten Parteien herbeizuführen, stellt sich folglich die Frage, wie die Provision der Vermittler bilanziell zu handhaben ist.

Der BFH hat dazu in seiner Rechtsprechung von 2011 ausgeführt, dass es sich bei der Provision an Spielervermittler um aktivierungspflichtige Anschaffungskosten handelt, soweit diese im Zusammenhang mit einem entgeltlichen Spielerwechsel erfolgen (vgl. BFH Urteil vom 14.12.2011 – I R 108/10 BStB 2012 II S. 244). Der BFH ordnet die Provisionszahlung als Nebenkosten ein, da sie eine erwerbsbezogene Aufwendung seitens des aufnehmenden Vereins darstellt, um die Nutzungsmöglichkeit am Spieler zu erlangen. Unerheblich ist dabei, dass der abgebende Verein nicht an den Provisionszahlungen beteiligt ist, da die Anschaffungskosten nicht zwangsläufig ein Bestandteil der Entschädigung sein muss, sondern auch an Dritte geleistet werden können (vgl. BFH Urteil vom 14.12.2011 – I R 108/10 BStB 2012 II S. 245). Zu berücksichtigen ist dabei, dass die Auszahlung der Vermittlungsprovision häufig unter aufschiebende Bedingungen gestellt wird, wie z.B. das ein Arbeitsvertrag mit dem Spieler nicht vorzeitig aufgelöst wird (vgl. Schiffers & Feldgen, 2015, S. 504). In diesem Fall sind die bedingten Beraterhonorare als aufschiebend bedingte Verbindlichkeiten anzusehen und sind bei Erfüllung der Konditionen zu passivieren sowie als nachträgliche Anschaffungskosten auszuweisen (vgl. Schubert, 2020, § 247, Rn. 224).

Als Anschaffungsnebenkosten können auch Handgelder sog. „Signing Fees“ an die Spieler in Betracht kommen. Diese dürfen jedoch nur aktiviert werden, wenn sie im Rahmen eines ablösepflichtigen Transfers dazu verwendet werden, den Spieler von einem Vertragsabschluss zu überzeugen, aber an den Spieler selbst keine sonstigen Gegenleistungen stellen (vgl. Schiffers & Feldgen, 2015, S. 502). Ist die Zahlung der Handgelder jedoch an eine Gegenleistung geknüpft oder auch als vorausgezahltes Gehalt anzusehen, steht das Handgeld nicht im direkten Zusammenhang mit dem Erwerb des Spielervermögens und ist somit vom Verein als aktiver Rechnungsabgrenzungsposten (RAP) anzusetzen (vgl. Schiffers & Feldgen, 2015, S. 502).

Anschaffungskosten können auch zeitlich nachgelagert zum Abschluss des Spielervertrages anfallen. Dieses kann im Profi-Fußball z.B. eine vertragliche Vereinbarung von Zusatzzahlungen sein, die sich bei einem festgelegten sportlichen Erfolg des Spielers nachträglich an den abgebenden Verein ergeben (vgl. Schröder & Specht, 2020, S. 961). Bei Eintritt dieser Vereinbarungen sind die nachträglich anfallenden Kosten rückwirkend zu aktivieren und auf den aktuellen Bilanzwert zuzuschreiben.

Hinsichtlich der Zugangsbewertung bei entgeltlich erworbenen Spielern zeigt sich ein umfassendes Konstrukt fußballspezifischer Anschaffungskosten. In Bezug auf ablösefreie Spieler wird zwar keine Ablösesumme als Kaufpreis gezahlt, es entstehen dem Verein aber dennoch Kosten ähnlich zum entgeltlichen Transfer. So fallen ebenfalls Vermittlerprovisionen und Handgeldzahlungen an die Spieler an. Eine Qualifikation dieser Kosten als Anschaffungskosten scheidet jedoch nach der Rechtsprechung allein deswegen aus, da sie keinen erforderlichen entgeltlichen Erwerb als Gegenleistung für die erlangte Nutzungsmöglichkeit darstellen (vgl. Schubert & Gadek, 2014, § 255, Rn. 72). So hat auch der BFH beschlossen, dass Spielervermittlerprovisionen im Zusammenhang mit einem ablösefreien Vereinswechsel nicht aktiviert werden dürfen und diese Aufwendungen allein nicht zu einem entgeltlichen Erwerb des Spielers führen (vgl. BFH Urteil vom 14.12.2011 – I R 108/10 BStB 2012 II S. 245). Die gezahlten Beraterhonorare und Handgelder sind somit mangels Anschaffungsvorgang als laufende Betriebsausgaben anzusetzen, wobei bei den gezahlten Handgeldern geprüft werden sollte, inwieweit ein RAP, d.h. eine Vorauszahlung für eine noch zu erbringende Leistung vorliegt (Schiffers & Feldgen, 2015, S. 502).

In Zusammenhang mit einer Spielerleihe, bei der ein Spieler temporär abgegeben wird, ist zu klären, wie mit einer etwaig gezahlten Leihgebühr in bilanzieller Hinsicht umzugehen ist. Bei der sog. Spielerleihe kommt es zu einer temporären Aussetzung des Arbeitsverhältnisses beim verleihenden Verein. Der ausleihende Verein übernimmt rechtlich die Arbeitgeberstellung und beantragt eine neue Spielerlaubnis, sie besitzt solange eine Gültigkeit, bis die vereinbarte Leihdauer abgelaufen ist und der Spieler zu seinem alten Verein zurückkehrt (vgl. Schiffers & Feldgen, 2015, S. 506). Wirtschaftlicher Eigentümer und somit die Herrschaftsmacht über den immateriellen Vermögensgegenstand bleibt der ausleihende Verein, da er die langfristigen Chancen einer Wertsteigerung, aber auch die Risiken einer Wertminderung trägt. Zudem verbleiben die Transferrechte und damit die selbstständige Verwertbarkeit beim verleihenden Verein (vgl. Weber, 2016, S. 250f.). Demnach bleibt der Spielerwert auch während der Spielerleihe ausschließlich beim verleihenden Verein bilanziert und ist dort planmäßig folgezubewerten. Die im Rahmen der Spielerleihe gezahlte Leihgebühr führt also nicht zu einem immateriellen Vermögensgegenstand beim leihenden Verein, sondern ist direkt aufwandswirksam als Transferaufwand in der GuV zu erfassen. Zu einem abweichenden Ergebnis kann es kommen, wenn innerhalb des Leihgeschäftes eine Kaufoption verankert ist. Diese räumt dem ausleihenden Verein die Möglichkeit ein, den Spieler nach Ende der Leihdauer gegen eine vorher festgelegte Ablösezahlung fest zu verpflichten. In diesem Fall würde der

Allgemeine Anschaffungskosten	Fußballspezifische Anschaffungskosten
Anschaffungspreis	Ablösezahlung
- Anschaffungspreisminderung	Verletzung
+ Anschaffungsnebenkosten	Provision Spielervermittler, Handgelder, Ausbildungsentschädigung
+ Nachträgliche Anschaffungskosten	Vereinbarungen hinsichtlich sportlicher Ziele
= Aktivierte Anschaffungskoten	= Aktivierte Anschaffungskosten eines Spielers

Abbildung 2: Anschaffungskosten im Profi-Fußball

Quelle: In Anlehnung an Wulf & Bosse, 2011, S. 84.

leihende Verein an einer Leistungssteigerung des Spielers partizipieren. Demgegenüber würde aber im Falle eines Leistungsabfalls der ausleihende Verein von der Kaufoption keinen Gebrauch machen und der verleihende Verein das Risiko tragen. Aufgrund der alleinigen Entscheidungsmacht über den Spielertransfer deutet vieles auf einen bilanziellen Ansatz beim ausleihenden Verein mit Anschaffungskosten in Höhe der vorher vereinbarten Leihgebühr hin. (vgl. Weber, 2016, S. 252). Um den letztendlichen wirtschaftlichen Eigentümer zu definieren, bedarf es einer individuellen Berücksichtigung der einzelnen Vertragsbestandteile des Leihgeschäftes.

Nachdem in diesem Kapitel aufgezeigt wurde, mit welchem Buchwert die Spielerwerte in die Bilanz eingehen, soll im folgenden Kapitel verdeutlicht werden, wie mit dem Buchwert im Zeitablauf umzugehen ist.

3.3.3. Folgebewertung von Spielerwerten

Im Hinblick auf die Folgebewertung ist zwischen einer planmäßigen und einer außerplanmäßigen Abschreibung auf das Spielervermögen zu unterscheiden.

Planmäßige Abschreibungen erfolgen grundsätzlich jährlich auf alle immateriellen Vermögenswerte, deren Nutzungsdauer begrenzt ist oder die abnutzbar sind (vgl. Schubert & Andrejewski, 2020, § 253, Rn. 384). Die Nutzungsdauer, über die die aktivierten Anschaffungskosten planmäßig abzuschreiben sind, bemisst sich anhand verschiedener individueller rechtlicher, wirtschaftlicher und sonstiger Faktoren. So kann z.B. der voraussichtlich begrenzte Zeitraum eines durch Konzession eingeräumten Nutzungsrechtes als Nutzungsdauer angesetzt werden (vgl. Schubert & Andrejewski, 2020, § 253, Rn. 387). In der Fußballbranche ergeben sich die Bestimmung der Nutzungsdauer und die damit zusammenhängenden Fragestellungen aus der Rechtsprechung des BFH. Dort wird ausgeführt, dass es sich bei der Spielerlaubnis um einen abnutzbaren Vermögenswert handelt, der mit Ablauf des Arbeitsverhältnisses erlischt und sich deshalb die betriebsgewöhnliche Nutzungsdauer nach der rechtlichen Laufzeit des Arbeitsvertrages richtet (vgl. BFH, Urteil vom 26.08.1992 – I R 24/91 BStBI. 1992 II, S. 983). Das aus der Spielerlaubnis resultierende Nutzungsrecht ist somit mit seinen Anschaffungskosten planmäßig über die Laufzeit des Arbeitsvertrages abzuschreiben. Als Abschreibungsmethode wird auf die lineare Abschreibung zurückgegriffen, da sich der Werteverzehr eines Spielers zwangsläufig nicht verlässlich bestimmen lässt (vgl. Schubert & Andrejewski, 2020, § 253, Rn. 385). Der Abschreibungsbeginn wird in der Regel auf den Anschaffungszeitpunkt festgelegt, bezogen auf den Spieler also der Zeitpunkt des Vertragsabschlusses (vgl. Schubert & Gadek, 2014, § 255, Rn. 32). Das jährliche Abschreibungsvolumen bemisst sich in Abhängigkeit von der Vertragsdauer aus der Höhe der gesamten Anschaffungskosten. In seinem Urteil aus dem Jahre 1992 hat der BFH zudem beschlossen, dass der aufnehmende Verein zukünftige Ansprüche auf Transferentschädigungen durch einen Verkauf des Spielers bei der Bestimmung der Nutzungsdauer und dem Abschreibungsvolumen außer Acht lassen soll. Nach Ansicht des BFH handelt es sich dabei um zukünftige Gewinne, die aufgrund des Realisationsprinzips noch nicht berücksichtigt werden dürfen (vgl. BFH, Urteil vom 26.08.1992 – I R 24/91 BStBI. 1992 II, S. 983). Grundsätzlich dürfen Restwerte aber durchaus in den Abschreibungsplan mit aufgenommen werden, solange sie die anfänglichen Anschaffungskosten nicht übersteigen. Nach geltender Rechtsprechung dürfen Restwerte dann außer Acht gelassen werden, wenn sie im Vergleich zu den Anschaffungskosten wertmäßig unbedeutend sind (vgl. Euler, 2020, S. 764). Ob dieses bei stetig steigenden Ablösesummen im Bereich des Profi-Fußball der Fall ist, scheint fraglich.

Bei der Bestimmung der Nutzungsdauer sind zudem etwaige Optionen auf eine im Arbeitsvertrag festgehaltene Vertragsverlängerung zu berücksichtigen. Gemäß der Rechtsprechung des BFH ist bei solchen Verlängerungsoptionen nach allgemeinen Wahrscheinlichkeitsgesichtspunkten zu prüfen,

ob die Option vom Verein gezogen wird (vgl. BFH, Urteil vom 26.08.1992 – I R 24/91 BStBI. 1992 II, S. 983). Diese Regelung überlässt den Vereinen einen erheblichen Spielraum bei der Auslegung der Nutzungsdauer und der Aufteilung des Abschreibungsaufwandes. Macht ein Verein von dieser Option Gebrauch, verlängert sich die Nutzungsdauer und der Restbuchwert wird auf die neue Nutzungsdauer verteilt abgeschrieben.

Die planmäßige Abschreibung teilt somit die Anschaffungskosten als gewinnmindernden Aufwand auf die einzelnen Vertragsjahre auf. Sie spiegelt auch gleichzeitig die Folgebewertung des Spielerwertes zu fortgeführten Anschaffungskosten wider. Diese stellen auch die Wertobergrenze dar, da auch im Falle eines höheren Marktwertes eines Spielers innerhalb der Nutzungsdauer keine Neubewertung erfolgen darf, die höher ist als der planmäßig abgeschriebene Buchwert (vgl. Hierl & Weiß, 2016, S. 66).

Bei der Betrachtung der planmäßigen Abschreibungen werden nur solche Aspekte berücksichtigt, die bei einer gewöhnlichen Nutzung des Vermögensgegenstandes zu einer Wertminderung führen. Kommt es zu außergewöhnlichen Abnutzungen, ist dieses im Rahmen der außerplanmäßigen Abschreibung zu berücksichtigen. So sind Unternehmen, aber auch die Mannschaften im Profi-Fußball dazu verpflichtet, jährlich zum Jahresabschluss einen Niederstwerttest durchzuführen. Dabei werden die Wertansätze ihrer Vermögensgegenstände auf Werthaltigkeit geprüft. Eine außerplanmäßige Abschreibung auf immaterielle Vermögenswerte ist gemäß § 253 Abs. 3 Satz. 5 HGB dann durchzuführen, wenn der beizulegende Wert niedriger ist als der sich am Abschlussstichtag ergebende Buchwert und es sich um eine voraussichtlich dauernde Wertminderung handelt (vgl. Schubert & Andrejewski, 2020, § 253, Rn. 396). Die Vereine haben demnach zu prüfen, inwieweit ihr Spielervermögen dauerhaft wertgemindert ist und ob sie dieses folglich zu einem niedrigeren Wert anzusetzen haben. Grundsätzlich ist dabei zu klären, was unter dem Begriff dauerhaft bei einer Wertminderung zu verstehen ist. Da der Begriff „dauerhaft“ keinen gesetzlich definierten Zeitraum darstellt, kam es in der Literatur sowie von Seiten der Rechtsprechung zu einer einheitlichen Auslegung dieses Begriffes. Von einer dauerhaften Wertminderung wird dann ausgegangen, wenn der beizulegende Wert für einen wesentlichen Teil der Restnutzungsdauer unterhalb des planmäßigen Restbuchwertes liegt (vgl. Littkemann, Schulte & Schaarschmidt, 2005, S. 663). Von einer dauerhaften Wertminderung wird daher zumeist bei einer über die Hälfte der Restnutzungsdauer, bezogen auf die Spieler die Restvertragslaufzeit andauernden Wertminderung ausgegangen (vgl. Schubert & Andrejewski, 2020, § 253, Rn. 396).

Die Gründe für eine solche dauerhafte Wertminderung und somit einer außerplanmäßigen Abschreibung sind nicht hinreichend konkretisiert. Im Fußballsektor lässt sich zwischen zwei möglichen Gründen unterscheiden. Eine schwerwiegende und langwierige Verletzung eines Spielers ist eine mögliche Ursache. Fällt der Spieler für eine längerfristige Zeit aus, ist die individuelle Leistungsfähigkeit des Spielers für den Verein wertlos. Bei schwerwiegenden Verletzungen folgen zumeist weitere Ausfälle und aus einer nur vorübergehenden Wertminderung kann sich gegebenenfalls eine dauerhafte Wertminderung ergeben, wenn dem Spieler die Sportinvalidität droht (vgl. Littkemann et al., 2005, S. 664). Für den Verein nimmt durch die Verletzungen der Wert der Nutzungsmöglichkeit ab und auch die in der Zukunft erzielbaren Ablösesummen verringern sich, sodass eine Abschreibung auf einen niedrigeren Buchwert als logische Schlussfolgerung erscheint. Als ein weiterer Grund kann eine Suspendierung des Spielers herangezogen werden. Wird ein Spieler vom Trainings- und Spielbetrieb ausgeschlossen, kommt es nicht nur zu einem Imageschaden des Spielers, sondern auch zu einer Verschlechterung bei Transferverhandlungen über einen möglichen Abgang (vgl. Littkemann et al., 2005, S. 665). Wird ein Einsatz des Spielers im Spielbetrieb aufgrund der Suspendierung unwahrscheinlich, so kann es gegebenenfalls von Seiten des Vereins zu einer dauerhaften Wertminderung des Spielervermögens kommen. Kommt es hingegen nur zu einer vorübergehenden Wertminderung, darf keine außerplanmäßige Abschreibung vorgenommen werden. Während bei einem Wegfall der Gründe für eine dauerhafte Wertminderung eine Zuschreibung erfolgt bis hin zu der Summe, die der Spielerwert ohne die außerplanmäßige Abschreibung entsprochen hätte (vgl. Köster & Ehemann, 2012, S. 156).

In der Fußballbranche führen der Niederstwerttest und die damit zusammenhängenden außerplanmäßigen Abschreibungen zu erheblichen Bewertungsproblemen. Dieses liegt daran, dass bei der Beurteilung der dauerhaften Wertminderung ein beizulegender Wert als Vergleichsmaßstab für den Buchwert erforderlich ist. Ein solcher Wert ist im Bereich des Profi-Fußball allerdings nicht einheitlich definiert und erschwert somit den Vergleich. In der Literatur werden sowohl die Betrachtung einzelner wertbildender Faktoren wie z.B. der Gesundheitszustand, die individuelle sportliche Leistungsfähigkeit und damit der Wert des Spielers für den Verein zur Ermittlung eines möglichen Zeitwertes in Betracht gezogen (vgl. Teschke, Knipping & Sundheimer, 2012, S. 1142).

Es wird aber auch vorgeschlagen, unter Rückgriff auf die Internet-Datenbank transfermarkt.de den beizulegenden Wert aus den dort veröffentlichten Marktwerten abzuleiten. Zwar stellen diese Marktwerte im Verhältnis zu den Buchwerten keine realen Werte dar, da sie nicht auf einer vollständig objektiv nachvollziehbaren Bewertungsmethode beruhen. Sie können aber zumindest als Indikator für die Entwicklung des Vermögenswertes eines Spielers herangezogen werden (vgl. Weber, 2020, S. 322).

Es lässt sich festhalten, dass die Folgebewertung und die damit einhergehenden Abschreibungen dazu dienen sollen, dass Spielervermögen mit ihrem tatsächlichen Wert zu erfassen und die Wertminderungen auf die Vertragslaufzeit zu verteilen. Aufgrund der fehlenden Konkretisierung außerplanmäßiger Abschreibungen ergibt sich auf Seiten der Vereine keine konsistente Bilanzierungspraxis und den Vereinen wird gleichzeitig ein beträchtlicher bilanzpolitischer Spiel-

raum gelassen.

3.4. Veräußerung des Spielers vor Vertragsende

Während in den vorangegangenen Kapiteln aufgezeigt wurde, wie die bilanzielle Behandlung eines Spielers beim Zugang erfolgt, soll in diesem Kapitel veranschaulicht werden, wie ein Abgang des Spielers und die damit verbundenen Einnahmen bilanziell zu würdigen sind. Grundsätzlich gehört das Spielervermögen als abnutzbarer immaterieller Vermögensgegenstand zum Anlagevermögen, da es dazu bestimmt ist, dem Verein dauerhaft zu dienen. Wird der Verkauf eines Spielers geplant, ist der Vermögenswert dem Umlaufvermögen zuzuordnen und gegebenenfalls auf einen gegenüber dem Buchwert niedrigeren Teilwert in Abhängigkeit von der zu erwartenden Ablösesumme abzuschreiben (vgl. Euler, 2020, S. 758). Generell stellt sich bei einem vorzeitigen Abgang die Frage, wie mit einem etwaigen Restbuchwert und möglichen Transfereinzahlungen aus bilanzieller Hinsicht beim abgebenden Verein umzugehen ist. Bei einem Abgang eines Spielers erlöschen die erteilte Spielerlaubnis- und Berechtigung beim abgebenden Verein. Es kommt im Rahmen des Transfervertrages zu einer Verhandlung über eine mögliche Ablösezahlung zwischen den Vereinen. Der aufnehmende Verein erwirbt damit die Möglichkeit, den Spieler in seiner Mannschaft einzusetzen. Die erhaltene Transferentschädigung wird beim abgebenden Verein in den Umsatzerlösen ausgewiesen. Dieses lässt sich in zweierlei Hinsicht begründen. Die Transfererlöse stellen eine regelmäßige bedeutende Einnahmequelle für die Vereine dar. Das Spielervermögen kann daher als Produkt angesehen werden, welches regelmäßig veräußert wird und somit als solches gemäß § 277 HGB ein Bestandteil der Definition der Umsatzerlöse ist (vgl. Schröder & Specht, 2020, S. 964). Zudem wird von Seiten des Lizenzgebers der Ausweis der Transfererlöse innerhalb des Jahresabschlusses gesondert innerhalb der Umsatzerlöse vorgeschrieben. So haben die Vereine die Erträge aus Transfergeschäften unter dem gesonderten Posten „Transfer- und Ausbildungsentschädigung“ auszuweisen (vgl. DFL, 2019a, S. 14). Der Zeitpunkt der Umsatzrealisierung hängt mit verschiedenen Bedingungen zusammen, wie etwa der Auflösung bzw. Abschluss eines neuen Arbeitsvertrages sowie die erfolgreiche Absolvierung des Medizinchecks. Sind diese Bedingungen kumulativ erfüllt, können die Umsatzerlöse in der GuV erfasst werden (vgl. Schröder & Specht, 2020, S. 964).

Wird der Spieler vor dem Ende der Laufzeit seines Arbeitsvertrages veräußert, so wird der Restbuchwert erfolgswirksam als sonstiger betrieblicher Aufwand ausgebucht. Der in der GuV ausgewiesene Gewinn bzw. Verlust eines Spielertransfers ergibt sich also aus der Differenz zwischen den vereinnahmten Transfererlösen und dem Restbuchwert des Spielers zum Zeitpunkt des Verkaufs (vgl. Pellens et al., 2019, S. 2476). Erfüllt der Spieler hingegen seine Vertragslaufzeit bei einem Verein, besitzt er am Vertragsende einen Restbuchwert von Null und kann den Verein ablösefrei verlassen.

4. Bilanzierung von Spielerwerten in der Praxis

4.1. Diskrepanz zwischen Buch- und Marktwerten

Im Profi-Fußball besteht eine ausgeprägte Diskrepanz zwischen den Buch- und Marktwerten. Der Buchwert eines Spielers spiegelt den Marktwert des Spielervermögens nur im Zugangszeitpunkt wider. Danach kommt es aufgrund der Folgebewertung mit fortgeführten Anschaffungskosten zu einer immer höheren wertmäßigen Abweichung (vgl. Elter, 2011, S. 387). Der Marktwert eines Spielers stellt in der Regel den Betrag dar, den er bei einem Verkauf einbringen würde. Während der Buchwert von den jeweiligen Rechnungslegungsvorschriften abhängt, wird der Marktwert eines Spielers zumeist anhand von verschiedenen individuellen Einflussfaktoren bestimmt. Zu diesen zählen bspw. die sportliche Leistungsfähigkeit, das Alter des Spielers, die Restvertragslaufzeit oder auch die Vermarktungsfähigkeit (vgl. Elter, 2011, S. 387). Der Marktwert stellt somit einen Indikator für die Leistungsfähigkeit eines Spielers dar. Die Mannschaften mit den höchsten Kadermarktwerten stellen nominell auch die sportlich erfolgreichsten Mannschaften dar.

