JUMS

Volume 7, Issue 2, June 2022

JUNIOR MANAGEMENT SCIENCE

Published by Junior Management Science e.V.

Junior Management Science 7(2) (2022) 267-288

Junior Management Science

journal homepage: www.jums.academy

The Impact of COVID-19 Policy Measures on European Companies – Empirical Evidence from Belgium, The Netherlands, Denmark and Norway

Heiko Hoppe

Technische Universität München

Abstract

This study investigates the economic consequences of COVID-19 policy measures in Belgium, The Netherlands, Denmark and Norway. Using panel data analysis, I examine the effects of various government interventions such as lockdowns or economic support measures on risk-adjusted stock returns of companies in these countries. The findings show that both lockdown-related measures and economic support measures have a positive influence on stock returns. This positive influence is robust against competing effects such as the financial situation of companies and the pandemic itself. This study further finds that the positive influence of policy measures is consistent for companies belonging to sectors that are severely or positively affected by the pandemic. Thereby, this study contributes to the understanding of how COVID-19 policy measures affect companies and closes a research gap by considering these effects in northern European countries. It shows that lockdown-related restrictions have a positive economic impact by hindering the spread of the disease and that economic support measures ease the burden of the pandemic and are thus beneficial.

Keywords: COVID-19; government interventions; financial markets; stock returns; Northern Europe.

1. Introduction

COVID-19 came as a surprise to the public, emerging in China in December 2019 and rapidly spreading to the world in early 2020.[1] The disease was declared a Public Health Emergency of International Concern by the WHO on the 30th of January, received the name COVID-19 on the 11th of February and was declared to be a pandemic on the 11th of March.[2] The virus spread to the globe swiftly, hitting South Korea and Iran with major outbreaks.[3] The wave of infections reached Europe in March 2020, causing high fatality rates in Italy.[4] The virus has continued to spread to this day, with the number of infections rising and falling in waves but staying on a high level.[5]

To slow the spread of COVID-19 and protect their citizens, many countries adopted stringency measures such as travel restrictions, workspace closings or stay-at-home-requirements. In Europe, the first stringency measures were taken in January 2020 and sharply tightened in March 2020, reacting to the dramatic rise in the number of infections.[6] Many restrictions have remained in place until this day.[7] These measures, together with the pandemic itself, lead to dramatic economic consequences.[8] To counter these, many governments have adopted measures supporting households and companies economically. This includes measures such as income support to households, debt and contract relief or corporate debt purchase.[9] Both the stringency and the economic support measures have a multitude of influences on companies not fully understood yet.

These economic consequences of policy measures can be observed particularly well on stock markets, as they quantify expectations regarding the future economic impact of

[1] See Ding, Levine, Lin, and Xie (2021), p. 4.
[2] See Ding et al. (2021), p. 4.
[3] See Zhang, Hu, and Ji (2020), p. 2.
[4] See Zhang et al. (2020), p. 2.
[5] See Johns Hopkins University and Medicine (2021).
[6] See Blavatnik School of Government, University of Oxford (2021).
[7] See Blavatnik School of Government, University of Oxford (2021).
[8] See Sheridan, Andersen, Hansen, and Johannesen (2020), p. 20468; Zaremba, Kizys, Aharon, and Demir (2020), p. 2.
[9] See Ding et al. (2021), p. 11.

DOI: https://doi.org/10.5282/jums/v7i2pp267-288

policy measures on a daily basis.[10] Stock returns thus offer insights into economic implications of the crisis that could hardly be quantified otherwise.[11] The pandemic and the policy measures have led to unprecedented stock market reactions, showing very high market volatilities especially at the beginning of the crisis.[12] After the initial shock in March 2020, the stock markets began to recover globally from April 2020 onwards.[13]

This thesis aims at understanding the economic consequences of COVID-19 policy measures in Belgium, The Netherlands, Denmark and Norway. It contributes to the ongoing discussion about how policy interventions affect companies, an important question not fully answered yet. The thesis especially fills gaps in the literature concerning the pandemic consequences in the countries of the sample. Utilizing the government response indices developed by Hale et al. (2021) at the University of Oxford,[14] I investigate the influence of stringency and economic support measures on stock returns while controlling for the pandemic itself, company financial performance indicators, attention to the pandemic and whether a company is classified as being an essential business or not. I investigate a timeframe of over a year, from January 2020 to February 2021 using a panel data structure. To focus on pandemic-related effects, I adjust stock returns using the Fama-French three factor model and look at the excess returns over returns predicted by that model.[15] In the beginning, I investigate how stringency measures affect stock returns in general. After that, I specifically consider different economic support measures, providing an understanding of the various effects these policy interventions have on companies. In the end, I look at sectors that are severely or positively affected by the pandemic and how their stock returns are influenced by the stringency and economic support measures. I also critically discuss my methodology and results and draw conclusions from the effect of policy measures on stock returns to their effect on companies in general.

2. Literature

In this section, I present the literature on the topic of economic consequences of COVID-19 policy measures and develop my hypotheses.

2.1. Literature overview

Research into the economic consequences of COVID-19 and various policy measures began very early after the outbreak of the pandemic and has produced a growing body of literature since. An important field of research in this aspect is the impact of the pandemic itself on stock returns. It appears to be clear that COVID-19 negatively influences stock returns,[16] but the nature of that influence is not obvious. One of the topics discussed in the literature in that regard is whether the stock markets react stronger to cases or to deaths related to the virus. In their study, Al-Awadhi et al. (2020) conclude that both daily growth in confirmed cases and in confirmed deaths negatively influences stock returns.[17] Heyden and Heyden (2021) find a stronger negative market reaction to the first death in a country than to the first case of COVID-19,[18] whereas Ashraf (2020c) discovers stock markets to react strongly to cases but not to deaths.[19]

As found by Alfaro, Chari, Greenland, and Schott (2020), not only actual cases negatively influence stock returns, but also cases predicted by pandemic models, with a decrease in predicted infections having a positive impact on stock returns and vice versa.[20]

Apart from the pandemic itself, the literature suggests that attention to COVID-19-related news negatively influences stock returns.[21] Cepoi (2020) finds media coverage of the pandemic negatively influencing stock returns in the US and European countries[22] while Engelhardt et al. (2020) conclude that news attention to the topic has an even larger negative effect on stock returns than rational investor's expectation.[23]

A different factor is investigated by Ashraf (2020a), finding that national cultures influence stock returns, especially a high uncertainty avoidance negatively affecting returns when COVID-19 cases increase.[24]

Apart from these factors, a major focus of interest for many scientists is investigating the effects COVID-19 policy measures have on stock markets. Comparing COVID-19 to previous pandemics, Baker et al. (2020) find government restrictions and voluntary social distancing to be the major causes of the unprecedentedly volatile stock market reaction to COVID-19 in the US, hitting a service-oriented economy especially hard.[25] Confirming these results in a study covering more countries, Zaremba et al. (2020) discover government interventions to increase stock market volatility even when controlling for the pandemic itself.[26]

Apart from market volatility, stock returns are also affected by the policy measures, found for example by Yang and Deng (2021), who investigate a time period of over half

[10] See Ramelli and Wagner (2020), p. 623.
[11] See Ramelli and Wagner (2020), p. 623.
[12] See Baker, Davis, Kost, Sammon, and Viratyosin (2020), p. 755; Baker et al. (2020), p. 743.
[13] See Ding et al. (2021), p. 7.
[14] See Hale et al. (2021), pp. 20-27.
[15] See Fama and French (1993), p. 5.
[16] See Al-Awadhi, Alsaifi, Al-Awadhi, and Alhammadi (2020), pp. 3f; Heyden and Heyden (2021), pp. 3f; Ashraf (2020c), pp. 4-6; See Ding et al. (2021), pp. 13f.
[17] See Al-Awadhi et al. (2020), pp. 3f.
[18] See Heyden and Heyden (2021), pp. 3f.
[19] See Ashraf (2020c), pp. 4-6.
[20] See Alfaro et al. (2020), p. 1.
[21] See Cepoi (2020), pp. 3f; Engelhardt, Krause, Neukirchen, and Posch (2020), p. 10.
[22] See Cepoi (2020), pp. 3f.
[23] See Engelhardt et al. (2020), p. 10.
[24] See Ashraf (2020a), p. 5.
[25] See Baker et al. (2020), pp. 755f.
[26] See Zaremba et al. (2020), p. 6.

a year and conclude that stringency measures increase the negative effects of the pandemic on stock returns.[27] Looking specifically at the travel and leisure sector in the US, Chen, Demir, García-Gómez, and Zaremba (2020) similarly find stringency measures negatively affecting Fama-French-adjusted stock returns when accounting for the influence of COVID-19.[28] But even economic support measures, aimed at cushioning the negative influences of the pandemic and the stringency measures, might not only have positive effects, as Zhang et al. (2020) find the pandemic leading to unprecedented stock market movements and unconventional policy interventions like unlimited quantitative easing creating more uncertainty, especially in the long run.[29] Shanaev, Shuraeva, and Ghimire (2020) go even further with arguing that the negative effect of COVID-19 itself is very small, but that an irrational panic surrounding the pandemic and especially national lockdowns and economic stimulus measures have severe negative consequences on stock markets.[30] The policy measure they recommend are regional lockdowns, as these do not lead to large market effects at all.[31]

However, not all papers conclude that policy measures have adverse effects on the stock markets. Investigating dividend futures, Gormsen and Koijen (2020) find that growth expectations are heavily influenced by the crisis and that, after strongly decreasing, long-term growth expectations were stabilized by the announcement of fiscal policies at the beginning of the crisis.[32] Even stringency or lockdown measures are found to have a positive influence on stock markets by mitigating the effect of the pandemic, as argued by Narayan, Phan, and Liu (2021) in their study of G7 stock market returns.[33] Similarly, Ding et al. (2021) discover a negative effect of COVID-19 cases and a positive effect of lockdown and stimulus measures on market returns when investigating the effects of country characteristics and policy measures.[34] When examining aggregate spending in Denmark and Sweden, Sheridan et al. (2020) also find COVID-19 itself to be responsible for more economic harm than the policy measures.[35] Heyden and Heyden (2021) come to a mixed conclusion by discovering that fiscal policy measures add uncertainty whereas monetary policy measures calm markets.[36]

This is a very interesting phenomenon, as the literature is clearly divided into some studies indicating negative and some studies indicating positive effects of policy interventions. Ashraf (2020b) addresses this phenomenon of policy measures having both positive and negative consequences, stating that stringency measures have both direct, negative effects on stock returns due to their adverse influence on economic activities and indirect, positive effects due to them reducing COVID-19 cases.[37] In his contribution, he largely finds economic support measures, testing and quarantining programs to exert a positive impact on stock returns.[38]

Similarly, economic support measures have a positive, intended effect on the economy by providing economic and financial support for companies. On the other hand, they may increase uncertainty about the crisis[39] and lead to negative consequences for the stock market.[40] Whether the positive or the negative effects are more important is not clear in the literature, both for stringency and economic support measures.

Another question addressed in the literature is which companies suffer most during the pandemic and which companies suffer less or are even positively affected. One of the factors in this regard is the industry in which the company operates. Considering the impact of COVID-19 on different sectors of the Chinese economy, He, Sun, Zhang, and Li (2020) find for example the transportation, mining and electric industries to be severely affected, whereas information technology, public management and entertainment are positively affected.[41] Apart from that, energy, apparel, real estate and the service industry are found to be severely affected and telecom, pharma/biotech and software are found to be positively affected by the crisis, amongst others.[42] Especially the travel and leisure industry is often investigated, as this sector is considered to be hit particularly hard by travel restrictions and social distancing measures.[43] Indeed, when looking specifically at the travel and leisure sector in the US, Chen et al. (2020) find the stringency measures negatively affecting Fama-French-adjusted stock returns, even when controlling for the pandemic itself.[44] Lin and Halk (2021) find similar results in their study of the Scandinavian (Denmark, Finland, Sweden) travel and leisure sector, concluding that this sector is affected by the pandemic, with especially international transport companies suffering while online casinos benefit from the crisis.[45]

Apart from the sector in which a company operates, different corporate characteristics can have an influence on how a company is affected by the pandemic. In their study, Ramelli and Wagner (2020) find the international orientation and financial position of companies influencing their stock returns during the early pandemic, expressing the positive impact of cash holdings and the negative impact of close relations with China in that period.[46] Ding et al. (2021) investigate the influence of corporate characteristics, including financial performance indicators, international orientation,

[27] See Yang and Deng (2021), p. 4.
[28] See Chen et al. (2020), p. 5.
[29] See Zhang et al. (2020), p. 5.
[30] See Shanaev et al. (2020), p. 42.
[31] See Shanaev et al. (2020), p. 42.
[32] See Gormsen and Koijen (2020), p. 574.
[33] See Narayan et al. (2021), p. 5.
[34] See Ding et al. (2021), p. 13.
[35] See Sheridan et al. (2020), p. 20471.
[36] See Heyden and Heyden (2021), pp. 3f.
[37] See Ashraf (2020b), p. 7.
[38] See Ashraf (2020b), p. 7.
[39] See Zhang et al. (2020), p. 5.
[40] See Shanaev et al. (2020), p. 42f.
[41] See He et al. (2020), p. 2206.
[42] Ramelli and Wagner (2020), p. 633; Baker et al. (2020), p. 752; Xiong, Wu, Hou, and Zhang (2020), p. 2236.
[43] See Chen et al. (2020), p. 1.
[44] See Chen et al. (2020), pp. 4f.
[45] See Lin and Halk (2021), p. 15.
[46] See Ramelli and Wagner (2020), pp. 637-643.

corporate governance and ownership structure on stock returns during the early days of COVID-19, similarly finding that a strong pre-pandemic financial position and less international orientation, amongst other characteristics, lead to better stock returns during the crisis.[47] Looking at the stock market in China, Xiong et al. (2020) also find companies having larger profits, a greater size and less fixed assets to show higher stock returns during the pandemic.[48]

So far, not much research has been done considering the economic consequences of COVID-19 specifically in Belgium, The Netherlands, Denmark and Norway. Hoekman, Smits, and Koolman (2020) investigate regional differences in the Netherlands, finding that the economic shock in this country was relatively mild in the early phase of the pandemic.[49] Lin and Halk (2021) examine the situation of the travel and leisure sector in Scandinavia (Denmark, Finland, Sweden)[50] and Sheridan et al. (2020) study economic activities during the crisis in Denmark and Sweden.[51]

2.2. Hypothesis formulation

Looking at the literature, it is obvious that the economic consequences of COVID-19 policy measures are not fully understood yet. As mostly China[52], the US[53] or larger sets of countries[54] have been investigated so far, studying a set of smaller countries in northern Europe might result in interesting findings and fills a gap in the literature.

As stated above, a major issue in which the literature is unclear is whether policy interventions, both stringency measures and economic support measures, have positive or negative effects on stock returns. I address this research question in my thesis. Following the results of Ding et al. (2021) and Narayan et al. (2021), I assume a positive relationship between stringency measures and adjusted stock returns, as it appears reasonable that these measures mitigate the negative effects of COVID-19.[55] This leads to the first hypothesis:

> H1: Stringency measures are positively correlated with adjusted stock returns.

Similarly, it appears sensible to assume that economic stimulus measures have a positive relationship with adjusted stock returns as these measures support companies and potentially calm markets, following the results of Gormsen and Koijen (2020), Ashraf (2020b) and Ding et al. (2021).[56] This leads to the second hypothesis:

> H2: Economic support measures are positively corelated with adjusted stock returns.

A noteworthy phenomenon during the COVID-19 pandemic is that different sectors of the economy are affected differently by the crisis, with some industries struggling and others performing well, as visualized by Ramelli and Wagner (2020).[57] A broad range of researchers finds different sectors to be severely or positively affected by the pandemic.[58] Here, it is interesting how the effect of the stringency measures on companies belonging to sectors regarded as severely or positively affected by the pandemic looks like. For companies belonging to sectors regarded as being severely affected by the pandemic, it appears logical to assume that the stringency measures have a negative relationship with adjusted stock returns, partly explaining the severe affection by the crisis. For companies belonging to sectors regarded as being positively affected by the pandemic, however, the stringency measures should have a positive relationship with adjusted stock returns, partly explaining the positive affection by the crisis. This leads to hypothesis 3, divided into two sub-hypotheses:

> H3a: For companies belonging to a sector regarded as being severely affected by the pandemic, the stringency measures are negatively correlated with adjusted returns.

> H3b: For companies belonging to a sector regarded as being positively affected by the pandemic, the stringency measures are positively correlated with adjusted returns.

The rest of this thesis is organized as follows: Section 3 explains the methodology of testing the hypotheses, section 4 shows the data sources, data collection and data treatment. Section 5 presents the results of the analyses and their interpretation and section 6 discusses the methods and results with their limitations. Section 7 concludes the thesis.

3. Methodology

In this section, I describe the methods I use to test the hypotheses. I explain the used variables and the regression models.

3.1. Variables

To examine the impact of COVID-19 policy measures on stock returns, I use multiple linear ordinary least squares (OLS) regressions on a panel data structure. The dependent variable are always risk-adjusted stock returns, using the Fama-French three factor model to adjust the returns.[59] The three factors of this model explain a part of the variance of stock returns by considering the relative performance of a stock compared to the overall market, the return differences between small and big companies and the influence of the

[47] See Ding et al. (2021), p. 25.
[48] See Xiong et al. (2020), p. 2240.
[49] See Hoekman et al. (2020), p. 620.
[50] See Lin and Halk (2021), pp. 2f.
[51] See Sheridan et al. (2020), pp. 20468f.
[52] See Xiong et al. (2020), p. 2234; He et al. (2020), p. 2202.
[53] See Ramelli and Wagner (2020), p. 631; Chen et al. (2020), p. 3.
[54] See Ding et al. (2021), p. 6; See Ashraf (2020a), pp. 2f.
[55] See Ding et al. (2021), p. 13; Narayan et al. (2021), p. 5.
[56] See Gormsen and Koijen (2020), p. 574; Ashraf (2020b), p. 7; Ding et al. (2021), p. 13.
[57] See Ramelli and Wagner (2020), p. 633.
[58] See Ramelli and Wagner (2020), p. 633; Baker et al. (2020), p. 752; Xiong et al. (2020), p. 2236; He et al. (2020), p. 2206.
[59] See Fama and French (1993), p. 5.

market-to-book ratio on stock returns.[60] These risk factors therefore account for influences on stock returns not related to the pandemic. By utilizing excess returns over returns predicted by the model, I leave the part of return variance that cannot be explained by Fama-French risk factors for the analyses. A definition of all used variables can be found in Appendix 1.

My main explanatory variables are the Stringency Index and the Economic Support Index from the Oxford COVID-19 Government Response Tracker, developed by Hale et al. (2021) at the University of Oxford.[61] I regard these Oxford indices as a very good measurement of policy responses, as they account for most, especially the most important, policy measures, use unified scales, cover the entire time period of the pandemic and are commonly used in research, for example by Alfaro et al. (2020), Ding et al. (2021), Ashraf (2020b) and Chen et al. (2020).[62] The Stringency Index measures policy interventions aimed at preventing the spread of the pandemic like workspace closing or travel restrictions and the Economic Support Index measures policy interventions aimed at financially supporting households.[63] I use the Stringency Index as the main explanatory variable in hypotheses 1, 3a and 3b and the Economic Support Index as the main explanatory variable in hypothesis 2. When an index is not the main explanatory variable, I use it as a control variable to account for its effect on adjusted stock returns. For some of the models testing hypothesis 2, I also use the individual economic support measures E1 income support, E2 debt and contract relief and E3 other fiscal measures to gain a more detailed insight into the effects of certain policy measures. That way, I can estimate the economic impact of individual policy measures and assess whether some measures have different effects than others.

In addition to these variables, I use control variables to account for effects on stock returns not explained by the Oxford indices. To control for the influence of a company's financial structure on its stock returns, I use several financial performance indicators that are employed by Ding et al. (2021) or Ramelli and Wagner (2020) or both:[64] Firm size as the logarithm of total assets, leverage, cash by assets, ROA and book-to-market equity. I include firm size and book-to-market equity although they are already considered by the Fama-French three factor model because the influence of these parameters on stock returns might have changed during the crisis.

During a global pandemic severely affecting multinational supply chains, the international orientation of a company is an important factor to consider.[65] To assess the international orientation of a company, I use the foreign sales ratio, defined as the percentage of total revenues that is generated in foreign countries by the company as a variable, as done by Ramelli and Wagner (2020).[66] As the data for this variable is very incomplete, with 160 companies lacking data on foreign sales completely, I use this variable as a robustness check rather than a major control variable to maintain a larger sample in the main analyses.

An important influence on stock returns is the pandemic itself, as shown by Ding et al. (2021), Ashraf (2020c) and Al-Awadhi et al. (2020).[67] To measure this influence, I utilize COVID-19 cases as a control variable in the analyses. Using deaths from COVID-19 would also be possible, as it measures the severe consequences of the pandemic, but Ashraf (2020c) finds that stock markets react stronger to cases than they do to deaths[68] and Al-Awadhi et al. (2020) find that both indicators are correlated.[69] I use two variables to estimate COVID-19 cases, both based on formulas used by Ding et al. (2021):[70] AdjustedCases is the main variable and measures the change in the ratio of positive test results and UnadjustedCases is a variable I use for robustness checks and measures the growth of confirmed cases. For both formulas, I use cases of the previous day and the day before, similarly to Ashraf (2020c),[71] because new cases of a day are unlikely to affect the stock markets on the same day.

Especially at the beginning of the crisis, when actual cases were low and few policy measures were in place, attention to COVID-19 had an influence on the stock markets, as Engelhardt et al. (2020), Baker et al. (2020) and Cepoi (2020) find.[72] To measure this attention, I use the Google Search Volume Index for the term "corona" at the beginning of the pandemic as a control variable, that measure being similarly used by Ramelli and Wagner (2020) and Engelhardt et al. (2020).[73]

An important difference between companies during the pandemic is whether they belong to essential industries or not. As essential industries are considered to provide services necessary for society, they are often excluded from lockdown measures or receive special support.[74] I therefore use a dummy variable as a control variable which is one if a company belongs to an essential industry.