Die Entstehung der Diskrepanz lässt sich anhand verschiedener Gründe aufzeigen. Das Spielervermögen als immaterieller Vermögensgegenstand wird über die Vertragsdauer abgeschrieben und besitzt am Ende der Vertragsdauer einen Restbuchwert von Null. Dieses ist aber nicht mit dem in der Realität entstehenden Werteverzehr gleichzusetzen, da insbesondere jüngere Spieler im Laufe der Vertragszeit ihren Marktwert steigern können (vgl. Weber, 2017, S. 292). Daneben ist mit den Nachwuchsspielern, ablösefreien Spielern sowie den Spielern, die einen Restbuchwert von Null besitzen und ihren auslaufenden Vertrag verlängern, ein erheblicher Teil der Mannschaft nicht in der Bilanz abgebildet. So ist z.B. Thomas Müller als Jugendspieler des FC Bayern München nicht als aktivierter Spielerwert in der Bilanz erfasst worden, obwohl er aus sportlicher und vermarktungstechnischer Sicht eine sehr bedeutende Rolle für den Verein einnimmt. Der Buchwert ist somit zwangsläufig kein guter Indikator für den Marktwert eines Spielers. Der Marktwert eines Spielers nimmt aufgrund der zum Teil geringen Aussagekraft der Buchwerte auch bei verschiedenen Bewertungsanlässen eine bedeutende Rolle ein. So kann eine individuelle Spielerbewertung bei der Bestimmung der Höhe der Ablösezahlungen hilfreich sein oder auch von den Rechnungslegungsadressaten herangezogen werden, um ein vollständiges Bild über das Wertpotenzial des gesamten Spielerkaders zu erhalten (vgl. Galli, 2012, S. 679). Ein allgemein akzeptiertes und angewendetes Verfahren zur Spielerbewertung existiert bislang allerdings nicht, wodurch die beschriebene Diskrepanz nicht vollumfänglich aufgelöst werden kann. Die Diskrepanz wird auf der nachfolgenden Seite mittels der Abbildung 3 noch einmal exemplarisch anhand der Vereine der 1. Fußball-Bundesliga aufgezeigt.

Die Tabelle veranschaulicht die Differenz zwischen den Buchwerten des Spielervermögens und den Marktwerten

von der Internetseite transfermarkt.de. Zum Vergleichszweck wurden die Buchwerte aus den im Bundesanzeiger veröffentlichten Abschlüssen der Bundesligavereine für das Geschäftsjahr 2018/2019 bzw. 2019 herangezogen. Mangels objektiver Bewertungsverfahren wurden die Marktwerte der einzelnen Vereine für die Saison 2018/2019 von der Internet-Datenbank transfermarkt.de als Wertmaßstab übernommen. Es zeigt sich, dass die Buchwerte zwischen den einzelnen Verein wertmäßig zum Teil deutlich abweichen. Dieses lässt sich dadurch begründen, dass ein Verein wie Bayern München mehr Umsatzerlöse erzielt als bspw. der 1. FC Köln und somit über mehr Finanzmittel verfügt, die er in den Spielerkader investieren kann. Ein hoher Buchwert kann häufig als Indiz für leistungsstarke Spieler in einem Verein angesehen werden, kann aber auch als Folge einer ungeschickten Transferpolitik verstanden werden. Im Verhältnis zur gesamten Bilanzsumme betrug das Spielervermögen bei den betrachteten Vereinen durchschnittlich 37,1 %. Die aktivierten Spielerwerte nehmen damit eine übergeordnete Stellung in den Bilanzen der Vereine ein. Vergleicht man jetzt die Buchwerte mit den geschätzten Marktwerten der Bundesligakader, wird ersichtlich, dass bei allen Vereinen eine erhebliche Diskrepanz aufgrund der oben genannten Gründe besteht. Bei einem Verein wie dem FC Bayern München mit einem geschätzten Marktwert des Lizenzspielerkaders in Höhe von 835.550.000 Millionen Euro beträgt der Anteil der bilanziell erfassten Spielerwerte nur knapp 12 % des kumulierten Marktwertes. Insgesamt betrug der Buchwert aller betrachteten Vereine lediglich ca. 32 % des aggregierten Marktwertes. In Anbetracht dieser hohen Abweichungen scheint es fragwürdig, ob das aktivierte Spielervermögen ein dem GoB entsprechendes tatsächliches Bild der Vermögenslage vermittelt. Dieses soll in den nachfolgenden Kapiteln diskutiert werden.

4.2. Kritische Würdigung der Bilanzierung von Spielerwerten

4.2.1. Bilanzansatz von Spielerwerten

Nach derzeit gültigen Bestimmungen werden nur die entgeltlich erworbenen Spieler nach den Vorschriften des HGB mit den gezahlten Anschaffungskosten als immaterielle Vermögensgegenstände in der Bilanz der Vereine aktiviert. Diese Regelung sorgt dafür, dass die im Kapitel 3.1 beschriebenen Zwecke der handelsrechtlichen Rechnungslegung im Bereich des Profi-Fußballs nicht vollständig erfüllt werden. Wie bereits im vorherigen Kapitel aufgezeigt, kommt es aufgrund der zahlreichen nicht bilanzierten Spielerwerte zu keiner verwertbaren Aussage über die tatsächliche Vermögenslage. Auch wenn diese Spieler nicht die Erfordernisse der Rechnungslegung zur Aktivierung erfüllen, resultiert aus ihnen eine exklusive zurechenbare Nutzungsmöglichkeit zum Vorteil des Vereins. Um den Rechnungslegungsadressaten einen vollständigen Überblick über die Vermögenslage und damit dem erzielbaren Erfolg zu vermitteln, sollte eine zusätzliche Aktivierung dieser Spieler erfolgen. Auch die Vereine würden von diesem Bilanzansatz profitieren. So kann die Aktivierung zusätzlicher Vermögensgegenstände zu einem höheren Eigenkapitalausweis im Verhältnis zur Aufwandsrechnung führen und somit zu einer positiveren Darstellung der Finanzlage des Vereins (vgl. Wulf & Bosse, 2011, S. 89). Ein hohes aktiviertes Spielervermögen ist aber nicht immer gleichbedeutend mit einem sportlichen Erfolg und einer finanziellen Absicherung. Die Spieler unterliegen als immaterielle Vermögensgegenstände dem Risiko einer Fehlinvestition sowie schwerer Verletzungen. Durch daraus entstehende Totalabschreibungen auf die Spielerwerte kann es zu einer erheblichen Verschlechterung der wirtschaftlichen und sportlichen Lage bei den Vereinen kommen (vgl. Weber, 2017, S. 292). Das aktivierte Spielervermögen bietet den Gläubigern auch nicht zwangsläufig einen vergleichbaren Einblick in die aktuelle und zukünftige wirtschaftliche Situation in den Vereinen. Ein Verein mit einem hohen Anteil an ausgewiesenen Spielerwerten in der Bilanz ist zwangsläufig nicht besser zu beurteilen als ein Verein mit geringeren Spielervermögen. Gerade Spieler aus der eigenen Jugend sowie ablösefrei verpflichtete Spieler weisen durch kurzfristige Verkäufe ein Liquiditätspotenzial auf, ohne dass dieses vorher in der Bilanz erfasst wurde (vgl. Weber, 2017, S. 293). Dieses erschwert es für die Adressaten, sich ein umfassendes Bild über die Chancen und Risiken einer Beteiligung am Verein zu machen.

Bei den Vereinen, die auf Spieler aus der eigenen Jugend setzten, fällt die Diskrepanz zwischen Buch- und Marktwerten höher aus als bei vergleichbaren Mannschaften. Die Vereine können zwar grundsätzlich seit dem BilMoG ihre Jugendspieler aktivieren, Anwendung findet dieses Wahlrecht in der Praxis aber eher weniger. Dieses liegt zum einen an den verbandsrechtlichen Einschränkungen zur Bilanzierung von Nachwuchsspielern und den dadurch gestiegenen Anforderungen an das Kostenrechnungssystem zur Zurechnung der Entwicklungskosten (vgl. Weber, 2017, S. 292). Zum anderen würde die Ausübung des Wahlrechtes gegen das im GoB verankerte Stetigkeitsprinzip des Bilanzansatzes sprechen. Demnach ist es nicht zulässig, dass Ansatzwahlrecht auszuüben, unabhängig davon, ob selbsterstellte immaterielle Vermögensgegenstände aktiviert werden dürfen (vgl. Schubert & Andrejewski, 2020, § 253, Rn. 397). Aus dieser Erklärung lässt sich folgern, dass allein schon aufgrund der handelsrechtlichen Vorschriften des HGB die Bilanz der Vereine kein den tatsächlichen Umständen entsprechendes Bild der Vermögenslage vermitteln kann. Die einheitliche Restriktion der Wahlrechtsausübung schränkt aber zu mindestens die willkürliche Ausgestaltung des Jahresabschlusses ein und fördert die Vergleichbarkeit für die Adressaten. Für die Vereine im deutschen Profi-Fußball stellt die Tatsache, dass sie zumindest einen Teil ihres Spielerkaders mit der gezahlten Ablösesumme aktivieren können, eine durchaus zufriedenstellende Lösung dar. Vor der Regelung durch den BFH wurden Transferausgaben direkt aufwandswirksam als Betriebsausgaben verbucht. Durch die Möglichkeit der Aktivierung können die immer höher werdenden Anschaffungskosten eines Spielers auf mehrere Perioden aufgeteilt werden und somit im Vergleich zur vorherigen Regelung ei-

	Vereine	Betrachteter Abschluss	Buchwert (in €)	Buchwert relativ zur Bilanzsumme (in %)	Marktwert von transfermarkt.de (in €)
1	FC Bayern München	Konzernabschluss (2018/2019)	102.789.341	14,05%	835.550.000
2	RB Leipzig	Jahresabschluss (2018/2019)	152.462.142	61,44%	314.450.000
3	Eintracht Frankfurt	Konzernabschluss (2019)	92.178.543	46,53%	139.550.000
4	Borussia Dortmund	Jahresabschluss (2018/2019)	182.136.000	35,45%	376.800.000
5	Borussia Mönchengladbach	Jahresabschluss (2019)	93.497.086	38,71%	191.530.000
6	VFB Stuttgart	Konzernabschluss (2019)	56.227.613	46,48%	159.450.000
7	TSG Hoffenheim	Konzernabschluss (2018/2019)	47.041.810	17,10%	188.630.000
8	SV Werder Bremen	Konzernabschluss (2018/2019)	35.734.324	47,13%	110.850.000
9	1. FC Köln	Konzernabschluss (2018/2019)	48.739.155	51,58%	99.830.000
10	Hertha BSC	Konzernabschluss (2018/2019)	17.222.856	12,91%	125.230.000
Σ			828.028.870		2.541.870.000

Abbildung 3: Buch- und Marktwerte der 1. Fußball-Bundesliga

Quelle: Eigene Darstellung, basierend auf Bundesanzeiger; Transfermarkt (vgl. Transfermarkt, 2021).

ne zusätzliche Sicherheit als Vermögenswert in der Bilanz darstellen.

Um den Adressaten eine wahrheitsgemäße Darstellung über die Vermögens-, Finanz- und Ertragslage zu liefern, bedarf es wie im vorherigen Kapitel beschrieben einer objektiv nachvollziehbaren Bewertung der einzelnen Spieler. Dadurch erhalten die Stakeholder mehr Transparenz und einen glaubwürdigeren Einblick in die Vermögenslage, auf deren Basis sie ihre Investitionsentscheidung treffen können. Die entstehenden stillen Reserven zwischen den aktivierten Buchwerten und den tatsächlichen Werten des Spielervermögens können dadurch aufgedeckt werden und gleichzeitig den Interessensgruppen einen Ansatzpunkt zur Wertermittlung bieten (vgl. Elter, 2012, S. 70f.). Dieses zeigt noch einmal die Bedeutung der Spielerbewertung bezogen auf die aus dem Jahresabschluss vermittelten Informationen. Die aktuelle Bilanzierungspraxis in der Fußball-Bundesliga wird aufgrund der handels- und verbandsrechtlichen Einschränkungen dem Zweck der handelsrechtlichen Rechnungslegung nicht vollkommen gerecht. Dem aktivierten Spielervermögen ist im Hinblick auf die Beurteilung der wirtschaftlichen Lage eines Vereins somit nur eine geringe substanzielle Aussagekraft zuzuordnen.

4.2.2. Zugangsbewertung von Spielerwerten

Bei der Zugangsbewertung stellt sich die Frage, inwieweit es gerechtfertigt ist, dass bei unentgeltlichen Transfers mangels eines Anschaffungsvorganges keine Anschaffungsnebenkosten aktiviert werden dürfen. Bei ablösefreien Transfers wird aus Sicht des BFH aufgrund des fehlenden entgeltlichen Erwerbsvorganges kein exklusives Nutzungsrecht erworben. Entgegen der Aussage des BFH resultiert aus der Lizenzerteilung an die Spieler seitens der DFL aber sehr wohl ein exklusives Nutzungsrecht. Die mit der Verpflichtung einhergehenden Kosten eines ablösefreien Spielers sind wie bei einem ablösepflichtigen Spielerwechsel erwerbsbezogen und einzeln zurechenbar. In den letzten Jahren sind die Zahlungen für Wechselprämien und Beraterhonorare immens gestiegen und würden deshalb zumindest wertmäßig eine bilanzielle Abbildung rechtfertigen. Die Vereine gaben in der Bundesligasaison 2019/2020 knapp 200 Millionen Euro für Beraterhonorare aus (vgl. DFL, 2020). Zwar stecken in diesem Gesamtbetrag auch die Beraterhonorare für die ablösepflichtigen Transfers, dennoch würde eine Einbeziehung der gesamten Summe in den Jahresabschluss der Vereine den Adressaten einen vollständigeren Einblick in die Geschäftstätigkeit ermöglichen.

Zudem besteht mit den nachträglich anfallenden Anschaffungskosten eine Position, die von den Adressaten nicht einsehbar ist. Werden bei einer Vertragsunterzeichnung etwaige Zusatzzahlungen für einen zukünftigen sportlichen Erfolg des Spielers vereinbart, wie etwa eine bestimmte Anzahl erzielter Tore oder die Qualifikation für den internationalen Wettbewerb, sind diese im Zeitpunkt des Anfalls zu verbuchen. Der Eintritt dieser Kosten ist für die Vereine bei Vertragsunterzeichnung aufgrund der begrenzten Einflussmöglichkeit nur schwer einkalkulierbar. Eine nachträgliche Erhöhung des Bilanzwertes in Höhe der anfallenden Kosten und der damit einhergehenden höheren Abschreibungsbeträge über die Restvertragslaufzeit würde den jährlichen Wertverlust des Spielervermögens erhöhen. Fraglich ist dann, ob diese Zahlungen noch in einem direkten Zusammenhang mit dem eigentlichen Erwerbsvorgang zu Vertragsbeginn stehen.

Festzuhalten bleibt, dass diese nachträglichen Kosten zumindest aus externer Sicht nicht nachprüfbar sind. Eine weitere Herausforderung besteht in der Zugangsbewertung der Leihspieler. Enthält der Leihvertrag eine Kaufoption, besteht in der Praxis häufig die Problematik der Bestimmung des wirtschaftlichen Eigentümers. Während der Marktwert eines Spielers im Regelfall dem ausleihenden Verein zugeschrieben wird, ist auf Grundlage individueller Vertragsgestaltung im Einzelfall zu entscheiden, bei welchem Verein eine Zuordnung des Buchwertes erfolgt. Hinzu kommt das die Vereine der Bundesliga in ihren veröffentlichten Abschlüssen keinen thematischen Schwerpunkt auf die bilanzielle Behandlung der Leihspieler setzen (vgl. Weber, 2016, S. 252f.). Zumindest zum Informationszweck im Sinn der handelsrechtlichen Rechnungslegung ist eine Übersicht der bilanziellen Zuordnung der Leihspieler zu den verschiedenen Vereinen für die Interessensgruppen wünschenswert.

4.2.3. Folgebewertung von Spielerwerten

Bei der Folgebewertung ergeben sich auch aufgrund der fehlenden Konkretisierung außerplanmäßiger Abschreibungen einige Diskussionspunkte.

Im Zuge der planmäßigen Abschreibung ergibt sich die Problematik der Bestimmung der Nutzungsdauer. Beinhaltet der Vertrag eine Verlängerungsoption, ist wie im Kapitel 3.3.3 veranschaulicht, nach allgemeinen Wahrscheinlichkeitsgesichtspunkten zu bestimmen, ob von der Option Gebrauch gemacht wird. Das die Vereine anhand von Verletzungsanfälligkeiten, Formschwankungen oder sonstigen Faktoren eines Spielers bestimmen sollen, ob sie die betriebsgewöhnliche Nutzungsdauer erhöhen, deutet auf eine subjektive und kaum prognostizierbare Entscheidung seitens der Vereine hin (vgl. Littkemann & Schaarschmidt, 2002, S. 377). Weiterhin ist bei der planmäßigen Abschreibung fraglich, ob der einzelne Spieler zum Ende der Vertragslaufzeit auf einen Restbuchwert von Null abzuschreiben ist. Wird nach Vertragsende mit dem Spieler ein neuer Arbeitsvertrag vereinbart, ergeben sich für den Verein neue Nutzungsrechte, die dem vollständig abgeschriebenen Buchwert des Spielers nicht zu entnehmen sind. Das Spielervermögen ist damit wie bereits in den vorherigen Kapiteln beschrieben nicht aussagekräftig bezüglich der tatsächlichen Vermögenslage des Vereins. Eine Einbeziehung eines Restwertes in Form einer möglichen künftigen Ablösezahlung in den Abschreibungsplan könnte hier Abhilfe schaffen. Zu berücksichtigen ist dabei aber, dass ein erzielbarer Veräußerungspreis im Zugangszeitpunkt nur schwer quantifizierbar ist und der Spieler nach Vertragsende den Verein ablösefrei ohne Zahlung einer Ablösesumme verlassen darf. Ein möglicher Restwert kann somit nicht einheitlich festgelegt werden.

Das die fortgeführten Anschaffungskosten unabhängig von der tatsächlichen Wertentwicklung die Wertobergrenze eines Spielers widerspiegelt, ist indes positiv zu beurteilen. Eine Aktivierung der Spieler zu ihren Marktwerten in der Bilanz kommt allein wegen der fehlenden objektiven Messbarkeit des Marktwertes nicht in Frage. Zudem sind die Marktwerte, wie in der Abbildung 3 dargestellt, deutlich höher als die Buchwerte. Die Bilanz würde durch diese fiktiven Werte künstlich aufgeblasen werden und wäre für Adressaten nicht mehr sachgerecht interpretierbar. Der Marktwert als objektiver Spielerwert dient daher eher als Vergleichsmaßstab für etwaige Bewertungsanlässe, wie z.B. der Werthaltigkeitsprüfung, aber nicht als Ersatz für den Buchwert.

Bei den außerplanmäßigen Abschreibungen stellt sich die Frage, inwieweit es gerechtfertigt ist, dass diese aufgrund von fehlenden Vergleichswerten häufig ausbleiben. Aufgrund der fehlenden einheitlichen Vorgaben zur Durchführung außerplanmäßiger Abschreibungen seitens der Rechtsprechung erhalten die Vereine einen erheblichen bilanzpolitischen Spielraum. Eine von Littkemann, Schulte und Schaarschmidt durchgeführte Umfrage bei den Fußball-Bundesligisten zu dieser Thematik ergab, dass sowohl bei der Bestimmung der Höhe und dem Zeitpunkt der Abschreibung sowie bei einer etwaigen Zuschreibung keine konsistente Anwendung bei

den Bundesligisten erfolgte (vgl. Littkemann et al., 2005, S. 665). Für die handelsrechtliche Rechnungslegung der Vereine gilt das Niederstwertprinzip, wonach sie bei einer dauerhaften Wertminderung zu einer außerplanmäßigen Abschreibung verpflichtet sind. Das Spielervermögen soll mittels dieses GoB möglichst vorsichtig bewertet werden, um den Gläubigerschutz zu stärken. Aufgrund der fehlenden einheitlichen Regelungen, wann und in welcher Höhe eine dauerhafte Wertminderung beim Spielervermögen vorliegt, kann es zwangsläufig nicht zu einer sachgerechten Durchführung des Niederstwerttests kommen. Die Adressaten sind dadurch nicht in der Lage, einen vergleichbaren und akkuraten Einblick in die Rechnungslegung der Vereine zu erlangen, sodass es folglich aus Sicht der Gläubiger zu Fehlinterpretationen der Bilanz kommen kann. Dazu sollte aber abschließend auch festgehalten werden, dass die Gründe für eine solche dauerhafte Wertminderung nur in Ausnahmefällen eintreten.

Zusammenfassend lässt sich festhalten, dass gerade die fehlende Konkretisierung der außerplanmäßigen Abschreibung die Zwecke der handelsrechtlichen Rechnungslegung beeinträchtigt. Für eine sachgerechte bilanzielle Behandlung erscheint die Notwendigkeit eines objektiv nachprüfbaren beizulegenden Zeitwertes als Vergleichsmaßstab unerlässlich.

5. Fazit

Die deutschen Bundesligavereine haben sich in den letzten Jahrzehnten von der Rechtsform eines e.V. zu Kapitalgesellschaften von erheblicher wirtschaftlicher Größe entwickelt. In diesem Zusammenhang sind auch die im Rahmen der Spielertransfers gezahlten Ablösesummen angestiegen und beeinflussen das Bilanzbild der Vereine beachtlich. Die Vereine agieren dabei in einer weltweit festgelegten hierarchischen Organisationsstruktur. Die deutsche Verbandsstruktur setzt sich dabei aus dem DFB und der DFL zusammen. Die DFL ist dabei verantwortlich für den ordnungsgemäßen Spielbetrieb in den deutschen Lizenzligen. Sie greift bei ihrer Arbeit auf ein umfassendes Regelungskonstrukt, bestehend aus der LO und der LOS zurück. Die Lizenzerteilung an Spieler und Vereine stellt dabei die rechtliche Grundlage zur Bilanzierung von Spielertransfers dar. Aufgrund der Lizenzerteilung entsteht ein besonderes Beziehungsgeflecht zwischen der DFL, den Vereinen und den Spielern. Zudem unterliegen die Vereine hinsichtlich der Auslegung ihrer Rechnungslegung aufgrund des Lizenzierungsverfahrens einigen Besonderheiten. So haben die Vereine neben den allgemeinen Vorschriften des HGB zusätzlich fußballspezifische Posten, wie etwa das Spielervermögen, mit in ihre Bilanz aufzunehmen. Das aus einem Spielertransfer überhaupt ein einzelner Bilanzposten resultiert, beruht auf einer Reihe von Rechtssprechungen seitens des BFH. Von einer zunächst unbestimmten Bilanzierungspraxis mit der Verpflichtung zur Zahlung einer Transfersumme entwickelte es sich zu einer von Seiten des BFH klar dargelegten Ansatzpflicht des Spielers in der Bilanz. Eine Ablöse ist seit dem Jahre 1995 zudem nur noch dann zu bezahlen, wenn ein Spieler aus einem laufenden Vertrag wechselt. Von den Rechtssprechungen des BFH wird auch heute noch bei den Bundesligavereinen und Verbänden Gebrauch gemacht. So kommt es im Zusammenhang mit der Verpflichtung eines entgeltlich erworbenen Spielers zu einem aktivierungsfähigen Vermögensgegenstand. Dabei bleibt aber festzuhalten, dass es nicht zum Ansatz des Spielers selbst kommt, sondern um eine durch den Verband abgesicherte exklusive Nutzungsmöglichkeit an dem Spieler. Die Zugangsbewertung erfolgt in Höhe der Ablösezahlung zuzüglich etwaiger Aufwendungen, die dem Spielerwechsel spezifisch zugeordnet werden können. Die Folgebewertung erfolgt planmäßig zu fortgeführten Anschaffungskosten, wobei insbesondere die Thematik der außerplanmäßigen Abschreibungen den Vereinen einen erheblichen bilanzpolitischen Spielraum überlässt. In der Praxis zeigt sich, dass zwischen den Buchwerten und den geschätzten Marktwerten eine erhebliche Diskrepanz besteht. Als Grund kann dazu aufgeführt werden, dass indes nur ein geringer Teil des Spielerkaders bilanziell abgebildet wird. Dieses spiegelt auch gleichzeitig die Problematik der derzeitigen bilanziellen Behandlung von Spielertransfers wider. Die Adressaten der Rechnungslegung interessiert, wie der Verein als Gesamtes wirtschaftlich dasteht. Die Aussagekraft der Bilanzposition Spielervermögen hinsichtlich der tatsächlichen wirtschaftlichen Verhältnisse im Verein ist dahingehend aber eher als ungenügend zu bezeichnen. Die Zwecke der handelsrechtlichen Rechnungslegung in Form der Informationsfunktion und des Gläubigerschutzes werden durch die derzeitige Handhabung der bilanziellen Abbildung von Spielerwerten im Profi-Fußball dagegen eher beeinträchtigt.

Um ein zutreffendes Bild über die wirtschaftliche Lage der Vereine zu erhalten, erfordert es ein bislang fehlendes objektives Bewertungsverfahren. Dieses könnte den Adressaten bei einer nachvollziehbaren Wertermittlung der Spieler unterstützen und dem Verein als Orientierungshilfe für zukünftige Kauf- und Verkaufspreise dienen. Die Entwicklung eines allgemein akzeptierten und einheitlich angewendeten Verfahrens zur Spielerbewertung stellt eine zukünftige Herausforderung dar. Um die Aussagekraft der derzeitigen Bilanzierungspraxis zu erhöhen, scheint die Erarbeitung einer solchen Bewertungsmethode unabdingbar. Da die Ablösesummen tendenziell auch zukünftig ansteigen und somit einen immer größer werdenden Einfluss auf das Bilanzbild ausüben werden.