In addition to these variables, I use fixed effects (FE) to account for influences not captured by the variables but of importance for the results. I use Industry FE to account for effects varying across industries but constant over time, like Heyden and Heyden (2021) and Ramelli and Wagner (2020).[75] Similar to the GICS industry groups used by

[60] See Fama and French (1993), p. 5.
[61] See Hale et al. (2021), p. 27.
[62] See Alfaro et al. (2020), p. 12; Ding et al. (2021), p. 11; Ashraf (2020b), p. 2; Chen et al. (2020), p. 3.
[63] See Hale et al. (2021), pp. 20-27.
[64] See Ding et al. (2021), p. 7; Ramelli and Wagner (2020), p. 635.
[65] See Ramelli and Wagner (2020), pp. 634f; Ding et al. (2021), pp. 7f.
[66] See Ramelli and Wagner (2020), pp. 634f.
[67] See Ding et al. (2021), pp. 14f; Ashraf (2020c), pp. 5f; Al-Awadhi et al. (2020), p. 3.
[68] See Ashraf (2020c), pp. 5f.
[69] See Al-Awadhi et al. (2020), p. 3.
[70] See Ding et al. (2021), pp. 4-6.
[71] See Ashraf (2020c), p. 2.
[72] See Engelhardt et al. (2020), p. 10; Baker et al. (2020), p. 749; Cepoi (2020), pp. 3f.
[73] See Ramelli and Wagner (2020), p. 630; Engelhardt et al. (2020), p. 3.
[74] See Wales (2020), pp. 3f; Heyden and Heyden (2021), p. 3.
[75] See Ramelli and Wagner (2020), p. 638; Heyden and Heyden (2021),

Ramelli and Wagner (2020), I use the first two numbers of the ICB code to classify the industries.[76] I use Country FE to account for effects varying across countries but constant over time, like Ashraf (2020b), Heyden and Heyden (2021) and Ding et al. (2021).[77] As done by Narayan et al. (2021) and Zaremba et al. (2020), I use Weekday FE to account for effects varying across days of the week but constant over companies and weeks.[78] Finally, I use Company FE in some models to account for effects varying across companies but constant over time. Upon adding Company FE, all other variables constant for a company drop out of the model. Industry and Country FE are automatically included in Company FE. As financial performance indicators are constant for some companies but have two values for others, the model would not completely drop them automatically, resulting in multicollinearity. When using Company FE, I therefore manually drop financial performance indicators.

3.2. Regression models

For each hypothesis, I run multiple regressions using different variables. I generally start with only the main explanatory variable or variables, then add COVID-19-related variables (Attention, AdjustedCases and essential) and then add company-related variables (financial performance indicators), before finally adding company fixed effects. This enables me to assess the effects various variables have on the regression results and how the effects of the main explanatory variables are changed by adding other variables.

To counter issues of heteroscedasticity, I use robust standard errors clustered by company in all regressions.

The formula for model 3 of hypothesis 1 looks as follows:

$$\begin{aligned}\alpha_{i,t} = {} & b_0 + b_1 SI_{c,t} + b_2 ESI_{c,t} + b_3 Size_{i,t} + b_4 Leverage_{i,t} \\ & + b_5 CashByAssets_{i,t} + b_6 ROA_{i,t} + b_7 BookToMarket_{i,t} \\ & + b_8 Attention_{c,t} + b_9 AdjustedCases_{c,t} + b_{10} essential_i \\ & + \mu_{ind} + \mu_c + \mu_w + u_i + \varepsilon \end{aligned} \tag{1}$$

where $\alpha_{i,t}$ is the adjusted return of company i on day t, $SI_{c,t}$ is the value of the Stringency Index for country c on day t, $ESI_{c,t}$ is the value of the Economic Support Index for country c on day t, $Size_{i,t}$ is the value of the variable Size for company i and day t, $Leverage_{i,t}$ is the value of the variable Leverage for company i and day t, $CashByAssets_{i,t}$ is the value of the variable CashByAssets for company i and day t, $ROA_{i,t}$ is the value of the variable ROA for company i and day t, $BookToMarket_{i,t}$ is the value of the variable BookToMarket for company i and day t, $Attention_{c,t}$ is the value of the variable Attention measuring attention to COVID-19 for country c and day t, $AdjustedCases_{c,t}$ is the value of the variable AdjustedCases measuring COVID-19 cases for country c and day t, $essential_i$ is the value of the binary variable indicating whether company i belongs to a sector classified as essential or not, μ_{ind}, μ_c, μ_w are Industry, Country and Weekday fixed effects, $u_i + \varepsilon$ is the error term for robust standard errors clustered by company, b_0 is the intercept and b_1 to b_{10} are the coefficients. A definition of the variables used can be found in Appendix 1. This is the model for hypothesis 1 containing most variables. All other models, for this and the other hypotheses, are just variations of this model, with some variables being added, removed, or replaced according to the regression tables. Therefore, this is the main model of the thesis.

For hypotheses 3a and 3b, I run the regressions on subsamples of companies, using companies belonging to industries regarded as being severely affected by the pandemic for hypothesis 3a and companies belonging to industries regarded as being positively affected by the pandemic for hypothesis 3b.

4. Data

In this section, I describe the data sources, data collection and data treatment for the data used in the analyses. I furthermore present descriptive statistics for the used variables. In the end, I briefly consider regression diagnostics for the main model.

4.1. Adjusted stock returns

The adjusted stock returns are excess returns, calculated as the difference between the actual, raw returns and the returns predicted by the Fama-French three factor model. To obtain raw returns, I download daily data for the return index (RI) from Refinitiv Datastream (2021) for all companies in the full sample beginning on the 26th of January 2018 and ending on the 26th of February 2021. I choose these timepoints for the following reasons: The 26th of February 2021 is the last day for which data necessary for calculating adjusted returns was provided on Kenneth French's website for the Fama-French three factor model at the time of the download.[79] As the 27th of January 2020 is the date on which one of the countries in the sample had a Government Response Index, the most comprehensive Oxford index,[80] larger than zero for the first time, I choose this date as the beginning of the observation period. To receive good estimates for the betas of the Fama-French three factor model, I use the two years before the observation period as the beta-calculating period, beginning on the 27th of January 2018. To calculate a meaningful value for the first raw return, I also include the 26th of January 2018 in the download of the return index.

To clean the data, I apply several screens based on the screens conducted by Hanauer and Windmüller (2020) and

p. 5.

[76] See Ramelli and Wagner (2020), p. 638.

[77] See Ashraf (2020b), p. 6; Heyden and Heyden (2021), p. 5; Ding et al. (2021), p. 13.

[78] See Narayan et al. (2021), pp. 3-5; Zaremba et al. (2020), pp. 4f.

[79] See French (2021).

[80] See Hale et al. (2021), p. 27.

Schmidt, von Arx, Schrimpf, Wagner, and Ziegler (2019).[81] The screens aim at deleting dead companies, abnormal returns, returns on holydays and other data potentially disturbing the results. To conduct some of the screens, I download the unadjusted prices (UP) for the entire time period for all companies in the sample after some of the screens. A detailed description of these screens is given in Table A.2. Static screens are mostly unnecessary, as I use the company list of Hanauer and Windmüller (2020) on which static screens had already been applied by these researchers.[82] I take care not to delete any company delisted during the observation period to exclude the possibility of survivorship bias.

I conduct the data preparation in Excel and R, using R version 4.0.4 for most of the preparation steps for all variables. The R file containing the data preparation code is provided in Attachment 14. However, it does not use the raw data files also attached, as I prepare some datasets in Excel before loading them into R.

I calculate the raw returns as the daily change of the return index relative to the return on the previous day. I run a regression over the beta-calculating period for each company to estimate the coefficients of the Fama-French three factor model using the regression equation of that model:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{i,1}\left(r_{m,t} - r_{f,t}\right) + \beta_{i,2} \times SMB_t + \beta_{i,3} \times HML_t + \varepsilon \quad (2)$$

[83]

where $r_{i,t}$ is the (raw) return of company i on day t, $r_{f,t}$ is the risk-free return on day t, $r_{m,t}$ is the market return on day t, SMB_t is the Small Minus Big factor and HML_t is the High Minus Low factor of the Fama-French three factor model on day t, ε is the error term (equals zero on average), α_i is the intercept and the three betas are the regression coefficients. The market return and the factors of the Fama-French three factor model are downloaded from Kenneth French's website as daily data for the European market for the entire time period.[84] Using the estimated betas, I calculate the adjusted returns by subtracting the returns predicted by the Fama-French three factor model from the raw returns for each company and each day of the observation period using the formula described above. A detailed description of the calculation of the adjusted returns is given in Appendix 2.

4.2. COVID-19 policy measures

I download the Oxford dataset, only keeping data for indicators, countries and time points I use in my analyses from the website of the Blavatnik School of Government of the University of Oxford.[85] The indicator M1, a wild card for policy measures not fitting into any other category, does not contain any information, so I do not regard this indicator. To adjust the values of E3, which are given in absolute US Dollar amounts, to the financial strength of a country, I divide the E3-value by the country's GDP of 2019. I download data on the GDP values, which are $ 533,097,455,830 for Belgium, $ 907,050,863,150 for The Netherlands, $ 350,104,327,660 for Denmark and $ 403,336,363,640 for Norway from the website of the World Bank.[86] I delete data from weekends for all indices, as they are also not included in the return data, adding E3-values of weekends to the next Monday. Further data manipulation is not necessary.

4.3. Control variables

The financial performance indicators are calculated as follows, following the approaches of Ding et al. (2021) or Ramelli and Wagner (2020) or both:[87] Size is the natural logarithm of total assets;[88] Leverage is the total debt divided by total assets;[89] CashByAssets is cash and short-term investments divided by total assets;[90] ROA is the net income before extraordinary items divided by total assets[91] and BookToMarket is the book value of equity divided by the market value of equity.[92] To do these calculations, I download the required data from Refinitiv Worldscope (2021), Refinitiv Datastream (2021) and Orbis (2021), using mainly Refinitiv Worldscope (2021) data and filling data gaps with Orbis (2021) data. To avoid extreme values from influencing the results, I winsorize all financial performance indicators at 1% and 99%. For all financial performance indicators, I use values calculated utilizing accounting data from the fiscal year 2019 until the end of the company's fiscal year 2020 and from the fiscal year 2020 thereafter if that data was already available at the time of the download, resorting to 2019 data if not. Similar to the data truncation done by Ramelli and Wagner (2020),[93] I set all Leverage values larger than 1 to 1 prior to winsorizing the data, as larger values are hardly possible and could potentially change the results of my analyses. The values of 7 companies are thus changed. I multiply all Leverage, CashByAssets and ROA values by 100 to obtain percentage values. Details on the data preparation steps for the company financial performance indicators can be found in Appendix 2.

I download the foreign sales in percent of total sales (WC07101) from Refinitiv Worldscope (2021) along with the company financial data for 2019 and 2020, this item being very similar to the foreign sales ratio used by Ramelli and Wagner (2020).[94] I apply the same data preparation steps as I do to the company financial performance indicators, winsorizing the data at 1% and 99% and using 2019 foreign sales data until the end of a company's fiscal year 2020 and

[81] See Hanauer and Windmüller (2020), p. 64; Schmidt et al. (2019), Online Appendix p. 19.
[82] See Hanauer and Windmüller (2020), pp. 61-63.
[83] See Fama and French (1993), p. 9f.
[84] See French (2021).
[85] See Blavatnik School of Government, University of Oxford (2021).
[86] See World Bank (2021).
[87] See Ding et al. (2021), p. 7; Ramelli and Wagner (2020), p. 635.
[88] See Ding et al. (2021), p. 7.
[89] See Ding et al. (2021), p. 7; Ramelli and Wagner (2020), p. 635.
[90] See Ding et al. (2021), p. 7; Ramelli and Wagner (2020), p. 635.
[91] See Ding et al. (2021), p. 7; Ramelli and Wagner (2020), p. 635.
[92] See Ramelli and Wagner (2020), p. 635.
[93] See Ramelli and Wagner (2020), p. 636.
[94] See Ramelli and Wagner (2020), pp. 634f.

2020 foreign sales data thereafter, if available. Similar to Leverage, I set all foreign sales values larger than 100 to 100 prior to winsorizing the data, changing the values of 8 companies.[95]

The formulas used to calculate the two variables measuring COVID-19 cases are the following, based on formulas used by Ding et al. (2021):[96]

AdjustedCases:

$$ac_{c,t} = \ln\left(1 + \frac{culcases_{c,t-1}}{tottest_{c,t-1}}\right) - \ln\left(1 + \frac{culcases_{c,t-2}}{tottests_{c,t-2}}\right)^{97} \quad (3)$$

UnadjustedCases:

$$uc_{c,t} = \ln\left(1 + culcases_{c,t-1}\right) - \ln\left(1 + culcases_{c,t-2}\right)^{98} \quad (4)$$

Where $culcases_(c,t)$ are the cumulative COVID-19 cases of country c on day t and $tottests_(c,t)$ is the number of all tests conducted in country c before and on day t. For the reason given above, I use cases of the previous day and the day before in both formulas. I download the required data on COVID-19 cases and tests for the countries in the sample from the website of the Foundation for Innovative New Diagnostics,[99] an organization having a partnership with the World Health Organization and the Bill & Melinda Gates Foundation, a source also used by Ding et al. (2021).[100] I delete the data for timepoints outside of the observation period and calculate the variables, replacing the fractions in the formula for the AdjustedCases with zero if the total number of tests is zero on a day. I remove data for weekends, multiply both variables by 100 to match them with the rest of the data and winsorize them at 1% and 99%, as done with the financial performance indicators and foreign sales.

I download the Google Search Volume Index for the term "corona" for the countries in my sample from the 27th of January 2020 until the 11th of May 2020 from Google Trends.[101] I use this term, as other terms associated with the pandemic were only created later on, like the novel disease being given the name COVID-19 on the 11th of February 2020 by the WHO.[102] As the pandemic and policy measures responding to it have vast effects on the economy, I assume that attention to the pandemic remains high even after the Search Volume Indices have reached their peaks. I therefore only keep the original Search Volume Indices until an index reaches 100 and replace all later values with 100. As done with the other variables, I delete all data for weekends.

I use the essential workforce classification by Wales (2020), issued by the US Cybersecurity and Infrastructure Security Agency[103] as basis for my essential industry classification as recommended by Heyden and Heyden (2021), because although every country defines essential businesses slightly differently, the classification is mostly done in a very similar way.[104] I classify companies as essential if the industry they belong to is mentioned as being an essential business by Wales (2020) using the companies' SIC codes and a website explaining SIC code meanings.[105] This classification is of course not perfectly precise, but should be sufficient for this purpose. A list of all industries I classified as essential and their corresponding SIC codes can be found in Table A.4.

To identify companies belonging to severely or positively affected industries, I use the results of Ramelli and Wagner (2020), Baker et al. (2020), Xiong et al. (2020) and He et al. (2020).[106] I therefore classify companies as belonging to sectors regarded as being severely or positively affected if the industries they belong to are described as being such, using SIC codes as done for the essential classification. Industries regarded as being severely affected include consumer services, tourism and hospitality and transportation, amongst others. Industries regarded as being positively affected include telecom, pharma/biotech and software companies, amongst others. A list of all industries I classify as severely or positively affected and their corresponding SIC codes can be found in Appendix 2, including Table A.5 and Table A.6.

4.4. Descriptive Statistics

Table 1 presents descriptive statistics for the used variables. It can be seen that the adjusted returns are centered very much around zero, with about half of the values being less than one percentage point away from that number. However, the large SD implies that more extreme values exist. This indicates that the Fama-French three factor model is able to explain a large part of the return variance, but not all of it. The values of the mean, the median and the percentiles show that the distribution of the variable is not alarmingly skewed.

In contrast, SI and ESI show a lot of variance with standard deviations and interquartile ranges between 20 and 50. As these indices can take values between 0 and 100,[107] the distributions indicate multiple changes of policy measures. Figure 1 confirms this observation, the figure presenting the development of stringency measures in the countries of the sample over time. The figure shows that stringency measures were sharply increased around March 2020 and remained high until the end of the observation period, only temporarily being eased during summer and autumn 2020 in the four countries. Interestingly, the four countries have very similar lines in the figure, tightening and easing stringency measures almost at the same time. The figure and the distribution of

[95] See Ramelli and Wagner (2020), p. 636.
[96] See Ding et al. (2021), pp. 4-6.
[97] See Ding et al. (2021), p. 6.
[98] See Ding et al. (2021), p. 4.
[99] See Foundation for Innovative New Diagnostics (2021).
[100] See Ding et al. (2021), p. 6.
[101] See Google LLC (2021).
[102] See Ding et al. (2021), p. 4.
[103] See Wales (2020), pp. 7-23.
[104] See Heyden and Heyden (2021), p. 2.
[105] See SIC-NAICS LLC (2021).
[106] See Ramelli and Wagner (2020), p. 633; Baker et al. (2020), p. 752; Xiong et al. (2020), p. 2236; He et al. (2020), p. 2206.
[107] See Hale et al. (2021), p. 29.

Table 1: Descriptive statistics

This table shows the descriptive statistics of the variables used in my analyses. The definitions and data sources for all variables can be found in Appendix 1. Reported are the number of non-missing observations for each variable (N), the mean and standard deviation (SD) of the variable and the value of the variable at the 25th percentile, 50th percentile (Median) and 75th percentile of the distribution of the variable.

Variable	N	Mean	SD	p(25)	Median	p(75)
AdjustedReturn	117,936	0.062	3.221	-1.053	-0.047	0.949
SI	129,675	52.244	20.366	40.7	56.0	66.7
ESI	129,675	57.762	30.962	37.5	62.5	87.5
E1	129,675	1.749	0.652	2	2	2
E2	129,675	0.899	0.908	0	1	2
E3	129,675	0.0002	0.004	0	0	0
AdjustedCases	129,675	-0.014	0.148	-0.02	-0.002	0.005
UnadjustedCases	129,675	2.742	7.718	0.2	0.7	1.6
Attention	129,675	89.629	29.079	100	100	100
Size	129,390	20.129	2.432	18.584	20.248	21.736
Leverage	129,100	27.798	21.753	10.031	24.399	40.564
CashByAssets	125,970	13.869	20.280	2.138	6.182	15.505
ROA	129,390	-2.013	24.360	0.090	1.919	5.152
BookToMarket	129,390	0.946	3.548	0.285	0.643	1.152
ForeignSales	82,855	46.669	37.793	1.000	50.890	81.900
essential	129,675	0.574	0.495	0	1	1
severelyAffected	129,675	0.312	0.463	0	0	1
positivelyAffected	129,675	0.141	0.348	0	0	0

the variable suggest a skewness of SI, as it contains many rather high values and a few very low values from the beginning of the observation period. ESI and its components E1 and E2 do not have a large temporal variation, as these measures often remain active for a long time.[108] The individual economic support indicators E1 and E2 can only take values between 0 and 2[109] and E3 can only be positive and less than 1 and is generally very small due to its definition. The variance of E1 and E2 is of course similar to the one of ESI, with E2 showing more variation.

Considering the control variables, the measurements for COVID-19-cases are logarithmic, changing the distribution compared to non-logarithmic variables. Attention is obviously skewed, as is to be expected regarding the definition of that variable. The financial performance indicators do not offer unexpected findings, the huge differences between the means and medians of CashByAssets and ROA suggest that outliers influence the distributions of these variables. However, mean and median are within one standard deviation from each other in both cases. The dummy variables show that a slight majority of companies is classified as being essential, about a third belongs to severely affected sectors and only a few companies belong to positively affected sectors.

Generally, the distribution of all variables is similar to their distribution in the literature I base the variable on. The distribution of the adjusted returns is similar to the distribution of CAPM-adjusted returns used by Ramelli and Wagner (2020).[111]

4.5. Regression diagnostics

To ensure the assumptions underlying multiple linear regressions are fulfilled, I perform regression diagnostics for the main model, the formula of which is given in section 3.2.

I consider the following assumptions, as provided on the website of the University of California, Los Angeles: Linearity, normality of residuals, homoscedasticity, independence, model specification, influential values, collinearity.[112] I further look at the randomness of the sample and possible issues of endogeneity, these being the regression assumptions described by Roberts and Whited (2013) that are not automatically fulfilled or already covered by the assumptions above.[113]

I check for linearity using the Residuals vs Fitted plot, shown in Figure A.1. A straight red line indicates no issues with the linearity assumption, which is the case.

I check for normality of residuals using a normal Q-Q-plot of the residuals, shown in Figure A.2. The plot shows that the residuals are not distributed normally, but follow a broader distribution, having long tails. However, no signs of skewness can be seen in the plot. This is not a problem for the calculated coefficients, which just rely on the residuals being identically and independently distributed, which should be

[108] See Blavatnik School of Government, University of Oxford (2021).
[109] See Hale et al. (2021), p. 22.
[110] See Blavatnik School of Government, University of Oxford (2021).
[111] See Ramelli and Wagner (2020), p. 636.
[112] See University of California, Los Angeles (2021).
[113] See Roberts and Whited (2013), pp. 497f.

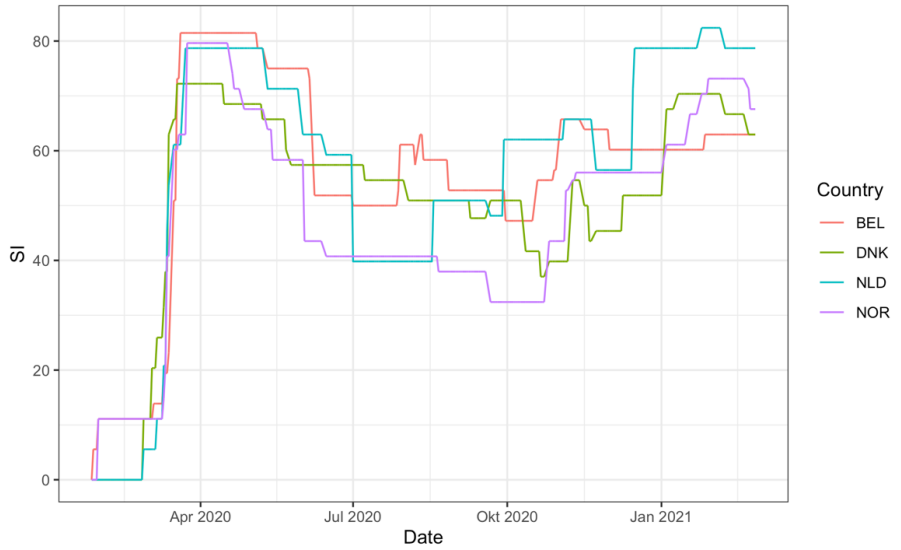

Figure 1: Stringency Index over time

This figure visualizes the values of the Stringency Index over time during the observation period for each country of the sample. Data source is the Oxford COVID-19 Government Response Tracker.[110] The time is displayed on the x-axis, the values of the Stringency Index (SI) are displayed on the y-axis. The colors represent the different countries. Variable definitions can be found in Appendix 1.

the case.[114] The t- and F-tests, however, are not fully valid.[115] I keep the data as it is, as changing the distribution would delete a large part of the data's variance, but p-values close to significance thresholds should be handled with care, which is a constraining factor.

I check for homoscedasticity using the Scale-Location plot, shown in Figure A.3. A straight red line indicates no issues with heteroscedasticity. In this case, the red line is slightly curved, but not alarmingly so. The error variance should therefore be relatively constant. All heteroscedasticity left is dealt with using robust standard errors.