Literatur

DFL. (2019a). *Anhang VII-Prüfung der wirtschaftlichen Leistungsfähigkeit vor einer Spielzeit*. Zugriff auf https://media.dfl.de/sites/2/2020/01/Anhang-VII-zur-LO-2019-12-04-Stand.pdf (Abfrage: 30.04.2021)

DFL. (2019b). *Satzung DFL Fußball Liga e.V.* Zugriff auf https://media.dfl.de/sites/2/2020/02/Satzung-DFL-e.V.-2019-08-21.pdf (Abfrage: 20.04.2021)

DFL. (2020). *Finanzkennzahlen: Clubs der Bundesliga in der Saison 2019-20*. Zugriff auf https://media.dfl.de/sites/2/2020/06/Bundesliga-Clubs-der-Spielzeit-2019-20-Geschftsjahresende-2019-Finanzkennzahlen-gem8-Nr-6k-LO.pdf (Abfrage: 02.05.2021)

DFL. (2021a). *DFL Wirtschaftsreport 2021*. Zugriff auf https://media.dfl.de/sites/2/2021/03/D_DFL_Wirtschaftsreport_2021_M.pdf (Abfrage: 19.04.2021)

DFL. (2021b). *Lizenzierungsordnung (LO)*. Zugriff auf https://media.dfl.de/sites/2/2021/03/Lizenzierungsordnung-LO-2021-03-05-Stand.pdf (Abfrage: 22.04.2021)

DFL. (2021c). *Lizenzordnung Spieler (LOS)*. Zugriff auf https://media.dfl.de/sites/2/2021/03/Lizenzordnung-Spieler-LOS-2021-03-05-Stand.pdf (Abfrage: 21.04.2021)

Elter, V. (2011). Bewertung von Fußballunternehmen. In J. Drukarczyk & D. Ernst (Hrsg.), *Branchenorientierte Unternehmensbewertung* (S. 382–392). München, Vahlen.

Elter, V. (2012). Bewertung von Fußballunternehmen. *Controlling & Management*, *56* (1), 66–76.

Euler, R. (2020). Transferzahlungen im Jahresabschluss von Unternehmen des Profisports. In *Perspektiven des Dienstleistungsmanagements* (S. 755–769). Wiesbaden, Springer Gabler.

Förschle, G. & Usinger, R. (2010). *Handelsgesetzbuch, Kommentar, 7. Aufl., München.*

Fross, I. (2012). Transferentschädigungen und entgeltlicher Erwerb. *Finanz-Rundschau Ertragsteuerrecht*, *94* (5), 193–205.

Galli, A. (2003). Individuelle finanzielle Spielerbewertung im Teamsport. *Finanz Betrieb*, *12*, 810–820.

Galli, A. (2012). Marktgestützte finanzielle Spielerbewertung im Teamsport. In Galli, A. and Elter, V.C. and Gömmel, R. and Holzhäuser, W. and Straub, W. (Hrsg.), *Sport-management* (S. 675–690). München, Vahlen.

Handschin, L. (2012). Bewertung eingeschränkt verkehrsfähiger Vermögenswerte. *NZG: Neue Zeitschrift für Gesellschaftsrecht*, *15*, 1281–1287.

Hierl, L. & Weiß, R. (2016). Zusammenfassung und Fazit. In *Bilanzanalyse von Fußballvereinen.* Wiesbaden, Springer Gabler.

Huwer, E. (2019). Rechnungslegung von Sportkapitalgesellschaften. In Walzel, S. and Römisch, V. (Hrsg.), *Teamsport Management* (S. 271–300). Wiesbaden, Springer Gabler.

Kaiser, T. (2004). Die Behandlung von Spielerwerten in der Handelsbilanz und im Überschuldungsstatus im Profifußball. *Der Betrieb*, *21*, 1109–1112.

Kicker. (2020). *Nur noch fünf eingetragene Vereine: Die Rechtsformen der Bundesligisten.* Zugriff auf https://www.kicker.de/nur_noch_fuenf_eingetragene_vereine_die_rechtsformen_der_bundesligisten-776662/slideshow (Abfrage: 20.04.2021)

Köster, O. & Ehemann, T. (2012). Ausgewählte Probleme der Rechnungslegung bei Kapitalgesellschaften im Sport. In Galli, A. and Elter, V.C. and Gömmel, R. and Holzhäuser, W. and Straub, W. (Hrsg.), *Sport-management* (S. 145–164).

Küting, K. & Strauß, M. (2010). „Die Wahrheit liegt auf dem Platz“ - oder etwa nicht? *Der Betrieb*, *15*, 793–802.

Küting, K. & Strauß, M. (2011). Financial Fair Play im Profifußball. *Der Betrieb*, *2*, 65–76.

Littkemann, J., Brast, C. & Stübinger, T. (2002). Neuregelung der Rechnungslegungsvorschriften für die Fußball-Bundesliga. *Steuern und Bilanzen*, *4* (24), 1196–1204.

Littkemann, J., Brast, C. & Stübinger, T. (2003). Neuordnung des verbandsrechtlichen Rahmens für die Fußball-Bundesliga. *WiSt-Wirtschaftswissenschaftliches Studium*, *32* (7), 415–418.

Littkemann, J. & Schaarschmidt, P. (2002). Probleme der bilanziellen Behandlung von Transferentschädigungen nach Handels-und Steuerrecht. *Unternehmensteuern und Bilanzen*, *8*, 372–380.

Littkemann, J., Schulte, K. & Schaarschmidt, P. (2005). Außerplanmäßige Abschreibungen auf Spielerwerte im Profifußball: Theorie und Praxis. *StuB*, *15*, 660–666.

Pellens, B., Küting, P. & Schmidt, A. (2019). Bilanzielle Erfassung von Transferentgelten im Profi-Fußball nach HGB und IFRS (Teil 1). *Der Betrieb*, *45*, 2473–2481.

Schäfer, S. (2019). *Handelsrechtliche Rechnungslegung: Jahresabschluss und Lagebericht mit Beispielen, Übungsaufgaben und Lösungshinweisen*. Stuttgart, Kohlhammer Verlag.

Schiffers, J. & Feldgen, R. (2015). Bilanzierung des exklusiven Nutzungsrechts an einem Lizenzspieler im Profisport. *StuB*, *13*, 500–507.

Schröder, D. & Specht, V. (2020). Spielertransfers im Profifußball. *WPg*, *16*, 959–965.

Schubert, W. J. (2020). *Handelsgesetzbuch, Kommentar, 12. Aufl., München.*

Schubert, W. J. & Andrejewski, K. C. (2020). *Handelsgesetzbuch, Kommentar, 12. Aufl., München.*

Schubert, W. J. & Gadek, S. (2014). *Handelsgesetzbuch, Kommentar, 9. Aufl., München.*

Schubert, W. J. & Huber, F. (2020). *Handelsgesetzbuch, Kommentar, 12. Aufl., München.*

Schüler, A. (2016). *Bewertungseinheiten nach HGB* (1. Aufl.). Wiesbaden, Springer Gabler.

Seip, M. (2017). *Vertragsrechtliche und ökonomische Analyse des Spielertransfervertrages im Profifußball* (1. Aufl.). Baden-Baden, Nomos.

Steiner, E. & Gross, B. (2005). Die Bilanzierung von Spielerwerten im Berufsfußball nach HGB und IFRS. *Unternehmensteuern und Bilanzen*, 531–536.

Teichmann, K. (2007). *Strategie und Erfolg von Fußballunternehmen* (1. Aufl.). Wiesbaden, Deutscher Universitäts-Verlag.

Teschke, M., Knipping, J. & Sundheimer, H. (2012). Bilanzierung von Spielererlaubnissen im Profi-Fußball vor dem Hintergrund des BFH-Urteils v. 14.12.2011, FR 2012, 213. *Finanz-Rundschau*, *24*, 1137–1144.

Thyll, A. (2004). Jahresabschluss und Prüfung nach der Lizenzierungsordnung: Grundlagen und Gegenüberstellung mit den handelsrechtlichen Vorschriften. In *Ökonomie des fußballs* (S. 163–192). Wiesbaden, Deutscher Universitäts-Verlag.

Transfermarkt. (2021). *Marktwerte der Verein in der Bundesliga 18/19*. Zugriff auf https://www.transfermarkt.de/1-bundesliga/startseite/wettbewerb/L1/saison_id/2018 (Abfrage: 29.04.2021)

Weber, C. (2016). Bilanzierung von Leihspielern im Profifußball - die Zurechnung des wirtschaftlichen Eigentums bei verliehenem bzw. ausgeliehenem Vermögen. *KoR*, 248–253.

Weber, C. (2017). Besonderheiten der Bilanzanalyse von Profifußballclubs. *Der Konzern*, *6*, 285–295.

Weber, C. (2020). Auswirkungen des Coronavirus SARS-CoV-2 auf die Rechnungslegung und Lizenzierung im deutschen Profifußball. *KoR*, *07-08*, 320–327.

Wulf, I. & Bosse, T. (2011). Der Einfluss der Bilanzierung von Spielervermögen auf das Finanzrating. *StuB*, *03*, 83–89.

Junior Management Science 7(4) (2022) 1164-1192

Junior Management Science

journal homepage: www.jums.academy

Innovation Performance in Healthcare M&A: An Empirical Analysis

Philipp S.R. Voss

Technische Universität Berlin

Abstract

The relationship between mergers and acquisitions (M&A) and innovation in the healthcare sector (pharmaceuticals, biotechnology, medical devices, and life sciences) is investigated using a new self-generated dataset of 41 firms. Patents are used as proxy for innovation performance of acquiring firms. This work can also be seen as an extended replication study of Ahuja and Katila (2001) and Cloodt, Hagedoorn, and Van Kranenburg (2006). The extension comprises of newly added variables relatedness of acquirer knowledge and acquisition experience. The findings are consistent with previous research. Non-technological M&A appear to have a negative impact on the acquiring firm's innovation performance. The absolute size of acquired knowledge has a small positive effect. The relative size of acquired knowledge has a negative effect on the acquiring firm's innovation performance. The relatedness of the target knowledge base has a curvilinear impact on innovative performance. The relatedness of acquirer knowledge has a negative effect on innovation performance. Finally, the effect of previous acquisition experience is ambiguous. The findings of this study indicate that the firms' innovation performance can benefit from M&A by carefully selecting targets that provide the appropriate amount of "innovative" input.

Keywords: Mergers & Acquisitions; Innovation; Innovation Performance; Patents; Knowledge based view.

1. Introduction

1.1. M&A and Innovation

Mergers and Acquisitions (M&A) continue to be an important part of the business world in both developed and emerging economies. Worldwide M&A transaction volume has peaked in 2015 and has remained high ever since, with a value of completed deals over 3T$ and over 40,000 deals in 2019 before the COVID-19 outbreak (Thomson Reuters Eikon, 2021).

This indicates that M&A is a very popular instrument in corporate strategy. However, M&A and especially the success of M&A as a strategy is a mystery for both researchers and managers: contrary to expectations, a considerable number of deals has lead to negative cumulative abnormal returns or no value creation for shareholders of the acquiring company (Agrawal, Jaffe, & Mandelker, 1992; Datta, Pinches, & Narayanan, 1992; Jarrell & Poulsen, 1989; King, Dalton, Daily, & Covin, 2004). Even managers in charge consider plenty of deals as failed. In one study, about 50% of executive managers evaluated their deals as failed (Cartwright & Schoenberg, 2006). There are multiple explanations for why executives choose to conduct M&A and why M&A is still a very common growth strategy. One reason for M&A are personal motives of executives. Acquisitions might help them to achieve their growth goals and in turn increase their compensation. Besides managerial hubris, other common motives are: acquiring new business units for new markets, acquiring technology, improving performance by using synergies, and acquiring a foreign firm to enter a domestic market or diversifying a firm. M&A can also be a response to exogenous shocks such as crises, changes in regulation or technology breakthroughs of competitors (Chen, Hsu, Officer, & Wang, 2020; Pidun, 2019). Given the occurrence of M&A and that many deals fail from a financial standpoint, other research metrics could provide interesting insights on M&A and why firms still choose to conduct M&A (Haleblian, Devers, McNamara, Carpenter, & Davison, 2009). Such metrics could include the innovation performance, accounting measurements or the long term financial performance. Most empirical studies on M&A stem from financial management research with a focus on shareholder value creation (King et al., 2004; Zollo & Meier, 2008). Only a fraction of M&A research has been devoted to the link of M&A and innovation, mainly in the area of strategic management (Cassiman, Colombo, Garrone, & Veugelers, 2005).

During the last 30 years, the speed of innovation has increased rapidly (Makri, Hitt, & Lane, 2010) and innovation has become an important way to achieve and maintain

DOI: https://doi.org/10.5282/jums/v7i4pp1164-1192

a competitive advantage (Hitt, Hoskisson, Johnson, & Moesel, 1996). Technological innovation and product innovation have become an important part and motive in the M&A landscape (Chen et al., 2020). M&A can be a valuable strategy to enhance the innovation performance: e.g., to reduce the time to market, to expand a product portfolio, to develop technological capabilities, or to gain access to intellectual property or technology (Cassiman et al., 2005; Goedhart, Koller, & Wessels, 2017; Schweizer, 2005b). This led many larger established corporations to rely on small innovative firms as a source of innovation. This strategy avoids uncertain and resource- and time-consuming processes of internal innovation sourcing (Granstrand & Sjölander, 1990; Puranam, Singh, & Zollo, 2006). The importance of innovation and the unilateral research focus on financial performance calls for more research in the area of M&A.

This thesis aims at providing insights on M&A innovation performance by analyzing a sample of 41 firms active in the healthcare industry. Competition in this industry is particularly driven by innovation and technology (Danzon, Nicholson, & Pereira, 2005; Schweizer, 2005b; Simoens, 2008). The motivation behind this thesis is two-fold: First, this thesis tries to replicate the findings of Ahuja and Katila (2001) - "*Technological Acquisitions and the Innovation Performance of Acquiring Firms: a Longitudinal Study*". A first replication study by Cloodt et al. (2006) has already confirmed selected findings. This study analyzes the post-M&A patenting output of M&A in the healthcare industry as a proxy for innovation performance. The healthcare sector is known for a high-patent intensity and technology and innovation play an important role in competition. It is therefore very suitable for a replication study (Cockburn & Long, 2015). Furthermore, as the previous studies were conducted more than 15 years ago, applying the model to current data could provide additional information. Second, this thesis extends Ahuja and Katila (2001) by differentiating between acquirer and target knowledge relatedness, inspired by Sears and Hoetker (2014). This additional variable allows for a deeper analysis of the interaction of the target and acquirer knowledge and the resulting combination of knowledge. Likewise, a variable for acquisition experience will be introduced to the model.

The research question is therefore:
How does M&A activity affect the innovation output (measured by the number of patents granted) in the healthcare industry?

This thesis is organized as follows: First, the characteristics of the healthcare industry are described. Second, the theoretical foundation and important theories are explained. A short section presents previous studies on M&A and innovation performance of firms. Then, the hypotheses for the empirical analysis are developed out of the conceptual framework. Third, the empirical analysis follows, using a modification of the model of Ahuja and Katila (2001), with additional variables focusing on knowledge overlap of the involved firms. Lastly, the author shows the limits of the research, discusses the findings and outlines recommendations for future research.

1.2. M&A in the Healthcare Industry

The healthcare industry was selected as a research setting for this thesis for multiple reasons. First, the healthcare industry, especially the pharmaceutical industry, has played a major role in some of the largest international M&A deals, such as the acquisition of Wyeth by Pfizer for 68B$ and of Shire Pharmaceuticals by Takeda for 62B$ (Chee, 2018; J. Hall & Krauskopf, 2009). Second, the healthcare industry relies heavily on intellectual property protection by patents. As the European Patent Office (EPO) recently stated: "innovation in healthcare drove patenting activity in 2020: Medical technology was the leading field for inventions in terms of volume, while pharmaceuticals and biotechnology were the fastest-growing areas" (European Patent Office, 2021, para.1). Patents grant a temporary monopoly to commercialize an invention and are therefore very valuable in this industry (Arrow, 1962). On the other hand, patent expirations constantly force these firms to stay innovative and introduce new technologies and products. Third, competition in this industry is mainly driven by innovation in technology, research and development (R&D) productivity and R&D outputs (Ornaghi, 2009). Technological learning and the introduction of new products are expected to be a key for achieving a competitive advantage (Cloodt et al., 2006). M&A are therefore a popular strategy to get access to new technologies, products, distribution channels and market positions, since "the pace and magnitude of technological change may not allow firms to internally develop all technologies and capabilities they need to remain competitive" (Schweizer, 2005b, p. 1052).

According to the Global Industry Classification Standard (GICS), the healthcare industry can be classified into different sub-industries: healthcare equipment providers, healthcare services and providers, biotechnology, pharmaceuticals, and life sciences tools and equipment. Each of these sub-industries has its own characteristics, considering industry dynamics, consolidation, regulation, industry structure, and market size (MCSI Inc., 2021). These industries are all technology and R&D intensive, except for the sub-industry healthcare services and providers, which mainly includes healthcare insurance and hospital chains. Before evaluating this work's contribution to current literature, the main sub-industries are further investigated regarding their specific structure and the subsequent impact on this analysis.

Biotechnology & Pharmaceuticals. The pharmaceutical industry has seen a considerable amount of consolidation (Pidun, 2019). Many of these large-scale mergers in the pharmaceutical industry were motivated by increased economies of scale in R&D or production (Ascher, Bansal, Dhankhar, & Kim, 2020; Graves & Langowitz, 1993; The Economist, 2008). The volume of pharmaceutical M&A deals has been on the rise in the last years and the industry has encountered some large deals. In recent years some of the largest firms in the pharmaceutical area merged (Chee, 2018; J. Hall & Krauskopf,

2009). The number and volume of deals has peaked in 2019 before the COVID-19 crisis, with a total deal value of 414B$ (Ascher et al., 2020).

The biotechnology and pharmaceutical industry is characterized by substantial R&D investments, which are a risky undertaking as many developed drugs do not prove to be effective (enough) and fail to get approval for sale by regulatory authorities. The average capitalized cost of a new product in the bio-pharmaceutical market was estimated to be about 1.2B$ in 2003 (DiMasi, Hansen, & Grabowski, 2003). On average, firms spend about 18% of their revenue on R&D (Danzon, Epstein, & Nicholson, 2007). Furthermore, the industry is shaped by rapid innovation, technological complexity and highly specialized skills (Schweizer, 2005b). M&A is often used by larger pharmaceutical firms to gain access to small biotechnology or pharmaceutical firms, providing them with cash and support for clinical trials as well as marketing and sales activities or production capabilities (Ascher et al., 2020; Powell, Koput, & Smith-Doerr, 1996; Schweizer & zu Knyphausen-Aufseß, 2008).

Pharmaceutical firms depend heavily on innovation and the introduction of new drugs: about 75% of revenue of a drug is made within five years after approval (Higgins & Rodriguez, 2006). A study by Danzon et al. (2007) found that pharmaceutical firms tend to acquire firms when patents tend to expire or they expect gaps in their product pipeline.

There is considerable evidence that larger pharmaceutical firms are less innovative than smaller biotechnology firms (Graves & Langowitz, 1993; Ornaghi, 2009). With an active M&A strategy, larger firms are not only able to acquire innovation, but also to avoid the risk of developing drugs in-house. For each new approved drug on the market, about five went into clinical trials. Therefore, buying promising smaller firms (often with drugs in advanced clinical trials) can be cheaper than in-house development (Danzon et al., 2007, 2005).

Furthermore, CEOs of large pharmaceutical firms have obligations to shareholders, which expect value creation and increased profitability. This might lead them to focus on certain capabilities that improve the firm performance and maximize the shareholder wealth, such as commercialization and production. To avoid risk and enhance firm performance, managers might lean towards M&A as a substitute for internal R&D (Higgins & Rodriguez, 2006). Additionally, portfolios are often reorganized and improved, therefore firms often choose to divest or invest in adjacent fields (Ascher et al., 2020; Schweizer, 2005b).

Healthcare & Life Science equipment. The healthcare equipment industry is facing fierce international competition in both their product markets and their level of technology (Lin & Jang, 2010). As the rest of the healthcare industry, healthcare equipment and healthcare technology firms are highly R&D and technology intensive and invested on average 11.4% of revenues in R&D in 2002 in the US (Panescu, 2006). The complexity of products is continuously increasing, as products incorporate new features in microtechnology and often use cutting-edge materials (Panescu, 2006). The healthcare equipment sector is characterized by short product life cycles and constant introduction of new products. It has been estimated that up to 80% of profits originate from new products, introduced in the last five years (Simoens, 2008).

Global markets for medical devices tend to be fragmented and sometimes suffer from a lack of transparency. Device prices are often determined by regulation. In some countries, healthcare investments are often part of public spending, therefore firms might be dependent on state budgets (Simoens, 2008). Products depend heavily on regulation and approval by government authorities and need to undergo a separate approval process in several different countries (Lin & Jang, 2010).

The life science sector tools & services sector mainly includes companies that manufacture equipment for drug discovery, development and production (MCSI Inc., 2021).

2. Theoretical Background

This thesis aims to replicate and expand previous research by Ahuja and Katila (2001), by analyzing the effects of M&A on innovation performance in the healthcare industry. In this chapter, a theoretical framework is developed out of the knowledge-based view (KBV), the combinative capability, the absorptive capacity, as well as M&A performance. Additionally, this chapter includes a review of existing empirical studies. Finally, this chapter states the main empirical hypotheses which are derived out of the theoretical considerations and which will be tested in the empirical part of this thesis.

2.1. M&A Performance, Innovation Performance and Value Creation

Previous research has used different metrics to analyze M&A performance, depending on the area of research. While research in corporate finance has mainly relied upon financial metrics such as shareholder value or accounting measures, strategic management research has often relied upon shareholder value or more subjective variables derived from questionnaires (e.g., synergy realization) (Zollo & Meier, 2008).

Typically, the M&A performance is considered positive when either the accounting or the financial performance of the combined firms is larger than the performance of the separate entities (Barney, 1988; King et al., 2004; Zollo & Meier, 2008). Many studies use different time horizons, ranging from a couple of days to multiple years (Zollo & Meier, 2008). In contrast to financial metrics, the innovation performance of an acquirer can already be evaluated as positive when the post-acquisition innovation performance of the acquiring firm has improved. M&A can affect both: innovation input and innovation output (Ahuja & Katila, 2001; Cassiman et al., 2005). Evaluating the post-acquisition innovation output provides evidence on the efficiency of the resource transfer and combination process (Ahuja & Katila, 2001). Metrics for the innovation output can be patents (Ahuja & Katila, 2001)

or the rate of new product introduction (Capron & Pistre, 2002). An increased innovation performance should lead to an increased economic rent in the future.

In the context of innovation input, improving performance could lie for example in R&D synergies. R&D synergies can be created in various ways. An example is a merger of two firms doing research in similar biotechnology fields. A combined R&D program can lead to a significant breakthrough and therefore to a competitive advantage in one market (Henderson & Cockburn, 1996). Additional input resources, such as the knowledge and capabilities of one R&D team, enable the firm to enhance its innovation performance. Even non-technological resources might enhance the innovation performance, as a target can bring better commercialization capabilities (Kaul, 2012). Admittedly, in most cases, the target will take over most routines and processes of the acquirer (Reus, 2012).

Post-acquisition integration affects the performance as well, as different degrees of post-acquisition integration lead to different outcomes. There are four different ways to integrate target firms: absorption, preservation, holding and symbiosis (Haspeslagh & Jemison, 1991; Pidun, 2019). Firms that remain independent might stay innovative, but the acquirer might not be able to fully benefit from it (Datta & Grant, 1990). "Acquisitions can increase innovation adoption by altering the subsidiary's leadership structure and removing obstacles to change" (Barden, 2012, p. 1269).

However, if the firm is deeply integrated into the acquirer, it might lose its ability to innovate. In an acquisition, the productivity of R&D employees usually falls sharply and the retention of key R&D personnel is often difficult when R&D departments experience a loss of autonomy (Paruchuri & Eisenman, 2012).

Innovation inputs can be affected in a negative way, e.g., when the funding of R&D operations is reduced (Hitt et al., 1996).

Multiple factors affect value creation in M&A. The M&A process involves multiple steps with various pitfalls, from target selection to post-acquisition integration (Pidun, 2019). According to the KBV key challenges lay in identifying, valuing and incorporating resources or knowledge of the target firm (Reus, 2012).

2.2. Knowledge-Based View

2.2.1. Key Concepts

The resource-based view (RBV) sees the source of economic rents and competitive advantage in the resources and capabilities of a firm (Barney, 1991). The level of analysis is "on the distinctive resource profiles of individual firms and the processes, both at the firm and industry level, that lead to specific new resource combinations and induce or reinforce heterogeneity among firms" (Rugman & Verbeke, 2002, p. 770). Firms can only gain a competitive advantage if they implement a value generating strategy by using their resources and capabilities in a way that other firms cannot copy (Barney, 1991; Pidun, 2019). The knowledge-based view (KBV) can be considered a variant of the RBV, viewing the firm as a bundle of knowledge instead of resources and capabilities. This knowledge is not rooted in individuals, but rather in the firm's specific mechanisms of exploring and exploiting knowledge. Knowledge is the primary source of competitive advantage and value creation (Grant, 1996; March, 1991; Reus, 2012). Thus, knowledge is considered the most strategically significant resource of a firm. To gain future rents, firms should invest in assets "that correspond to a combination of current capabilities and expectations regarding future opportunities" (Kogut & Zander, 1992, p. 385). Innovations are the result of the firms' ability to reconfigure knowledge or combine existing and new knowledge within an organizational context (Grant, 1996; Henderson & Cockburn, 1996; Kogut & Zander, 1992). In general, the development of knowledge in a firm is idiosyncratic, reflecting the firm's particular history and experience (Cohen & Levinthal, 1989). Industry competition is driven on the basis of the firms' ability to generate and utilize knowledge efficiently and effectively for innovation (Nonaka, 1994). This idea follows Schumpeter's theory of economic development and innovation. Innovation is the main driver of progress and economic development, as well as competition among firms (Schumpeter, 1981). Firms combine knowledge to create innovation and therefore economic development:

"To produce other things, or the same things by a different method, means to combine these materials and forces differently ... Development in our sense is then defined by the carrying out of new combinations" (Schumpeter, 1981, pp. 65-66).

According to Kogut and Zander (1993), firms are superior in comparison to markets in organizing and transferring knowledge. While firms specialize in generating and transferring knowledge, the mechanism of M&A enables them to transfer knowledge across entities (Kogut & Zander, 1993).

Knowledge has certain properties that make it a strategically relevant resource and allow it to be a foundation for a competitive advantage. An important characteristic is tacitness (Kogut & Zander, 1992). Tacitness depends on the degree of codifiability of knowledge. Tacitness is based on the idea that human knowledge is beyond explicit understanding (so called Polyani's puzzle). This form of knowledge is especially relevant for a competitive advantage, as this kind of knowledge is hard to imitate or replicate by others. In an organizational context, this often comprises of know-why, know-how, and certain skills (Grant, 1996). Furthermore, knowledge is socially embedded into the networks of interpersonal relations among employees and other agents.

M&A can enable the firm access to new knowledge. However, acquiring knowledge, technology or capabilities is not enough. Firms need to be able to explore and exploit new knowledge to develop products or to innovate. This requires two forms of organizational capabilities: the combinative capability of a firm and the absorptive capacity of a firm (Reus, 2012).

2.2.2. Combinative Capability

Kogut and Zander (1992) introduce the concept of combinative capability. This concept refers to the firm's capability to exploit existing knowledge and the unexplored potential of technology, which is dependent on organizing and recombining knowledge. The combinative capability is the essential capability to innovate.

An important presumption is that the knowledge of a firm is socially constructed, resting in the organization of human resources (Kogut & Zander, 1992).

Therefore, the combinative capability is especially crucial in M&A, in order to learn from the acquired knowledge. "The knowledge of the firm has an economic value over market transactions when firm identity leads to social knowledge that supports coordination and communication" (Kogut & Zander, 1996, p. 502). Thus, the combinative capability depends on the firms' social community and its ability to organize, share and distribute knowledge. Firms that have a beneficial culture for knowledge sharing might be able to better exploit their knowledge (Bresman, Birkinshaw, & Nobel, 1999; Haspeslagh & Jemison, 1991).

Knowledge is of little value if it does not result in products. "The ability to indulge in a forward-looking development of knowledge is strongly contingent on the selection environment. Long-term survival involves a complex tradeoff between current profitability and investing in future capability. Future capabilities are of little value if the firm does not survive" (Kogut & Zander, 1992, p. 393). Apart from the combinative capability, the absorptive capacity of an organization determines the ability to generate value from knowledge.

2.2.3. Absorptive Capacity

Cohen and Levinthal (1990) define the absorptive capacity of a firm as the "ability to recognize the value of new information, assimilate it, and apply it to commercial ends" (Cohen & Levinthal, 1990, p. 128). Thus, the absorptive capacity of an organization depends on the ability to transfer knowledge between and within units and to exploit knowledge. Knowledge needs to be passed to the right location within the firm to be exploited. Cohen and Levinthal (1990) and Cohen and Levinthal (1989) see the absorptive capacity as a byproduct of R&D activities. Absorptive capacity is thus a function of prior knowledge. As learning is cumulative, the learning performance is highest when learning items are related to prior knowledge. In consequence, diversity of knowledge is very important, in order to relate new knowledge to what is already known (Cohen & Levinthal, 1990).

Correspondingly, the absorptive capacity is a dynamic capability, which develops and changes over time with the firm operating in different environments and markets (Zahra & George, 2002). Consequently, the absorptive capacity of a firm is (1) depending on the absorptive capacity of the individual members within the organization, (2) developing over time cumulatively and path-dependent and (3) depending on the ability to identify the correct points of contact for external and internal knowledge within the organization (Cohen & Levinthal, 1989, 1990; Reus, 2012).

The concept of absorptive capacity has been extended into potential and realized absorptive capacity. Potential absorptive capacity covers knowledge acquisition and assimilation capabilities, while realized capacity describes the ability of knowledge transformation and exploitation (Zahra & George, 2002).

In empirical research, the absorptive capacity has been operationalized in mainly two ways: R&D intensity and patents (Lane, Koka, & Pathak, 2006).

The absorptive capacity is a rather complex, multilayered concept, involving different factors, such as individual capabilities, firm capabilities and external factors (Lane et al., 2006).

There is a difference between the absorptive capacity and the combinative capability. The absorptive capacity describes how firms can explore their knowledge and absorb new knowledge, while combinative capability describes how firms can exploit their knowledge. Both capabilities are essential for a firm's ability to build knowledge-based resources.

2.2.4. Knowledge-Based View and M&A Innovation Performance

M&A marks a disruptive point in the lifetime of an organization and can therefore disrupt internal processes, which can also affect the process of post-merger integration and knowledge transfer. Post-merger integration either hinders or enables acquirers to use the technology (Bresman et al., 1999; Haspeslagh & Jemison, 1991). Scientists, engineers and other types of "knowledge workers" have socially embedded routines for conducting R&D. These routines can be disturbed by an acquisition, which leads to a decline in productivity (Paruchuri, Nerkar, & Hambrick, 2006). As strategically important knowledge is often tacit (Cohen & Levinthal, 1990), it is difficult to transfer it between two separate entities, as this kind of knowledge is socially embedded. There is even a risk of breaking these kind of resources (Larsson & Finkelstein, 1999).

Successful integration of large bundles of knowledge-based resources is difficult following M&A deals, sometimes even leading to a loss of knowledge-based resources (Reus, 2012). There is plenty of literature on why integration fails (see for example Capron, Dussauge, and Mitchell (1998); Haspeslagh and Jemison (1991); Jemison and Sitkin (1986); Larsson and Finkelstein (1999); Pidun (2019)).

"One of the central dilemmas in managing acquisitions — and perhaps the pivotal factor in affecting employee disruption — is the decision about whether to integrate the newly acquired firm and the acquiring firm" (Puranam et al., 2006, p. 545). A target firm that remains independent might keep a high degree of inventiveness, but the acquirer will receive no benefit for paying the high acquisition premium. On the other hand, a full integration might lead to a significant reduction of inventiveness of the target, by causing "organiza-

tional trauma" or damage of the desired resources (Zollo & Singh, 2004).

"Technologically rich acquisition targets provide opportunities for organizational learning by exposing the acquirer to new and diverse knowledge" (Makri et al., 2010, p. 603). The success of post-acquisition integration depends on the combinative capability and absorptive capacity of both firms.

An acquisition can be seen as a unification of knowledge, which can act as a trigger event for new innovations as both firms start to explore their knowledge. Furthermore, acquisitions enable firms to develop capabilities and expertise by acquiring new and unfamiliar knowledge developed by target employees (Kogut & Zander, 1992).

2.3. Previous Empirical Studies

There is a small stream of research concerning at M&A and innovation in strategic management research. Previous studies rely on different theoretical frameworks, draw their ideas from different schools of thought and use different methodologies. The results are mixed: Some studies show positive effects, while others show neutral or even negative effects. Research has mainly relied upon a couple of proxies for measuring the innovation performance, which include R&D indicators, patenting and new product introduction (Ahuja & Katila, 2001; Cassiman et al., 2005; Freeman & Soete, 1997; Stuart, 2000). Finding a unified measurement for innovation is still an unresolved challenge in the area of management research.

2.3.1. Pre 2000

The first studies in this area are based on the framework of a market for resources within the firm. The firms' managers face a trade-off between growth by acquisitions and investments in other areas of the firm, such as commitment to innovation. Managers will shift their attention away from innovation and internal R&D in favor of the process of integrating target firms or other M&A related activities, therefore affecting the innovation performance negatively (Hitt, Hoskisson, & Ireland, 1990).

Hitt, Hoskisson, Ireland, and Harrison (1991) found significant negative effects of M&A deals on R&D inputs (expenses) and R&D outputs (patents) of acquiring firms using a sample of 191 US based manufacturing firms. This indicates that managers use M&A as a substitute for internal innovation sourcing. Furthermore, the reduction in R&D output can lead to a decline in innovativeness and therefore a decline in financial performance in the long run. The decline in R&D inputs also suggests that acquirers are not fully exploiting acquired technologies, as the investments are eventually not sufficient to foster innovation.

In a subsequent study, Hitt et al. (1996) found a significant negative relationship between a high M&A frequency and the frequency of internal innovations as well as the amount of R&D spending. They attribute this effect to transaction costs and managerial attention devoted to M&A, instead of internal innovation processes. The researchers used a sample of 250 US based firms and evaluated surveys of managers and detailed firm data.

To conclude, the research of Hitt et al. suggests that firms with a M&A strategy might shift their innovation sourcing away from internal innovation towards external innovation, which is associated with a decline in post-acquisition R&D spending and lower rates of internal innovation: "Acquisitions can decrease innovation adoption by taxing resources and the managerial attention" (Barden, 2012, p. 1280).

2.3.2. Post 2000

Cassiman et al. (2005) conducted interviews with key personnel from 31 firms on 62 acquisitions to analyze the relationship between M&A, technological and market relatedness and R&D expenses. They found that firms with complementary technologies have higher R&D expenses after acquisitions and tend to increase R&D efficiency. Technological substitutive firms have lower R&D expenses following an acquisition. This might be due to the technological motive of scope M&A deals in complementary M&A. Firms then try to use synergies to become more innovative. On the other hand, R&D efficiency increases when the two firms operate in technologically complementary areas. Their findings indicate a predominant scope motive of technological M&A, rather than a scale motive (Henderson & Cockburn, 1996).

Desyllas and Hughes (2010) investigated the effect of acquisitions on R&D-intensity (R&D expenditures / total assets) and R&D productivity (patents granted / R&D expenditures) in high-tech industries using a sample of 573 firms and 2624 acquisitions. They found neutral effects in the first year of acquisitions on R&D-intensity, which turn to positive effects after that. The overall effect is positive. The effect of acquisitions on R&D productivity is negative in the first year, which diminishes to an overall neutral effect. On top of that, they analyze the effect of absorptive and financial capacity. They found a positive relationship between the size of an acquirer's knowledge base and the post-acquisition R&D productivity, supporting the hypothesis that "superior absorptive capacity will be better positioned to select appropriate targets and exploit the acquired knowledge base" (Desyllas & Hughes, 2010, p. 1118). They found a positive effect of relatedness in terms of product-relatedness on R&D productivity, supporting the findings of Cassiman et al. (2005).

Both studies show that M&A can improve the R&D efficiency, depending on the circumstances of the involved firms. Related acquisitions seem to have a positive effect on R&D efficiency and therefore have a positive effect on innovation.

Another research stream emerged from the RBV and the KBV. These studies mainly analyze the size of knowledge bases and technological relatedness of knowledge of the involved entities, using patents as a proxy for innovation.

Ahuja and Katila (2001) employed a KBV approach, viewing patents as a proxy for innovation and knowledge of firms. They analyze the innovation performance of acquiring companies of 71 firms active in the chemical industry. The dimensions studied were the absolute and relative size of the

acquired knowledge base as well as the relatedness of knowledge. They found a small positive effect of acquiring a large patent portfolio and a negative of a large relative knowledge base of the target compared to the acquirer. Furthermore, they found a curvilinear effect of the relatedness of the knowledge bases. Low and high relatedness of knowledge lead to a small positive increase of innovation performance, while a moderate amount of relatedness of knowledge leads to a bigger increase of innovation performance. Non-technological acquisitions have a negative effect on innovation output.

Cloodt et al. (2006) did an extended replication study of Ahuja and Katila (2001). They examined aerospace and defense, computers and office machinery, pharmaceuticals, and electronics and communications industry with over 347 acquiring firms and 2429 M&A deals. They reported similar findings to Ahuja and Katila (2001), besides the effect of the absolute size of acquired knowledge. The effect is slightly negative, implying that it might be challenging to integrate a large amount of knowledge.

Ahuja and Katila (2001) and Cloodt et al. (2006) used the same methodology and tested the same hypotheses, but did not find the same results. According to their research, performance is the lowest for deals in which both firms operate in similar or very unrelated technology or knowledge areas and the highest when there is a moderate degree of technology or knowledge relatedness. This indicates that there needs to be enough new knowledge, but also enough similar knowledge to generate value from M&A. Consequently, firms should choose targets that provide technological input with a degree of differentiation in knowledge and technology. This knowledge possibly enriches the innovation performance.

Building upon previous research, Makri et al. (2010) consider science similarity or complementarity as drivers of innovation in M&A. They analyzed 95 M&A deals in knowledge intensive high-technology industries. Their results suggest that complementary scientific knowledge and complementary technological knowledge contributed to enhanced post-acquisition innovation performance. Knowledge complementarity leads to a higher quality and quantity of inventions. In their study, the authors constructed different metrics using patent data. Moreover, knowledge similarity had no effect on innovation quantity and quality.

Sears and Hoetker (2014) split the knowledge relatedness into two parts: acquirer knowledge relatedness and target knowledge relatedness. Apart from this, they also used the measure of patent forward citations as a proxy for technological capabilities. These factors affect value creation from the firms' capabilities differently due to absorptive capacity, knowledge redundancy and organizational disruption. They found a negative relationship between both target relatedness and cumulative abnormal returns (CAR) and patents. As relatedness increases, less value can be generated from each unit (patent) of knowledge. Interestingly, they had similar findings using patents instead of CAR.

To conclude, Makri et al. (2010) and Sears and Hoetker (2014) add to previous research by analyzing not only knowledge relatedness, but also the quality of knowledge. They suggest that there is a "sweet spot", where the complementarity of knowledge enables firms to increase their innovation performance from M&A.

Bena and Li (2014) showed that firms with low R&D expenses and a comparatively large patent portfolio are more likely to become acquirers, while firms that have high R&D expenses and innovation are more likely to become targets. Furthermore, knowledge relatedness between firms had a positive effect on the likelihood of merger pair formation, which is reduced when firms operate in the same product markets.

Colombo and Rabbiosi (2014) analyzed the coherence of R&D restructuring activities and innovation performance. The results of their empirical analysis imply that technological similarity between firms has a large and direct negative effect on post-acquisition innovation performance. The effect was mediated by restructuring the R&D department of the target firm. To the author's knowledge, this is the only study that analyzes such activities.

In one noteworthy study, Higgins and Rodriguez (2006) focused on M&A as a means to outsource R&D in the pharmaceutical industry. Using the variables 'expected years of patent life', 'marketed products' and 'upcoming products', they construct a 'desperation index'. Firms with a high desperation index were more likely to acquire another firm. In the year following a deal, 71% of acquirers had at least maintained or even improved their product pipeline. On top of that, they report that firms which engaged in a strategic alliance with the target before the deal enjoy higher CAR. All analyzed firms formed a strategic alliance during or before the sample. This suggests that alliances might be a tool to reduce the information asymmetry between the acquirer and target.

2.3.3. Review of Previous Studies

To summarize, much of the previous research has addressed different problems in M&A and has built upon different theoretical frameworks. The author has identified a couple of research streams that build on the same frameworks and have coherent empirical findings. The overall literature, however, has many different granular findings, which are quite difficult to generalize due to the complexity of M&A, the challenge of measuring innovation and the interrelation of factors affecting innovation. To deal with this complexity, most studies focus on certain aspects and characteristics of M&A deals. On top of that, different studies have different sampling approaches, some exclude large M&A deals (e.g., Makri et al. (2010)), while others exclude small "hands-on" deals (e.g., Ahuja and Katila (2001), Cloodt et al. (2006) and Sears and Hoetker (2014)). Furthermore, research has been mainly concerned with high-tech industries.

Some trends consistently appear in literature: M&A seem to have a negative effect on R&D expenses (Cassiman & Veugelers, 2006; Hitt et al., 1996). The relatedness (complementarity or similarity) of knowledge bases is an important predictor for the post-acquisition innovation (measured by

patents) (Ahuja & Katila, 2001; Cloodt et al., 2006; Makri et al., 2010; Sears & Hoetker, 2014). As a result, acquired knowledge and technology should be complementary to the acquirer's knowledge base (Kogut & Zander, 1992), which then can lead to a "surplus" over the value that the target's and acquirer's knowledge could provide. Although previous research has provided important insights, there is still an incomplete understanding of the effect of M&A on corporate innovation.

2.4. Hypotheses

As this work intends to replicate the findings of Ahuja and Katila (2001), the first four hypothesis are based on Ahuja and Katila (2001).

There are multiple ways how acquisitions can affect the post-acquisition innovation performance. First of all, not all acquisitions are motivated by technology. Motives for acquiring other companies can be access to markets, sales channels, vertical or horizontal integration, diversification, to gain synergies or to increase market power (Cassiman & Veugelers, 2006; Schweizer, 2005b).

Acquisitions that do not involve any patents or do not have a technological motive are unlikely to improve the post-acquisition performance, as these acquisitions will likely not provide any technological input for the acquirer. The knowledge of a target provides little to none technological value to the acquiring firm (Ahuja & Katila, 2001; Reus, 2012). In this case, an acquisition should not increase the innovation performance of an acquiring firm.

Even improved routines or processes for innovation of the target are unlikely to be transferred to the acquiring firm, as post-acquisition integration decision tends to be housed in the acquiring firm. The target will likely converge to similar innovation routines as the acquiring firm (Kapoor & Lim, 2005).

Acquisitions need considerable managerial attention and consequently consume managerial resources. This can lead to a decline in innovation performance after an acquisition, because managerial attention is devoted to the acquisition instead of internal innovation projects (Haspeslagh & Jemison, 1991; Hitt et al., 1996). Moreover, acquisitions are characterized by significant disruptions of organizations, which can affect all dimensions of productivity (Bresman et al., 1999; Haspeslagh & Jemison, 1991).

Therefore, one would expect that non-technological acquisitions, that do not involve any patents, affect the innovation performance either negatively or non-significantly.

H1: *The post-acquisition innovation performance will be affected either negatively or non-significantly by non-technological M&A* (Ahuja & Katila, 2001).

Deals that are motivated by technology and include a certain amount of technological knowledge in the form of patents can have different effects on the post-acquisition innovation performance of the acquiring firm. An acquisition of a firm can be seen as an acquisition of a knowledge base. The acquired knowledge base needs to be integrated and then combined with the knowledge base of the acquiring firm, which can lead to an improved innovation performance.

On the other hand, acquisitions can also have a negative impact on innovation, as they mark a disruptive event in a firm's lifetime. To investigate whether the impact of an acquisition is positive, negative or non-significant, the following dimensions need to be considered here: the size of the acquired knowledge base, the size of acquirer knowledge base and the relatedness of the knowledge bases (Ahuja & Katila, 2001; Lubatkin, 1983). The interplay of these factors should determine the post-acquisition innovation performance.

An acquisition can improve the innovative output of the acquiring firm by three possible mechanisms: the (successful) merging of two distinct knowledge bases can lead to improved economies of scale and/or scope and improved knowledge recombination (Henderson & Cockburn, 1996): "M&A can realize economies of scale through the spreading of fixed costs over more output and the elimination of common inputs for the production of the same output" (Cassiman et al., 2005, p. 200). In the healthcare industry, for example, an acquisition allows more employees to use expensive R&D equipment for virtually no additional costs for new purposes (e.g., lab equipment). Additionally, larger R&D bases can lead to more specialization and increased economies of scope. Combination of new and existing knowledge can yield new ideas, which then lead to new technologies and innovations. Exchange between both R&D units can stimulate the further development of ideas among colleagues. With an increased size of the combined knowledge bases of target and acquirer, the extended range of possible combinations of knowledge can foster an increased patent output (Galunic & Rodan, 1998; Kogut & Zander, 1992). An acquirer has more opportunity for learning when buying a technologically rich target (Makri et al., 2010). Therefore, the absolute size of acquired knowledge gives the acquirer more opportunity to combine and recombine knowledge.

For the reasons stated above, the author expects a positive relationship between the absolute size of the acquired knowledge base and the post-acquisition patent output of the acquiring firm.

H2: *A larger absolute size of the acquired knowledge base increases the post-acquisition innovation performance* (Ahuja & Katila, 2001).

Knowledge is primarily transferred through interactions between individuals in both the acquired and acquiring entity. Resources are bound in the integration process in the form of integration teams, meetings and considerable communication between both firms. The transfer of knowledge strongly depends on this process (Haspeslagh & Jemison, 1991). If the two knowledge bases are relatively equal in size, most of the firm's knowledge resources are devoted to the task of unifying the two distinct knowledge bases. This can lead to a decrease in innovation performance, as the re-

sources for innovating are absorbed by the process of integration. If the acquiring knowledge base is much larger, however, the integration should be easier, as fewer resources of the acquiring unit are bound to the integration.

On the other hand, if the acquired knowledge base is far larger, the integration should be more difficult, as there are fewer resources on the acquirer side to handle integration and knowledge absorption. Furthermore, the entire acquiring organization needs to adapt its structures and processes to handle the integration (Haspeslagh & Jemison, 1991; Pidun, 2019).

Thus, the author expects a negative relationship between the acquired size of knowledge and the acquirer size of knowledge.

H3: *A large relative size of the acquired knowledge base will have a negative impact on the post-acquisition innovation performance* (Ahuja & Katila, 2001).

The third dimension is the relatedness of knowledge bases. The knowledge of both firms should be related, so that they can exchange knowledge and are familiar with their knowledge. However, there should also be room for new ideas. Closely related knowledge might be easier to integrate, but provides smaller room for new ideas. This kind of knowledge provides little to no value to the acquirer and should therefore not affect the patent output. Very different knowledge bases are difficult to integrate and to commercialize, since there is no connection to what is already known to the acquirer. In this case, the acquirer might not know how to commercialize the target knowledge. Therefore, a mixture of new and familiar knowledge should provide the best basis for improved post-acquisition innovation performance (Ahuja & Katila, 2001; Cloodt et al., 2006; Cohen & Levinthal, 1990; Reus, 2012).

H4: *The relatedness of the acquired knowledge base will be curvilinearly (inverted U-shape) related to the post-acquisition innovation performance of the acquiring firm* (Ahuja & Katila, 2001).

The relatedness of acquirer knowledge describes the knowledge that is part of the knowledge of the acquirer and known by the target. The higher the relatedness, the larger the share of already familiar knowledge. An increase in acquirer relatedness should have a negative effect on post-acquisition innovation performance, as the acquirer does not have the opportunity to learn anything new. In addition, acquirer employees might even retain their knowledge in the integration process, because of the risk of becoming redundant. Therefore, acquirer employees have an incentive to resist the target integration and the resulting combination of knowledge will be affected in a negative way (Sears & Hoetker, 2014). This, in turn, should have a negative effect on innovation performance.

H5: *The larger the relatedness of the acquirer knowledge, the smaller the post-acquisition innovation performance*.

The model of Ahuja and Katila (2001) falls short in capturing the effect of absorptive capacity on innovation performance. Absorptive capacity refers to the ability to recognize, assimilate and apply new external knowledge. It depends on prior knowledge and diversity of knowledge (Cohen & Levinthal, 1989, 1990). Knowledge transfer between firms depends on how easily knowledge can be communicated, interpreted and absorbed (Kogut & Zander, 1992, 1993). With more cumulative and repetitive experience in M&A, firms should be able to better leverage the acquired knowledge (Cohen & Levinthal, 1990; Lane & Lubatkin, 1998; Zahra & George, 2002). Experience with M&A could build absorptive capacity that facilitates knowledge development. This knowledge is explicitly codified into systems, routines, and procedures of the acquirer. Thus, acquisition experience should have a positive effect on innovation performance.