Independence, meaning that the errors associated with one observation are not corelated with errors of any other observation, is addressed by clustering standard errors by company.[116]

Model specification refers to the model including all relevant and excluding all irrelevant variables.[117] A statistical test for the model specification is the Ramsey RESET test, which creates new predictor variables and checks whether any of them are significant.[118] I perform this test using both a power of 2 and 3 on the model without clustered robust standard errors (as using them gave an error) and both were significant, indicating that the model has a specification error. Further considerations of omitted variables will be given in the paragraph discussing endogeneity.

I check for influential values using the Residuals vs Leverage and the Cook's distance plot, Figures A.4 and A.5. Observations are considered influential if they have large residuals and deviate far from the mean, with the most common measurement for influential values being the Cook's distance.[119] A Cook's distance greater than 4/n with n being the number of observations is considered especially large, in this model the critical distance is $3.495*10^{-5}$.[120] Many observations exceed that value in the model. Considering the three observations having the highest Cook's distances, their values are not extremely different from the mean and I cannot find any reason for these observations to be abnormal. I therefore keep all observations in the sample, but the generally large Cook's distance is a minor constraining factor.

I check for collinearity by calculating the variance inflation factors (VIFs) for the model. Looking at the VIFs, no variable has a value larger than 10, indicating that collinearity is not an issue in this model.[121]

Concerning the randomness of the sample, it is sufficient to assume "that the error term is independent of the sample selection mechanism conditional on the covariates."[122] As the sample is given, that assumption is not to be tested.

No real check for endogeneity exists, but three main reasons for this phenomenon can be considered: Omitted variables, simultaneity and measurement errors.[123] The model specification test implies that some important variables might be omitted. However, when carefully checking the literature, I could not find important variables which I do not include in the model. Usage of fixed effects also takes care of some

[114] See University of California, Los Angeles (2021).
[115] See University of California, Los Angeles (2021).
[116] See University of California, Los Angeles (2021).
[117] See University of California, Los Angeles (2021).
[118] See University of California, Los Angeles (2021).
[119] See University of California, Los Angeles (2021).
[120] See University of California, Los Angeles (2021).
[121] See University of California, Los Angeles (2021).
[122] Roberts and Whited (2013), p. 497.
[123] See Roberts and Whited (2013), pp. 498-501.

factors otherwise omitted. Furthermore, when looking at the regression results (Table 2), the effect of SI remains significant when adding control variables and fixed effects, reducing concerns of omitted variable bias. Omitted variable bias therefore possibly exists, but I am probably not missing any important variables commonly used in this research area.

To estimate whether simultaneity is an issue in the model, I lag SI ten working days into the future and look at the regression results. Comparing the coefficients of SI, a change from 0.009 to -0.003 can be noted, that being a reduction in magnitude and a change in sign. Both coefficients are significant at 1% and the R^2 slightly decreases. These results show that simultaneity should not be a large issue, as the previously seen effect of SI on the adjusted returns vanishes.

To minimize the possibility of measurement errors occurring, I use variables carefully selected and successfully used in previous research. Only Attention is significantly altered in comparison to the literature, but not to a great extent and it is not a major variable. I can therefore say that all variables should be good at measuring what they are supposed to measure, but the possibility of measurement errors can never be fully ruled out.

5. Results

In this section, I present, discuss and interpret the results of the analyses. The R code used to generate these results is provided in Attachment 15, using the dataset in Attachment 13 as an input. This code can also generate the descriptive statistics and regression diagnostics tests.

5.1. Hypothesis 1

Hypothesis 1 concerns the relationship between stringency measures and adjusted stock returns, hypothesizing that a positive correlation exists. Table 2 reports the results of the analysis of this hypothesis. Model 1 is the baseline specification, only including SI as a predictor variable. In model 2, COVID-19 related control variables are added, as are industry, country and weekday fixed effects. Model 3 is the main model, including pandemic-related and company-related control variables and ESI as a control variable. Model 4 includes company fixed effects, therefore dropping the essential dummy and financial performance variables. Figure 2 illustrates the relationship between SI and the adjusted returns, it visualizes model 1.

The results report a significant, positive correlation between SI and the adjusted returns in all models, thereby providing support for hypothesis 1. This finding is in line with the results of Narayan et al. (2021) and Ding et al. (2021).[124] In the main model (model 3), for an increase of SI by one SD (20.366), the adjusted returns on average rise by 0.1833 percentage points (20.366*0.009), which is 0.0569 standard deviations (0.1833/3.221) of the adjusted returns. This shows that the correlation is positive and significant, but of small magnitude. This is also shown in Figure 2, where the adjusted returns are distributed around zero for all values of the Stringency Index. The generally positive slopes of SI can be explained by the reasoning Ashraf (2020b) provides, concluding that the positive effect of stringency measures on stock returns stems from stringency measures mitigating the negative effects of COVID-19 itself, while the direct effects of stringency measures on the economy are adverse.[125] The operations of companies might therefore be negatively affected by the stringency measures, but their stock returns can show a positive correlation for this reason, as the economic situation would be worse without these measures in the eyes of the market. For that reason, the stringency measures are potentially beneficial for companies.

Interestingly, the slope of SI is positive and significant even when not controlling for COVID-19 cases, that variable having a negative and significant slope in all models. However, the slope of SI is lower in the first model compared to the models where COVID-19 cases are included. This further supports the suggestion made by Ashraf (2020b) that stringency measures mitigate the negative effects of COVID-19.[126]

The slopes of ESI will be subject to investigation in the second hypothesis. Concerning the company financial performance variables, only Size and BookToMarket have significant coefficients at conventional significance levels, even after adjusting for the Fama-French three factor model. This implies that company financial performance, measured by Leverage, CashByAssets and ROA does not have a significant effect on adjusted stock returns in this sample, in contrast to the results of Ding et al. (2021).[127] Size has a negative coefficient and BookToMarket has a positive one, suggesting that the market prefers small companies and companies with a relatively higher book value of equity compared to their market value during the crisis. Attention has a small, but negative and significant correlation with adjusted stock returns where an increase of Attention by one SD (29.079) relates to a decrease of the adjusted returns of 0.1454 percentage points (-0.005*29.079) in the main model. Rising attention to the pandemic is thereby suggested to have a negative impact on stock returns, although not a strong one. Not being significant in any model and with a changing sign, the essential classification does not seem to exert important influence on adjusted stock returns. This is in line with the results of Heyden and Heyden (2021), who also find essential companies not to react differently to policy measures than companies on average.[128]

Special attention should be given to the adjusted cases of COVID-19, that variable being the primary measurement for the influence of the pandemic itself. As stated above, the coefficients for this variable are negative and significant in all models it is included in. For an increase of AdjustedCases by one SD (0.148), the adjusted returns decrease by 0.05831

[124] See Narayan et al. (2021), p. 5; Ding et al. (2021), p. 13.

[125] See Ashraf (2020b), p. 7.

[126] See Ashraf (2020b), p. 7.

[127] See Ding et al. (2021); pp. 14f.

[128] See Heyden and Heyden (2021), p. 3.

Table 2: Effect of stringency measures on adjusted stock returns

This table shows the regression results of how adjusted returns react to measures captured in the Oxford Stringency Index while controlling for economic support measures, company financial performance indicators, attention to COVID-19, COVID-19 cases and whether a company is classified as an essential business. Variable definitions and data sources can be found in Appendix 1. Fixed effects are included when stated as such. Robust standard errors clustered by company are denoted in parentheses. ***, ** and * report statistical significance levels at 1%, 5% and 10%, respectively, using clustered robust standard errors.

	Dependent variable: AdjustedReturn			
Variable	(1)	(2)	(3)	(4)
SI	0.006***	0.009***	0.009***	0.009***
	(0.0004)	(0.001)	(0.001)	(0.001)
ESI			0.004***	0.004***
			(0.001)	(0.001)
Size			-0.024***	
			(0.006)	
Leverage			-0.0001	
			(0.001)	
CashByAssets			0.001	
			(0.001)	
ROA			0.0004	
			(0.001)	
BookToMarket			0.011***	
			(0.001)	
Attention		-0.002***	-0.005***	-0.005***
		(0.0005)	(0.001)	(0.001)
AdjustedCases		-0.391***	-0.394***	-0.406***
		(0.081)	(0.083)	(0.078)
essential		-0.020	0.010	
		(0.024)	(0.026)	
Industry FE	No	Yes	Yes	Yes
Country FE	No	Yes	Yes	Yes
Weekday FE	No	Yes	Yes	Yes
Company FE	No	No	No	Yes
Observations	117,936	117,936	114,440	117,936
Adj. R squared	0.00153	0.00321	0.00386	-0.00044
F Statistic	215.91***	62.459***	57.485***	34.833***
	(df = 1)	(df = 30)	(df = 36)	(df = 8)

percentage points (-0.394*0.148) in the main model, on average. It has to be kept in mind that this variable is logarithmic, altering the slope of the coefficients in comparison to other variables and making the magnitude of its impact more difficult to assess. Such a negative correlation is in line with the results of Ding et al. (2021) and Ashraf (2020b).[129] This implies that the pandemic itself has a negative effect on companies and that the market takes this negative effect into consideration. A possible reasoning behind this is that the economic consequences of the pandemic become worse when more cases are reported, thus leading to decreased stock returns. This could also explain the negative effect of Attention, as attention to the pandemic and its adverse effects can have a negative influence on the stock market as well. The suggested negative impact of the pandemic itself on companies also supports the explanation of the positive slopes of SI, mitigating these negative influences.[130]

A factor that has to be addressed is the low adjusted R^2 of all models, which is at most 0.0386% (in the main model) and is even negative in model 4. This is to be expected, however, as I use risk-adjusted returns, leading to the adjusted R^2 only measuring what percentage of return variance, which was not explained by the Fama-French factors, the model additionally explains. As the Fama-French three factor model explains a large part of the return variance,[131] small adjusted R^2 are not an issue. The models are all valid in explaining a part of the return variance because the F-statistics measuring whether the models explain anything at all are all significant.

[129] See Ding et al. (2021), p. 14; Ashraf (2020b), pp. 6f.

[130] See Ashraf (2020b), p. 7.

[131] See Fama and French (1993), p. 5.

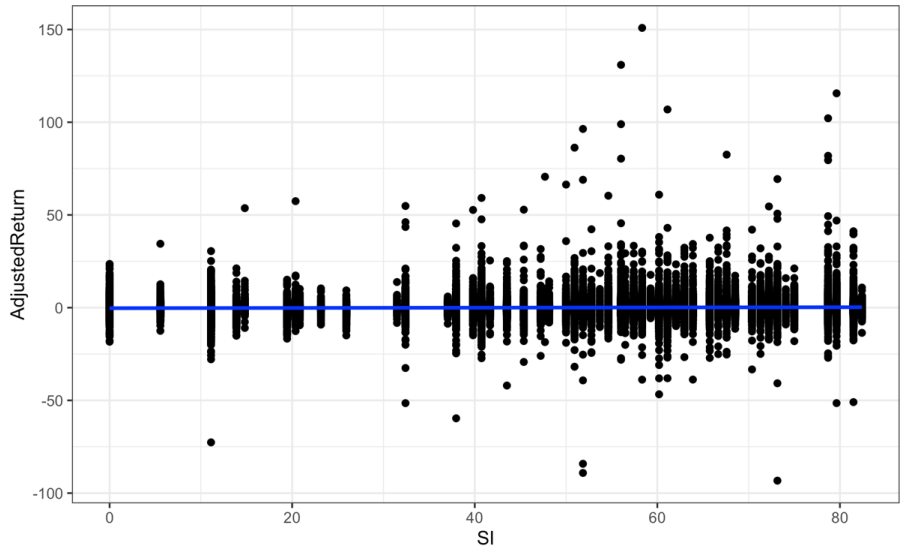

Figure 2: Relation of stringency measures and adjusted stock returns

This figure visualizes the relationship between stringency measures and adjusted stock returns without using any control variables, essentially column (1) of table 2. The blue line represents the regression line of the relationship. The values of the Stringency Index (SI) are displayed on the x-axis, the values of the adjusted returns (AdjustedReturn) are displayed on the y-axis. Variable definitions and data sources can be found in Appendix 1.

These results contradict the findings of various researchers, including Yang and Deng (2021), Chen et al. (2020) and Shanaev et al. (2020).[132] The different results might be due to a different empirical setting, considered timeframe or the sample consisting of different companies from different countries.

5.2. Hypothesis 2

Hypothesis 2 concerns the relationship between economic support measures and adjusted stock returns, hypothesizing that a positive correlation exists. To analyze this hypothesis, I consider the Economic Support Index and E3 in table 3 and the individual support measures E1, E2 and E3 in table 4 in order to gain specific insights into the effects of the individual policy measures. In both tables, control variables and fixed effects are added following the same sequence as in table 2, with the third model always containing most control variables. The only difference is the earlier addition of policy measure control variables, with the Stringency Index, the effect of which already having been investigated, being added in the second model already. As it is possible that ESI and E3 are collinear, I calculate the VIFs for model 3 of table 3 to exclude the possibility of multicollinearity falsifying the results. As all VIFs are below 10, multicollinearity is not an issue.

The results partly support hypothesis 2 by reporting a significant, positive correlation between ESI and the adjusted returns in all models. This is in line with a broad range of literature, including Narayan et al. (2021), Ding et al. (2021), Gormsen and Koijen (2020) and Ashraf (2020b).[133] For an increase of ESI by one SD (30.962), the adjusted returns increase by 0.09289 percentage points (30.962*0.003), on average in the third model. This is again much less than a standard deviation of the adjusted returns (3.221), showing that the effect, although significant, has a rather small magnitude. A possible explanation for this is presented by Ashraf (2020b), describing that the Economic Support Index measures support given to households and not to businesses, which results in stock market reactions to the measures not being very strong.[134] Still, a positive and significant correlation is observed, implying that economic support measures can indeed be beneficial for companies.

The situation looks differently for E3. The coefficient is negative and statistically significant in all models, with an increase of E3 by one SD (0.004) being related to a decrease of -0.03842 percentage points (-9.607*0.004) of the adjusted returns in the third model, on average. This finding contradicts the previous results showing a positive correlation between economic support measures and adjusted returns. However, these adverse effects of economic support measures on stock returns are also shown in the literature, where Zhang et al. (2020) and Shanaev et al. (2020) arrive at similar results.[135] Heyden and Heyden (2021) find a similar ambivalent relationship, reaching the conclusion that fiscal policies add uncertainty whereas monetary policies calm markets.[136] Another factor possibly explaining these results are the differ-

[132] See Yang and Deng (2021), p. 4; Chen et al. (2020), pp. 3f; Shanaev et al. (2020), p. 42.

[133] See Narayan et al. (2021), p. 5; Ding et al. (2021), p. 13; Gormsen and Koijen (2020), p. 574; Ashraf (2020b), p. 7.

[134] See Ashraf (2020b), p. 5.

[135] See Zhang et al. (2020), p. 5; Shanaev et al. (2020), pp. 42f.

[136] See Heyden and Heyden (2021), pp. 3f.

Table 3: Effect of economic support measures on adjusted stock returns

This table shows the regression results of how adjusted returns react to economic support measures captured in the Economic Support Index and E3 of the Oxford Indices while controlling for stringency measures, company financial performance indicators, attention to COVID-19, COVID-19 cases and whether the company is classified as an essential business. Variable definitions and data sources can be found in Appendix 1. Fixed effects are included when stated as such. Robust standard errors clustered by company are denoted in parentheses. ***, ** and * report statistical significance levels at 1%, 5% and 10%, respectively, using clustered robust standard errors.

	Dependent variable: AdjustedReturn			
Variable	(1)	(2)	(3)	(4)
ESI	0.002***	0.003***	0.003***	0.003***
	(0.0003)	(0.001)	(0.001)	(0.001)
E3	-9.644***	-9.826***	-9.607***	-9.594***
	(3.312)	(3.367)	(3.451)	(3.363)
SI		0.009***	0.009***	0.009***
		(0.001)	(0.001)	(0.001)
Size			-0.024***	
			(0.006)	
Leverage			-0.0001	
			(0.001)	
CashByAssets			0.001	
			(0.001)	
ROA			0.0004	
			(0.001)	
BookToMarket			0.011***	
			(0.001)	
Attention		-0.004***	-0.005***	-0.005***
		(0.001)	(0.001)	(0.001)
AdjustedCases		-0.383***	-0.392***	-0.404***
		(0.081)	(0.083)	(0.078)
essential		-0.020	0.010	
		(0.024)	(0.026)	
Industry FE	No	Yes	Yes	Yes
Country FE	No	Yes	Yes	Yes
Weekday FE	No	Yes	Yes	Yes
Company FE	No	No	No	Yes
Observations	117,936	117,936	114,440	117,936
Adj. R squared	0.0005	0.0036	0.00398	-0.00031
F Statistic	29.477***	59.374***	56.512***	31.42***
	(df = 2)	(df = 32)	(df = 37)	(df = 9)

ences between ESI and E3. Whereas ESI measures policies of income and debt relief active over a long period of time, E3 measures additional monetary support on individual days.[137] It is therefore possible that the stock market reaction to these interventions differs. Furthermore, the effect has very small magnitude, even smaller than SI or ESI. Still, the negative and significant correlation implies that not all economic support measures are favored by the market.

Even when stock market reactions to economic support measures are negative, this does not automatically mean that these measures are negative for companies. As Zhang et al. (2020) and Heyden and Heyden (2021) point out, the major reason behind negative stock market reactions to economic support measures is that these measures add uncertainty to the market.[138] Even when the support measures are beneficial for companies, the increased uncertainty can lead to negative correlations between support measures and stock returns.

Concerning the other variables, no real change takes place in comparison to table 2. Similarly, the R^2 are generally small, which is not a problem as long as the F-statistics are significant, which is the case for all models.

The results of this analysis provide evidence that economic support measures generally are beneficial for compa-

[137] See Hale et al. (2021) pp. 22-27.

[138] See Zhang et al. (2020), p. 5; Heyden and Heyden (2021), pp. 3f.

nies and that the market perceives these measures positively. As Zhang et al. (2020) point out, such measures are also necessary to calm stock markets.[139] However, certain policy measures add uncertainty to the market and are therefore possibly not perceived well by investors.

In the following table, the results of regressions using E1 and E2 instead of ESI are presented. Otherwise, the models are exactly the same as in table 3. E1 measures income support given to households negatively affected by the pandemic and E2 measures debt and contract relief for households.[140] As it is possible that E1, E2 and E3 are collinear, I calculate the VIFs for model 3 to exclude the possibility of multicollinearity falsifying the results. As all VIFs are below 10, multicollinearity is not an issue.

The results offer contradicting insights into the relationship between economic support measures and adjusted stock returns. Whereas ESI has a positive and significant coefficient in all models, as shown in table 3, the situation is not so clear for its components E1 and E2. E1 has a positive and significant correlation with the adjusted returns in all models, with an increase of E1 by one SD (1.749) being related to an increase of the adjusted returns by 0.6786 percentage points (0.388*1.749) in model 3. Compared to coefficients of other policy indices, this is relatively high, suggesting that income support can be a policy measure beneficial for companies. For E2, only the coefficient in the first model is significant, showing a negative slope. Addition of control variables and fixed effects therefore takes explanatory power from this variable. In the first model, for an increase of E2 by one SD (0.899), adjusted returns decrease by -0.03506 percentage points (-0.039*0.899). The negative sign of this coefficient indicates that an adverse effect of debt and contract relief on companies might exist. The coefficients and significance levels of E3 are largely unchanged in comparison to table 3.

These results are interesting, as only income support is positively correlated with adjusted returns, but debt relief and other fiscal measures are not. A possible explanation could be the different nature of the policy measures captured by E1 and E2. Whereas income support measures aim at replacing the income of households who lost their income due to the crisis, debt and contract reliefs aim at freezing financial obligations for households, for example by banning evictions or stopping loan repayments.[141] Income support can therefore be beneficial both to households and companies, mitigating the loss of purchasing power of consumers and perhaps even helping companies with paying salaries. Debt and contract reliefs, on the other hand, can have adverse effects on companies having a business model linked to these contracts, for example real estate companies which receive less rent payments when such policy measures are active.

A look on the adjusted R^2 reveals that these values are higher than the corresponding adjusted R^2 in table 3. All adjusted R^2 are still small but using the individual indices E1 and E2 leads to the models explaining more variance in the adjusted returns. This is most likely due to ESI capturing the positive correlation of E1 and the negative or insignificant correlation of E2 in one index, resulting in a lower overall correlation compared to the individual indices.

The other variables have not changed dramatically compared to table 3.

Generally, economic support measures do not have a clear effect on adjusted stock returns. While the correlation as a whole, using the Economic Support Index, is positive and significant, some individual measures have negative or insignificant correlations. This is in line with the results of Heyden and Heyden (2021), finding ambivalent relationships between economic support measures and stock returns.[142] Income support measures show the strongest positive correlation with adjusted stock returns, providing evidence that this policy measure can be especially beneficial.

5.3. Hypothesis 3

Hypothesis 3 investigates the relationship between stringency measures and adjusted stock returns for specific sets of industries regarded as being severely or positively affected by the crisis, where H3a investigates the severely affected industries and H3b investigates the positively affected industries. A list of industries regarded as being severely or positively affected in the literature and which industries I classify as being severely or positively affected can be found in Appendix 2.

Table 5 reports the results of the analysis of hypothesis 3a. The models are similar to the ones previously applied, using the same variables in the same order as in table 2. The only difference is the absence of industry fixed effects, as their inclusion does not make sense when considering a subset of specific industries.

The results of this analysis contradict hypothesis 3a, assuming a negative correlation between stringency measures and adjusted stock returns for companies belonging to severely affected industries. In all models, the coefficients for SI are positive and significant, suggesting that, although the companies belong to severely affected industries, their adjusted stock returns are not severely affected by the stringency measures. Comparing these results to the results presented in table 2, using the entire sample, the coefficients and significance levels for SI are almost the same, with only its coefficient in the fourth model being slightly smaller in table 5 (by 0.001). This means that the correlation between SI and the adjusted returns is close to identical for companies on average and companies belonging to severely affected industries.

While this sounds counterintuitive, it can be explained by the reasoning Ashraf (2020b) formulates, stating that the positive effect of stringency measures on stock returns stems from stringency measures mitigating the negative effects of COVID-19 itself, while the direct effects of stringency measures on the economy are adverse.[143] If the market perceives

[139] See Zhang et al. (2020), p. 5.
[140] See Hale et al. (2021), p. 22.
[141] See Hale et al. (2021), p. 22.
[142] See Heyden and Heyden (2021), pp. 3f.
[143] See Ashraf (2020b), p. 7.

Table 4: Further analyses for the effect of economic support measures on adjusted returns

This table shows the regression results of analyses on the effect of individual economic support measures, using E1 and E2 instead of ESI, on adjusted returns. Variable definitions and data sources can be found in Appendix 1. Fixed effects are included when stated as such. Robust standard errors clustered by company are denoted in parentheses. ***, ** and * report statistical significance levels at 1%, 5% and 10%, respectively, using clustered robust standard errors.