H6: *Previous acquisition experience has a positive effect on post-acquisition innovation performance.*

3. Empirical Analysis

The previous chapter described and evaluated important theories and previous empirical studies on innovation performance in M&A. Six hypotheses were derived from the presented theories, four of them have been suggested by Ahuja & Katila, the other two have been suggested by the author. The following chapter contains a description of the sample, data, data processing, model, empirical methodology, limitations, and variables. Finally, the empirical results will be discussed.

3.1. Sample and Data Description

The data was drawn mainly from two sources: the Thomson Reuters Eikon Database and the PATSTAT database of the EPO. To construct a variable on cultural differences, the data provided by Geert Hofstede was used (Hofstede, 2015).

3.1.1. Sample

The sample was drawn from the Standard and Poor's (S&P) Global 1200 Health Care index (Reuters Instrument Code (RIC): SPGLOHC), which contains all healthcare firms in the S&P Global 1200.[1] The S&P Global 1200 measures the performance of large-cap stocks from all major markets and represents about 70% of world market capitalization (Standard & Poor's, 2021).

S&P uses the Global Industry Classification Scheme (GICS), which classifies firms according to their revenue to specific industries and sub-industries. The Healthcare industry is further classified in the GICS sub-industries "Health

[1]The S&P Global 1200 consists of the different component indices from all over the world: S&P 500 (US), S&P All Australian 50, S&P Europe 50, S&P Asia 50, S&P TSX 60 (Canada), S&P Latin America 40, S&P TOPIX 150 (Japan).

Care Equipment", "Health Care Technology", "Pharmaceuticals", "Life Science Tools & Services", and "Biotechnology". In this work, firms in the industry "Health Care Providers & Services" (351020) were excluded, due to the non-technological nature of this sub-industry. Firms in this industry were mainly based in the United States and active in the area of health insurance.

The distribution of the different industries in the sample can be seen in figure 1. "Pharmaceuticals" have the largest share with about 50% in the sample, followed by "Health Care Equipment", "Biotechnology" and lastly "Life Science Tools & Services".

The sample consists of 41 of the largest players in healthcare. The full list of firms can be found in the appendix in table A4.

This approach was chosen, because the author did not have a way to draw a random sample of firms. Such an approach is fairly consistent with previous research of Ahuja and Katila (2001) and Cloodt et al. (2006). However, Ahuja and Katila (2001) did not specify how they drew their sample. They identified leading players within the chemicals industry and put them together as their sample. The variance of firm size in this sample is smaller than in previous research. The range is from about 7,000 to over 130,000 employees, compared to Ahuja and Katila (2001) with 2,300 to 181,000 employees.

This study analyzes a ten year period between January/01/2006 and December/31/2017. The end of the study period on December/31/2017 was selected, so that it takes into account the time between application and grant of patents. The observation period was set to ten years, which is consistent with previous research (Ahuja & Katila, 2001; Cloodt et al., 2006). Due to the distributed lag in the regression model, the observation period reduced by four years to 2010-2017. The sample contains 41 firms and 328 (41 times per 8 years) observations. The number of M&A deals is 368.

The sample can be classified as a balanced longitudinal data set, meaning that no firm exited the sample. The data set is complete, no data is missing.

3.1.2. Data

The Thomson Reuters Eikon database was used to obtain data on M&A deals and the financial data of the sample firms (R&D expenses and revenues from 2006-2016). The database is among the leading banking databases and includes a vast amount of financial and firm data (Thomson Reuters Eikon, 2021).

The M&A deals were obtained using Deals Screener of Thomson Reuters Eikon. The deals were filtered using the following criteria: the deals had to be completed, larger than 5M$, effective between January/01/2006 and January/01/2017. Additionally, the acquirer needed to own a majority stake in the target (>50% owned after the transaction, including pre-deal shares). The M&A type was set to Disclosed Dollar Value Deal, Undisclosed Dollar Value and Stake Purchases Deal (filter codes: DI, UN, SP; the codes RE and ST are self-tenders and recapitalization and were therefore excluded). This approach is consistent with previous research (Ahuja & Katila, 2001; Cloodt et al., 2006; Hitt et al., 1996). The deals include small size acquisitions and large mergers, the deal size range was from 8M$ to 63B$.

Acquisitions of only certain assets or business units of other firms were excluded manually from the list of deals (e.g., the acquisition of Novartis AG Biologics Manufacturing Facility from Bayer Schering Pharma AG in 2007). Acquisitions of business units had to be excluded, as it was not possible for the author to correctly attribute patents to business units. Further research could address this problem, as it is common in this industry to acquire or divest business units (Ascher et al., 2020).

Not all deals were listed in the Deals Screener app. The author conducted a manual search for each firm using news reports of Thomson Reuters. This problem arose due the name change of Mylan N.V. (now: Viatris Inc.), Valeant Pharmaceuticals Inc. (now: Bausch Health Inc.) and the frequently used name of Waters Corp. Additional 19 deals were identified. Another problem occurred with the merger of "Thermo Electron" and "Fisher Scientific" to "Thermo Fisher Scientific" (TMO) in April 2006. The patents of both firms, before the merger, were retrospectively assigned to the new entity "Thermo Fisher Scientific". As the firms merged in the beginning of 2006, the firm was treated as one entity for the whole sample.

Patent data was drawn from the PATSTAT database of the European Patent Office (EPO). The PATSTAT database was specifically developed for use by government organisations or academic institutions. PATSTAT contains data from more than 100 countries and over 110 million patents. The newest version number was used (Autumn 2020). The patents of the firms were downloaded from PATSTAT using a SQL query.[2]

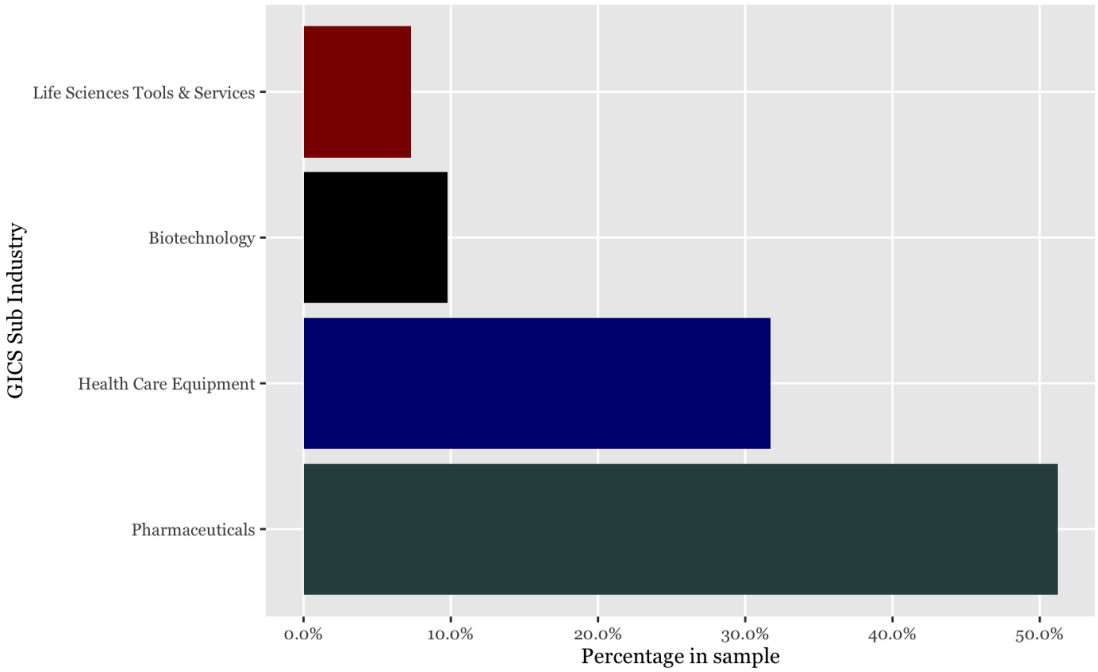
When the patent search did not specify any patents, the acquisition was marked as a non-technological acquisition. This approach is consistent with Cloodt et al. (2006), but not with Ahuja and Katila (2001), as Ahuja and Katila (2001) additionally screened news reports for the motives of the acquisition. Contrary to previous research, patent citations were not included in this research. This would have exceeded the time frame of this research.

3.1.3. Data processing and merging

Patent data of acquiring firms were obtained from January/01/2005-December/31/2017. Patent data of target firms were obtained four years preceding the date of the respective deal. The data from PATSTAT was then matched to

[2]The information included in the download were: appln_id (unique ID of the patent), appln_auth (application au thority), appln_nr (application number), appln_kind (kind of application), appln_filing_date (date of application filing), granted (selector for only granted patents), nb_applicants (number of applicants), nb_inventors (number of inventors), person_id (unique person id for the applicant), person_name, psn_name, person_ctry_code, ipc_class_symbol, ipc_version, and ipc_gener_auth. The variable appln_kind was set to "A", which means that all other document types other than patents were excluded.

Figure 1: Distribution of Global Industry Classification Standard (GICS) Sub Industries in the sample (own figure)



the firm data from Thomson Reuters Eikon using the Python libraries pandas and fuzzywuzzy (pandas development team, 2020; Van Rossum & Drake, 2009). Legal forms of the firms were removed using pandas. For each appln_id a ratio of similarity was calculated. The ratio compares the name of the firm from the list of M&A deals obtained from Thomson Reuters Eikon and from the list of patents from PATSTAT, using the Levensthein distance metric (Levenshtein et al., 1966). If the match was higher than 90%, the patents were attributed to the firm.

The dataset was constructed using the "tidyverse" package of the R programming language (R Core Team, 2019; Wickham et al., 2019). The panel data set was organized in the long format: one observation subject per row and the columns contain the variables.

3.2. Model Specification and Econometric Methodology

The empirical model follows Ahuja and Katila (2001), a random effects Poisson regression model with a distributed lag. The regression coefficients are approximated by a generalized estimating equation (GEE) (Ahuja & Katila, 2001; Liang & Zeger, 1986).

This type of model is appropriate for non-negative count values (number of patents granted per year per firm) and can handle data aggregated over multiple periods. With repeated observations, the correlation among yearly measurements in a generalized linear model would make that model useless. GEE models have been introduced to handle this deficit (Liang & Zeger, 1986; Lipsitz & Fitzmaurice, 2008).

GEE models avoid assumptions about a multivariate distribution and instead rely on a marginal distribution (a subset of random variables) at each time. It relies on the independence of different subjects (clusters) to estimate the variance of the proposed estimators (Liang & Zeger, 1986). In this research setting, each firm represents one cluster.

To account for the post-acquisition integration period, the model includes four distributed lag terms. The earliest possible application for a patent is one year after the M&A deal happened. Patent output of a given year is determined by the input of the preceding years (Ahuja & Katila, 2001). Following Ahuja and Katila (2001) four distributed lag terms are introduced, as they have found good results using four distributed lag terms. They also tested three and five lag terms respectively, but did not find significant differences. Thus, the model will have four distributed lag terms $A_{it-year}$. The regression equation has the following form:

$$P_{it} = exp(X_{it-1}\gamma + A_{it-1}\beta_1 + A_{it-2}\beta_2 + A_{it-3}\beta_3 + A_{it-4}\beta_4)$$

P_{it} is the dependent variable: the number of patents obtained by firm i in year t. X_{it-1} is a vector of control variables. $A_{it-year}$ are the vectors for independent variables (Ahuja & Katila, 2001). The variables are explained in section 3.3.

To calculate the total impact of a variable, the sum of the distributed lag coefficients can be summed up. The standard errors are then determined by the variances plus the covariances between the periods. This allows for hypothesis testing (Ahuja & Katila, 2001; Cloodt et al., 2006).

Since the observations in the longitudinal data set are cluster correlated over time, a Generalized Linear Model (GLM) is not applicable. A GLM would lead to an inaccurate estimation of coefficients and standard errors. Therefore,

Ahuja and Katila (2001) opted for a Generalized Estimating Equation (GEE). GEE models are used for population-average estimations and provide a regression approach for GLMs with longitudinal data, when the different cluster measurements are not independent. This allows for inferences about the population, accounting for the within-cluster correlation. In this research setting each firm is a cluster, with a correlation between the annual measurements. Similar to GLMs, GEE uses a link function that models the relationship between the expected value of the dependent variable to the predictors in the model (Agresti, 2003). In the case of the Poisson regression, the link function is the natural logarithm.

The GEE approach has two advantages: it is flexible and robust. Neither linearity between the predictors and the dependent variables, nor homogeneity of variance for the range of the dependent variable are assumed (Liang & Zeger, 1986).

Regression coefficients are approximated by a quasi-likelihood approach instead of a maximum likelihood algorithm. This approach uses the expected value and the variance of the dependent variables instead of a likelihood function. The quasi-likelihood approach consists of two steps: in the first step, coefficients are estimated using a specified working correlation matrix. Results of this estimate are then used in the next iteration, which then estimates the coefficients and standard errors accounting for the observed correlation of subjects and residuals (Ahuja & Katila, 2001).

To account for possible correlation between the residuals of observations for the same firm, the GEE method requires a working correlation structure to be specified. This working correlation captures the relationship between observations. Examples for such correlation structures are: independent, exchangeable or auto-regressive. In this work, the covariance structure was set to independent, as this is the most robust structure (Stokes, Davis, & Koch, 2012). This covariance structure was used by Ahuja and Katila (2001) as well. In case of a falsely assumed working correlation, GEE is still able to find robust results, by updating the working correlation matrix with each iteration (Lipsitz & Fitzmaurice, 2008).

As GEE uses a quasi-likelihood approach, typical regression fitting criteria that use likelihood functions, such as Akaike's information criterion, are not applicable to GEE models (Liang & Zeger, 1986; Pan, 2001). Thus, Pan (2001) proposed the "quasi-likelihood under the independence model criterion" (QIC). QIC uses a similar approach than Akaike's information criterion, but is based on a quasi-likelihood approach. QIC allows model comparison and model selection.

As previously mentioned, the underlying econometric model is a Poisson regression, which is an appropriate model for discrete, non-negative count variables such as patents (Agresti, 2003; Stokes et al., 2012). The dependent variable is equal to the mean of a Poisson process. The Poisson distribution requires that the variance is equal to the mean. If this assumption is violated, the standard errors can be underestimated (Agresti, 2003). The typical case is when the conditional variance exceeds the conditional mean, which is called overdispersion. There are plenty of sources for overdispersion: variability of subjects, correlation or dependence between individual sample responses, a clustered structure of the population, small sample size and more (Stokes et al., 2012). Table 3 shows the summary statistics. It can be seen that the conditional mean of the variable patents is lower than the variance. This indicates overdispersion (Cameron & Trivedi, 2013).

Furthermore, unobserved effects might lead to a serial correlation of the residuals of observations from the same firm over the years (Ahuja & Katila, 2001). This effect can also lead to an underestimation of coefficients. To solve these problems, this work uses a pre-sample approach. With this approach, unobserved heterogeneity is modeled with an additional covariate (Blundell, Griffith, & Reenen, 1995). For that reason, the *pre-sample patents* variable is included in the regression model, to account for unobserved differences in knowledge stocks between the sample firms (Ahuja & Katila, 2001). This variable is constructed from the values of the dependent variable in the period immediately preceding the observation period. The information provides the basis for controlling for unobserved heterogeneity, which then leads to an reduction in overdispersion and serial correlation (Blundell et al., 1995).

Another possibility to handle overdispersion is a Negative binomial regression. The underlying distribution of this model can be considered a generalization of the Poisson distribution. This lifts the restriction that the variance is restricted to the mean. The Negative binomial distribution has the following properties:

$$E(Y) = \mu \qquad Var(Y) = \mu + D\mu^2$$

The parameter D models the empirical variance, which corrects for overdispersion (or underdispersion). Both regression models are reported in this thesis. The replication study by Cloodt et al. (2006) uses a Negative binomial model. This study includes both models, the Poisson results are reported in this chapter, the Negative binomial results can be found in the appendix.

To find the regression parameters, the Python package "statsmodels" was used (Seabold & Perktold, 2010; Van Rossum & Drake, 2009).

3.3. Variables

Tables 1 and 2 give an overview of all variables included in the model. The following subsection will provide a comprehensive overview of the variables used in this work, as well as the underlying calculations and assumptions made to derive them. Table 4 shows the correlation matrix.

3.3.1. Dependent variable

Patents. The dependent variable is the patent output of company i in year t, which is a non-negative count value.

Researchers in management and economics have used multiple indicators to measure innovation. Typical metrics

Table 1: Overview of control variables

Variable	Description	Source
R&D Expenses	The absolute amount of R&D spending per year.	Thomson Reuters Eikon One
Natural Log. of Revenue	The natural log of firm revenue.	Thomson Reuters Eikon One
Foreign Acquisitions	Control variable for cultural differences of firms.	Homepage of Geert Hofstede
Pre-sample patents	Sum of patents in the three years prior to the entry into each period.	PATSTAT
Country dummy	Categorical dummy variable to capture different countries of origin	-
Year dummy	Categorical dummy variable to capture yearly differences.	-
GICS Industry dummy	Categorical dummy variable to capture industry differences.	-

Table 2: Overview of independent variables

Variable	Description	Source
Number of non-technological acquisitions	The absolute number of non-technological acquisitions.	PATSTAT
Absolute size of acquired knowledge	The absolute number of unique target patents numbers.	PATSTAT
Relative size of acquired knowledge	The absolute number of unique target patents numbers, divided by the absolute size of unique patents of the acquiring firm.	PATSTAT
Relatedness of target knowledge	The overlap of patent classification codes, divided by all patent classification codes of the target.	PATSTAT
Relatedness of acquirer knowledge	The overlap of patent classification codes, divided by all patent classification codes of the acquirer.	PATSTAT
Acquisition experience	Number of acquisitions from 01.01.2006 until the focal year.	Thomson Reuters Eikon One

are R&D expenditures, patents, patent citations or the introduction of new products or new product lines. Other approaches of measuring innovation performance include surveys of executives (Hagedoorn & Cloodt, 2003).

In this research setting patents are appropriate, because in the healthcare industry patents and "comprehensive patent protection is a prerequisite for subsequent commercial success" (Henderson & Cockburn, 1996, p. 38). Additionally, previous studies in the pharmaceutical sector have shown that patents are highly correlated with multiple measures of economic success (Griliches, 1998).

Patents are required to describe something novel and not obvious, and thus provide a good measure of technologically new knowledge as defined above. They have been used as a proxy for innovation in previous studies (Ahuja & Katila, 2001; Cloodt et al., 2006; B. H. Hall, Jaffe, & Trajtenberg, 2005; Sears & Hoetker, 2014). Weaknesses of using patents as a proxy for innovation will be discussed later in section 3.4.2.

3.3.2. Independent variables

The independent variables are included in the vector $A_{it-year}$ with $year \in \{1,2,3,4\}$. This vector contains the *absolute size of the acquired knowledge* base, the *relative size of the acquired knowledge* base, the *relatedness of the target and acquirer knowledge* base and the *number of technological acquisitions* where the patents were unavailable. The variable *acquisition experience* was only included with a lag of one period.

Number of non-technological acquisitions. The first independent variable *number of non-technological acquisitions* was constructed using the data from Thomson Reuters Eikon of

Table 3: Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	Median	Max.
Patents	264.28	284.31	0.00	178.00	1,516.00
No. non-tech. acquisitions t-1	0.17	0.44	0.00	0.00	3.00
No. non-tech. acquisitions t-2	0.18	0.45	0.00	0.00	3.00
No. non-tech. acquisitions t-3	0.18	0.46	0.00	0.00	3.00
No. non-tech. acquisitions t-4	0.19	0.48	0.00	0.00	3.00
Abs. size of acquired knowledge t-1	47.27	190.40	0.00	0.00	2,478.00
Abs. size of acquired knowledge t-2	50.10	198.98	0.00	0.00	2,478.00
Abs. size of acquired knowledge t-3	51.17	206.20	0.00	0.00	2,478.00
Abs. size of acquired knowledge t-4	54.69	217.89	0.00	0.00	2,478.00
Rel. size of acquired knowledge t-1	0.03	0.10	0.00	0.00	0.98
Rel. size of acquired knowledge t-2	0.04	0.10	0.00	0.00	0.98
Rel. size of acquired knowledge t-3	0.04	0.10	0.00	0.00	0.98
Rel. size of acquired knowledge t-4	0.04	0.11	0.00	0.00	0.98
Relatedness of target knowledge t-1	0.33	0.59	0.00	0.00	5.18
Relatedness of target knowledge t-2	0.34	0.61	0.00	0.00	5.18
Relatedness of target knowledge t-3	0.34	0.57	0.00	0.00	5.18
Relatedness of target knowledge t-4	0.34	0.57	0.00	0.00	5.18
Relatedness of target knowledge^2 t-1	0.43	1.74	0.00	0.00	26.82
Relatedness of target knowledge^2 t-2	0.40	1.94	0.00	0.00	26.82
Relatedness of target knowledge^2 t-3	0.46	1.42	0.00	0.00	26.82
Relatedness of target knowledge^2 t-4	0.45	1.37	0.00	0.00	26.82
Relatedness of acquirer knowledge t-1	0.03	0.07	0.00	0.00	0.57
Relatedness of acquirer knowledge t-2	0.03	0.07	0.00	0.00	0.57
Relatedness of acquirer knowledge t-3	0.03	0.07	0.00	0.00	0.57
Relatedness of acquirer knowledge t-4	0.03	0.07	0.00	0.00	0.57
R&D expenses t-1	2,273.05	2,594.85	47.58	1,018.53	10,740.22
Log. revenue t-1	3.99	0.47	2.88	4.00	4.87
Pre-sample patents t-1	879.53	880.94	6.00	601.50	4,420.00
Foreign acquisitions t-1	0.93	0.75	0.00	1.00	4.00
Foreign acquisitions t-2	0.93	0.78	0.00	1.00	4.00
Foreign acquisitions t-3	0.94	0.78	0.00	1.00	4.00
Foreign acquisitions t-4	0.94	0.80	0.00	1.00	4.00
Acquisition experience t-1	7.23	5.04	0.00	6.00	26.00

the M&A activity of the sample firms from 2006-2017 and combining it with the patent data from PATSTAT.

Whenever a target firm had no patenting activity in the preceding five years before the acquisition, the deal was marked as a non-technological deal (Cloodt et al., 2006). Firms therefore do not provide any significant technological input to the acquirer.

Absolute size of acquired knowledge. The second variable, *absolute size of acquired knowledge base* was obtained by summing up the number of patents of target firms of the five years preceding the acquisition (Ahuja & Katila, 2001).

Relative size of acquired knowledge. The third variable, *relative size of acquired knowledge base*, was constructed using the *absolute size of acquired knowledge base* and the number of unique patents of the acquiring firm of the five years preceding the acquisition (Ahuja & Katila, 2001). Subsequently, the variable was calculated dividing the *absolute size of acquired knowledge base* by the number of unique patents of the acquiring firm:

$$relative_size_of_acquired_knowledge = \frac{absolute_size_of_acquired_knowledge}{absolute_size_of_acquirer_patents}$$

As proposed by Ahuja and Katila (2001), when the *absolute size of acquired knowledge base* was larger than the number of patents of the acquirer, the reciprocal fraction was calculated.

Relatedness of target and acquirer knowledge. The fourth and fifth variable, are calculated using the International Patent Classification (IPC) codes. Each patent has multiple classification codes. For each patent, a list of classification codes was prepared. In order to compare the patents, the first four digits of the codes were used. The first four digits allow for a comparison of subclasses. Subclasses are the third hierarchical level of patent classification (World Intellectual Property Organisation, 2020). The use of more digits would lead to an overly granular analysis (Kogler, Rigby, & Tucker, 2013; Leydesdorff, Kogler, & Yan, 2017).

For the variable *target relatedness*, the relatedness between patent codes of both firms (redundancy of codes) was calculated and then divided by the total number of IPC codes of the target's patents, representing the target knowledge.

$$relatedness_of_target_knowledge = \frac{knowledge_overlap}{target_knowledge}$$

The variable *acquirer relatedness* was calculated in a similar way, but using the patent codes of the acquiring firm in the denominator. Note that the target knowledge is different to the absolute size of acquired knowledge. This variable is also different from *relative size of acquired knowledge*, as one patent can have multiple IPC codes.

Acquisition experience. The variable *acquisition experience* was calculated by the sum of all acquisitions until the current observed year from the beginning of 2006. This approach is consistent with previous research (Haleblian & Finkelstein, 1999; Hayward, 2002; Zollo & Singh, 2004).

3.3.3. Control variables

To control for unobserved heterogeneity among firms, a few control variables were introduced to the model (Ahuja & Katila, 2001).

Foreign Acquisitions. Since international acquisitions have higher rates of failure than domestic mergers and involve greater integration challenges (Heimeriks, Schijven, & Gates, 2012) a control variable for acquisition involving foreign firms was introduced. The control variable for foreign acquisitions was calculated following the approach of Kogut and Singh (1988) based on the theory of cultural differences by Hofstede (1980).