	Dependent variable: AdjustedReturn			
	(1)	(2)	(3)	(4)
E1	0.199***	0.400***	0.388***	0.397***
	(0.016)	(0.069)	(0.071)	(0.070)
E2	-0.039***	0.013	0.016	0.018
	(0.011)	(0.016)	(0.017)	(0.016)
E3	-10.513***	-9.709***	-9.505***	-9.478***
	(3.303)	(3.321)	(3.406)	(3.316)
SI		0.008***	0.009***	0.008***
		(0.001)	(0.001)	(0.001)
Size			-0.024***	
			(0.006)	
Leverage			-0.0001	
			(0.001)	
CashByAssets			0.001	
			(0.001)	
ROA			0.0003	
			(0.001)	
BookToMarket			0.011***	
			(0.001)	
Attention		-0.010***	-0.010***	-0.010***
		(0.001)	(0.001)	(0.001)
AdjustedCases		-0.416***	-0.423***	-0.435***
		(0.081)	(0.084)	(0.079)
essential		-0.021	0.010	
		(0.024)	(0.026)	
Industry FE	No	Yes	Yes	Yes
Country FE	No	Yes	Yes	Yes
Weekday FE	No	Yes	Yes	Yes
Company FE	No	No	No	Yes
Observations	117,936	117,936	114,440	117,936
Adj. R squared	0.00156	0.00428	0.00459	0.00034
F Statistic	52.907***	58.185***	55.936***	28.926***
	(df = 3)	(df = 33)	(df = 38)	(df = 10)

these mitigating effects to be more important than the economically adverse effects, SI can be positively correlated with the adjusted stock returns. This is especially interesting for companies belonging to industries severely affected by the crisis, where such a perception would not be expected to the same extent as for companies on average. The findings, however, show that this can be the case.

Concerning economic support measures, similar results can be observed. In both models containing ESI as a variable, the coefficient of ESI is significant and positive and has exactly the same value as in table 2. In the third model, for an increase of ESI by one SD (30.962), adjusted returns increase by 0.1238 percentage points (30.962*0.004) on average. This suggests that the effect of economic support measures on adjusted stock returns is very similar for companies on average and companies belonging to sectors severely affected by the pandemic. The results for SI and ESI in this analysis generally imply that the effects of policy measures are beneficial even on companies severely affected by the pandemic.

The control variables offer more diversity. The first observation here is a loss in significance for Size and BookToMarket. This means that no company financial performance indicator has significant explanatory power any more for this subset of companies. These variables are therefore unfit to estimate adjusted stock returns for companies belonging to

Table 5: Effect of stringency measures on adjusted stock returns for companies belonging to sectors severely affected by the pandemic

This table shows the regression results of how adjusted returns of companies belonging to sectors regarded as being severely affected by the pandemic react to measures captured in the Oxford Stringency Index while controlling for economic support measures, company financial performance indicators, attention to COVID-19, COVID-19 cases and whether the company is classified as an essential business. Variable definitions and data sources can be found in Appendix 1. Which industries are classified as being severely affected can be seen in Table A.5. Fixed effects are included when stated as such. Robust standard errors clustered by company are denoted in parentheses. ***, ** and * report statistical significance levels at 1%, 5% and 10%, respectively, using clustered robust standard errors.

	Dependent variable: AdjustedReturn			
Variable	(1)	(2)	(3)	(4)
SI	0.006***	0.009***	0.009***	0.008***
	(0.001)	(0.002)	(0.002)	(0.002)
ESI			0.004***	0.004***
			(0.001)	(0.001)
Size			-0.013	
			(0.010)	
Leverage			0.001	
			(0.001)	
CashByAssets			0.002	
			(0.001)	
ROA			-0.0002	
			(0.003)	
BookToMarket			-0.001	
			(0.024)	
Attention		-0.002*	-0.005***	-0.005***
		(0.001)	(0.002)	(0.002)
AdjustedCases		-0.427***	-0.420***	-0.433***
		(0.130)	(0.128)	(0.128)
essential		-0.025	0.008	
		(0.026)	(0.033)	
Country FE	No	Yes	Yes	Yes
Weekday FE	No	Yes	Yes	Yes
Company FE	No	No	No	Yes
Observations	35,968	35,968	35,968	35,968
Adj. R squared	0.00146	0.00349	0.00399	-0.00007
F Statistic	59.549***	8.6739***	6.4718***	11.069***
	(df = 1)	(df = 11)	(df = 17)	(df = 8)

sectors severely affected by the pandemic. In model 2, the significance of Attention is slightly decreased compared to table 2, but the coefficient is not changed. Interestingly, the slope of COVID-19 cases is significant and of greater negative magnitude in table 5 than in table 2 for all models, implying that companies belonging to sectors regarded as being severely affected by the pandemic really are affected more severely by the pandemic than companies on average.

Table 6 reports the results of the analysis of hypothesis 3b. The models use the same variables and fixed effects as in table 5 in the same order.

In contrast to the previous analysis, the result of this analysis supports hypothesis 3b. Looking at SI, all models have a significant and positive slope. The magnitude of these slopes is very similar to the magnitudes in the previous findings, but a bit higher in model 2,3 and 4 of table 5 compared to table 2. This suggests that the mitigating effect of the stringency measures on adjusted stock returns is similar or slightly increased for companies belonging to sectors positively affected by the pandemic in comparison to companies on average.

For ESI, all coefficients are the same as in table 2, but the significance is slightly decreased. As both coefficients are still significant at 5%, this is not a problem. The coefficients for ESI suggest that the influence of economic support measures on adjusted stock returns is similar for companies belonging to sectors positively affected by the pandemic and companies on average.

Considering the control variables, Size and BookToMarket and, for the first time, ROA are significant at 5%. While the coefficients of Size and BookToMarket again suggest that smaller companies and companies with more book equity in comparison to market equity are preferred by investors dur-

ing the crisis, the significant and positive slope of ROA implies that investors also prefer more profitable companies in these industries. This is in line with the findings of Ding et al. (2021), who also find companies with larger profits to be more resilient to the crisis, although they come to this conclusion investigating companies from all industries.[144] The other company financial performance indicators are again not significant at 10%, making Leverage and CashByAssets insignificant in every analysis. All coefficients for attention are significant at 1% and slightly more negative than in table 2. However, no real economic reason for this very small effect exists and it might be due to Attention only having meaningful values at the beginning of the pandemic, when stock market volatility was especially high.[145]

The COVID-19 cases should be given special attention, as they are not significant at 10% in any model they are included in. This is a major contrast to all other analyses, where COVID-19 cases are always significant at 1%. However, this does not come unexpectedly when analyzing companies belonging to sectors regarded as being positively affected by the pandemic, implying that the analyzed companies really are less affected by the pandemic than companies on average. When looking at the coefficients, although they are not significant, all slopes are less negative than in table 2, further supporting the conclusion that these companies are less affected by the pandemic.

In general, this analysis provides support for hypothesis 3b. It suggests that the companies belonging to sectors regarded as positively affected by the pandemic really are affected less by the pandemic. This analysis further implies that, for these companies, economic support measures are beneficial and stringency measures are mitigating the effects of COVID-19. As the effect of the pandemic itself is assumed to be small, at least for some companies, the stringency measures might be beneficial for certain firms, for example supporting the business model of IT corporations. These results are generally to be expected for companies belonging to positively affected industries. As the coefficients for SI are partly higher than in table 2, the positive or mitigating effects of stringency measures are possibly increased for companies belonging to positively affected industries. It is noteworthy that the slopes of ESI are the same in table 2, 5 and 6, suggesting that all companies profit similarly from economic support measures.

6. Discussion

In this section, I present limitations of my used methods and data and critically discuss my methodology and results.

6.1. Limitations

Several biases and factors left out of consideration can limit the validity of my results. For a start, I cannot rule out other factors having influenced stock returns during the observation period. Events like the ongoing Brexit, the US presidential elections, or drastic developments in the oil market occurred during the observation period and potentially had an influence on stock returns, as pointed out by Ramelli and Wagner (2020).[146] As these events could have affected certain sectors or companies especially, they might exert unseen influence on the results.

Another factor possibly influencing the findings is the choice of companies in this sample. Only companies listed on stock markets are considered, leaving many companies, especially small and medium-sized firms, out of the analyses.[147] Many restaurants, barbers or other small enterprises are not considered, although they might be affected by the policy measures. Therefore, the sample does not exactly represent the economies of the countries. However, as most of the large companies are dependent on smaller companies through various links[148] and the listed companies represent a large portion of the economies, this restriction does not overly affect the validity of the results.

Taking companies with their headquarters being located in a certain country as a sample incurs another limitation, as the operations of a company might not be conducted at the place of the headquarter. The sites of operations could be in other countries and can have varying influences on the company as a whole. In extreme cases, all operations might be conducted abroad with only the headquarters being located in a country. Different pandemic developments and policy measures in other countries can therefore be an unconsidered factor influencing the stock returns of these companies. Considering foreign sales in the robustness checks addresses the international orientation of a company in general and does not cover this issue.

Apart from the stringency measures aimed at slowing the spread of COVID-19, some companies and individuals adopt voluntary social distancing measures, reducing their economic activities. As Baker et al. (2020) conclude, these measures have an important effect on the economy, especially on service-oriented companies.[149] The only approximation for this effect in my models is H1 public information campaigns, a constituent of the Stringency Index, which measures information campaigns raising awareness of the pandemic.[150] These campaigns potentially result in more voluntary social distancing.

Checking the regression assumptions in section 4.5 also reveals limitations of the used models: The residuals are not normally distributed, leading to the t- and F-tests not being fully valid.[151] Large Cook's distance values suggest that some influential observations exist.[152] Endogeneity can also not be ruled out.

[144] See Ding et al. (2021), pp. 14f.
[145] See Baker et al. (2020), p. 743.
[146] See Ramelli and Wagner (2020), pp. 630f.
[147] See Lin and Halk (2021), p. 15.
[148] See Lin and Halk (2021), p. 15.
[149] See Baker et al. (2020), p. 756.
[150] See Hale et al. (2021), pp. 23-27.
[151] See University of California, Los Angeles (2021).
[152] See University of California, Los Angeles (2021).

Table 6: Effect of stringency measures on adjusted stock returns for companies belonging to sectors positively affected by the pandemic

This table shows the regression results of how adjusted returns of companies belonging to sectors regarded as being positively affected by the pandemic react to measures captured in the Oxford Stringency Index while controlling for economic support measures, company financial performance indicators, attention to COVID-19, COVID-19 cases and whether the company is classified as an essential business. Variable definitions and data sources can be found in Appendix 1. Which industries are classified as being positively affected can be seen in Table A.6. Fixed effects are included when stated as such. Robust standard errors clustered by company are denoted in parentheses. ***, ** and * report statistical significance levels at 1%, 5% and 10%, respectively, using clustered robust standard errors.

	Dependent variable: AdjustedReturn			
Variable	(1)	(2)	(3)	(4)
SI	0.006***	0.010***	0.010***	0.011***
	(0.001)	(0.002)	(0.002)	(0.002)
ESI			0.004**	0.004**
			(0.002)	(0.002)
Size			-0.037**	
			(0.015)	
Leverage			0.001	
			(0.001)	
CashByAssets			0.001	
			(0.001)	
ROA			0.004**	
			(0.002)	
BookToMarket			0.128**	
			(0.065)	
Attention		-0.003***	-0.006***	-0.006***
		(0.001)	(0.002)	(0.002)
AdjustedCases		-0.246	-0.219	-0.239
		(0.202)	(0.200)	(0.200)
essential		-0.100**	-0.054	
		(0.046)	(0.046)	
Country FE	No	Yes	Yes	Yes
Weekday FE	No	Yes	Yes	Yes
Company FE	No	No	No	Yes
Observations	16,504	16,504	16,504	16,504
Adj. R squared	0.00152	0.0036	0.00472	-0.00021
F Statistic	35.17***	6.1055***	8.169***	6.7943***
	(df = 1)	(df = 11)	(df = 17)	(df = 8)

Using Fama-French three factor model-adjusted returns also incurs some limitations. As the adjusted returns are excess returns over the predicted returns, including the market excess return times market beta,[153] movements of the entire market can hardly be observed. An influence affecting the entire European market is therefore less visible than an influence affecting only certain industries. Furthermore, as Schmidt et al. (2019) point out, the Fama-French three factor model leaves out some influences on stock returns, for example momentum.[154] These influences can also explain parts of the variance of stock returns but are too complicated to be included here.

6.2. Critical discussion of methodology and results

The limitations section above shows some boundaries to the validity of my results. To assess further constraining factors, I perform several robustness checks on the main model (the formula is given in section 3.2). Applying robustness checks on other models, being variations of this one or applied to a subset of the data, is not necessary. I check the robustness against four influences and changes: First, I replace testing-adjusted cases with testing-unadjusted cases to assess how they change the results, similar to Ding et al. (2021).[155] Second, I include foreign sales to estimate the influence they exert on the results. Third, I exclude data from April 2020 and before in order to estimate whether only this period

[153] See Fama and French (1993), p. 5.
[154] See Schmidt et al. (2019), p. 214.
[155] See Ding et al. (2021), p. 23.

causes the results, as the market volatility was especially high at that time.[156] Last, I exclude financial companies to check that they are not the main drivers of the results, as they have a different financial structure and are often excluded in financial research.[157] None of the robustness checks changes the main findings, a positive and significant correlation of stringency and economic support measures with adjusted stock returns. The only notable discovery is the coefficient of AdjustedCases changing its sign when excluding April 2020 and before. The negative correlation between COVID-19 cases and adjusted stock returns can therefore only be observed during the early pandemic, implying that coefficients of AdjustedCases should be handled with care. As a negative influence of the pandemic itself on stock returns is found by many researchers,[158] the validity of the main results is not threatened. Detailed results of the robustness checks can be found in Appendix 4.

To mitigate the effects of unknown factors influencing stock returns during the observation period, I include fixed effects in the regressions, controlling for parameters constant within each industry, country, weekday and, in some models, company. Furthermore, considering a long timeframe of over a year of daily observations should reduce the influence of other events on stock returns. Still, these influences can never really be excluded using a panel data structure. Despite having some limitations, panel data analyzed using a similar approach to the one I applied leads to valid results for many researchers investigating stock market reactions to COVID-19.[159] As most of the variables I use are also used successfully by other researchers in the same or a similar way, their effectiveness is empirically tested. Therefore, the methodology has some limitations, but should generally deliver valid results.

When drawing conclusions from the results, the real-world-significance of findings is a major issue to address. Most importantly, correlations do imply causal relationships and statistical significance is no guarantee for real-world-significance. A potential problem in this regard is reverse causality, meaning that not the independent variables explain the dependent variable, but vice versa. From an economic perspective, it seems very unlikely that stock returns of companies affect cases of COVID-19, the classification of essential companies or the financial performance of a company of the previous year. Similarly, stringency measures are probably not influenced by stock returns on the same day and lagging SI ten days into the future as a simultaneity check reveals no signs of a reversely causal relationship. Attention to the pandemic could possibly be caused by attention to volatile stock market movements, but Engelhardt et al. (2020) find news attention having a negative effect on stock returns and not vice versa.[160] Although economic support measures are introduced to counter adverse economic effects of the pandemic,[161] they are hardly affected by stock returns on the same day. Reverse causality is therefore likely to be no issue for all used variables.

I further validate my results by checking them with the literature, as can be seen in chapter 5. All major findings are in line with the results of other researchers and no result contradicts economic reasoning or cannot be explained convincingly. Therefore, my findings should provide sufficient real-world-significance to draw conclusions. Arriving at causal conclusions is still hardly possible in quantitative research and I only make causal statements when they are backed by the literature or in line with previous conclusions.

The last step when investigating the impact of policy measures on companies is inferring from the impact on stock returns to the impact on companies in general. Stock returns do not measure the current performance of companies, but the expectations of the market regarding their future performance, as pointed out by Ramelli and Wagner (2020).[162] Stock returns should therefore be treated with care when drawing conclusions about short-term economic consequences of policy measures but can generally be used as a valid approximation for their economic impacts, especially their future influences. Furthermore, using stock returns when investigating economic consequences of COVID-19 is a common approach in the literature that leads to valid results.[163]

7. Conclusion

In this thesis, I investigate the impact of COVID-19 policy measures on companies in Belgium, The Netherlands, Denmark and Norway. Using a panel data structure, I utilize daily stock returns adjusted by the Fama-French three factor model and the indices developed by the University of Oxford to examine this impact.[164] Controlling for testing-adjusted growth of COVID-19 cases, attention to the pandemic, company financial performance indicators and whether a company is classified as being essential or not, I find both stringency and economic support measures to have a small, positive impact on adjusted stock returns.

While the policy measures have a generally positive influence on stock returns, growth in testing-adjusted cases and attention to the pandemic have a generally negative one in the considered countries, especially during the early phase of the pandemic. The positive impact of stringency measures can be explained by these measures mitigating the economically negative influences of COVID-19 itself, as reasoned by

[156] See Baker et al. (2020), p. 743.
[157] See Ramelli and Wagner (2020), pp. 631f.
[158] See Al-Awadhi et al. (2020), pp. 3f; Heyden and Heyden (2021), pp. 3f; Ashraf (2020c), pp. 4-6; See Ding et al. (2021), pp. 13f.
[159] See Ding et al. (2021), pp. 12-25; Ashraf (2020b), pp. 5-7; Chen et al. (2020), pp. 4-6; Al-Awadhi et al. (2020), pp. 2-4.
[160] See Engelhardt et al. (2020), p. 10.
[161] See Ramelli and Wagner (2020), pp. 650f; Heyden and Heyden (2021), p. 1.
[162] See Ramelli and Wagner (2020), p. 623.
[163] See Ashraf (2020b), p. 1f; Al-Awadhi et al. (2020), pp. 1f; Ding et al. (2021), pp. 2f; Chen et al. (2020), p. 2.
[164] See Hale et al. (2021), pp. 20-27.

Ashraf (2020b).[165] Although these measures have adverse economic effects, the market perceives them to be overall beneficial, as the pandemic would be worse otherwise. The robustness checks, such as excluding financial companies and using testing-unadjusted COVID-19 cases do not limit the validity of the main findings.

When looking specifically at economic support measures, I find income support for households to have a strong positive impact on adjusted stock returns. Debt and contract relief have no impact when controlling for the pandemic and other fiscal measures have a negative influence on adjusted stock returns. The negative effect of fiscal measures can be explained by the uncertainty they add to the market, as pointed out by Zhang et al. (2020) and Heyden and Heyden (2021).[166] When using the Economic Support Index, support measures have a small, positive influence on adjusted stock returns. This leads to the conclusion that economic support measures are generally beneficial for companies but can also increase uncertainty in the market.

Considering companies belonging to severely affected sectors, I find these companies to be more negatively affected by the pandemic itself but find both stringency measures and economic support measures to have a positive influence on adjusted stock returns. The magnitude of these effects is similar for these companies and companies on average. Interestingly, stringency measures therefore have a positive influence even on severely affected companies, as do economic support measures. When investigating companies belonging to positively affected sectors, I also find stringency and economic support measures positively affecting adjusted stock returns, while the pandemic itself has no significant impact on these companies. Stringency measures have a more positive influence on positively affected companies than on companies on average while economic support measures have a positive effect of similar magnitude.

Many questions regarding the influence of COVID-19 policy measures on companies are still unanswered, leaving room for future research. For a start, different phases of the pandemic could be investigated, as I only look at the entire timeframe of over a year in my analyses. Furthermore, looking at different stringency measures and their effect on companies could be interesting, as could be a more detailed investigation of different sectors. I mainly consider absolute values of indices, but the influence of changes of indices could be interesting as well. Finally, the anticipation of policy measures and the effects of anticipated policy measures might reveal noteworthy insights.

[165] See Ashraf (2020b), p. 7.
[166] See Zhang et al. (2020), p. 5; Heyden and Heyden (2021), pp. 3f.

References

Al-Awadhi, A. M., Alsaifi, K., Al-Awadhi, A., & Alhammadi, S. (2020). Death and contagious infectious diseases: Impact of the COVID-19 virus on stock market returns. *Journal of Behavioral and Experimental Finance*, *27*(100326).

Alfaro, L., Chari, A., Greenland, A. N., & Schott, P. (2020). Aggregate and firm-level stock returns during pandemics, in real time. *NBER Working Paper Series, Working Paper 26950*.

Ashraf, B. N. (2020a). Economic impact of government interventions during the COVID-19 pandemic: International evidence from financial markets. *Journal of Behavioral and Experimental Finance*, *27*(100371).

Ashraf, B. N. (2020b). Stock markets' reaction to COVID-19: Cases or fatalities? *Research in International Business and Finance*, *54*(101249).

Ashraf, B. N. (2020c). Stock markets' reaction to COVID-19: Moderating role of national culture. *Finance Research Letters*.

Baker, N., S. R. / Bloom, Davis, S. J., Kost, K., Sammon, M., & Viratyosin, T. (2020). The Unprecedented Stock Market Reaction to COVID-19. *The Review of Asset Pricing Studies*, *10*(4), 742–758.

Blavatnik School of Government, University of Oxford. (2021). *Oxford COVID-19 Government Response Tracker.* Retrieved from https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker

Cepoi, C. O. (2020). Asymmetric dependence between stock market returns and news during COVID-19 financial turmoil. *Finance Research Letters*, *36*(101658).

Chen, M. H., Demir, E., García-Gómez, C. D., & Zaremba, A. (2020). The impact of COVID-19 policy responses on U.S. travel and leisure companies. In *Annals of Tourism Research Empirical Insights* (p. 100003).

Ding, W., Levine, R., Lin, C., & Xie, W. (2021). Corporate immunity to the COVID-19 pandemic. *Journal of Financial Economics*.

Engelhardt, N., Krause, M., Neukirchen, D., & Posch, P. (2020). What drives Stocks during the Corona-Crash? News Attention vs. Rational Expectation. *Sustainability*, *12*(12), 5014.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Foundation for Innovative New Diagnostics. (2021). *SARS-COV-2 Test Tracker.* Retrieved from https://www.finddx.org/covid-19/test-tracker/

French, K. R. (2021). *European 3 Factors.* Retrieved from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Google LLC. (2021). *Search Volume Index for "corona" for Belgium, The Netherlands, Denmark and Norway.* Retrieved from https://trends.google.com/trends/?geo=DE

Gormsen, N. J., & Koijen, R. S. J. (2020). Coronavirus: Impact on Stock Prices and Growth Expectations. *The Review of Asset Pricing Studies*, *10*(4), 574–597.

Hale, T., Anania, J., Angrist, N., Boby, T., Cameron-Blake, E., Ellen, L., … Zhang, Y. (2021). Variation in government responses to COVID-19. *Blavatnik School of Government Working Paper Series, Version 11.0*.

Hanauer, M. X., & Windmüller, S. (2020). *Enhanced Momentum Strategies.* Retrieved from https://ssrn.com/abstract=3437919

He, P., Sun, Y., Zhang, Y., & Li, T. (2020). COVID-19's Impact on Stock Prices Across Different Sectors – An Event Study Based on the Chinese Stock Market. *Emerging Markets Finance and Trade*, *56*(10), 2198–2212.

Heyden, K. J., & Heyden, T. (2021). Market reactions to the arrival and containment of COVID-19: An event study. *Finance Research Letters*, *38*(101745).

Hoekman, L. M., Smits, M. M. V., & Koolman, X. (2020). The Dutch COVID-19 approach: Regional differences in a small country. *Health Policy and Technology*, *9*(4), 613–622.

Johns Hopkins University and Medicine. (2021). *COVID-19 Dashboard by the Center for Systems Science and Engineering (CCSE) at Johns Hopkins University (JHU).* Retrieved from https://coronavirus.jhu.edu/map.html

Lin, X., & Halk, M. T. (2021). Nordic Stock market performance of the travel and leisure industry during the first wave of the COVID-19 pandemic. *Tourism Economics*.

Narayan, P. K., Phan, D. H. B., & Liu, G. (2021). COVID-19 lockdowns, stimulus packages, travel bans, and stock returns. *Finance Research Letters*, *38*(101732).

Orbis. (2021). *Bureau van Dijk Electronic Publishing Ltd. (2021): Orbis. Database.*

Ramelli, S., & Wagner, A. F. (2020). Feverish Stock Price Reactions to COVID-19. *The Review of Corporate Finance Studies*, *9*(3), 622–655.

Refinitiv Datastream. (2021). *Refinitiv Inc., an LSEG business (2021): Datastream Macroeconomic Analysis. Database.*

Refinitiv Worldscope. (2021). *Refinitiv Inc., an LSEG business (2021): Worldscope Fundamentals. Database.*

Roberts, M. R., & Whited, T. M. (2013). Endogeneity in Empirical Corporate Finance. *Handbook of the Economics of Finance*, *2*(A), 493–572.

Schmidt, P. S., von Arx, U., Schrimpf, A., Wagner, A. F., & Ziegler, A. (2019). Common risk factors in international stock markets. *Financial Markets and Portfolio Management*, *33*(3), 213–241.

Shanaev, S., Shuraeva, A., & Ghimire, B. (2020). *The Financial Pandemic: COVID-19 and Policy Interventions on Rational and Irrational Markets.* Retrieved from https://ssrn.com/abstract=3589557

Sheridan, A., Andersen, A. L., Hansen, E. T., & Johannesen, N. (2020). Social distancing laws cause only small losses of economic activity during the COVID-19 pandemic in Scandinavia. *Proceedings of the National Academy of Sciences on the United States of America*, *117*(34), 20468-20473.

SIC-NAICS LLC. (2021). *SIC Codes Lookup – Standard Industrial Classification.* Retrieved from https://siccode.com/sic-code-lookup-directory

University of California, Los Angeles. (2021). *Regression with STATA chapter 2 – regression diagnostics.* Retrieved from https://stats.idre.ucla.edu/stata/webbooks/reg/chapter2/stata-webbooksregressionwith-statachapter-2-regression-diagnostics/

Wales, B. (2020). *Advisory memorandum on ensuring essential critical infrastructure workers' ability to work during the COVID-19 response, in: Guidance on the Essential Critical Infrastructure Workforce by the Office of the Director of the Cybersecurity & Infrastructure Security Agency of the U.S. Department of Homeland Security, Version 4.0 as of 2020-12-16.* Retrieved from https://www.cisa.gov/publication/guidance-essential-critical-infrastructure-workforce

World Bank. (2021). *World Bank Open Data.* Retrieved from https://data.worldbank.org/

Xiong, H., Wu, Z., Hou, F., & Zhang, J. (2020). Which Firm-specific Characteristics Affect the Market Reaction of Chinese Listed Companies to the COVID-19 pandemic? *Emerging Markets Finance and Trade*, *56*(10), 2231–2242.

Yang, H., & Deng, P. (2021). The Impact of COVID-19 and Government Intervention on Stock Markets of OECD countries. *Asian Economics Letters*, *1*(4).

Zaremba, A., Kizys, R., Aharon, D. Y., & Demir, E. (2020). Infected Markets: Novel Coronavirus, Government Interventions, and Stock Market Volatility around the Globe. *Finance Research Letters*, *35*(101597).

Zhang, D., Hu, M., & Ji, Q. (2020). Financial markets under the global pandemic of COVID-19. *Finance Research Letters*, *36*(101528).

Junior Management Science 7(2) (2022) 289-337

Junior Management Science

journal homepage: www.jums.academy

The Mission Comes First: Exploring the Mechanisms of Organizational Sponsorship for the Acceleration of Social Start-Ups

Niklas Manhart

Technische Universität München

Abstract

A significant number of incubators and accelerators have emerged to support start-ups aiming to solve societal or environmental problems. However, there is still limited understanding of how these ventures perceive the value proposition of incubators and accelerators – and whether their support needs differ from conventional start-ups. This study uses the framework of organizational sponsorship to explore the acceleration of social start-ups. It is based on in-depth interviews with the founders of 10 start-ups from an impact-oriented incubator in Duisburg, Germany. Through an inductive case study, this research generates three main insights. First, the social-mission focus of these ventures leads to significant differences as compared to commercial ventures. Second, social start-ups profit more from intangible resources such as social capital than from tangible resources such as seed funding. Third, incubators and accelerators need to adapt their offerings to address the needs of social start-ups. This study presents the first systematic assessment of incubator and accelerator services from the perspective of social start-ups. Its main theoretical contribution is to extend the organizational sponsorship framework by proposing a novel support mechanism: impact acceleration.

Keywords: Business incubation; start-up accelerators; social entrepreneurship; social start-ups; organizational sponsorship.

1. Introduction

The history of social entrepreneurship dates back to the 1980s – as do initiatives to support it. The Ashoka Fellowship, called "the pioneer of the accelerator model" (Pandey, Lall, Pandey, & Ahlawat, 2017, p. 3), started to support social entrepreneurs in India in 1981. The world's first impact accelerator Echoing Green opened in London in 1987 (Casasnovas & Bruno, 2013, p. 185). In recent years, the popularity of social start-ups – early-stage ventures performing a commercial activity in pursuit of social goals (Doherty, Haugh, & Lyon, 2014, p. 420) – has increased significantly. Several new programs, academies, and coworking spaces have been created to support the incubation of early-stage social ventures (Miller & Stacey, 2014).

The shift towards social impact has been particularly noticeable in the German start-up ecosystem. The German Social Entrepreneurship Network spun off from the German Start-up Association in 2017 (www.send-ev.de). Germany's primary start-up conference Bits & Pretzels chose "impact" as its motto in 2019, with social entrepreneurs on stage during its opening ceremony (Bruckschlögl, 2019, September 29). New accelerator programs such as Respond (www.respond-accelerator.com), F-Lane (www.f-lane.com), and the Impact Factory (www.impact-factory.de) have launched to support social start-ups.[1] Even the Catholic Church has opened a "social hub" in Frankfurt (www.villa-gruendergeist.de).

These developments are indicative of a broader trend. In 2019, half the incubators and accelerators in Germany (46%) supported organizations with significant social or environmental impact; across Europe, the share of incubators partially or only supporting social start-ups was 57% (SIM, 2020). The growing popularity of social venture incubation has had a transformative effect on the entrepreneurial support landscape.

Despite this flurry of activity, there is still comparatively limited research on incubators and accelerators targeting so-

[1] Disclosure notice: The author of this thesis has been involved in launching the Impact Factory, the research setting of this study, in his role at the Beisheim Foundation, which is the main funder of the program together with the KfW Foundation and Franz Haniel & Cie. GmbH. This thesis has been authored in a personal capacity and all views expressed are the author's own.

DOI: https://doi.org/10.5282/jums/v7i2pp289-337

cial start-ups (Crisan, Salanță, Beleiu, Bordean, & Bunduchi, 2019; J. Hausberg & Korreck, 2017). The existing studies on the acceleration of social start-ups primarily offer descriptions and typologies, often by practitioners or consultancies. By contrast, the social incubation process remains unexplored: Aside from a study by Pandey et al. (2017), little is known about how social start-ups assess the value of incubators and accelerators.

This uncertainty has practical implications. Public and private funders are currently investing considerable resources in supporting social start-ups. However, whether these organizations require a different support model than traditional start-ups remains unclear (J. Hausberg & Korreck, 2017, p. 13). This uncertainty can affect the outcomes of these support programs, as the design of accelerators influences the performance of their ventures (C. S. R. Chan, Patel, & Phan, 2020; Cohen, Bingham, & Hallen, 2019).

Studying this novel phenomenon can also contribute to the theoretical understanding of incubators and accelerators in the following two ways: The first regards the process of acceleration. For decades, researchers have treated incubators as a black box (Hackett & Dilts, 2008). Multiple authors have called for less focus on their form (organizational features) and more focus on their mechanisms and activities (Colombo, Rossi-Lamastra, & Wright, 2018; Crisan et al., 2019; Shankar & Clausen, 2020). The second regards their specialization. As incubators and accelerators are becoming more popular, there is growing interest in the organizational contexts or sectors in which they operate (J. P. Hausberg & Korreck, 2021;Lall, Chen, & Roberts, 2020). The context of an impact-oriented accelerator is a rich opportunity to "examine accelerators more in depth across different groups of participants, contexts, and periods of time" (Crisan et al., 2019, p. 23).

The purpose of this thesis is to explore how incubators and accelerators can support the acceleration of start-ups aiming to solve societal or environmental challenges. Therefore, this thesis examines how these nascent ventures perceive the activities and services of incubators and accelerators. To gain a holistic understanding of the acceleration process, this thesis explores how the characteristics of social start-ups affect their support needs – and why they join a support program in the first place.

Two fundamental research decisions helped this study to address these questions. The first was the use of social start-ups as the units of analysis. Following Colombo et al. (2018, p. 195), who call it "one of the most promising research avenues in the field of accelerators", this study adopted the perspective of the beneficiaries of support activities. Thereby, it explored how incubators and accelerators create value for start-ups. The second research decision was to employ the emerging theory of organizational sponsorship. Originally developed by Flynn (1993a, 1993b) and popularized by Amezcua, Grimes, Bradley, and Wiklund (2013), organizational sponsorship describes how sponsors can support the establishment and growth of young organizations. By illustrating the mechanisms incubators and accelerators use to provide resources and mediate between start-ups and their environment, it can serve as a theoretical lens to assess how social start-ups perceive entrepreneurial support activities. Therefore, the framework seems well suited to study the acceleration of social start-ups (for a recent review, see Breivik-Meyer, 2020).

Furthermore, this study employed an inductive case study design based on Eisenhardt (1989). The primary data sources were in-depth interviews with 10 founders of social start-ups from the Impact Factory, an impact-oriented incubator in Duisburg, Germany. The structured interviews included a combination of open-ended questions and a questionnaire about the most relevant incubator and accelerator services. These activities were identified by reviewing 26 studies of incubators and accelerators for conventional and social start-ups. The qualitative interview data were coded on a within-case and cross-case basis to identify novel findings and emerging concepts.

With the aid of this research approach, this study suggests four major findings. First, it systematically describes the services provided to social start-ups by incubators and accelerators. Second, it explores how social start-ups at two stages of development perceive the value of these interventions in addressing their resource needs. It demonstrates that social start-ups highly value intangible resources such as social capital and knowledge, whereas the provision of tangible resources, such as seed capital, has a lower priority than for conventional start-ups. Third, it proposes a novel support mechanism called impact acceleration, which specifically addresses their support needs. Finally, this thesis explores how social start-ups determine the balance between the benefits of sponsorship and entrepreneurial self-reliance.

This study contributes to three streams of literature: 1) Regarding social entrepreneurship, it illustrates how the social-mission focus of these start-ups is manifested through hybridity, prioritization of the purpose, and a focus on measurable impact. 2) Regarding incubators and accelerators, it suggests that the services provided by these institutions are not sufficient for social start-ups. As a result, this study argues that the acceleration of social start-ups requires organizational sponsors to tailor their services to the unique characteristics of these ventures. 3) Regarding organizational sponsorship, it observes how the predominant focus on survival fails to capture the reality of resilient ventures such as social start-ups.

The study is structured in four parts. The following section introduces the theoretical background and its key concepts by drawing on three literature streams. The subsequent section illustrates the research design and the approach for collecting and analyzing data, leading up to the analytical model for the acceleration of social start-ups (Figure 1 in Section 3.5). The results are then presented along four thematic dimensions. The final section discusses how these findings contribute to the existing literature – as well as limitations and opportunities for future research.

Finally, this study covers the period from June 2019 to October 2020. How the social start-ups experienced the Im-

pact Factory was therefore affected by the outbreak of the Covid-19 pandemic in March 2020, although the interviewees mentioned its effects surprisingly rarely. This aspect is further discussed in the limitations section at the end of the study.

2. Theoretical Background

This section draws on three emerging streams of literature to establish the theoretical framework for this study: First, it defines social start-ups in contrast to conventional start-ups. Second, it reviews the existing research on incubators and accelerators – including both conventional ones and those that support social start-ups. Third, it introduces the mechanisms of organizational sponsorship as a theoretical lens to study the acceleration of social start-ups. Finally, it compiles an exemplary services portfolio to help structure the data collection and data analysis.

2.1. Social Start-Ups

Social start-ups are the primary units of analysis of this study. Before exploring the role incubators and accelerators play in accelerating their growth, this section addresses two questions: What distinguishes social start-ups from conventional start-ups, and what obstacles do social start-ups face when scaling their impact and business models?

2.1.1. Defining Social Start-Ups

Social entrepreneurship has become a prominent phenomenon in response to growing societal and environmental challenges. The EU Commission has estimated that Germany is home to between a few hundred and more than 100,000 social enterprises (Wilkinson, 2015, p. 29). Another study has estimated their number in Germany to be between 1,700 and 70,000 (evers & jung, 2016, p. 5). This broad range reflects conceptual ambiguity: Although the recent academic interest in social entrepreneurship has been considerable, its definition remains disputed (Dacin, Dacin, & Matear, 2010; Gupta, Chauhan, Paul, & Jaiswal, 2020; Mair & Marti, 2006; Zahra, Gedajlovic, Neubaum, & Shulman, 2009). Counting no less than 37 definitions, a literature review recommended focusing on mission- or outcome-based definitions – rather than on the individual characteristics of social entrepreneurs or their operating sectors, processes, or resources – to avoid a "debate that has no resolution" (Dacin et al., 2010, p. 42).

Consequently, the boundaries of social start-ups in this study are defined by two constituent elements: a "prime strategic focus on social impact" and an innovative and entrepreneurial approach to achieving their mission (Nicholls, 2006, p. 13). Another useful way to conceptualize the hybrid nature of social enterprises is the double bottom line: Social enterprises aim at mission accomplishment, or "social value creation," and financial sustainability, or "economic value creation" (Alter, 2003, p. 8). This duality of objectives – and the performing tension (W. K. Smith & Lewis, 2011, p. 388) inherent in pursuing a social mission and exploiting commercial opportunities – is a central theme in analyzing the support needs of social start-ups.

A related – and contested – question is the distinction between social entrepreneurship and conventional, or commercial, entrepreneurship. While there are strong indications for a "continuum ranging from purely social to purely economic" (Austin, Stevenson, & Wei-Skillern, 2012, p. 372), the literature is divided on this issue (see Gupta et al., 2020, for a review). Clearly, there are parallels, such as opportunity recognition, risk tolerance, innovation, network-building capabilities, and continuous learning (Dees, 1998; Perrini & Vurro, 2006). Simultaneously, most researchers regard the social mission as the key difference between conventional entrepreneurs, who seek to generate economic profits and shareholder wealth, and social entrepreneurs, who apply business principles to achieve a social mission (Dacin et al., 2010, p. 44).

It may appear tautological to refer to the social mission of social entrepreneurs to distinguish them from their commercial counterparts. However, Santos (2012) has argued that all entrepreneurs face a trade-off between "value creation" at the societal level and economic "value capture" at the unit level because organizations can only maximize one of the two dimensions in the same organizational unit (see also Mair & Marti, 2006). This characteristic of social enterprises is reflected by the start-ups in this study, which all share a "clear social purpose [as] the driving force for the inception of the enterprise," resulting in a "social mission that is integral, not tangential to, the enterprise" (Wilson & Post, 2013, p. 723).

The centrality of this self-defined "social mission coupled with a market-based method," creates "a context of intention pervading all other design decisions" of the nascent enterprise (Wilson & Post, 2013, p. 726) – with a direct impact on its resource needs and the design of appropriate support programs. Consequently, this thesis employs a broad definition of social start-ups as early-stage ventures that pursue social value creation through innovative and market-oriented solutions (adapted from Casasnovas & Bruno, 2013, p. 177).

The first part of the definition, "early-stage," refers to the development stage of ventures, rather than to their age. As explained in the methods section, this study focused on start-ups beyond the ideation stage, which have already validated that their product or service fulfills a real societal need (so-called proof of concept) but have not yet significantly increased their headcount or revenue. As to the other parts of the definition ("social value," "innovative," and "market-oriented"), it is beyond the scope of this study to define them. Rather, these concepts are illustrated empirically through a case study of 10 nascent ventures that all target a societal or environmental challenge with a novel product or service.

Regarding terminology, in the literature and public discourse, "social" and "impact" are often used interchangeably to describe organizations of this kind. Given that "impact" is an even broader term than "social," this study refers to such organizations as "social start-ups."

2.1.2. Scaling Social Start-Ups

All start-ups face challenging conditions in the first years of their existence, causing them to fail at a higher rate than incumbents (Triebel, Schikora, Graske, & Sopper, 2018). The vulnerability of new organizations has been attributed to the liability of newness (Stinchcombe, 1965) and the liability of smallness (Aldrich, 1986). These challenges are especially evident for social start-ups, which have been called "a very peculiar and fragile breed of start-ups" (J. Hausberg & Korreck, 2017, p. 2).

The resource needs of social start-ups (evaluated in Section 4.2) become apparent when social entrepreneurs attempt to scale. Although it can be argued that "maximum impact may best be achieved by staying small and local" (Nicholls, 2006, p. 21), most social enterprises seek to achieve impact on a wider scale. A poll of social entrepreneurs in Germany indicated that 87% intended to scale, 9% were undecided, and only 3% wished to stay small (DSEM, 2020, p. 43). However, only few social start-ups manage to expand their operations, build their teams, and raise the funds necessary to scale – a phenomenon called the "pioneer gap" (Lall, Bowles, & Baird, 2013, p. 15) or the "valley of death" (Branscomb & Auerswald, 2002, p. 36). Studies have identified multiple barriers to the growth of social start-ups: The issues they seek to solve are often systemic and wicked (Dorado & Ventresca, 2013, p. 69); they work in resource-constrained environments and focus on vulnerable target groups, reducing their customer base (Pandey et al., 2017, p. 8). Social start-ups often lack access to markets and capital because of their reduced earning potential (Gianoncelli, Gaggiotti, Miguel, & Charro, 2020, p. 27).

In the face of these challenges, social start-ups benefit from stakeholders – multilateral agencies, governments, or foundations – that are resource-rich with the "potential to sponsor and support social entrepreneurship" (Pandey et al., 2017, p. 2). Consequently, new programs and institutions have emerged to support social entrepreneurs in growing their ventures, addressing the pioneer gap, and driving social change worldwide (Casasnovas & Bruno, 2013; Lall et al., 2013; Miller & Stacey, 2014; Yang, Kher, & Newbert, 2020).

2.2. Incubators and Accelerators

Among the entities supporting social start-ups in Germany, Austria, and Switzerland, incubators and accelerators are the most prominent, according to a recent review of the support landscape for social entrepreneurship (Leirich, 2020, p. 48). To review the multifaceted research that has accompanied the emergence of incubators and accelerators, this section focuses on three questions: What are incubators and accelerators? Is distinguishing between them critical? What do researchers know about those explicitly supporting social start-ups?

2.2.1. The Emergence of Incubators and Accelerators

The establishment of the first incubator for technology start-ups, the Stanford Research Park, took place in 1959 in the United States (Galbraith, McAdam, & Cross, 2019), and its first review was a 1985 study by Allen and Rahman. The number of incubators has increased to around 7,000 worldwide (Van Weele, van Rijnsoever, & Nauta, 2017). Business incubators "have become an integral part of the modern entrepreneurial ecosystem" (J. P. Hausberg & Korreck, 2021, p. 152). Their popularity has sparked a rich research stream, reviewed by Hackett and Dilts (2004) and more recently by Mian, Lamine, and Fayolle (2016) and J. P. Hausberg and Korreck (2021). A bibliometric analysis was conducted by Albort-Morant and Ribeiro-Soriano (2016).

Incubators have evolved significantly since the 1950s. Mian et al. (2016) have described three waves of incubation models: Before 1980, science parks or technology gardens aimed at economic restructuring and job creation. The second wave in the 1980s and 1990s also offered value-adding services such as mentoring or networking. The third wave, since 2000, has seen the emergence of specialized incubators, innovation centers, and accelerators. Owing to this history, which aligns with the three generations of incubators described by Bruneel, Ratinho, Clarysse, and Groen (2012), older definitions often emphasize their physical collocation. Hackett and Dilts (2004, p. 55) called them "enterprises that facilitate the early-stage development of firms by providing office space, shared services and business assistance". More recent definitions tend to reference their goals or behaviors, rather than their resources. J. P. Hausberg and Korreck (2021, p. 163) have reconciled these views by defining them as "organizations that support the establishment and growth of new businesses with tangible and intangible resources during a flexible period".

A newer but no less popular incubation model emerged as part of the third wave to support the rapid growth of start-ups: the accelerator. Y Combinator, widely considered the first accelerator for technology start-ups, launched in 2005 (www.ycombinator.com). Between 2009 and 2018, the number of accelerators grew fivefold from 560 to 2,616, according to research by Roland Berger (Bioulac, Ditsche, & Dujacquier, 2019, p. 3). Hochberg (2016, p. 26) has provided a comparable estimate of over 3,000 accelerator programs worldwide.

The rise of accelerators has prompted a wealth of research, reviewed by Colombo et al. (2018) and Crisan et al. (2019). The majority of these studies fall into two categories: conceptual descriptions or empirical studies on the impact on venture performance (Hochberg, 2016). Although recent studies have indicated positive effects of accelerators on ventures (C. S. R. Chan et al., 2020; Hallen, Cohen, & Bingham, 2020), there is still no consensus definition of accelerators, despite pioneering work by Cohen (2013) and Cohen and Hochberg (2014). Consequently, it is necessary to ask whether incubators and accelerators are conceptually different entities.

2.2.2. Incubators and Accelerators – Same or Distinct?

Following the first definition of accelerators, or seed accelerators (Adkins, 2011; Miller & Bound, 2011), schol-

ars have argued that they constitute a distinct organizational form from incubators (Pauwels, Clarysse, Wright, & Van Hove, 2016). The most common definition of an accelerator, that by Cohen and Hochberg (2014, p. 4), lists five features in which it differs from an incubator and other models of entrepreneurial assistance, such as angel investors and coworking environments. An accelerator is defined as a "fixed-term, cohort-based program, including mentorship and educational components, that culminates in a public pitch event".