The cultural dimensions were first developed by Hofstede (1980) using data on IBM-employees. They provide a metric for a quantitative comparison of cultural differences between countries. The cultural dimensions were subsequently extended by Minkov (2007) and Hofstede and Minkov (2010), to six cultural variables. As Ahuja and Katila (2001) used only the original four dimensions, the formula was adapted to fit the newest framework with six variables.

$$foreign_acquisitions = \sum_{i=1}^{I} (I_{ij} - I_{iu})^2 / V_i / 6$$

I_i is the value of cultural dimensions, V_i is the variance of the dimension. The dimensions are Power Distance, Individualism, Masculinity, Uncertainty Avoidance, Time Orientation, and Indulgence. Each country has a value between 0 and 100 for each dimension (Hofstede, 1980; Minkov, 2007).

Firm Size. Another factor that needs to be controlled for was firm size, since larger firms tend to be less innovative. Firm size was measured as the natural log of revenue (Cloodt et al., 2006).

Ahuja and Katila (2001) use the natural log of employees. Due to missing data availability in Thomson Reuters Eikon, the author decided to discard this variable. The natural logarithm of revenue is an appropriate replacement and used in many similar studies (Cloodt et al., 2006; Sears & Hoetker, 2014).

R&D expenditures. As larger firms enjoy economies of scale and scope in R&D (Henderson & Cockburn, 1996), the control variable *R&D expenses* was introduced to the model. R&D expenses affect the patent output (Ahuja & Katila, 2001).

Pre-sample Patents. To control for firm patent output differences at the start of each period, the variable for pre-sample patents was introduced. The variable was the sum of patents by the acquiring of the last three years for each period. It also serves as a fixed-effect variable to correct for serial correlation and overdispersion (Ahuja & Katila, 2001).

Country dummy. This categorical variable controls for differences in patenting behavior of firms from different countries (Ahuja & Katila, 2001). The reference was set to "ISO Country US"

Year dummy. This categorical variable controls for differences in patenting behavior of each year in the sample, as patenting is also affected by macroeconomic trends and the current economic situation of the firms (Ahuja & Katila, 2001). The reference was set to "Year 2016".

Industry dummy. This categorical variable controls for different characteristics of the GICS Sub-Industries. The reference was set to "Pharmaceuticals".

To avoid the dummy variable trap, one-hot encoding was used (Draper & Smith, 1998).

3.4. Econometric Limitations

3.4.1. Sample

The sample only includes large firms, which are listed in the S&P 1200. Therefore, this research does not cover small or medium sized businesses. The innovation performance of small or medium sized firms in M&A might be very different. Firms were required to stay in the sample for the whole period.

Only firms that share the same name were included. Subsidiaries with different names were excluded, due to different names of patent applicants (e.g., Ethicon Inc., a subsidiary of Johnson & Johnson). In spite of a different name, knowledge transfer or spillovers could happen between those firms and are not taken into account in this study.

Acquisitions of business units were excluded, because patents could not be reliably attributed to individual business units of larger firms. Acquisitions of business units of competitors happened frequently, as it is common to restructure, divest or acquire business units (Ascher et al., 2020). Business units are often described as a bundle of strategically relevant resources and could therefore provide relevant insights (Eschen & Bresser, 2005).

Table 4: Correlation Matrix

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1	Patents	1																																
2	No. non-tech. acquisitions t-1	0.07	1																															
3	No. non-tech. acquisitions t-2	0.04	0.23	1																														
4	No. non-tech. acquisitions t-3	0.03	0.07	0.22	1																													
5	No. non-tech. acquisitions t-4	-0.01	0.01	0.07	0.21	1																												
6	Abs. size of acquired knowledge t-1	0.12	0.12	0.02	0.07	0	1																											
7	Abs. size of acquired knowledge t-2	0.27	0.09	0.08	0	0.04	0	1																										
8	Abs. size of acquired knowledge t-3	0.23	-0.03	0.07	0.09	0.02	-0.01	0	1																									
9	Abs. size of acquired knowledge t-4	0.19	0.09	-0.03	0.07	0.07	0.01	-0.01	0	1																								
10	Rel. size of acquired knowledge t-1	-0.08	0.21	0.25	0.08	0.12	0.66	0.02	0.01	0.04	1																							
11	Rel. size of acquired knowledge1 t-2	-0.01	0.15	0.2	0.25	0.08	0.06	0.62	0.02	0.01	0.1	1																						
12	Rel. size of acquired knowledge t-3	-0.01	0.07	0.14	0.2	0.22	0.04	0.04	0.62	0.01	0.16	0.09	1																					
13	Rel. size of acquired knowledge t-4	-0.04	0.09	0.06	0.12	0.16	0.07	0.01	0.03	0.61	0.22	0.15	0.05	1																				
14	Relatedness of target knowledge t-1	0.27	0.09	0.04	0.07	-0.02	0.39	0.02	0.03	-0.01	0.22	0	0.02	-0.03	1																			
15	Relatedness of target knowledge t-2	0.33	0.11	0.05	0.02	0.06	0.07	0.48	0.02	0.02	0	0.26	-0.02	0	0.15	1																		
16	Relatedness of target knowledge t-3	0.36	0.05	0.05	0.06	0.02	0.05	0.24	0.45	0.02	0.01	0.05	0.24	-0.03	0.11	0.24	1																	
17	Relatedness of target knowledge t-4	0.3	-0.02	0.06	0.04	0.06	-0.01	0.01	0.25	0.45	-0.04	0.01	0.06	0.25	0.01	0.07	0.24	1																
18	Relatedness of target knowledge^2 t-1	0.2	0.05	-0.01	0.02	-0.03	0.37	0	0	-0.02	0.16	-0.03	-0.02	-0.04	0.82	0.06	0.03	-0.03	1															
19	Relatedness of target knowledge^2 t-2	0.27	0.09	0.01	-0.02	0	0	0.53	-0.01	-0.01	-0.04	0.21	-0.04	-0.03	0.07	0.84	0.18	0	0.02	1														
20	Relatedness of target knowledge^2 t-3	0.39	0	0.01	0.03	-0.02	0	0.27	0.56	0	-0.03	0.04	0.22	-0.05	0.07	0.25	0.88	0.25	0.02	0.22	1													
21	Relatedness of target knowledge^2 t-4	0.35	0.01	0.01	0.01	0.04	-0.04	-0.01	0.28	0.57	-0.06	-0.03	0.05	0.22	0.01	0.05	0.24	0.87	-0.02	0	0.31	1												
22	Relatedness of acquirer knowledge t-1	0.04	0.21	0.23	0.22	0.09	0.71	0.02	0.09	-0.01	0.76	0.18	0.25	0.17	0.52	0.05	0.07	-0.02	0.37	-0.02	0.01	-0.05	1											
23	Relatedness of acquirer knowledge t-2	0.15	0.18	0.16	0.2	0.19	0.05	0.75	0.01	0.08	0.09	0.74	0.15	0.21	0.08	0.59	0.17	0.03	0	0.49	0.17	-0.01	0.18	1										
24	Relatedness of acquirer knowledge t-3	0.13	0.05	0.15	0.15	0.18	0.02	0.06	0.75	0.01	0.1	0.09	0.73	0.11	0.03	0.08	0.58	0.18	-0.02	0.01	0.55	0.18	0.2	0.18	1									
25	Relatedness of acquirer knowledge t-4	0.1	0.08	0.04	0.12	0.11	0	0	0.05	0.76	0.06	0.1	0.05	0.73	0	0.01	0.07	0.6	-0.04	-0.03	0.02	0.55	0.08	0.17	0.15	1								
26	R&D expenses t-1	0.49	0.02	0.01	0.01	0.01	0.15	0.21	0.23	0.22	0	0.03	0.05	0.04	0.17	0.2	0.18	0.16	0.17	0.21	0.21	0.21	0.08	0.11	0.14	0.12	1							
27	Log. revenue t-1	0.23	-0.03	-0.01	0.02	0.05	0.08	0.13	0.14	0.13	-0.02	0.02	0.05	0.04	0.13	0.18	0.19	0.16	0.09	0.13	0.16	0.13	0.02	0.08	0.1	0.09	0.37	1						
28	Pre-sample patents t-1	0.91	0.03	0.04	0.05	0.01	0.11	0.22	0.21	0.2	-0.09	-0.05	-0.04	-0.06	0.26	0.32	0.36	0.31	0.2	0.27	0.38	0.36	0.02	0.12	0.12	0.1	0.57	0.28	1					
29	Foreign acquisitions t-1	0.02	-0.01	0.09	0.11	0.11	-0.07	0.04	-0.04	-0.06	-0.1	0.07	0.05	-0.09	-0.1	-0.07	-0.04	0	-0.06	-0.08	-0.02	0	0	0.05	-0.01	-0.08	-0.05	-0.03	0.02	1				
30	Foreign acquisitions t-2	0.08	-0.05	0.01	0.09	0.12	-0.09	0.02	0.04	-0.04	-0.06	-0.08	0.06	0.03	-0.02	-0.04	-0.04	-0.03	-0.02	0	-0.01	-0.01	-0.01	0.04	0.06	-0.02	-0.03	-0.01	0.07	0.26	1			
31	Foreign acquisitions t-3	0.07	-0.02	-0.05	-0.01	0.07	0.06	-0.05	0.03	0.04	-0.03	-0.07	-0.04	0.08	-0.11	0.02	0.01	-0.03	-0.04	0.02	0.05	0	-0.03	-0.01	0.08	0.08	0.02	0	0.08	0.21	0.22	1		
32	Foreign acquisitions t-4	0.05	0.03	0	-0.05	0.02	-0.04	0.05	-0.05	0.01	-0.05	-0.03	-0.12	-0.07	-0.06	-0.08	-0.02	0.01	-0.03	-0.04	-0.01	0.04	-0.1	-0.02	-0.07	0.01	0.02	0	0.06	0.15	0.27	0.18	1	
33	Acquisition experience t-1	0.15	0.24	0.31	0.33	0.35	0.13	0.17	0.19	0.21	0.22	0.28	0.29	0.28	0.22	0.26	0.31	0.28	0.11	0.13	0.22	0.22	0.27	0.32	0.33	0.3	0.22	0.32	0.26	-0.02	-0.03	-0.06	-0.12	1

For the variable *Number of non-technological acquisitions* Ahuja and Katila (2001) analyzed newspaper reports to find the motive behind each M&A deal. In this work, a M&A deal was marked as non-technological if the target firm did not have any patenting activity in the five years preceding the deal. This may lead to different results, as some deals involving patents may not be motivated by acquiring a specific technology and, on the other hand, deals without any patents may be motivated by acquiring a technology. Following the approach of Ahuja and Katila (2001) would have exceeded the frame of this research.

For GEE, some researchers recommend at least 50 research subjects (Stokes et al., 2012), while others recommend at least more than 10, possibly 30 (Norton, Bieler, Ennett, & Zarkin, 1996). This work includes 41 firms with 368 deals. Due to the distributed lag the sample was reduced to 328 observations. More clusters and more data should lead to more robust results (Ghisletta & Spini, 2004; Stokes et al., 2012).

3.4.2. Patents as a Proxy for Innovation

Patents can be seen as a distinct set of elements of knowledge (Ahuja & Katila, 2001). This knowledge can be united in a combinational way to generate new knowledge and innovation. The larger the set of elements, the more combinations are possible and a firm can enjoy more innovation. As previously mentioned, innovations are the result of a firm's knowledge (Kogut & Zander, 1992).

Patents as a proxy for innovation performance suffer from the following weaknesses: not all inventions are patented (Henderson & Cockburn, 1996; Nagaoka, Motohashi, & Goto, 2010) and not all patents yield innovations or significant inventions. Patents can also be used strategically to block competitors from technology areas: one study found that about 40% of patents were filed to block competitors from certain technology areas (Blind, Cremers, & Mueller, 2009; Grimpe & Hussinger, 2014b). On top of that not all patents are published, especially in the US: In 2009 about 7% of the patents were not published. Patents in other countries are usually published within 18 months after the filing date. (Nagaoka et al., 2010). Therefore, patent counts can underestimate the effect of innovation activity in the US, compared to other countries.

Patents are not the only means of protection of intellectual capital. For example, pharmaceutical firms enjoy an additional protection in the United States: the Food and Drug Administration (FDA) grants exclusivity rights for new chemical entities (Higgins & Rodriguez, 2006). Firms also have explicit knowledge, such as "technical know-how", that is transferred between firms in an acquisition (Bresman et al., 1999). Patents fail to measure this effect, thus this research underestimates this effect.

Measuring innovation remains a key challenge in management research. Moreover, patents are only one possible dimension to measure innovation. The capacity to innovate is determined by multiple different factors and can be measured by multiple factors as well. Due to the nature of innovation, it is unlikely that a unified metric will be found (Adams, Bessant, & Phelps, 2006; Hagedoorn & Cloodt, 2003). Besides this, it is hard to quantify a monetary value of patents or of innovation (Griliches, 1998).

In this research approach, knowledge relatedness is measured by the overlap of patent codes. Even though two patents might share the same patent codes, the underlying knowledge is not necessarily related.

Contrary to the work of Ahuja and Katila (2001), this work does not include cited patents. This may critically affect the research. Following the idea of absorptive capacity, Ahuja and Katila (2001) write: "By creating a patent that builds on these prior patents, the firm provides evidence that the knowledge contained in those past patents is a part of the firm's knowledge set" (Ahuja & Katila, 2001, p. 202). A short discussion on patent citations will be in section 3.5.1.

3.4.3. Methodological Weaknesses

As with all econometric and statistical procedures, it cannot be ruled out that the correlation is not accompanied by a causal effect of M&A on innovation: "A convincing identification of causality will be always hindered by the fact that econometricians cannot observe most of the information that the merging firms employ in their decision" (Ornaghi, 2009, p. 78).

As M&As are hugely complex undertakings, a lot of factors affect the outcome. Not all factors can be considered in an empirical analysis. Similar to Ahuja and Katila (2001) this study ignores the acquisition management process, e.g., different degrees of integration are not controlled for. As previously mentioned, there are multiple strategies to integrate firms, each having a different impact on the innovation output of the acquirer (Pidun, 2019). For instance, if the target is left alone (preservation strategy) a non-significant impact on innovation performance of the acquirer can be expected (Haspeslagh & Jemison, 1991).

Innovation performance or the realization of resource recombination depends upon the flow of competency-related knowledge between competence areas (Cohen & Levinthal, 1990). Each deal in the sample has its unique properties and motives. Further analysis and details could provide more insights on acquisitions that involve technology.

Time series analysis allows to analyze the effects of M&A on firm performance, especially the use of distributed lags (Ahuja & Katila, 2001; Sears & Hoetker, 2014). However, as the span of the distributed lag is limited, it falls short of capturing long-term effects.

3.4.4. Poisson regression with pre-sample approach vs Negative binomial Regression

The original study by Ahuja and Katila (2001) used a Poisson regression with a pre-sample approach to control for unobserved heterogeneity, following a suggestion by Blundell et al. (1995). The Poisson regression requires the variance to be equal to the mean, which is often violated. A common cause for this is variance among subjects (Agresti, 2003).

In this case, the computed standard errors in the Poisson regression are then underestimated. The hypothesis testing can then be invalidated. For that reason, a *pre-sample patents* variable is included in the regression model, to account for unobserved differences in knowledge stocks between the sample firms (Ahuja & Katila, 2001; Blundell et al., 1995; Cameron & Trivedi, 2013).

This variable includes the values of the dependent variable in the period immediately preceding the observation period. This information provides the basis for controlling for unobserved heterogeneity, which can lead to a reduction in overdispersion and in serial correlation (Blundell et al., 1995).

An alternative to using a Poisson regression is a Negative binomial regression, which the replication study by Cloodt et al. (2006) uses. The Negative binomial distribution can be considered a generalization of the Poisson distribution, which lifts the restriction that the variance is restricted to the mean. Researchers should report both Poisson and Negative binomial models, when the software allows for it (Cameron & Trivedi, 2013). As the Negative binomial models also account for over- or underdispersion, the results should be more reliable.

3.5. Empirical Results: Estimating the impact of M&A on Innovation

Table 5 shows the results of a Poisson regression with the model, following Ahuja and Katila (2001), table 7 shows the summed model. The author's model is shown in table 6 and the sum of results in table 8. Results of both models with a Negative binomial regression can be found in the appendix, in tables A2 and A3 respectively. The results of a Poisson regression with just the control variables can be found in the appendix in table A1.

First, the results of the replication model of Ahuja and Katila (2001) are discussed, then the results of the authors adapted model.

H1 predicts a negative relationship between the *number of non-technological acquisitions* and the post-acquisition innovation performance. In the estimated model the first three coefficients are negative, followed by a positive coefficient in period four. The coefficients in the second and third period are significant. The summed coefficient, which reflects the total impact is negative as well. Overall these findings support the results of Cloodt et al. (2006) and therefore support **H1** as well. Ahuja and Katila (2001) find a negative, but non-significant effect here. This might be due to the industry nature, as the chemicals industry is less of a high-tech industry, than the healthcare industry (Cloodt et al., 2006). This also supports the idea of Hitt et al. (1996) that M&A has a negative effect on innovation as resources of the firm are consumed and bound to the M&A process, instead of innovation.

The coefficients *absolute size of acquired knowledge base* are all positive besides period four. The overall effect is positive in both models, although of small magnitude. Only the estimates for the first and third period are significant. The summed coefficient over all periods is positive. Ahuja and Katila (2001) and Cloodt et al. (2006) show different results regarding this variable. The results of this study support the findings of Ahuja and Katila (2001) and therefore support **H2**. The target firm's knowledge base will lead to increased economies of scale and scope for the acquirer and therefore an improved innovation performance.

The *relative size of acquired knowledge* has a negative and significant effect on post-acquisition patent output in the first three periods, as predicted in **H3**. This is consistent with both previous studies. **H3** is therefore also corroborated. When the acquirer has to devote a large amount of resources to the integration of a large knowledge base, the post-acquisition innovation performance will suffer (Hitt et al., 1996). The majority of deals in the sample were acquisitions of firms with a small patent portfolio. Only a handful of deals included large mergers.

H4 predicts a curvilinear relationship between innovation performance and target relatedness. This hypothesis can be supported by the results. The coefficients of *relatedness of target knowledge* are all positive, while the coefficients of the squared term are all negative. The knowledge base of both firms should be similar to some degree, but also provide enough new knowledge to fertilize post-acquisition firm learning and knowledge creation (Cloodt et al., 2006).

Unsurprisingly, *R&D expenses* have a positive effect on patent output. The control variable for firm size *Log. revenue* is negative and significant, as larger firms tend to be less innovative (Cloodt et al., 2006).

Foreign acquisitions have a mixed and non-significant effect. The first two periods are positive, while the last two periods have a negative effect. Previous research by Penner-Hahn and Shaver (2005) found that Japanese pharmaceutical firms enjoy greater international R&D productivity, when they already have some previous knowledge in a domain prior to the international expansion. This indicates that there should be a positive effect. Additional research could shed some more light on the relationship between international M&A, R&D productivity, and innovation.

The variable for the control of heterogeneity *pre-sample patents* is positive and significant as well. The "Health Care Equipment" sector seems to be more active in patenting than the other sectors, while "Biotechnology" and "Life Sciences Tools & Services" have a lower patenting activity. All control variables have a small correlation with the dependent variable, besides the pre-sample variable with a correlation of 0.91. This correlation was expected, as the pre-sample variable contains information of the dependent variable about the previous periods, although the correlation is lower than in previous research.

The author's model yields similar findings for the first four hypothesis. All of the hypotheses are supported as well. The newly introduced variables do, however, not show any significant effects.

H5 predicts a negative relationship between *relatedness of acquirer knowledge* and innovation performance. The signs of

Table 5: GEE Poisson regression based on Ahuja and Katila (2001) model

Year	Coef.	P-value	S.E.
No. non-tech. acquisitions t-1	−0.0202	0.155	0.041
No. non-tech. acquisitions t-2	−0.1356 **	0.015	0.072
No. non-tech. acquisitions t-3	−0.0822 *	0.06	0.07
No. non-tech. acquisitions t-4	0.0258	0.145	0.047
Abs. size of acquired knowledge t-1	0.0004 **	0.011	0.0
Abs. size of acquired knowledge t-2	0.0001	0.126	0.0
Abs. size of acquired knowledge t-3	0.0003 *	0.056	0.0
Abs. size of acquired knowledge t-4	−0.0001	0.139	0.0
Rel. size of acquired knowledge t-1	−1.7254 ***	0.001	0.557
Rel. size of acquired knowledge t-2	−1.0617 **	0.021	0.611
Rel. size of acquired knowledge t-3	−0.8687 **	0.042	0.626
Rel. size of acquired knowledge t-4	0.0202	0.24	0.42
Relatedness of target knowledge t-1	0.1363 ***	0.005	0.058
Relatedness of target knowledge t-2	0.1707 ***	0.003	0.069
Relatedness of target knowledge t-3	0.276 ***	0.002	0.102
Relatedness of target knowledge t-4	0.1195 **	0.037	0.083
Relatedness of target knowledge^2 t-1	−0.0174 *	0.062	0.015
Relatedness of target knowledge^2 t-2	−0.0152	0.102	0.018
Relatedness of target knowledge^2 t-3	−0.124 ***	0.0	0.039
Relatedness of target knowledge^2 t-4	−0.0352	0.103	0.043
ISO Country AU	−0.9575 ***	0.0	0.113
ISO Country BE	−0.4211 ***	0.002	0.149
ISO Country CA	−2.3768 ***	0.0	0.256
ISO Country CH	−0.6243 ***	0.002	0.214
ISO Country DE	−0.0927	0.143	0.087
ISO Country FR	−0.5504 ***	0.005	0.211
ISO Country DK	−0.3913 ***	0.004	0.144
ISO Country GB	−0.1794	0.188	0.202
ISO Country JP	−0.1177	0.326	0.26
ISO Country NL	−1.7438 ***	0.0	0.191
Biotechnology	−0.2066 **	0.04	0.117
Health Care Equipment	0.0179	0.435	0.109
Life Sciences Tools & Services	−0.6666 ***	0.001	0.202
Year 2009	1.0468 ***	0.0	0.131
Year 2010	1.1158 ***	0.0	0.143
Year 2011	0.9992 ***	0.0	0.128
Year 2012	0.8965 ***	0.0	0.115
Year 2013	0.8549 ***	0.0	0.107
Year 2014	0.6898 ***	0.0	0.1
Year 2015	0.4835 ***	0.0	0.088
R&D expenses t-1	0.00003 **	0.024	0.0
Log. revenue t-1	−0.0133	0.46	0.133
Foreign acquisitions t-1	0.0384	0.156	0.038
Foreign acquisitions t-2	0.0149	0.345	0.037
Foreign acquisitions t-3	−0.0363	0.142	0.034
Foreign acquisitions t-4	−0.0178	0.195	0.021
Pre-sample patents	0.0007 ***	0.0	0.0
Intercept	3.924 ***	0.0	0.571

P-values originate from two-sided t-test for control variables, from one-sided t-test for independent variables.
Note: *p<0.1; **p<0.05; ***p<0.01

the coefficients are all negative. However, all coefficients are non-significant. Therefore, the *relatedness of acquirer knowledge* has no significant effect on innovation performance.

Surprisingly, *acquisition experience* has a negative, but non-significant coefficient. Further analysis with a squared term indicates a curvilinear relationship, as the squared term is positive (u-shaped). Previous research, specifically on acquisition experience, has indeed found an u-shaped relationship (see Barkema and Schijven (2008); Haleblian and Finkelstein (1999)). The coefficient is non-significant. Using the acquisition experience as a proxy for absorptive capacity might not be the best approach, as the acquisition management process also determines the acquisition outcome (Pidun, 2019). Furthermore, the variable has been calculated using the sum of all acquisitions from the beginning of the sample until the focal acquisition following Haleblian and Finkelstein (1999) and Hayward (2002). Research has operationalized this variable in different ways. Other approaches could lead to different results, e.g., using a moving value of acquisition experience (Sears & Hoetker, 2014) as knowledge also faces depreciation (Sampson, 2005).

The effect of *R&D expenses* is positive and significant.