Although most scholars now regard accelerators as a "distinct form of innovation intermediary" (Crisan et al., 2019, p. 10), this study follows Mian et al. (2016) and Sansone, Andreotti, Colombelli, and Landoni (2020) in treating them as a form of incubator. This approach is justified on the following four grounds:

First, both entities essentially pursue the same goal, namely to "support rapid growth and rapid scaling up of entrepreneurial ventures" (Pandey et al., 2017, p. 18). Second, the definition by Cohen and Hochberg is modeled narrowly on U.S. technology accelerators such as Y Combinator and TechStars. By contrast, accelerators working in the social enterprise space "tend to work across a fairly wide spectrum of enterprise development stages, perhaps reflecting the relatively limited pipeline of firms" (Lall et al., 2013, p. 115). Third, the distinction does not hold empirically. There is "significant heterogeneity even among groups that meet the formal definition" (Hochberg, 2016, p. 35), with entities that could be defined as incubators referring to themselves as accelerators, and vice versa. This observation is exemplified by the research setting of this study, the Impact Factory, which does not fit into either of the two categories, as Table 3 in Section 3.2 shows. The final and most compelling argument is that the predominant focus on the organizational form is a "constraint on advancements in [the] field," that would "benefit by moving the focus of study to the level of the mechanism (i.e., acceleration)" (Shankar & Clausen, 2020, p. 102174). Likewise, Crisan et al. (2019, p. 20) have focused on mechanisms to "open the accelerator's black box" and explain how accelerators "pursue different interventions in different contexts".

This discussion has two implications for this study. First, the study incorporates the framework of organizational sponsorship in its research design to shed light on accelerator mechanisms. Second, it does not distinguish between incubators and accelerators. Rather, it refers to them interchangeably, or as "entrepreneurial support programs," to mean organizational entities aiming to support early-stage entrepreneurial ventures through the provision of resources or services. The current research on a particular type of accelerator – with the aim of supporting social start-ups – is presented in the following section.

2.2.3. The Acceleration of Social Start-Ups: What is Known

As the number of incubators and accelerators has expanded, so have their specializations. These entities have emerged in various organizational contexts, such as governments, corporations, and universities. Lall et al. (2020, p. 3) have distinguished three subtypes of the accelerator model: seed, corporate, and impact-oriented. Among these, the acceleration of social start-ups has experienced particularly dynamic growth. According to a 2018 survey, 15% of incubators and accelerators in Europe primarily targeted social start-ups, and 42% partially targeted such start-ups (SIM, 2020). The European Venture Philanthropy Association has counted 62 impact incubators and accelerators in Europe (Gianoncelli et al., 2020). TechStars, one of the most prominent accelerators in the world, has even announced a program for "for-profit, mission-driven founders building technologies to solve our most pressing social and environmental needs" (Shieber, 2017).

Considering the popularity of impact-oriented incubators, there is a remarkable scarcity of academic studies on these entities – even though "social incubators" were mentioned for the first time over 15 years ago (Aernoudt, 2004). Even in 2014, a study called social accelerators "quite rare" and "experimental" (Dempwolf, Auer, & D'Ippolito, 2014, p. 25). The review by Crisan et al. (2019, p. 14) included one study (out of 81) about accelerators aiming to "support social entrepreneurship", and the only review on social incubators to date is a working paper by J. Hausberg and Korreck (2017). The scholarly interest in this novel phenomenon has only grown in recent years.

To categorize the extant literature on the incubation of social start-ups, one can use the three categories suggested by J. P. Hausberg and Korreck (2021) in their review of conventional business incubators. Most publications have provided "definitions and typologies," including case studies (Nicolopoulou, Karataş-Özkan, Vas, & Nouman, 2017; Sonne, 2012), surveys (Casasnovas & Bruno, 2013; King et al., 2015; Lall et al., 2013; Miller & Stacey, 2014), and consultancy reports (Aspen Network, 2014; SIM, 2020). Studies on their "performance" have demonstrated that social-impact-oriented accelerators also improve the revenues and funding of the incubated ventures (Lall et al., 2020; Roberts & Lall, 2018) and that they are as efficient as other types of incubators (Sansone et al., 2020). By contrast, their "incubation process" remains largely unexplored – except for a study on the appeal of social accelerator benefits (Pandey et al., 2017) and one on social accelerator selection (Yang et al., 2020).

This lack of research makes it difficult to answer an important question raised by (J. Hausberg & Korreck, 2017, p. 13): If social businesses face different challenges, do they also require different support models? Recent empirical evidence has indicated that impact-oriented incubators and accelerators differ from their conventional counterparts. A study of incubators in Italy has suggested, for instance, that different types of incubators value different services: Business incubators considered physical spaces more important than social incubators, which in turn valued services linked to social impact (Sansone et al., 2020, p. 132). Similarly, a survey of incubators in Germany has found that 20% of all

incubators and 33% of social incubators offered specific services for ventures with significant social impact (SIM, 2020, p. 51). However, both studies reflected the views of incubator managers, and not of start-ups. Moreover, they failed to explain causality. The same applies to a quantitative study by Pandey et al. (2017, p. 1), who conceded that "little is known about how social entrepreneurs – the primary intended beneficiaries – assess the value-proposition of social accelerators".

Based on the above findings, do incubators and accelerators supporting social start-ups need to tailor their services to these ventures? Or can they rely on best practices learned from the acceleration of conventional start-ups, as the support needs of social start-ups are comparable? To address these questions, this study uses the framework of organizational sponsorship.

2.3. Organizational Sponsorship

One reason that this study does not differentiate between incubators and accelerators based on their form is that this factor indicates little about their functioning – and their effects. As Shankar and Clausen (2020, p. 2) have argued, "knowledge about the form (accelerator) is incomplete without knowledge about the mechanism (acceleration)". Hence, multiple authors have called for a better understanding of the acceleration process (Colombo et al., 2018; Crisan et al., 2019). However, attempts to do so are complicated by the diversity and fragmentation of the entrepreneurial support landscape.

The framework of organizational sponsorship helps to overcome this hurdle and structure the study of entrepreneurial support. Although this study does not rely on theory to develop and test hypotheses, employing a theoretical lens can still help with identifying and discussing relevant issues. Therefore, organizational sponsorship is introduced in three steps: by defining its original framework, by extending it to the acceleration of social start-ups, and by viewing incubators and accelerators as a type of organizational sponsor.

2.3.1. Original Framework

Combining perspectives from population ecology and resource dependence, Flynn (1993b, p. 51) originally defined sponsorship as "the intervention by government agencies, business firms, and/or universities to create an environment conducive to the birth and survival of organizations". Sponsors can strengthen nascent organizations by making resources available to them in their early stages, when they are most exposed to external liabilities (Stinchcombe, 1965). Although Flynn (1993b, p. 51) already mentioned "university and private industry sponsored business incubators" as an example of sponsorship, his concept was popularized 20 years later by Amezcua et al. (2013, p. 1628), who defined organizational sponsorship as "attempts to mediate the relationship between new organizations and their environments by creating a resource-munificent context intended to increase survival rates among those organizations".

In their effort to "better understand why and how different attempts to assist new organizations might succeed or fail", Amezcua et al. (2013, p. 1628) found that "resource munificence is not necessarily predictive of organizational survival", as the effect of sponsorship is contingent on geographic-based founding density. Moreover, they argued that sponsorship influences the survival of new organizations through two mechanisms: buffering and bridging (p. 1629). Buffering helps new organizations engage in formational and developmental activities without being exposed to external threats. Sponsors can also function as a bridge between organizations and their environment, and thereby provide legitimacy and social capital to new ventures. Amezcua et al. (2013, p. 1633) considered business incubators an ideal setting to observe how providing resources, social connections, and management advice supports the creation and growth of new businesses.

2.3.2. Extended Framework

In a recent literature review, Breivik-Meyer (2020, p. 174) called organizational sponsorship "an emerging theory that bridges the conversation between scholars of different types of sponsorship". Although the framework is suitable for studying start-up acceleration, it has not been clearly defined yet. Pandey et al. (2017, p. 8) have argued, for example, that social accelerators also engage in "bolstering" mechanisms by offering "mentoring, opportunities for additional fundraising and adding to an early-stage social venture's credibility and awareness". However, it remains unclear how bolstering is conceptually different from building and bridging, as Breivik-Meyer (2020, p. 182) has noted. A recent study has proposed that business incubators engage in "curating" by selectively directing entrepreneurs to the best available provider of a given resource (Amezcua, Ratinho, Plummer, & Jayamohan, 2020, p. 3). Yet, this mechanism also appears redundant to the two original mechanisms, in particular bridging.

Autio and Rannikko (2016, p. 43) noted that the concept of sponsorship, "while informing survival, has paid less attention to new venture growth". They argued that sponsorship is not only about passively insulating new ventures against market realities, but also about "boosting" their capacities to affect growth. Such policies may include "emphasizing strong growth motivations," "controlling milestone achievement," and "promoting the exchange of experiential insights".

Adding the boosting mechanism to building and bridging to study the acceleration of social start-ups leads to the framework presented in Table 1. Importantly, Table 1 omits the activities and services associated with each sponsorship mechanism. These are specified in the services portfolio in the final part of this section. It is first necessary to summarize what is – and is not – known about incubators and accelerators as organizational sponsors.

Table 1: The Extended Organizational Sponsorship Framework

	Mechanisms of Organizational Sponsorship		
	Buffering	Bridging	Boosting
Entrepreneurial resources	Focus on the development of internal resources	Focus on the acquisition of external resources	Focus on boosting the organizational capacities for growth
The role of sponsorship	Maintaining a protective environment	Serving as a connective intermediary	Formation and achievement of milestones and serving as intermediary between firms
The goal of sponsorship	Developing internal resources while minimizing resource dependencies	Acquiring social capital and legitimacy to build sustainable competitive advantage	Affecting the capacity for growth

Note. Adapted from Breivik-Meyer (2020).

2.3.3. Incubators and Accelerators as Organizational Sponsors

Organizational sponsorship is still evolving as a theoretical perspective. Most contributions using this framework were published after 2016 (Breivik-Meyer, 2020, p. 176). While these recent studies show promise for advancing the study of start-up incubation, "the actual content of those mechanisms is somewhat unclear and may differ across sponsorship phenomena" (Breivik-Meyer, 2020, p. 185). A comparison of five intermediaries underlined the importance of different types of sponsors, suggesting that every support organization "leaves a fading yet indelible mark" on nascent entrepreneurial firms (Clayton, Feldman, & Lowe, 2018, p. 117). Thus, this study focuses on sponsorship in the context of incubators and accelerators.

Nevertheless, this process still requires further exploration. For example, a study on incubators in Norway has indicated that buffering and bridging can facilitate the development of new firms by increasing their resource access and capability development – but also that its quantitative research design provided "little explanation as to why tenant firms choose to use these services or why they do not" (Breivik-Meyer, Arntzen-Nordqvist, & Alsos, 2019, p. 29). The provision of resources by organizational sponsors can even be counterproductive, as certain conditions can "inhibit or reverse the intended outcomes of organizational sponsorship" (Amezcua et al., 2020, p. 3). A study of U.S. accelerators has revealed that their design choices influence how new ventures process available information – they can help new firms overcome issues of bounded rationality by concentrating consultations, practicing disclosure, and standardizing activities (Cohen, Bingham, & Hallen, 2019). Incubators also struggle to determine which services supported companies need, as nascent entrepreneurs are often unaware of their resource gaps (Van Weele et al., 2017). Therefore, the incubator–incubatee interaction is essential for the codevelopment of service offerings (Vanderstraeten, van Witteloostuijn, & Matthyssens, 2020). Another factor is the identity of their sponsor: While government-sponsored incubators provide a combination of services, private, academic, and NGO-sponsored incubators tend to specialize in certain services (Dutt et al., 2016).

In summary, the effectiveness of organizational sponsorship by incubators seems to depend on a multitude of factors, including their design, their funder, and their interaction with ventures. This study focuses on the process of supporting social start-ups, which in their early stage are often "devoid of markers of quality in the market, financial and social resources to generate growth, and sustained competitive advantages" (Amezcua et al., 2020, p. 3). Hence, this study assesses the resource needs of social start-ups – and how incubators and accelerators leverage the mechanisms of buffering, bridging, and boosting to support them.

This study simultaneously considers the potential disadvantages of organizational sponsorship. Resource munificence can, for example, decrease survival rates among new organizations (Amezcua et al., 2013). It can prevent the early adaptation of new ventures (Cohen, 2013) and adversely affect firm performance by hampering the incentivizing effects of market exchanges (Jourdan & Kivleniece, 2017). Consequently, this study also asks whether – and why – social start-ups regard accelerator intervention as detrimental to their development.

2.4. Compiling an Exemplary Services Portfolio

Reviewing the existing literature on incubators, accelerators, and organizational sponsorship has illustrated why scholars from these fields have called for more studies on the activities of incubators and accelerators. Particular areas of further interest are the relationships between incubatees and sponsors (J. P. Hausberg & Korreck, 2021, p. 170), the services portfolio accelerators offer (Crisan et al., 2019, p. 2), and how these services contribute to the development of tenant firms (Breivik-Meyer et al., 2019, p. 7). These questions are especially relevant for social accelerators, as "no large sample studies" have examined their relationship with social entrepreneurs (Pandey et al., 2017, p. 1).

Due to the lack of research on incubator and accelerator activities, it is difficult to draw on existing studies to compile an overview of their services portfolio. The following examples from widely cited studies on incubator activities demonstrate this challenge: In a review, Hackett and Dilts (2004)

mentioned "selection," "monitoring and assistance," and "resource infusion", whereas Bergek and Norrman (2008) listed "selection," "business support," and "mediation" as the components of incubation. Additional examples include "mentorship," "connectivity," and "brand enhancement" (Wise & Valliere, 2014), along with "business support," "infrastructure," "access to networks," and "access to external resources, knowledge and legitimacy" (Bruneel et al., 2012).

This cursory list brings two conceptual issues to light: the need to differentiate between activities and outcomes, and arbitrariness in the selection of categories (e.g., the vague term "business support"). To address the first point, this study follows Crisan et al. (2019, pp. 16), who distinguished among "interventions" (services and activities offered), "outcomes" (achievements in specific contexts or bundles of services), and "mechanisms" (processes that transform interventions into outcomes). This study uses the framework of organizational sponsorship to address the second point and systematically analyze what incubators and accelerators do.

To apply this framework to the present research question, an intermediary step is required: identifying the most relevant services and activities of incubators and accelerators and matching them to the mechanisms of organizational sponsorship. To compile this exemplary services portfolio, this study reviewed 26 studies in three categories: 1) 10 on social incubators and accelerators, 2) seven on conventional incubators, and 3) nine on conventional accelerators.

These studies were mostly peer-reviewed, although practitioner and research reports had to be additionally consulted for accelerators and social accelerators. The services or activities mentioned in these studies were subsequently grouped and assigned to the mechanisms of organizational sponsorship. Conceptually similar concepts were aggregated to identify the most prevalent services for each mechanism (see Appendix for the full results and the used sources).

The resulting portfolio in Table 2 contains nine services, ranked by the frequency of their mentions in each category: "education and training," "internal mentoring," "seed funding," and "coworking space" for the buffering mechanism; "external networking," "access to external funding," and "validation and visibility" for bridging; and "peer support" and "milestones and progress tracking" for boosting.

Importantly, the portfolio in Table 2 is based on statistical considerations but also theoretical salience. Table 2 indicates that services associated with the boosting mechanism were barely mentioned in the reviewed studies. However, the boosting mechanism was still included to assess the relevance for start-up acceleration as suggested by Autio and Rannikko (2016).

Moreover, this exercise did not constitute a systematic literature review, and the statistical results should be interpreted with caution. Regardless, Table 2 indicates the services that were mentioned most frequently in the reviewed literature, which helped to ensure that the most relevant services of incubators and accelerators were addressed in the interviews – in combination with open-ended questions to allow novel concepts to emerge. Such "a priori specification of constructs" can facilitate "the initial design of theory-building research" and "permits researchers to measure constructs more accurately" (Eisenhardt, 1989, p. 536).

Regardless of its statistical validity, Table 2 offers interesting observations. There is, for example, consistency regarding the two most prominent services across the three categories ("education and training" and "external networking"), while funding services ("seed funding" and "access to funding") have a mid-level to low ranking. There is simultaneous variation between the categories (e.g., the high prominence of "coworking" in incubator studies or "internal mentoring" in accelerator studies).

The most relevant comparison is between social incubators and accelerators, on the one hand, and their conventional counterparts, on the other hand. Table 2 suggests that their service offerings differ, for example, the higher prominence of the bridging mechanism for social incubation. Therefore, one primary question explored in this study is whether social start-ups expect social incubators and accelerators to offer support services tailored to their needs.

3. Research Design

The previous section argued that recent research on incubators and accelerators has advanced the understanding of entrepreneurial support programs for nascent ventures. However, it also maintained that researchers have not sufficiently investigated the support for start-ups aiming to achieve societal or environmental goals. For this reason, this study utilizes the mechanisms of organizational sponsorship to study the acceleration of social start-ups.

This research combines a study of multiple cases with an inductive research approach to generate additional insights. Following Eisenhardt (1989) roadmap for building theory from case study research, this section introduces the case study method, the research setting, the case selection, and the approach for the collection and analysis of the data.

3.1. Case Study Method

A case study has been defined by Yin (2003, p. 13) as an "empirical inquiry that investigates a contemporary phenomenon within its real-life context, when the boundaries between phenomenon and context are not clearly evident". It has a distinct advantage when three conditions are met: a causal research question (how, why), no control of behavioral events, and a focus on contemporary events. All these conditions hold in the case of the present research question.

Using multiple cases, as opposed to a single case, typically provides a more potent base for theory building (Yin, 2003, p. 33). It permits a replication logic in which multiple cases are considered experiments to replicate or contradict an initial set of propositions (Yin, 2003, p. 47). Multiple cases are a "bridge from rich qualitative evidence to mainstream deductive research" (Eisenhardt & Graebner, 2007, p. 25) and allow to look for generalizability of constructs across cases, helping to detect rational or causal patterns (Dooley, 2002, p. 342).

Table 2: The Exemplary Services Portfolio

Sponsorship mechanism	Services (prominence in %)	Category (# of sources) Social incubators and accelerators (10)	Conventional incubators (7)	Conventional accelerators (9)
Buffering	Education and training	100%	100%	90%
	Internal mentoring	80%	0%	90%
	Seed funding	60%	40%	40%
	Coworking space	60%	90%	60%
Bridging	External networking	90%	70%	70%
	Validation and visibility	80%	60%	60%
	Access to external funding	50%	30%	40%
Boosting	Peer support	40%	0%	10%
	Milestones and progress tracking	0%	0%	30%

Note. Based on a review of 26 sources in three categories. The "prominence" refers to the mentions in % (rounded) in all sources in that category. see Appendix for the full table and all sources.

The novelty of social start-ups and social accelerators motivates an inductive approach – three-quarters of social enterprises in Germany were founded in 2014 or later (DSEM, 2020, p. 17). The paucity of existing research complicates the use of a deductive methodology, i.e., using theories to formulate hypotheses and evaluate them against empirical evidence. Hackett and Dilts (2008, p. 440) cautioned that most research on business incubation is anecdotal, fragmented, and should be "used with caution". Scholars have applied theoretical lenses to the study of incubators only recently (Mian et al., 2016). In such an under-theorized field of research, an inductive approach is better suited to explore patterns and causal relationships – especially in a complex system such as an incubator, in which the interaction of multiple stakeholders complicates the distinction between internal and external factors.

3.2. Research Setting

The primary units of analysis of this study are social start-ups associated with the Impact Factory, a program launched in 2019 in Duisburg, Germany, to support the foundation and growth of start-ups aiming to create positive social change. The Impact Factory describes itself as a "unique program that offers social entrepreneurs a collaborative space to develop scalable innovations for solving complex social and environmental challenges" (Impact Factory, 2020a).

According to the taxonomy by Cohen and Hochberg (2014), the Impact Factory can be described as a hybrid between an incubator and an accelerator. On the one hand, Table 3 indicates that the Impact Factory resembles an incubator by having a permanent location and not investing in its tenants. On the other hand, it acts as an accelerator with fixed cohorts, a limited duration, an education portfolio, and regular demo days. Consequently, the Impact Factory demonstrates the difficulty of separating incubators from accelerators by their design features – one of the reasons why this study treats the terms as interchangeable.

The Impact Factory has three main funders: two nonprofit foundations, the Beisheim Foundation and the KfW Foundation, and the family-owned investment holding Franz Haniel & Cie. GmbH, on whose premises it is located. A nonprofit organization called Anthropia gGmbH runs the Impact Factory, accepting applications from for-profit, nonprofit, and hybrid ventures without a specific industry or thematic focus. The main requirement for a start-up is to pursue one of the 17 Sustainable Development Goals (SDGs) of the United Nations. Start-ups can enter the program twice a year through a selective application process that involves a two-day assessment.

Between its launch in January 2019 and October 2020, the Impact Factory accepted 65 social start-ups out of 240 applications, which amounted to an acceptance rate of 27% (Impact Factory, 2020b). In each cohort, the selected start-ups joined one of three programs, depending on their development stage: "create-up," "ramp-up" or "fellow." Since these three programs played a key role in the case selection, Section 3.3 describes them in more detail. Once the start-ups entered the create-up or ramp-up stage, they embarked on a learning journey that included workshops, seminars, and peer-to-peer formats, accompanied by regular mentoring and coaching sessions. Both programs at the end culminated in a pitching challenge in front of an expert jury.

Table 4 compares the Impact Factory with the average German incubator or accelerator to provide more context about the research setting. It is based the Social Innovation Monitor 2020 that surveyed 51 German incubators and accelerators in 2018. This comparison shows that the Impact Factory largely corresponded to the benchmark, at least in terms of mean values.

Overall, the Impact Factory seemed a suitable setting for researching the acceleration of social start-ups due to its broad program range, its exclusive focus on start-ups aiming to achieve social impact, and its hybrid nature between an incubator and an accelerator.

Table 3: Impact Factory – Incubator or Accelerator?

Program features	Incubator (I)	Accelerator (A)	Impact Factory	(I or A)
Duration	1-5 years	3 months	5 months (create-ups and ramp-ups); 1 year (fellows)	(I + A)
Cohorts	No	Yes	Yes	(A)
Business model	Rent; nonprofit	Investment; nonprofit	Subsidised; nonprofit	(Neither)
Selection frequency	Non competitive	Competitive, cyclical	Competitive, cyclical	(A)
Venture stage	Early, or late	Early	Early	(A)
Education offered	Ad-hoc, HR/legal	Seminars	Ad-hoc and seminars	(I + A)
Venture location	On-site	Usually on-site	Usually on-site, but remote during Covid-19	(I + A)
Mentorship	Minimal, tactical	Intense, by self and others	Intense (except for Fellows)	(A)

Note. Taxonomy based on Cohen and Hochberg (2016, p. 9). Data for the Impact Factory from Impact Factory (2020a, 2020b) and discussions of the author with the management of the Impact Factory.

Table 4: Benchmarking the Impact Factory

	Incubator or accelerator in Germany (2018)		The Impact Factory (2019)
Full time employees	4.5 (mean)	3 (median)	4
Applications received	118 (mean)	50 (median)	117
Teams supported	23.8 (mean)	12 (median)	21*
Average incubation time	13.8 months (mean)	9.0 months (median)	Ca. 5 months**
Fee requirement	Yes: 22.2% No: 77.80%		No
Equity stake in tenants	Yes: 8% No: 92.0%		No
Revenue sources (top 3)	Subsidies (41.0%) Rent (17.0%) Services provided (12.7%)		Subsidies (100%)

Note. German averages from the SIM (2020). Data for the Impact Factory from Impact Factory (2020a, 2020b) and discussions of the author with the Impact Factory management.
* Only the formal create-up and ramp-up programs are included. Fellows are excluded.
** Subject to change, as 2019 was the first full year of operations.

3.3. Case Selection

The purpose of a case study is to develop theory, not to test it; consequently, theoretical sampling is appropriate, i.e., selecting cases because they are "particularly suitable for illuminating and extending relationships and logic among constructs" (Eisenhardt & Graebner, 2007, p. 27). While researchers often select single cases for their uniqueness, multiple cases can yield better theory, including "to replicate previous cases or extend emergent theory" or "to fill theoretical categories and provide examples of polar types" (Eisenhardt, 1989, p. 537). Importantly, the goal of the case selection is not to produce a representative sample but to allow for comparison while ensuring variation to improve the reliability and generalizability of findings.

Consequently, the definition of the population is crucial, as it "defines the set of entities from which the research sample is to be drawn" (Eisenhardt, 1989, p. 537). Restricting the population of this study to the Impact Factory (2020a, p. 537) controlled extraneous variation and set the limits for generalizing its findings. The total population consisted of 65 social start-ups accepted by the Impact Factory in three cohorts between July 2019 and October 2020. This study applied the following four selection criteria to select the final sample of 10 start-ups.

3.3.1. Selection Based on Development Stage

The Impact Factory accepted start-ups at distinct stages of development, as Table 5 indicates – from create-ups in the ideation phase with few available resources, to fellows looking to scale their business and impact model. Given that the support needs of start-ups at an early stage differ strongly from more mature ventures (Isabelle, 2013), the inclusion of all the stages would have created an excessive disparity in support needs. Consequently, the 15 participants of the create-up program were excluded from the total of 65 start-ups, leaving 50 start-ups in the sample.

3.3.2. Selection Based on Legal Structure

A crucial distinction between social start-ups is their legal structure, which in turn influences their funding. Reflecting the variety of funding models of social ventures (Nicholls, 2006, p. 12), the start-ups of the Impact Factory ranged from nonprofits to self-funded social enterprises. This study excluded nonprofit organizations for two reasons: First, because nonprofits, which in Germany rely primarily on donations and philanthropic grants, display vastly different support needs compared to for-profit start-ups. Second, because less than 2% of organizations supported by incubators in Germany had a nonprofit status (SIM, 2020, p. 39). Accordingly, 15 additional start-ups were removed, including all start-ups that had not been legally established yet. This decision left a sample of 35 start-ups with either a for-profit or hybrid (for-profit and nonprofit) legal form, consisting of nine ramp-ups and 26 fellows.

3.3.3. Selection Based on Theoretical Considerations

Due to time and capacity constraints, the maximum number of cases was set at 10. This decision was a compromise between informative value and feasibility in the available time frame. The 10 start-ups were not to be selected at random, however. Instead, the first subgroup of five start-ups was selected from the ramp-up program and the second subgroup of five start-ups from the fellows program (see Table 5 for the three different programs of the Impact Factory).

Taking advantage of the different programs of the Impact Factory for the case selection served three theoretical purposes. First, the selection of five start-ups from different programs yielded enough cases to replicate observations within a subgroup of start-ups at similar stages of development – similar to conducting multiple experiments under the conditions of the original experiment (Yin, 2003, p. 47). Second, choosing cases that were alike and analyzing findings across similar cases enhanced "generalizability relative to a single case" and demonstrated the issues "across a more varied range of circumstances" (Chmiliar, 2010, p. 582).

Variation between subgroups strengthens the external validity of findings (Yin, 2003, p. 54) and maximizes opportunities for developing hypotheses or theories (Bleijenbergh, 2010, p. 63). In the case of the Impact Factory, the two subgroups based on the programs in Table 5 could be considered contrasting "polar types" (Eisenhardt, 1989, p. 537). While all ramp-ups took part in an intensive learning journey that included workshops, seminars, and individual mentoring, fellows did not participate in the educational program or mentoring activities (Impact Factory, 2020b). Their affiliation to the program was more fluid, and their support was restricted to outward-facing activities, such as networking and increasing their visibility.

As a third consideration, the two subgroups fulfilled distinctive theoretical categories (Eisenhardt, 1989, p. 537) based on the organizational sponsorship framework of Amezcua et al. (2013): The program activities for ramp-up start-ups focused on building internal capabilities and helping their go-to-market, as captured by the buffering mechanism; start-ups in the fellows program focused on strengthening external relationships, aligned with the bridging mechanism.

3.3.4. Selection Based on Active Participation

The final criterion for selecting 10 cases out of the remaining 35 start-ups was active participation in the Impact Factory. It made sense to choose cases "in which the process of interest is observable," i.e., likely to replicate or extend emergent theory (Eisenhardt, 1989, p. 537). Identifying active participation was straightforward for the nine ramp-ups, given that it was a requirement to stay in the program. After discarding two start-ups that dropped out early, five of the remaining seven ramp-ups were contacted in random order. After one declined to participate, a sixth start-up was successfully approached.

Establishing active participation was more difficult for the 26 fellows due to their larger number, but also because their participation varied significantly, as the author observed on multiple occasions on-site. Hence, the program managers of the Impact Factory were asked to suggest founders who could provide rich insights into the program. They recommended nine teams, five of which were approached at random, as well as a sixth after one fellow declined to participate.

Table 6 presents a list of the final 10 start-ups, as well as key information about their stage, industry, commercial activities, and social mission. To protect their anonymity, all start-ups are only referred to by an acronym. Subsequent sections provide further information, for instance regarding their funding, educational background, and professional experience, together with relevant findings of the case studies.

Although the 10 start-ups operated in a broad range of industries, they all shared the definition of a social start-up established in Section 2.1 by pursuing both a social mission and an entrepreneurial activity. Digital platforms were prevalent, as half the start-ups operated a platform business model. The development stage ranged from those with a minimum viable product (MVP) to those with a first customer base, although three fellows were already seeking to grow their revenues. The start-ups were all founded between 2018 and 2020, except for two founded in 2015.

3.4. Data Collection

Consistent with most qualitative research, the primary data sources of this study were semi-structured interviews with the founders of 10 start-ups from the Impact Factory. Interviews are "a highly efficient way to gather rich, empirical data, especially when the phenomenon of interest is highly episodic and infrequent" (Eisenhardt & Graebner, 2007, p. 28), such as the perception of entrepreneurial support programs by social ventures. In the case of multiple founders, a single founder was interviewed depending on the team's availability.

The 10 interviews took place in November 2020, lasting between 45 and 60 minutes each. Due to Covid-19 restrictions, the interviews were conducted remotely over video.

Table 5: The Three Programs of the Impact Factory

	Programs of the Impact Factory		
	Create-up	Ramp-up	Fellows
Average cohort size	5 start-ups	6 start-ups	12 start-ups
Entry requirement	Valid idea	Marketable product or service	Proof of concept and early customer base
Program duration	5 months	5 months	Flexible; up to 1 year
Program focus	Prototype development/ customer discovery	Go-to-market/ customer validation	Company building/ business model scaling
Frequency of interaction	Every two weeks	Every two weeks	Occasional
Program participation	Mandatory	Mandatory	Voluntary (except to provide personal input at least once)
Program goal	Problem-solution fit	Product-market and product-channel fit	Organization-market fit

Note. Data for the Impact Factory from Impact Factory (2020a, 2020b) and discussions of the author with the management of the Impact Factory.

The author recorded, transcribed, and translated the interviews from German into English to facilitate the coding process. He used an interview guide containing four sections and 16 questions. The interview guide is attached as Appendix B.

The first section of the interview contained introductory questions about the founders and their start-ups. The second addressed their resource needs at different points in time and their motivation for joining a support program. The third focused on the services and resources of incubators and accelerators – first with open questions, and then by addressing the nine most relevant services identified a priori in the literature (as explained in Section 2.4) unless they had already been addressed. The fourth section asked the founders to evaluate how the support they received impacted the overall development of their start-up.

Following the interviews, all founders were asked to rank the nine most common incubator and accelerator services in their order of importance for a social start-up – not for statistical purposes but to enrich their assessment of incubator services. During the coding process, the founders were occasionally contacted again for clarifications or follow-up questions – what Eisenhardt (1989, p. 538) described as an overlap of data collection and data analysis.

An important method to increase the credibility of case study data is triangulation, or the use of multiple sources of data (Tracy, 2010, p. 843). Triangulation can also mitigate the bias inherent in interviews as the primary source of information (Eisenhardt & Graebner, 2007, p. 28).

Consequently, secondary data sources were also consulted, such as the start-ups' websites, as well as investor presentations, pitch decks, public interviews, and additional materials. The author also had the opportunity to visit the Impact Factory several times since its launch in 2019 to gather observational data. In combination, these data sources provided a rich picture of 10 social start-ups and their experiences of entrepreneurial support programs.

3.5. Data Analysis

After completing the interviews and gathering the data, it was necessary to analyze them by "focusing on some and disregarding other parts of it" and "aggregating data into a small number of themes" (Creswell, 2014, p. 245). This process involved two steps. The first was a detailed study of each case to look for inter-case similarities and differences. The purpose of this so-called within-case analysis is to be "intimately familiar with each case as a stand-alone entity" and to "allow the unique patterns of each case to emerge before investigators push to generalize patterns across cases" (Eisenhardt, 1989, p. 540).

The second step was a cross-case analysis, searching for patterns and causal relations across cases in multiple rounds of comparative analysis – first within the two subgroups of ramp-ups and fellows, and then across the entire sample. In keeping with the exploratory nature of the research, the coding process – i.e., the segmentation and labeling of data into categories (Creswell, 2014, p. 248) – was inductive. So instead of using pre-determined codes, the constructs emerged from a close observation of the data during coding. This inductive process was occasionally structured by assumptions based on the theoretical framework to maximize coherence.

Following the approach recommended by Eisenhardt (1989, p. 540), the cross-case analysis was preceded by the selection of categories or dimensions "suggested by the research problem or by existing literature". The idea of using such "structured and diverse lenses on the data" is to "go beyond initial impressions," to improve "close fit with the data," and to capture novel findings in the data.

The coding process was iterative, as it involved going back and forth between the data, the emerging propositions, and constant comparisons to relevant literature. The interview data was split, merged, and narrowed down to develop higher-level concepts, generating a smaller number of first-order codes with representative quotes by the respondents for each theoretical concept and emerging theme. These

Table 6: Case Descriptions

Acronym	Legal form	Team size	Year founded	Stage*	Industry**	Commercial activity	Social mission
Start-ups of fellows program							
F1	For-profit	11	2019	Growth	Health Services	Develop AI-based software for dynamic tour management	Improve quality of outpatient care with efficient tour planning
F2	For-profit	8	2018	Growth	Commerce	Import and sell sustainable home accessories	Foster appreciation and fair prices for handmade products
F3	For-profit	8	2020	MVP	Utilities	Develop and distribute solar home systems and IoT software	Widen access to electricity in rural Sub-Saharan Africa
F4	For-profit	2	2015	Growth	Food and drink	Import, process, and sell sustainable cocoa, coffee and chocolate	Protect biodiversity and strengthen ecosystems in Peru
F5	For-profit	7	2018	Go-To-Market	Education	Develop and market platform to measure and train digital skills	Promote social participation and reduce digital skills gap
Start-ups of ramp-up program							
R1	Hybrid (for/non-profit)	9	2015	Go-To-Market	Health Services	Develop and operate platform for recruiting voluntary caregivers	Combat lack of specialist caregivers for elderly living at home
R2	Hybrid (for/non-profit)	1 + volunteers	2019	Go-To-Market	Food and drink	Produce packaged foods based on consumer surveys	Improve animal welfare and empower farmers and consumers
R3	For-profit	5	2020	MVP	Financial Services	Develop platform for efficient B2B payment processing	Give donations to social projects with cashback scheme
R4	For-profit	12	2020	Go-To-Market	Food and drink	Develop and operate system for reusable takeaway packaging	Reduce waste created by disposable food packaging
R5	For-profit	6	2020	Go-To-Market	Tourism	Develop online platform for eco-friendly travel	Raise funds for aid organizations by renting their unused space

Note. All information as of the date of the interviews in November 2020.
*Stage taxonomy: Idea, Launch, Proof of Concept, MVP, Go-To-Market, Growth, Maturation.
**Based on the Industry and Sectors Taxonomy of the International Labour Organization (2020).

first-order codes were subsequently aggregated to identify second-order codes for each theme.

This process produced the four thematic dimensions that structure the results of this study. Three of these dimensions – "founding motivations," "resource needs," and "sponsorship mechanisms," were specified a priori and addressed purposefully in the interviews. A fourth dimension – "entrepreneurial self-reliance" as an alternative to joining a start-up support program – emerged during the coding process, just as "impact acceleration" was identified as an additional support mechanism for social start-ups. Figure 1 shows the resulting analytical model and the relationships between the four dimensions, the support mechanisms, and the corresponding interventions by an incubator or accelerator.

A separate table was created for each dimension and service in accordance with the approach recommended by Miles

& Huberman, 1994 for summarizing and presenting case evidence. Such construct tables help to "indicate how the focal construct is 'measured', thus increasing the 'testability' of the theory and creating a particularly strong bridge from the qualitative evidence to theory-testing research" (Eisenhardt & Graebner, 2007, p. 29). They were used to link the case evidence from primary and secondary data to emerging theoretical concepts.

4. Results

After having introduced the theoretical framework and research design, the following section presents the results from the case interviews and relates them to the extant literature. As illustrated in Figure 1 in Section 3.5, the results are structured in four dimensions: The social-mission focus and founding motivations form the basis for the resource needs of social start-ups. Building on these two dimensions, the third part discusses the mechanisms of organizational sponsorship. In addition to the three mechanisms of the extended framework, it identifies a novel support mechanism – impact acceleration – that is highly beneficial to social start-ups. The final dimension emphasizes a contrasting motive that emerged in this study: Entrepreneurial self-reliance – or, put differently, reasons for not joining a formal support program.

4.1. Social Mission and Founding Motivations

Founding motivations are at the heart of every entrepreneurial venture. They determine why founders embark on the perilous journey of launching a start-up – after all, 50% of companies fail in their first five years (Triebel et al., 2018, p. 121). Founding motivations also affect how founders pursue opportunities and acquire resources. Studying the motives and backgrounds of their founders helps to explore the support needs of social start-ups.

The purpose of this section is not, however, to provide a full characterization of the personality and motivations of social entrepreneurs, which have already been thoroughly studied (Gupta et al., 2020; Zahra et al., 2009). Social entrepreneurs share many qualities with founders of conventional start-ups, such as "leadership, vision, drive and opportunism" (Nicholls, 2006, p. 20), and it is difficult to identify characteristics unique to them.

Nonetheless, this study maintains that a social-mission focus drives social entrepreneurs and "affects how [they] perceive and assess opportunities" (Dees, 1998, p. 2). Additionally, the literature suggests that founders of social start-ups are motivated by other factors than founders of conventional ventures (Germak & Robinson, 2014). For this reason, it is helpful to explore what their social-mission focus means in practice – and how it influences their receptiveness to and perception of incubator and accelerator benefits.

Table 7 on the following two pages identifies six ways in which the social start-ups in this study demonstrated a social-mission focus. Or in other words: How they prioritized the pursuit of a social mission. In addition, Table 7 presents information about the educational background and professional experience of the founders, given that these characteristics strongly influence how social entrepreneurs approach their enterprises (Germak & Robinson, 2014).

Overall, Table 7 shows six manifestations of the social-mission focus that emerged in the interviews. "Sustainability," "independence," "long-time thinking," "hybridity," "credible impact," and "priority of purpose" all differ in the degree to which they are unique to social start-ups. The most common among founders is an ambition to be independent. Sustainability and long-time thinking are also significant for social as well as conventional entrepreneurs. By contrast, the latter three themes are more specific to the social start-up context.

Hybridity refers to aligning the social and economic logics of a venture, such as using business efficiency to maximize social impact (F1) or leveraging impact through additional revenue streams (R4). Nine of the 10 start-ups pursued a social mission through their key economic activities and not as an add-on. The only exception (R3) was a fintech company that generated donations for social projects. However, even R3, like R2, anchored its mission with a novel legal form called steward-ownership that preserves its essential purpose (Ventures, 2020).

With the concept of credible impact, the founders signaled an intention to prove their impact in a trustworthy and transparent way, for instance by including impact metrics in business monitoring (F3) or by influencing all steps of the value chain (F4). Multiple respondents noted that referring to "impact" has become so prevalent among start-ups that it has almost lost its meaning. Indeed, in a large poll among U.S. internet start-ups, 68% of founders cited "impact" as their main motivation, before "experience" at 27% and "money" at 5% (Marmer et al., 2011, p. 59).

The priority of purpose supports Santos (2012) claim that entrepreneurship involves a trade-off between societal value generation and economic value capture, with most respondents signaling that they were not driven by profit motives – a motivation shared with 84% of social entrepreneurs in Germany (DSEM, 2020, p. 39). This effect was not consistent across the cases, but it did not appear to correlate with age. Although experienced founders such as F4, F5, R1, and R2 expressed a stronger purpose motive than younger respondents, the most inexperienced founder (F5) also claimed to put their purpose ahead of other motivations.

The results of Table 7 show a clear intention to question, if not reject (F2 and R2), the practices of conventional business life. Interestingly, all ramp-ups and three fellows had a background in business or management, but only two founders had previously pursued a social mission – running a bio-diversity consultancy (F4) and an education nonprofit (F5). Moreover, Table 7 shows that most founders were experienced, up to the C level, with only two teams (F1 and R5) launched by young professionals. The maturity of the founders needs to be considered, given that "the organizing behaviors and decision making of individuals is dependent on their knowledge structures," as well as on their work experience and background (Katre & Salipante, 2012, p. 972).

Table 7: Social-Mission Focus

	Educational background	Professional experience	Manifestations of the social-mission focus	Representative quotes
			Start-ups of the fellows program	
F1	Mathematics and Computer Science, PhD	Research, 4 years	1. Hybridity	People are not packages. We have to consider more than pure business efficiency. At the same time, business efficiency plays a major role. There is no other way. (I)
			2. Sustainability	I believe those who work sustainably have universally higher success rates. But I don't really see a need to do anything different than what a normal start-up does. (I)
F2	Business Administration	Pharmaceuticals, 10 years	3. Independence	If I would even think about working with a business angel or with an external investor, then only if the social impact component is safeguarded. (I)
			4. Priority of purpose	I have experienced corporate life for too long. I have always said I don't want to discuss purchase prices with an investor. Those are the cornerstones of our business, and I can't touch them. (I)
			2. Sustainability	We set a high standard of socially and ecologically sustainable cooperation. We source natural materials and work directly with our producers to protect their traditions. (W)
F3	Information Management	IT consulting, <1 year	5. Long-time thinking	Investors are exit-driven. It is about growing quickly and then selling. We are looking for somebody who wants to give patient capital and share in the profits. Not just put money in and bang, away with it. (I)
		Co-founder: Solar energy, 8 years	6. Credible impact	We want to include impact KPIs in our balanced scorecard, in our monitoring. We haven't done that yet, but as soon as this loan is safe, it will be on the list. (I)
			2. Sustainability	We develop products with a positive carbon footprint that are fun, durable, locally produced, and repaired and recycled locally. (P)
F4	Tropical Biology, PhD	Researcher / Founder of a consultancy, >15 years	3. Independence	The vast majority of start-ups are looking for money. We have been asked time and again by investors whether they could join us. And we have always refused. (I)
			4. Priority of purpose	Our aim is not to achieve the fastest and largest possible monetary profit, but to make a good business for everyone in the true sense of the word...(W)
			6. Credible impact	From cultivation to the finished product, we are directly involved and can influence all production steps. Our internal rules exceed the criteria of any certification. (W)
			2. Sustainability	The start-up scene is dominated by people who think they have a really cool idea and need 1.5 million. Often it's just hot air with a fancy website. To be honest, that's a bit strange to us. (I)
F5	Education	Education, >25 years	4. Priority of purpose	You can present what we do as a company one way or another. It would be easy to get financing. This is really a good story. But we didn't want to play this card at all. No, actually we are a purpose company. (I)
	Co-founder: IT, MBA	Co-founder: IT (C Level), >20 years	6. Credible impact	With social start-ups, you have to have [impact measurement], that's your legitimation for saying I really want to achieve outcome and impact. And I am quite rigorous. You have to be serious about that. (I)

(Continued)

Table 7—continued

				Start-ups of the ramp-up program
R1	Business Administration	Logistics (C Level), >20 years	3. Independence	I don't want to have to give account to somebody and pay eight percent interest. You need investors who actually think about this solution and are convinced of it and who don't really care about the money. (I)
			4. Priority of purpose	We simply lack a profit motive. When I go to the bank, they say: what do you want here with us. I have built up this organization with 60, 70 hours a week for three years completely free of charge. (I)
			5. Long-time thinking	I could open a platform and create a sort of Tinder. The expenditure is not much. But that doesn't solve the problem. I don't want to charge 35 euros per hour to look after a poor mother, it has to be cheap. (I)
R2	Management	Food and drink (C Level) >15 years	3. Independence	I don't want to depend on external people. If I don't know them or their intentions, then that is too risky for me. But if you spend a little money, it goes faster and in the end, it may be more efficient. (I)
			4. Priority of purpose	A start-up can only be successful if it exists on the market for a long time and fulfills a real purpose. To make money quickly, you need a short-term focus, which I used to hate in normal business life. (I)
			5. Long-time thinking	I have problems with short-termism. And that's why I'm in the impact business and not in the classic money-making business. It's about the purpose and not about the money. (I)
R3	Business Informatics	Banking 9 years	1. Hybridity	We use profits and growth to do this as a means to social and environmental goals. This social cashback creates the positive impact we need as a society. We put purpose before profit maximization. (W)
			4. Priority of purpose	It is quite clear the profit will never be the main focus for us. We have built the company completely differently. We have committed our company to reinvesting all profits and donating the rest. (I)
R4	Marketing and Sales	Digital Marketing 10 years	2. Sustainability	Sustainability is the most important thing for us and the reason we founded the company. We want to look at ourselves and say: we made a big contribution to climate protection and against plastic flood. (M)
			1. Hybridity	For us sustainability also includes the economic aspect. Therefore, we were looking for a solution where purpose and profit are "streamlined", i.e. the more sustainable impact, the more revenue we generate. (P)
			6. Credible impact	No matter if you talk about profit or not, you always need take your impact into account. Personally, I would say that no start-up or company today would be harmed by taking this topic seriously. (I)
R5	International Business	None (University) <1 year	3. Independence	We didn't want to take the classic start-up path of directly taking up an investment and getting an investor in who is not 100 percent committed to our development and the impact idea. (I)
			4. Priority of purpose	When we decided to found this company, we said to ourselves: Okay, the impact must always come first. If that means that we can't pay out a cent for the first two or three years, then that's perfectly fine. (I)

Note.The personal information refers to the main interviewee. Co-founders have been mentioned when their profile differs materially from the interviewee. Sources: (I)nterview with the author; (M)edia articles; (P)itch decks or (P)resentations; (W)ebsite. All German quotes translated by the author.