Table 6: GEE Poisson regression based on author's model

	Coef.	P-value	S.E.
No. non-tech. acquisitions t-1	−0.0224	0.132	0.036
No. non-tech. acquisitions t-2	−0.115 **	0.018	0.064
No. non-tech. acquisitions t-3	−0.0547 *	0.066	0.049
No. non-tech. acquisitions t-4	0.0418 *	0.084	0.043
Abs. size of acquired knowledge t-1	0.0003 **	0.025	0.0
Abs. size of acquired knowledge t-2	0.0001	0.152	0.0
Abs. size of acquired knowledge t-3	0.0004 **	0.017	0.0
Abs. size of acquired knowledge t-4	−0.0001	0.122	0.0
Rel. size of acquired knowledge t-1	−2.0113 ***	0.003	0.781
Rel. size of acquired knowledge t-2	−0.8634 *	0.052	0.686
Rel. size of acquired knowledge t-3	−0.4662	0.118	0.646
Rel. size of acquired knowledge t-4	0.2436	0.146	0.444
Relatedness of target knowledge t-1	0.1169 **	0.028	0.074
Relatedness of target knowledge t-2	0.2003 ***	0.001	0.069
Relatedness of target knowledge t-3	0.331 ***	0.0	0.097
Relatedness of target knowledge t-4	0.1449 **	0.019	0.081
Relatedness of target knowledge^2 t-1	−0.0172 *	0.065	0.015
Relatedness of target knowledge^2 t-2	−0.0195 *	0.078	0.019
Relatedness of target knowledge^2 t-3	−0.1191 ***	0.0	0.034
Relatedness of target knowledge^2 t-4	−0.0315 *	0.092	0.035
Relatedness of acquirer knowledge t-1	0.8929 *	0.085	0.933
Relatedness of acquirer knowledge t-2	−0.28	0.185	0.838
Relatedness of acquirer knowledge t-3	−1.2317 **	0.029	0.777
Relatedness of acquirer knowledge t-4	−0.3645	0.151	0.706
Year 2009	0.9912 ***	0.0	0.141
Year 2010	1.0641 ***	0.0	0.149
Year 2011	0.9529 ***	0.0	0.14
Year 2012	0.8571 ***	0.0	0.124
Year 2013	0.8205 ***	0.0	0.114
Year 2014	0.6689 ***	0.0	0.109
Year 2015	0.4819 ***	0.0	0.094
Biotechnology	−0.1936 *	0.05	0.117
Health Care Equipment	0.0324	0.391	0.116
Life Sciences Tools & Services	−0.6558 ***	0.001	0.207
ISO Country AU	−0.975 ***	0.0	0.113
ISO Country BE	−0.4071 ***	0.003	0.147
ISO Country CA	−2.2891 ***	0.0	0.283
ISO Country CH	−0.5651 ***	0.008	0.233
ISO Country DE	−0.1243 *	0.081	0.088
ISO Country FR	−0.3913 *	0.087	0.287
ISO Country DK	−0.3733 ***	0.005	0.142
ISO Country GB	−0.1478	0.22	0.191
ISO Country JP	−0.1294	0.313	0.266
ISO Country NL	−1.8004 ***	0.0	0.218
R&D expenses t-1	0.00003 **	0.042	0.0
Log. revenue t-1	0.0466	0.367	0.137
Foreign acquisitions t-1	0.032	0.179	0.035
Foreign acquisitions t-2	0.0151	0.347	0.038
Foreign acquisitions t-3	−0.0332	0.155	0.033
Foreign acquisitions t-4	−0.0146	0.253	0.022
Pre-sample patents	0.0007 ***	0.0	0.0
Acquisition experience t-1	−0.0112	0.183	0.012
Intercept	3.7579 ***	0.0	0.578

P-values originate from two-sided t-test for control variables, from one-sided t-test for independent variables.
Note: *p<0.1; **p<0.05; ***p<0.01

Table 7: Summed GEE regression results based on Ahuja and Katila (2001)

	Coef.	P-value	S.E.
No. non-tech. acquisitions	−0.2122	0.104	0.0284
Abs. size of acquired knowledge	0.0007	0.136	0.000
Rel. size of acquired knowledge	−3.636**	0.021	3.198
Relatedness of target knowledge	0.702***	0.001	0.0541
Relatedness of target knowledge^2	−0.191***	0.008	0.0064
Foreign acquisitions	−0.0008	0.995	0.032

P-values originate from two-sided t-test for control variables, from one-sided t-test for independent variables.
Note: *p<0.1; **p<0.05; ***p<0.01

Due to the scale of R&D expenses the coefficient is very small (0.00003742). *Log. revenue* has a negative, but non-significant effect.

A comparison of results of Ahuja and Katila (2001), Cloodt et al. (2006) and this study can be found in table 9.

To compare both models, QIC is an appropriate metric. The smaller the QIC value, the better the model fit. The QIC value of the author's model is lower (48,979.30) compared to the model after Ahuja and Katila (2001) (49,066.74).

The p-values of the models have to be regarded with caution. As the model violates the Poisson assumptions, the standard errors can be underestimated. Even though the pre-sample approach tries to correct for that, hypothesis testing can be invalidated to some extent. For this reason a Negative binomial regression was also estimated, as it yields better estimates of the standard errors and therefore more robust results of hypothesis testing. The results of the Negative binomial models can be found in the appendix in tables A2 and A3. The results and p-values are similar, indicating that the findings are robust.

3.5.1. Sensitivity analysis

To ensure validity of results, a sensitivity analysis was performed using different random sub-samples of the sample (Ahuja & Katila, 2001). The results are quite similar.

Interestingly, the number of patents granted per year per firm shows a negative trend, while the R&D expenditures continue to rise. The trend lines can be seen in figure 2. Besides, the variance of patents granted is also decreasing over time, as shown by the black dots in figure 2. This decrease is noticeable. This finding is consistent with previous research by Griliches (1998), which found that patent output gradually decreases with higher R&D expenditures.

This decrease in patenting might be also due to the increased product complexity in the medical device area or the increased challenge in finding molecules that prove to be effective (Danzon et al., 2005; Lin & Jang, 2010). However, the speed of innovation has increased in the last years, which should be mirrored in the number of patents. To further ensure the reliability of the results, the patent data should be checked again.

Also surprising is the lower correlation between R&D expenses and patents. Ahuja and Katila (2001) report a correlation of 0.89 between R&D expenses and pre-sample patents, Cloodt et al. (2006) report 0.784, while this study finds 0.57. The correlation with the dependent variable is also lower: 0.49 versus 0.89. The reduced strength of the correlation coefficient might be the result of a lower R&D expenditures to patent efficiency (Danzon et al., 2005; Ornaghi, 2009). However, the difference could be the result of a different sample and a different variance in the sample. As this idea is only based on a correlation, further analysis is required to confirm or reject it.

As this study does not include patent citations, the variable *relatedness of target knowledge* and *relatedness of acquirer knowledge* should be less interpretable. Previous research has included patent citations, e.g., Ahuja and Katila (2001); Cloodt et al. (2006); Makri et al. (2010); Sears and Hoetker (2014). Ahuja and Katila (2001) ran a sensitivity analysis, comparing patents cited by the firm and patents owned by the firm. Results indicate that own patents are reliable for measuring the *absolute* and *relative size of acquired knowledge*. However, Ahuja and Katila (2001) report that using only owned patents for measuring the *relatedness of target knowledge* does not capture the relationship well. The findings of this study are different, as the variable *relatedness of target knowledge* is significant. This difference might be a result of a different construction of the variable. Ahuja and Katila (2001) use the list of "patent numbers" to calculate the *relatedness of target knowledge*. It is unclear to the author, whether Ahuja and Katila (2001) used all digits in the "patent numbers" or only a subsection of digits. Furthermore, their analysis might be based on a different patent classification system. This work used the first four digits of the IPC system, similar to previous research on patent classification and similarity (Kogler et al., 2013; Leydesdorff et al., 2017).

4. Discussion

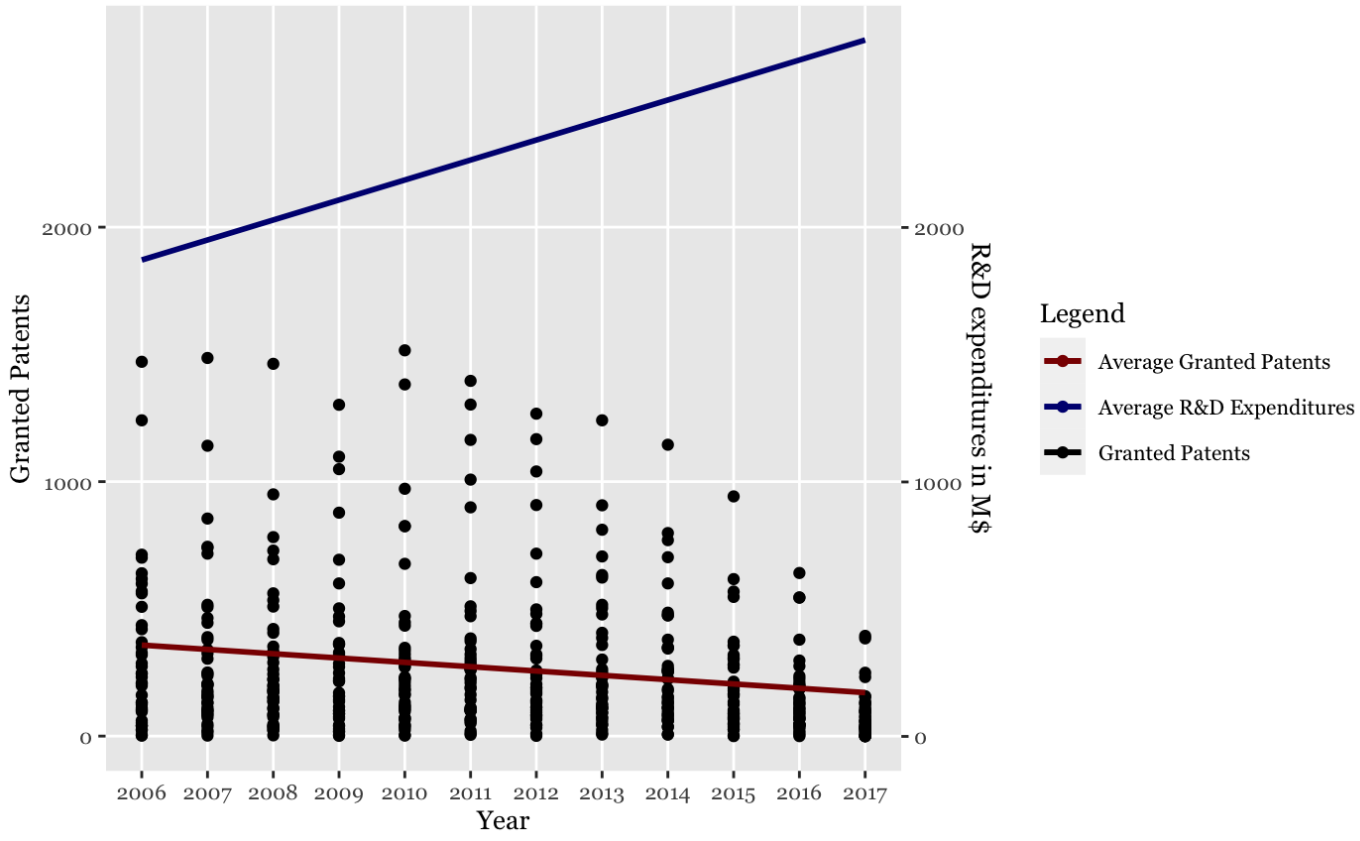
This work extends the research stream in strategic management research of M&A in several ways. First, this study focuses on an understudied post-acquisition outcome: the innovation performance of acquiring firms in high-technology (Zollo & Meier, 2008). Second, it can be seen as a replication study of Ahuja and Katila (2001) and Cloodt et al. (2006) using newer data on another high-tech industry: the healthcare industry. Third, it further extends the econometric model of Ahuja and Katila (2001) with the additional variables relatedness of acquirer knowledge, as a measurement of techno-

Table 8: Summed GEE regression results of the author's model

	Coef.	P-value	S.E.
No. non-tech. acquisitions	−0.150	0.126	0.017
Abs. size of acquired knowledge	0.0007*	0.095	0.000
Rel. size of acquired knowledge	−3.097*	0.083	5.012
Relatedness of target knowledge	0.793***	0.001	0.063
Relatedness of target knowledge^2	−0.187***	0.003	0.005
Relatedness of acquirer knowledge	−0.983	0.351	6.598
Foreign acquisitions	−0.0007	0.497	0.009

P-values originate from two-sided t-test for control variables, from one-sided t-test for independent variables.
Note: *p<0.1; **p<0.05; ***p<0.01

Figure 2: R&D expenditures vs granted patents (own figure)



logical relatedness in an acquisition and the acquisition experience of the acquirer, as a measure of absorptive capacity.

An overall description of the different sub-industries of the healthcare industry was given. Important theories of strategic management were explained, such as the knowledge-based view, absorptive capacity, and combinative capability. Previous empirical studies were described and set into context. The hypotheses draw on ideas from these theories and previous empirical research. Finally, the empirical analysis used a GEE approach for a Poisson regression and a Negative binomial regression. The findings of this work are very similar to those of Ahuja and Katila (2001) and Cloodt et al. (2006): (1) Non-technological acquisitions affect the post-acquisition innovation performance in a negative way. In the case of technological acquisitions, the following dimensions have been examined: (2) absolute size of acquired knowledge, (3) relative size of acquired knowledge, and (4) relatedness of target knowledge. The (2) absolute size of acquired knowledge has a positive effect on post-acquisition performance. Large deals provide a lot of knowledge and therefore a lot of opportunity to generate new knowledge. Firms seem to be able to successfully integrate this knowledge and benefit from it. However, a (3) large relative size of acquired knowledge has a negative effect. Integrating large knowledge bases is a challenging undertaking and therefore has a negative effect on post-acquisition innovation performance.

An inverted u-shape models the relationship between

Table 9: Comparison of Ahuja and Katila (2001), Cloodt et al. (2006), and this study

Study	Hypothesis	Results
Ahuja and Katila (2001) - Technological acquisitions and the innovation performance of acquiring firms: A longitudinal study	**H1**: Non-technological acquisitions will affect the post-acquisition innovation output of acquiring firms either negatively or non-significantly.	(ns)
	H2: The greater the absolute size of the acquired knowledge base, the greater the subsequent innovation output of the acquiring firm.	(+)
	H3: The greater the relative size of the acquired knowledge base, the less the subsequent innovation output of the acquiring firm.	(+)
	H4: The relatedness of the acquired knowledge base will be curvilinearly (inverted U) related to the subsequent innovation output of the acquiring firm.	(+)
Cloodt et al. (2006) - Mergers and acquisitions: Their effect on the innovative performance of companies in high-tech industries	**H1**: Non-technological acquisitions will affect the post-acquisition innovation output of acquiring firms either negatively or non-significantly.	(+)
	H2: The greater the absolute size of the acquired knowledge base, the greater the subsequent innovation output of the acquiring firm.	(-)
	H3: The greater the relative size of the acquired knowledge base, the less the subsequent innovation output of the acquiring firm.	(+)
	H4: The relatedness of the acquired knowledge base will be curvilinearly (inverted U) related to the subsequent innovation output of the acquiring firm.	(+)
Study	Hypothesis	Results
This study	**H1**: The post-acquisition innovation performance will be affected either negatively or non-significantly by non-technological M&A.	(+)
	H2: A larger absolute size of the acquired knowledge base increases the post-acquisition innovation performance.	(+)
	H3: A large relative size of the acquired knowledge base will have a negative impact on the post-acquisition innovation performance.	(+)
	H4: The relatedness of the acquired knowledge base will be curvilinearly (inverted U-shape) related to the post-acquisition innovation performance of the acquiring firm.	(+)
	H5: The larger the acquirer overlap, the smaller the post-acquisition innovation performance.	(ns)
	H6: Previous acquisition experience has a positive effect on post-acquisition innovation performance.	(-)

the (4) relatedness of knowledge of the target. Too similar knowledge does not spur innovation, nor does too different knowledge. Similar to the previous findings (Cloodt et al., 2006), this study suggests that there is a "sweet spot" of knowledge similarity. The overlapped knowledge serves as a baseline of absorptive capacity for handling and integrating the complexity of the target firm's high-quality knowledge.

The additional variables introduced by the author show a negative effect of (5) relatedness of acquirer knowledge and acquisition experience on post-acquisition innovation perfor-

mance. The effect of knowledge of relatedness of acquirer is negative, but non-significant. The higher the share of acquirer knowledge, that is also known by the target, the less the acquirer can learn from the target.

The effect of (6) acquisition experience is negative, contrary to expectation. Previous research has found different effects, either positive (Barkema & Schijven, 2008) or u-shaped (Barkema & Schijven, 2008; Haleblian & Finkelstein, 1999). Additional analysis with a squared term indicates a u-shaped relationship. Further research might reveal moderating effects.

Patents were used as a metric to analyze the effect of M&A on innovation performance. The knowledge of a firm can be approximated by its patent base (Ahuja & Katila, 2001). Innovations are the result of knowledge (Grant, 1996). Patents provide detailed insight into the knowledge base of a firm and they permit quantification of knowledge and thus innovation (Ahuja & Katila, 2001; Hagedoorn & Cloodt, 2003).

There are some differences in motives for M&A in the healthcare industry. As the revenue of pharmaceutical firms depends on the introduction of new drugs, M&A is often used to increase the number of potential drugs in the pipeline and to close a potential earnings gap (Danzon et al., 2007, 2005; Schweizer & zu Knyphausen-Aufseß, 2008). The healthcare equipment manufacturers are not as depended as pharmaceutical firms on M&A, as their innovations tend to be more incremental (Lin & Jang, 2010).

4.1. Theoretical Implications

M&A research has mainly been concerned with financial performance. Only a small research stream has focused on M&A and innovation performance (Zollo & Meier, 2008). This work is clearly an extension of this research stream. The most important finding is that knowledge, especially knowledge relatedness, plays an important role in technological M&A. As competition in the healthcare industry is driven by innovation and technology many firms rely on M&A as a source of innovation (Lin & Jang, 2010; Schweizer, 2005b). M&A is therefore a popular strategy and can be frequently observed in the industry. Firms rely on external knowledge to combine it with internal knowledge to compete in the respective market (Danzon et al., 2007; Grant, 1996; Kogut & Zander, 1992; Schweizer & zu Knyphausen-Aufseß, 2008).

This research shows that there is a clear difference between technological M&A and non-technological M&A, whereas non-technological M&A seems to hinder the innovative capabilities of acquiring firms. Technological M&A can have a positive impact on innovation performance. Firms can use M&A to innovate in two ways: they can leverage what the target knows or rely on the target as an independent source of innovation (Puranam & Srikanth, 2007; Schweizer, 2005a), depending on the degree of target integration (Bresman et al., 1999; Schweizer, 2005a). Regardless of the degree of integration, the ability to gain value from a transaction depends on the absorptive capacity of firms: the capability to successfully identify, assimilate and commercialize knowledge made by external entities (Cohen & Levinthal, 1990). Besides the absorptive capacity, the combinative capability plays an important role, which enables firms to combine explored knowledge with unexplored knowledge and exploit it (Kogut & Zander, 1992). Following both concepts, the knowledge of both firms should be related. The empirical results show that there is an inverted u-shape of target and acquirer knowledge relatedness. The higher the amount of "new" knowledge for the acquirer the better, until a certain tipping point, where the innovation performance starts to suffer. This indicates that both the absorptive capacity and the combinative capability of the acquirer have a limit.

Both firms should have an overlap in R&D projects (Cassiman et al., 2005) and in their patent portfolio. Based on the relationship of relatedness of knowledge and the absorptive capacity of a firm, an acquiring firm with a diversified knowledge base should be able to benefit more from acquisitions, as the firm is likely to have many similarities with the target's knowledge base. In order to improve the innovation performance by M&A, firms should aim for having a diverse R&D and patent portfolio and target firms that have complementary knowledge. However, this does not necessarily lead to an improved financial performance.

Firms that acquire regularly should benefit from an increased absorptive capacity, such as the enhanced identification and absorption of important target knowledge (Cohen & Levinthal, 1990; Desyllas & Hughes, 2010; Grant, 1996). The result of this study shows a negative effect of acquisition experience on innovation performance. To proxy the absorptive capacity with acquisition experience might not be the best approach, as the acquisition management process also determines the outcome (Bresman et al., 1999; Pidun, 2019). In addition to that, the absorptive capacity is a complex and multilayered concept (Lane et al., 2006), which cannot be easily proxied by one variable. Therefore, more research on the relationship of innovation performance, knowledge transfer, absorptive capacity, and other theories of organizational learning should be considered.

Similar to previous studies, this work has analyzed a different high-tech industry: the healthcare industry. Therefore, a generalization of these findings to low- or medium tech industries is questionable (Cloodt et al., 2006).

4.2. Managerial Implications

M&A can be a valuable tool to improve the innovation performance, as long as the target provides a moderate and right amount of technological input. A dedicated M&A strategy should help identify specific M&A themes and dealscreening criteria that support corporate innovation (Ascher et al., 2020; Pidun, 2019). Harrison, Hitt, Hoskisson, and Ireland (2001) argue that an active acquisition strategy does not necessarily reduce managerial commitment to innovation. Paying close attention when search for a complementary target with an emphasis on innovation in the integration period can enhance success (Desyllas & Hughes, 2010; Paruchuri & Eisenman, 2012). An improved due diligence process, that also considers the knowledge complementari-

ties and similarities of the firms, could have a positive impact on the innovation performance as well (Pidun, 2019).

Among pharmaceutical companies, internal R&D has often become a secondary source of drug innovation as the negative correlation between R&D cost and productivity is simply unbearable for many firms (Danzon et al., 2007; Ornaghi, 2009). For pharmaceuticals, acquiring smaller innovative firms might be a solid strategy to avoid the risk of in-house drug development and to improve the product pipeline. Of course, M&A is not the only way for firm collaboration. Strategic Alliances, Joint Ventures, or licensing agreements can be valuable alternatives as well, e.g., the case of Biontech and Pfizer (Danzon et al., 2007; Deeds & Hill, 1996; Hagedoorn & Duysters, 2002; Schweizer & zu Knyphausen-Aufseß, 2008; Thomson Reuters, 2020).

Previous research has also shown that pharmaceutical firms acquire smaller firms, when their product pipeline is empty (Danzon et al., 2007; Schweizer & zu Knyphausen-Aufseß, 2008). This study shows a positive effect of M&A on patenting, indicating that firms can revitalize their patenting activities and therefore their product pipeline and avoid inertia (Cefis & Marsili, 2015; Vermeulen & Barkema, 2001). Yet, managers tend to be biased towards more incremental innovations, as they provide short-term returns and are less risky (Hoskisson, Hitt, Johnson, & Grossman, 2002). This bias needs to be considered in an acquisition process.

4.3. Limitations

In order to not overstretch the framework of this work, this research left out multiple important deal characteristics: e.g., the motivation of acquirers, the type of deal (e.g., vertical, horizontal), the type of financing and the degree of post-acquisition integration. All of these factors can affect the post-acquisition innovation performance in various ways. The deals in the sample have a large range in terms of the monetary volume. Previous research indicates that large deals are typically not motivated only by technology and that often large deals do not lead to an improvement in innovation performance (James, Georghiou, & Metcalfe, 1998; Schweizer & zu Knyphausen-Aufseß, 2008). In this research, there is no consideration of the motivation behind the deal.

Compared to previous research (e.g., Ahuja and Katila (2001) and Cloodt et al. (2006)), the statistical power of this work is lower, with only 41 firms observed over a period of 10 years. In addition, the author did not specifically analyze the different sub-industries. As the GEE method gives population average estimators, different characteristics in the sub-industries were "canceled out".

In this study, the time frame to analyze effects of M&A was four years. This might not be enough, as the average time to develop a new drug is about 12 years and the development of medical equipment usually takes about 3-7 years (Van Norman, 2016). Research on technological deals should also observe long-term effects. These long-term effect are usually underestimated in empirical M&A research (Chakrabarti, Hauschildt, & Süverkrüp, 1994), but can play a significant role in technological M&A. Especially synergistic effects in R&D might take time to develop. Improved innovation performance eventually leads to a better financial performance in the long run. However, there is research that suggests that — depending on the circumstances — a successful integration of the target should be a comparatively fast process (Homburg & Bucerius, 2006). Thus, the effect of a short and successful period of integration should be reflected in this research.

As knowledge and innovation are very hard to quantify, this research tries to analyze it by relying on patents to measure it. Patents as a proxy are limited in several ways. First of all, "not all patents represent innovation, nor are all innovations patented" (Nagaoka et al., 2010, p. 1084). Patents are an extremely noisy indicator of innovation, since they are subject to the variance in the significance, quality, and value of individual patents (B. H. Hall et al., 2005). Using granted patents as a dependent variable, has the advantage of filtering low quality or trivial patents (Hagedoorn & Cloodt, 2003). As patents are simply used as a count variable, knowledge quality is not considered in this research (Ahuja & Katila, 2001; Han, Jo, & Kang, 2018). Knowledge relatedness is calculated by using patent classification codes. These codes were developed for other purposes and do not fully reflect knowledge overlap. Patents with the same code can still have very different underlying knowledge (Makri et al., 2010).

Additionally, patents are often used strategically. For instance, some firms patent new inventions simply to block other firms' patents or to deter entry into certain products, instead of capitalizing them. New patents often require significant novelty to existing ones and can therefore be used to block competitors (Grimpe & Hussinger, 2014a). Additionally, as there is a lot of M&A activity in the sample and many firms have multiple subsidiaries with different names, not every patent is considered (Valentini, 2012). Despite their limitation patents are used as a proxy in most empirical studies in this area (Ahuja & Katila, 2001; Cloodt et al., 2006; Desyllas & Hughes, 2010; Makri et al., 2010; Sears & Hoetker, 2014). Of course, there are also other ways to protect intellectual property in the healthcare industry, such as regulatory exclusivity and manufacturing challenges (Lakdawalla, 2018).

The variables *relatedness of target knowledge* and *relatedness of acquirer knowledge* might be constructed in a different way than in previous research of Ahuja and Katila (2001) and Cloodt et al. (2006). Hence, a direct comparison should be done carefully.

4.4. Further research

Further research could investigate if and how firms use M&A as a substitute for internal innovation programs and R&D efforts or how external and internal innovation and R&D are used. Similarly, research should investigate if and how internal and external R&D efforts are combined after M&A to enhance innovation. Especially with the involvement of highly specific knowledge in M&A, the integration process becomes a crucial part, as knowledge is not easily trans-

ferred and capitalized (Capron & Pistre, 2002; Schweizer & zu Knyphausen-Aufseß, 2008). Furthermore, different motives behind the deal require different integration strategies (Bauer, Matzler, & Wolf, 2016; Pidun, 2019).

Future work could also analyze deal outcomes and the interplay of different deal characteristics (financing, degree of integration, etc.) and the interplay with acquirer characteristics (R&D expenditures, acquisition experience, knowledge/patent portfolio). Research into the acquisition management process such as post-acquisition integration and restructuring could provide insights on additional factors that affect post-acquisition innovation performance. Additional research could also use other metrics to measure the innovation performance, as most contemporary studies have relied on patents (Ahuja & Katila, 2001; Cloodt et al., 2006; Makri et al., 2010; Sears & Hoetker, 2014). Metrics such as the rate of product introduction or productivity of inventors and R&D personnel could help understand the effect of M&A on the innovation performance better. Also, combining innovation performance with other metrics, such as financial metrics, to gain a more holistic view of M&A and M&A performance could provide new insights (King et al., 2004; Zollo & Meier, 2008).

The combinative capability and the absorptive capacity determine how firms are able to interact with new knowledge. Firms that acquire on a regular basis should have an improved absorptive capacity compared to their peers. For this reason, the variable *acquisition experience* was introduced. The results of the effects of this variable in the study are inconclusive. Previous research has found different effects of *acquisition experience*. For example, Haleblian and Finkelstein (1999) found different effects depending on the similarity of the current acquisition compared to previous acquisitions. The relation of absorptive capacity, acquisition experience and innovation performance seems to be a fruitful research direction.

Previous research has overlooked the power of preemptive patenting. Acquiring the right target firms to block competitors from entering into certain technologies might be a common strategy, especially in healthcare. Apart from this, firms might also protect themselves by acquiring targets that could block the other (potential) acquirers from capitalizing certain intellectual property (Grimpe & Hussinger, 2014b). There has been very few research on M&A and preemptive patenting in strategic management research.

Higgins and Rodriguez (2006) found that many pharmaceutical acquirers already had strategic alliances with their targets and thus avoid the "winner's curse" of paying too much for the target by having insider knowledge. Research into knowledge complementarity, insider information, target valuation, and post-acquisition performance could also be fruitful.

References

Adams, R., Bessant, J., & Phelps, R. (2006). Innovation management measurement: A review. *International Journal of Management Reviews*, *8*(1), 21–47.

Agrawal, A., Jaffe, J. F., & Mandelker, G. N. (1992). The post-merger performance of acquiring firms: a re-examination of an anomaly. *Journal of Finance*, *47*(4), 1605–1621.

Agresti, A. (2003). *Categorical data analysis* (Vol. 482). John Wiley & Sons, Hoboken, New Jersey.

Ahuja, G., & Katila, R. (2001). Technological acquisitions and the innovation performance of acquiring firms: A longitudinal study. *Strategic Management Journal*, *22*(3), 197–220.

Arrow, K. J. (1962). *Economic welfare and the allocation of resources for invention*. Princeton, New Jersey: Princeton University Press, Princeton, New Jersey.

Ascher, J., Bansal, R., Dhankhar, A., & Kim, E. (2020). *A new prescription for M&A in pharma.* https://www.mckinsey.com/business-functions/m-and-a/our-insights/a-new-prescription-for-m-and-a-in-pharma. (Date accessed 15.06.2021)

Barden, J. Q. (2012). The influences of being acquired on subsidiary innovation adoption. *Strategic Management Journal*, *33*(11), 1269–1285.

Barkema, H. G., & Schijven, M. (2008). How do firms learn to make acquisitions? A review of past research and an agenda for the future. *Journal of Management*, *34*(3), 594–634.

Barney, J. (1988). Returns to bidding firms in mergers and acquisitions: Reconsidering the relatedness hypothesis. *Strategic Management Journal*, *9*(S1), 71–78.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of management*, *17*(1), 99–120.

Bauer, F., Matzler, K., & Wolf, S. (2016). M&A and innovation: The role of integration and cultural differences—A central European targets perspective. *International Business Review*, *25*(1), 76–86.

Bena, J., & Li, K. (2014). Corporate innovations and mergers and acquisitions. *The Journal of Finance*, *69*(5), 1923–1960.

Blind, K., Cremers, K., & Mueller, E. (2009). The influence of strategic patenting on companies' patent portfolios. *Research Policy*, *38*(2), 428–436.

Blundell, R., Griffith, R., & Reenen, J. V. (1995). Dynamic count data models of technological innovation. *The Economic Journal*, *105*(429), 333–344.

Bresman, H., Birkinshaw, J., & Nobel, R. (1999). Knowledge transfer in international acquisitions. *Journal of International Business Studies*, *30*(3), 439–462.

Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (Vol. 53). Cambridge University Press, Cambridge.

Capron, L., Dussauge, P., & Mitchell, W. (1998). Resource redeployment following horizontal acquisitions in europe and north america, 1988–1992. *Strategic Management Journal*, *19*(7), 631–661.

Capron, L., & Pistre, N. (2002). When do acquirers earn abnormal returns? *Strategic Management Journal*, *23*(9), 781–794.

Cartwright, S., & Schoenberg, R. (2006). Thirty years of mergers and acquisitions research: Recent advances and future opportunities. *British Journal of Management*, *17*(S1), S1–S5.

Cassiman, B., Colombo, M. G., Garrone, P., & Veugelers, R. (2005). The impact of M&A on the R&D process: An empirical analysis of the role of technological-and market-relatedness. *Research Policy*, *34*(2), 195–220.

Cassiman, B., & Veugelers, R. (2006). In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition. *Management Science*, *52*(1), 68–82.

Cefis, E., & Marsili, O. (2015). Crossing the innovation threshold through mergers and acquisitions. *Research Policy*, *44*(3), 698–710.

Chakrabarti, A., Hauschildt, J., & Süverkrüp, C. (1994). Does it pay to acquire technological firms? *R&D Management*, *24*(1), 047–056.

Chee, F. Y. (2018). Takeda to win eu approval for $62 billion shire deal. *Reuters*. https://www.reuters.com/article/us-shire-m-a-takeda-pharma-eu-exclusive-idUSKCN1NE21B. (Date accessed 22.05.2021)

Chen, I.-J., Hsu, P.-H., Officer, M. S., & Wang, Y. (2020). The oscar goes to…: High-tech firms' acquisitions in response to rivals' technology breakthroughs. *Research Policy*, *49*(7), 104078.

Cloodt, M., Hagedoorn, J., & Van Kranenburg, H. (2006). Mergers and acquisitions: Their effect on the innovative performance of companies in high-tech industries. *Research Policy*, *35*(5), 642–654.

Cockburn, I., & Long, G. (2015). The importance of patents to innovation: updated cross-industry comparisons with biopharmaceuticals. *Expert Opinion on Therapeutic Patents*, *25*(7), 739-742.

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and learning: the two faces of R & D. *The Economic Journal*, *99*(397), 569–596.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 128–152.

Colombo, M. G., & Rabbiosi, L. (2014). Technological similarity, post-acquisition R&D reorganization, and innovation performance in horizontal acquisitions. *Research Policy*, *43*(6), 1039–1054.

Danzon, P. M., Epstein, A., & Nicholson, S. (2007). Mergers and acquisitions in the pharmaceutical and biotech industries. *Managerial and Decision Economics*, *28*(4-5), 307–328.

Danzon, P. M., Nicholson, S., & Pereira, N. S. (2005). Productivity in pharmaceutical–biotechnology R&D: the role of experience and alliances. *Journal of Health Economics*, *24*(2), 317–339.

Datta, D. K., & Grant, J. H. (1990). Relationships between type of acquisition, the autonomy given to the acquired firm, and acquisition success: An empirical analysis. *Journal of Management*, *16*(1), 29–44.

Datta, D. K., Pinches, G. E., & Narayanan, V. K. (1992). Factors influencing wealth creation from mergers and acquisitions: A meta-analysis. *Strategic Management Journal*, *13*(1), 67–84.

Deeds, D. L., & Hill, C. W. (1996). Strategic alliances and the rate of new product development: An empirical study of entrepreneurial biotechnology firms. *Journal of business venturing*, *11*(1), 41–55.

Desyllas, P., & Hughes, A. (2010). Do high technology acquirers become more innovative? *Research Policy*, *39*(8), 1105–1121.

DiMasi, J. A., Hansen, R. W., & Grabowski, H. G. (2003). The price of innovation: new estimates of drug development costs. *Journal of Health Economics*, *22*(2), 151–185.

Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (Vol. 326). John Wiley & Sons, Hoboken, New Jersey.

Eschen, E., & Bresser, R. K. (2005). Closing resource gaps: toward a resource-based theory of advantageous mergers and acquisitions. *European Management Review*, *2*(3), 167–178.

European Patent Office. (2021). Healthcare innovation main driver of European patent applications in 2020. https://www.epo.org/news-events/news/2021/20210316.html. (Date accessed 18.08.2021)

Freeman, C., & Soete, L. (1997). *The economics of industrial innovation*. Routledge, London.

Galunic, D. C., & Rodan, S. (1998). Resource recombinations in the firm: Knowledge structures and the potential for schumpeterian innovation. *Strategic Management Journal*, *19*(12), 1193–1201.

Ghisletta, P., & Spini, D. (2004). An introduction to generalized estimating equations and an application to assess selectivity effects in a longitudinal study on very old individuals. *Journal of Educational and Behavioral Statistics*, *29*(4), 421–437.

Goedhart, M., Koller, T., & Wessels, D. (2017). The six types of successful acquisitions. https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/the-six-types-of-successful-acquisitions. (Date accessed 09.01.2021)

Granstrand, O., & Sjölander, S. (1990). The acquisition of technology and small firms by large firms. *Journal of Economic Behavior & Organization*, *13*(3), 367–386.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*(S2), 109–122.

Graves, S. B., & Langowitz, N. S. (1993). Innovative productivity and returns to scale in the pharmaceutical industry. *Strategic Management Journal*, *14*(8), 593–605.

Griliches, Z. (1998). Patent statistics as economic indicators: a survey. In *R&d and productivity: the econometric evidence* (pp. 287–343). Chicago: University of Chicago Press.

Grimpe, C., & Hussinger, K. (2014a). Pre-empted patents, infringed patents and firms' participation in markets for technology. *Research Policy*, *43*(3), 543–554.

Grimpe, C., & Hussinger, K. (2014b). Resource complementarity and value capture in firm acquisitions: The role of intellectual property rights. *Strategic Management Journal*, *35*(12), 1762–1780.

Hagedoorn, J., & Cloodt, M. (2003). Measuring innovative performance: is there an advantage in using multiple indicators? *Research Policy*, *32*(8), 1365–1379.

Hagedoorn, J., & Duysters, G. (2002). External sources of innovative capabilities: the preferences for strategic alliances or mergers and acquisitions. *Journal of Management Studies*, *39*(2), 167–188.

Haleblian, J., Devers, C. E., McNamara, G., Carpenter, M. A., & Davison, R. B. (2009). Taking stock of what we know about mergers and acquisitions: A review and research agenda. *Journal of Management*, *35*(3), 469–502.

Haleblian, J., & Finkelstein, S. (1999). The influence of organizational acquisition experience on acquisition performance: A behavioral learning perspective. *Administrative Science Quarterly*, *44*(1), 29–56.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics*, 16–38.

Hall, J., & Krauskopf, L. (2009). Pfizer to buy wyeth for $68 billion. *Reuters*. https://www.reuters.com/article/us-wyeth-pfizer-idUSTRE50M1AQ20090126. (Date accessed 22.05.2021)

Han, J., Jo, G. S., & Kang, J. (2018). Is high-quality knowledge always beneficial? Knowledge overlap and innovation performance in technological mergers and acquisitions. *Journal of Management & Organization*, *24*(2), 258–278.

Harrison, J. S., Hitt, M. A., Hoskisson, R. E., & Ireland, R. D. (2001). Resource complementarity in business combinations: Extending the logic to organizational alliances. *Journal of Management*, *27*(6), 679–690.

Haspeslagh, P. C., & Jemison, D. B. (1991). *Managing acquisitions: Creating value through corporate renewal* (Vol. 416). Free Press, New York.

Hayward, M. L. (2002). When do firms learn from their acquisition experience? Evidence from 1990 to 1995. *Strategic Management Journal*, *23*(1), 21–39.

Heimeriks, K. H., Schijven, M., & Gates, S. (2012). Manifestations of higher-order routines: The underlying mechanisms of deliberate learning in the context of postacquisition integration. *Academy of Management Journal*, *55*(3), 703–726.

Henderson, R., & Cockburn, I. (1996). Scale, scope, and spillovers: the determinants of research productivity in drug discovery. *RAND Journal of Economics*, 32–59.

Higgins, M. J., & Rodriguez, D. (2006). The outsourcing of R&D through acquisitions in the pharmaceutical industry. *Journal of financial economics*, *80*(2), 351–383.

Hitt, M. A., Hoskisson, R. E., & Ireland, R. D. (1990). Mergers and acquisitions and managerial commitment to innovation in m-form firms. *Strategic Management Journal*, 29–47.

Hitt, M. A., Hoskisson, R. E., Ireland, R. D., & Harrison, J. S. (1991). Effects of acquisitions on R&D inputs and outputs. *Academy of Management Journal*, *34*(3), 693–706.

Hitt, M. A., Hoskisson, R. E., Johnson, R. A., & Moesel, D. D. (1996). The market for corporate control and firm innovation. *Academy of Management Journal*, *39*(5), 1084–1119.

Hofstede, G. (1980). Culture and organizations. *International Studies of Management & Organization*, *10*(4), 15–41.

Hofstede, G. (2015). Cultural Dimension Data Matrix. https://geerthofstede.com/research-and-vsm/dimension-data-matrix/. (Date accessed 03.02.2021)

Hofstede, G., & Minkov, M. (2010). Long-versus short-term orientation: new perspectives. *Asia Pacific Business Review*, *16*(4), 493–504.

Homburg, C., & Bucerius, M. (2006). Is speed of integration really a success factor of mergers and acquisitions? An analysis of the role of internal and external relatedness. *Strategic Management Journal*, *27*(4), 347–367.

Hoskisson, R. E., Hitt, M. A., Johnson, R. A., & Grossman, W. (2002). Conflicting voices: The effects of institutional ownership heterogeneity and internal governance on corporate innovation strategies. *Academy of Management Journal*, *45*(4), 697–716.

James, A. D., Georghiou, L., & Metcalfe, J. S. (1998). Integrating technology into merger and acquisition decision making. *Technovation*, *18*(8-9), 563–591.

Jarrell, G. A., & Poulsen, A. B. (1989). The returns to acquiring firms in tender offers: Evidence from three decades. *Financial Management*, 12–19.

Jemison, D. B., & Sitkin, S. B. (1986). Corporate acquisitions: A process perspective. *Academy of Management Review*, *11*(1), 145–163.

Kapoor, R., & Lim, K. (2005). The impact of acquisitions on the innovation performance of inventors at semiconductor companies. *Academy of Management Journal*, *50*(5), 1133-1155.

Kaul, A. (2012). Technology and corporate scope: Firm and rival innovation as antecedents of corporate transactions. *Strategic Management Journal*, *33*(4), 347–367.

King, D. R., Dalton, D. R., Daily, C. M., & Covin, J. G. (2004). Meta-analyses of post-acquisition performance: Indications of unidentified moderators. *Strategic Management Journal*, *25*(2), 187–200.

Kogler, D. F., Rigby, D. L., & Tucker, I. (2013). Mapping knowledge space and technological relatedness in US cities. *European Planning Studies*, *21*(9), 1374–1391.

Kogut, B., & Singh, H. (1988). The effect of national culture on the choice of entry mode. *Journal of International Business Studies*, *19*(3), 411–432.

Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science*, *3*(3), 383–397.

Kogut, B., & Zander, U. (1993). Knowledge of the firm and the evolutionary theory of the multinational corporation. *Journal of International Business Studies*, *24*(4), 625–645.

Kogut, B., & Zander, U. (1996). What firms do? Coordination, identity, and learning. *Organization Science*, *7*(5), 502–518.

Lakdawalla, D. N. (2018). Economics of the pharmaceutical industry. *Journal of Economic Literature*, *56*(2), 397–449.

Lane, P. J., Koka, B. R., & Pathak, S. (2006). The reification of absorptive capacity: A critical review and rejuvenation of the construct. *Academy of Management Review*, *31*(4), 833–863.

Lane, P. J., & Lubatkin, M. (1998). Relative absorptive capacity and interorganizational learning. *Strategic Management Journal*, *19*(5), 461–477.

Larsson, R., & Finkelstein, S. (1999). Integrating strategic, organizational, and human resource perspectives on mergers and acquisitions: A case survey of synergy realization. *Organization Science*, *10*(1), 1–26.

Levenshtein, V. I., et al. (1966). Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet physics doklady* (Vol. 10, pp. 707–710).

Leydesdorff, L., Kogler, D. F., & Yan, B. (2017). Mapping patent classifications: portfolio and statistical analysis, and the comparison of strengths and weaknesses. *Scientometrics*, *112*(3), 1573–1591.

Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13–22.

Lin, C.-H., & Jang, S.-L. (2010). The impact of M&As on company innovation: evidence from the US medical device industry. *Scientometrics*, *84*(1), 119–131.

Lipsitz, S., & Fitzmaurice, G. (2008). Generalized estimation equations for longitudinal data analysis. *Longitudinal Data Analysis*, 43–78.

Lubatkin, M. (1983). Mergers and the performance of the acquiring firm. *Academy of Management Review*, *8*(2), 218–225.

Makri, M., Hitt, M. A., & Lane, P. J. (2010). Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions. *Strategic Management Journal*, *31*(6), 602–628.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*(1), 71–87.

MCSI Inc. (2021). The global industry classification standard. https://www.msci.com/gics#:~:text=GICS%20is%20a%20four%2Dtiered,to%20its%20principal%20business%20activity. (Date accessed 20.01.2021)

Minkov, M. (2007). *What makes us different and similar: A new interpretation of the world values survey and other cross-cultural data*. Klasika i Stil Publishing House, Sofia.

Nagaoka, S., Motohashi, K., & Goto, A. (2010). Patent statistics as an innovation indicator. In *Handbook of the economics of innovation* (Vol. 2, pp. 1083–1127). Amsterdam: Elsevier.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science*, *5*(1), 14–37.

Norton, E. C., Bieler, G. S., Ennett, S. T., & Zarkin, G. A. (1996). Analysis of prevention program effectiveness with clustered data using generalized estimating equations. *Journal of Consulting and Clinical Psychology*, *64*(5), 919.

Ornaghi, C. (2009). Mergers and innovation in big pharma. *International Journal of Industrial Organization*, *27*(1), 70–79.

Pan, W. (2001). Akaike's information criterion in generalized estimating equations. *Biometrics*, *57*(1), 120–125.

pandas development team, T. (2020). *Pandas.* Zenodo.

Panescu, D. (2006). Medical device industry. *Wiley Encyclopedia of Biomedical Engineering*.

Paruchuri, S., & Eisenman, M. (2012). Microfoundations of firm R&D capabilities: A study of inventor networks in a merger. *Journal of Management Studies*, *49*(8), 1509–1535.

Paruchuri, S., Nerkar, A., & Hambrick, D. C. (2006). Acquisition integration and productivity losses in the technical core: Disruption of inventors in acquired companies. *Organization Science*, *17*(5), 545–562.

Penner-Hahn, J., & Shaver, J. M. (2005). Does international research and development increase patent output? An analysis of Japanese pharmaceutical firms. *Strategic Management Journal*, *26*(2), 121–140.

Pidun, U. (2019). *Corporate strategy*. Springer, Wiesbaden.

Powell, W. W., Koput, K. W., & Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, 116–145.

Puranam, P., Singh, H., & Zollo, M. (2006). Organizing for innovation: Managing the coordination-autonomy dilemma in technology acquisitions. *Academy of Management Journal*, *49*(2), 263–280.

Puranam, P., & Srikanth, K. (2007). What they know vs. what they do: How acquirers leverage technology acquisitions. *Strategic Management Journal*, *28*(8), 805–825.

R Core Team. (2019). R: A language and environment for statistical computing. https://www.R-project.org/.

Reus, T. H. (2012). A knowledge-based view of mergers and acquisitions revisited: Absorptive capacity and combinative capability. Emerald Group Publishing Limited.

Rugman, A. M., & Verbeke, A. (2002). Edith penrose's contribution to the resource-based view of strategic management. *Strategic Management Journal*, *23*(8), 769–780.

Sampson, R. C. (2005). Experience effects and collaborative returns in R&D alliances. *Strategic Management Journal*, *26*(11), 1009–1031.

Schumpeter, J. (1981). *Theory of economic development (social science classics series)*. Routledge, London, originally published in 1912 (in German).

Schweizer, L. (2005a). Knowledge transfer and R&D in pharmaceutical companies: A case study. *Journal of Engineering and Technology Management*, *22*(4), 315–331.

Schweizer, L. (2005b). Organizational integration of acquired biotechnology companies into pharmaceutical companies: The need for a hybrid approach. *Academy of Management Journal*, *48*(6), 1051–1074.

Schweizer, L., & zu Knyphausen-Aufseß, D. (2008). Mergers and acquisitions in the biotechnology industry. In *Handbook of bioentrepreneurship* (pp. 133–148). Springer, New York.

Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. In *9th python in science conference.*

Sears, J., & Hoetker, G. (2014). Technological overlap, technological capabilities, and resource recombination in technological acquisitions. *Strategic Management Journal*, *35*(1), 48–67.

Simoens, S. (2008). Health economics of medical devices: opportunities and challenges. *Journal of Medical Economics*, *11*(4), 713–717.

Standard & Poor's. (2021). S&P Global 1200. https://www.spglobal.com/spdji/en/indices/equity/sp-global-1200/#overview. (Date accessed 27.01.2021)

Stokes, M. E., Davis, C. S., & Koch, G. G. (2012). *Categorical data analysis using sas*. SAS institute, Cary, North Carolina.

Stuart, T. E. (2000). Interorganizational alliances and the performance of firms: a study of growth and innovation rates in a high-technology industry. *Strategic Management Journal*, *21*(8), 791–811.

The Economist. (2008). Economies of scale and scope. https://www.economist.com/news/2008/10/20/economies-of-scale-and-scope. (Date accessed 13.05.2021)

Thomson Reuters. (2020). *Pfizer, biontech to co-develop potential coronavirus vaccine.* https://www.reuters.com/article/us-health-coronavirus-pfizer-biontech-idUSKBN2140LM. (Date accessed 23.05.2021)

Thomson Reuters Eikon. (2021). *Deals Screener*. https://eikon.thomsonreuters.com/index.html. (Date accessed 15.02.2021)

Valentini, G. (2012). Measuring the effect of M&A on patenting quantity and quality. *Strategic Management Journal*, *33*(3), 336–346.

Van Norman, G. A. (2016). Drugs, devices, and the fda: part 2: an overview of approval processes: Fda approval of medical devices. *JACC: Basic to Translational Science*, *1*(4), 277–287.

Van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual*. Scotts Valley, CA: CreateSpace.

Vermeulen, F., & Barkema, H. (2001). Learning through acquisitions. *Academy of Management Journal*, *44*(3), 457–476.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686.

World Intellectual Property Organisation. (2020). Guide to the international patent classification. , *2020*.

Zahra, S. A., & George, G. (2002). Absorptive capacity: A review, reconceptualization, and extension. *Academy of Management Review*, *27*(2), 185–203.

Zollo, M., & Meier, D. (2008). What is M&A performance? *Academy of Management Perspectives*, *22*(3), 55–77.

Zollo, M., & Singh, H. (2004). Deliberate learning in corporate acquisitions: post-acquisition strategies and integration capability in us bank mergers. *Strategic Management Journal*, *25*(13), 1233–1256.