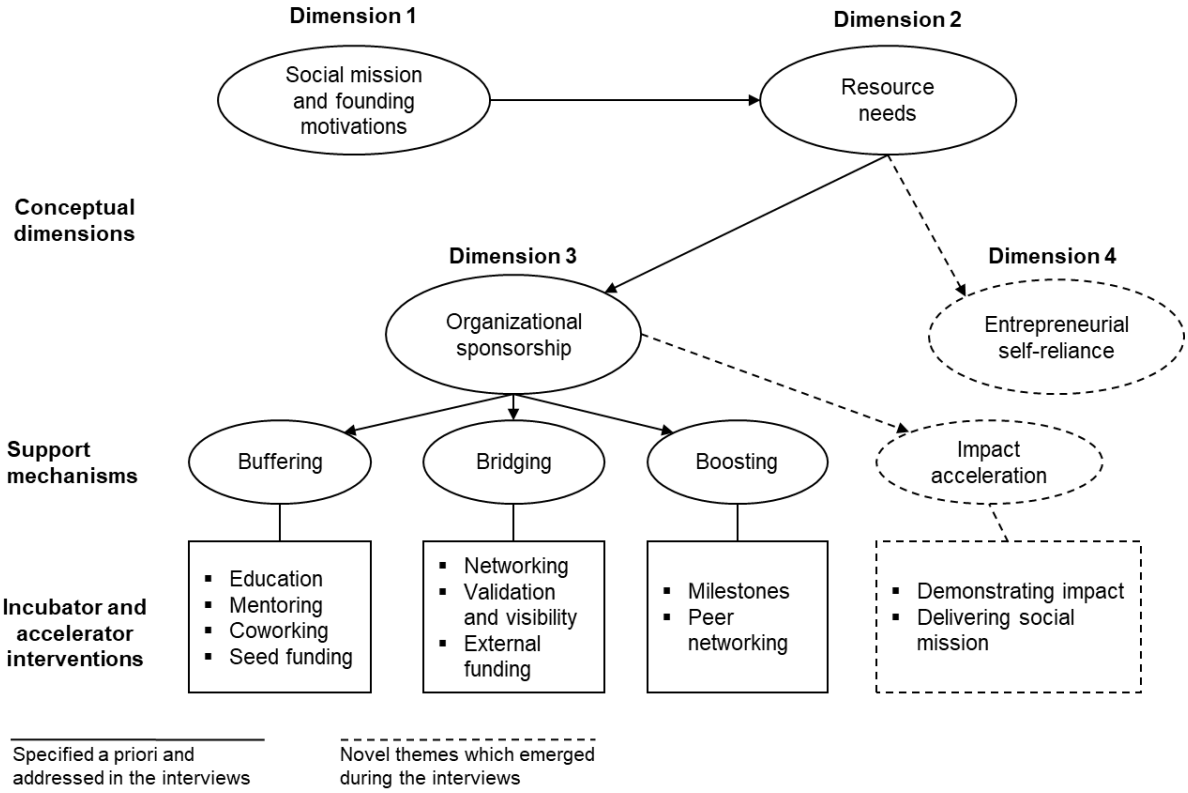

Figure 1: Analytical Model for the Acceleration of Social Start-Ups

Finally, Table 7 signals no clear difference in social-mission focus between ramp-ups and fellows, which suggests that the advanced stage of fellows had not led to mission drift.

To explore the social-mission focus further, the founders were asked why they started their ventures in the first place. Table 8 groups their founding motivations into four common themes.

In their study of social entrepreneurship motivations, Germak and Robinson (2014, p. 18) found "a unique blend of motivational components in nascent social entrepreneurs". Table 8 confirms that a combination of four factors motivated the founders in this study.

Achieving societal impact seemed to be a prime motivation, with nine founders explicitly mentioning their impact on direct or indirect beneficiaries, and six aiming to achieve a systemic impact. This is not entirely surprising – in a poll 97% of German social entrepreneurs expressed a desire to solve societal challenges, and only 3% a desire to "become rich" (DSEM, 2020, p. 55).

The same survey, in line with Table 8, indicated that seeking personal fulfillment is a strong motivator for social entrepreneurs. Germak and Robinson (2014, p. 13) also found that the needs of social entrepreneurs are "at a higher level of personal fulfillment" than seen with "necessity-based" entrepreneurs. Observing this factor more closely, the primary sources of personal fulfillment in this study were seizing opportunities, tackling complexity, and achieving autonomy, whereas no founder talked about material gains.

Nonetheless, accomplishing significant achievement is also a strong motivator for social entrepreneurs, "not entirely dissimilar from what one would expect of commercial entrepreneurs" (Germak & Robinson, 2014, p. 16). In the case of the start-ups in this study, this desire focused on developing innovative solutions through innovative technologies and business practices but also by drawing on experiences from other sectors or countries – reflecting their maturity and diverse professional experiences.

A fourth motivational factor cited in the literature is personal closeness to a social problem (Germak & Robinson, 2014, p. 17). Katre and Salipante (2012, p. 977) even found that starting with an "initial concept of social change based on personal, family, or community experiences" and then developing an economic opportunity distinguished successful from struggling social entrepreneurs. Accordingly, half the founders reported being personally affected by a societal issue. More surprising is another source of personal closeness: Six out of 10 explicitly cited traveling or working abroad as a personal motivation for founding their venture.

In combination, the findings of Table 7 and Table 8 confirm Nicholls (2006, p. 13) definition that social start-ups are characterized by a "prime strategic focus on social impact" and an innovative and entrepreneurial approach to achieving it. For the start-ups in this study, the primary social purpose "is not a difference in net profits, but a net difference in total value creation" (Wilson & Post, 2013, p. 723) – regardless of their development stage. Having established their motiva-

Table 8: Founding Motivations

Second order codes	First order codes	Representative quotes	All mentions	
Developing innovative solutions	new research	It appeared on my desk and I thought how wonderful, because my dissertation was on this area. (I) (F1)	F1, F4	R3
	new technology	You have to combine pedagogical thinking and sophisticated technical skills and make it operational. (I) (F5)	F1, F3, F5	
	new business practices	The core of our innovation is that we have developed a valid procedure to measure digital skills. (I) (F5)	F5	R3, R4
	business experience	I decided to bring 20 years of logistics expertise and digitization into the social sphere. (I) (R1)	F4, F5	R1, R2
	replication	We first tried our approach in Cambodia, and step by step we added more and more countries. (M) (R5)		R2, R5
Seeking personal fulfillment	meaningful work	The decisive factor for me was that I was looking for more meaning. (I) (R2)	F2	R2
	achieving autonomy	It started as a student organization and we decided to spin it off as an impact start-up. (I) (R5)		R3, R4, R5
	having fun	It is very time-consuming, especially because none of us earns anything from it, but it is simply fun. (M) (R5)		R5
	tackling complexity	Mathematical algorithms are incredibly diverse and adaptable. This is what excites me about them. (M) (F1)	F1, F4, F5	R1
	seizing opportunities	There was a call for tender by the EU to operationalize this, but they couldn't find anyone to do it. (I) (F5)	F1, F5	R4
Achieving social impact	on direct beneficiaries	I don't want to charge 35 euros per hour to look after a poor mother, it has to be cheap. (I) (R1)	F1, F2, F3, F4, F5	R1, R2
	on indirect beneficiaries	There are many NGOs/NPOs which lack the necessary capital to keep their engines running. (W) (R3)		R3, R5
	on economic stakeholders	You as a consumer or as a a restaurant owner can make a personal contribution to environmental relief. (W) (R4)	F2, F4, F5	R2, R3, R4
	for everyone	Our vision is to create a social and environmental impact for millions of people and our planet. (P) (F3)	F1, F3	R1, R3, R4, R5
Personal closeness	from being affected	Nursing care and especially the current nursing emergency affect us all, sooner or later. (W) (F1)	F1, F2	R1, R2, R4
	from working abroad	We discovered our passion for solar energy and African culture during two years on the ground. (P) (F3)	F2, F3, F4	
	from travelling	After six months of travelling and searching through Southeast Asia, we opened our online store. (M) (F2)	F2, F3	R2, R3, R4, R5

Sources: (I)nterview with the author; (M)edia articles; (P)itch decks or (P)resentations; (W)ebsite of the start-up. All German quotes translated by the author.

tional drivers, the following section explores which resources social start-ups require to pursue a double bottom line – and how incubators and accelerators can best support them.

4.2. Resource Needs

One way to conceptualize the relationship between incubators and social start-ups is the resource-based view of the firm (RBV). This theory is used "to investigate how the deployment of key resources," such as business and social support, "changes during the lifecycle development of the small entrepreneurial firm" (McAdam & McAdam, 2008, p. 278). According to the RBV, organizations require unique resources and capabilities, both tangible and intangible, to create a sustained competitive advantage (Barney, 1991; Grant, 1991; Wernerfelt, 1984). Strategic resources are defined as basic inputs owned or controlled by the firm that are valuable, rare, hard to imitate, and difficult to substitute (Barney, 1991, pp. 105-106), while capabilities involve "complex patterns of coordination between people and between people and other resources" (Grant, 1991, p. 122). When it comes to acquiring resources, the entrepreneurship literature has primarily focused on the ability of firms to leverage them internally (Dacin et al., 2010, p. 48).

However, young organizations facing the liabilities of

newness (Stinchcombe, 1965) and smallness (Aldrich, 1986) find it difficult to control and expand their internal resource base. The RBV suggests that creating a resource-rich environment can address these liabilities. Incubators, for example, can support start-ups by providing a flow of tangible and intangible resources (Carayannis & Von Zedtwitz, 2005; McAdam & McAdam, 2008; Rothaermel & Thursby, 2005). The literature on sponsorship describes this process as deliberately increasing the level of resources available to new firms (Flynn, 1993b, p. 57) and mediating the relationship between resources and their founding environment (Amezcua et al., 2013, p. 1632).

Yet, studies have also shown that the provision of resources is not a unilateral process. Placing resources at the disposal of start-ups is not sufficient to support their growth. Cultural incompatibility or difficulty to absorb intangible resources can complicate resource acquisition (Becker & Gassmann, 2006). Rice (2002) called the process of business assistance in incubators a coproduction, whereas Van Weele et al. (2017) found that entrepreneurs fail to take full advantage of incubator resources when they are unaware of resource gaps. In short, it makes sense to explore how social start-ups define their resource needs before exploring the support mechanisms.

To do this in a structured way, this study used the classification by Van Weele et al. (2017, p. 19), who distinguish between two tangible resources ("physical and financial capital") and three intangible ones ("knowledge," "social capital," and "legitimacy"). Based on the interview results, the classification was adapted by adding "training" to the "knowledge" category, as well as "personal support" as an intangible resource to include coaching and critical sparring. "Legitimacy," which was not mentioned in the interviews, was dropped as a second-order code.

Two further explanations regarding Table 9: First, the interview question specifically addressed the support expected by an incubator (see Appendix B), not the overall resource needs of a start-up. Second, only ramp-ups were asked how their resource needs changed after the end of the program since fellows enjoyed a fluid relation to the accelerator with no definite end date.

To lay the foundation for the detailed assessment of sponsorship mechanisms in the following section, four aspects of Table 9 are discussed: within-group commonalities, reliability, longitudinal variation, and receptiveness to external support.

The first aspect are the shared characteristics of the start-ups in this study, which presumably affected their resource needs: a for-profit legal form, a social mission combined with commercial activities, and in most cases an experienced founding team with a business background. Hence, it was of interest to examine how the resource needs expressed in Table 9 compared to the wider start-up population. In fact, the two intangible resources most frequently mentioned in Table 9 – "knowledge and training" and "social capital" – are also often mentioned in studies on start-up resource needs: Van Weele et al. (2017) have found that business knowledge is one of the most significant resources provided by an incubator. Similarly, relational assets, or "social capital," are highly relevant resources for social ventures (Dacin et al., 2010; Mair & Marti, 2006). Within the "knowledge" category in Table 9, the resource needs were heterogeneous, except "marketing and sales" with six mentions. This was not surprising: Three-quarters of accelerator companies claim that "not understanding their target market" and "difficulties reaching their customers" are the greatest obstacles for a new venture, next to funding (Radojevich-Kelley & Hoffman, 2012, p. 64). Surprisingly, the founders mentioned "financial capital" and "physical capital" less frequently. By contrast, German social entrepreneurs named funding as a key resource constraint in a recent survey (DSEM, 2020, p. 65), just as other studies have identified accessing tangible resources as the most important reason for joining an incubator (McAdam & McAdam, 2008; Van Weele et al., 2017). Rather than indicating overall resource needs, it appears that Table 9 signals the resources expected from a support program. Furthermore, this result may reflect a broader trend in the incubator landscape, given that their value proposition is shifting from tangible resources to networks, knowledge, and legitimacy (Bruneel et al., 2012).

The discrepancies between Table 9 and other studies lead to a second aspect: How reliable are the views of founders in determining the value of support mechanisms? Start-ups sometimes struggle to determine their resource needs, which can reduce the effectiveness of incubator programs. They are "hesitant to step out of their comfort zone" (Van Weele et al., 2017, p. 26) and tend to experience problems of bounded rationality, such as incomplete and inaccurate information, while their decision making can be affected by cognitive biases (Cohen, Bingham, & Hallen, 2019). These factors may explain why the resource needs articulated in Table 9 diverged from the assessment of incubator and accelerator services presented subsequently in Section 4.3.1.

Third, there was longitudinal variation between the resource needs expressed by the ramp-ups after the formal program (except for R1, who decided not to pursue the venture he entered the program with) compared to their responses at the start. Tables 9 indicates a need for more specialized support as start-ups mature, for example regarding international expansion (R4) or employment laws (R5). Similarly, R2 and R3 expressed a desire for more targeted networking. These responses suggest a change in resource needs over the duration of the program (five months) – a finding echoed by Casasnovas and Bruno (2013) and Drori and Wright (2018), who have noted that support needs of social start-ups progress with their stage of development.

The final aspect is the relation between the resource needs of social start-ups and their receptiveness to external support. In theory, the combination of resource constraints and the complexity required to pursue a social and economic mission in parallel should make the founders of social start-ups "likely to be receptive to assistance from external parties such as social accelerators" (Pandey et al., 2017, p. 8). But although ramp-ups and fellows displayed similar resource

Table 9: Resource Needs

Second order code	First order code	Representative quotes	All mentions	
		At the beginning of the program		
Knowledge and training	business model	We hoped to receive constructive feedback on our business model. (F1)	F1	R3, R5
	founding	What all three of us haven't done yet is to set up our own company.		R4
	legal issues (setting up)	For us there were many questions at the time, such as corporate law issues and which legal form to choose. (R5)	F4	R3, R5
	marketing and sales	We looked at a different market when we went into the Impact Factory, but we validated the market and came to a pivot. (R3)	F1, F3, F4	R1, R3, R5
	operations	We plan to develop cloud services in the sales and after-sales area.	F3	
	product	We want an app to make our service more user-friendly. (R3)		R3, R5
	strategy	We hope for new ideas for the strategic direction of our start-up.	F4	
Personal support	coaching	We hope that coaching will give us valuable advice. (R5)	F3	R4, R5
	critical sparring	I am often on my own. I need sparring partners, I need people who make me want to get better. (R2)	F2, F5	R2, R3
Social capital	networking	The biggest need at the beginning was actually to find contacts, contacts and contacts. (F1)	F1, F2, F3, F5	R1, R2, R4
	partnerships	The establishment of sustainable partnerships of mutual benefit is desirable for us, especially from a sales perspective. (F3)	F2, F3, F4	R1, R5
	peer exchange	We would like to use the Impact Factory ecosystem to meet like-minded people and learn from the experiences of others. (F5)	F2, F3	R3, R5
Financial capital	seed financing	The number one topic is financing. How do I get in touch with investors? Door-knocking won't be sufficient. (F3)	F1, F3	R1, R3
Physical capital	office space	We need places to work.	F1	
		At the end of the program (ramp-ups only)		
Knowledge and training	international expansion	We need an accelerator that is active in various markets, because our entire business model is built on expanding relatively quickly.	R4	
	industry updates	I would be interested in being kept up to date, more than in a larger program in which perhaps a third of the content is duplicated. (R2)	R2, R3, R4	
	legal issues (tax, labor)	We face new issues such as tax law or labor law. You just develop further and these matters become more pressing.	R5	
Social capital	targeted networking	Today the network is much more important to us, but we also know how to get intros ourselves or get in touch with other founders. (R3)	R2, R3	
Financial capital	growth financing	We also know that growth capital plays an important role in our platform model. (R4)	R3, R4, R5	

Note. All quotes are taken from the interviews and translated from German by the author. Second order codes adapted from Van Weele et al. (2017).

needs, the first group decided to join a formal program, whereas the second chose an affiliation that provided them with a fraction of incubator benefits.

The observation that some start-ups opt for an entrepreneurial support program (or multiple ones), while others with comparable resource needs prefer to go it alone,

is further underlined by Table 10, which shows no clear pattern even within the two subgroups. Among the different factors affecting a start-up's decision to join an incubator, such as venture stage, mission fit, selection policies, services provided, and the partner network (Chmiliar, 2010), this study focused on one aspect in particular – the services provided. Consequently, the upcoming section explores how the mechanisms of sponsorship address the resource needs of social start-ups shown in Table 9. The trade-off between organizational sponsorship and entrepreneurial self-reliance is examined subsequently in Section 4.4.

4.3. Organizational Sponsorship Mechanisms

As illustrated in the analytical model for the acceleration of social start-ups (Figure 1 in Section 3.5), this study first presented the social-mission focus and founding motivations of 10 social start-ups. It then explored the resource needs required to pursue their social missions and concurrent economic activities. The following section now discusses how incubators and accelerators address the resource needs of social start-ups.

In addition to the three mechanisms of organizational sponsorship and their underlying interventions, this study suggests that a fourth mechanism, impact acceleration, is particularly relevant to social start-ups. Before studying the individual mechanisms, this section explores how social start-ups rank the benefits of incubator support services.

4.3.1. Ranking of Sponsorship Interventions

This study aims to create a better understanding of incubator and accelerator interventions, in particular from the perspective of social start-ups. Therefore, it compiled an exemplary services portfolio based on studies of conventional incubators and accelerators and their social counterparts (Table 2 in Section 2.4; see Appendix for the full table). The resulting list of nine services was used in the interviews to discuss the activities of incubators and accelerators. The bulk of the interviews used open-ended questions to explore how specific services were perceived. Yet, the founders were also asked to rank the services according to their overall importance for social start-ups. Due to the small size of the sample and its selection based on theoretical considerations, the purpose of this exercise was not to provide a quantitative assessment of sponsorship services. Rather, the aim was to highlight how the founders in this sample perceived and prioritized them.

As Table 11 reveals, the resulting ranking looked surprisingly similar for both subgroups. "Access to external networks" was rated as the most valuable service by ramp-ups and fellows alike. "Internal mentoring" and "peer networking" were also ranked highly across the sample, whereas "milestones and progress tracking" ended in the bottom ranks. In accordance with the resource needs expressed in Table 9, the three services related to tangible resources – "access to external funding," "coworking space," and "seed funding" –, scored lower than the ones providing intangible resources.

The largest discrepancy between the two subgroups was observed for "validation and visibility," which ranked second for fellows and last for ramp-ups. By contrast, ramp-ups attributed greater value to "education and training" than fellows. In their assessment of these two services, the two subgroups reflected the different priorities and contents of the ramp-up and fellows programs of the Impact Factory.

Regarding the mechanisms of organizational sponsorship, Table 11 suggests a slight preference of fellows for the bridging mechanism, with two outward-facing interventions in the top three. Ramp-ups, perhaps owing to their reduced internal resource base, rated the buffering mechanism slightly higher. Overall, the mechanisms look evenly distributed. The most striking result concerns the provision of tangible and intangible resources.

Comparing the results in Table 11 to the extant literature is challenging for three reasons already mentioned in Section 2: Studies on incubators and accelerators usually focus on their definitions and effects, and only rarely on their activities. If they do, they often take the perspective of incubator managers, and not of start-ups. Moreover, there have been few attempts to generalize the activities of incubation and acceleration beyond individual case studies.

Studies that do attempt to study incubator services usually employ a rough classification, as three widely cited articles demonstrate: A study on the effect of accelerator services on venture performance compared "basic services of funding and coworking space" with "entrepreneurial schooling" (Gonzalez-Uribe & Leatherbee, 2018). Research by Bruneel et al. (2012) showed that the usage of "business support" increased in recent incubator generations, while Bergek and Norrman (2008) distinguished between "strong intervention" and a "laissez-faire regime".

The broadness of these terms reflects the empirical diversity of incubators and accelerators. The review by Crisan et al. (2019, p. 12) counted 45 "typical interventions" and 36 "extended interventions," the latter referring to "additional services based on participants' needs". These interventions were aggregated into five outcomes (p. 13): "Funding" was mentioned most frequently (in 52% of the reviewed studies), followed by "validation" (40%), "product development" (37%), "network" (33%), and "knowledge" (32%). While these figures refer to prevalence, and not to their value for start-ups, it is interesting that the literature mentions tangible resources such as "funding" more frequently than the respondents in this study.

In contrast, Crisan et al. confirmed the importance of "networking," which the respondents ranked as the most valuable service – in line with a study of five U.S. accelerators that cited "networking" as their most significant benefit (Radojevich-Kelley & Hoffman, 2012). A study of 88 Italian incubators also found that "networking," with "managerial support" and "physical spaces," was among the most important incubator and accelerator services (Sansone et al., 2020). Another relevant data set confirmed the importance of networking: The application data of 23,368 early-stage ventures suggested that entrepreneurs placed the highest

Table 10: Experience with Additional Support Programs

Ramp-ups	Program(s)	Fellows	Program(s)
R1	None	F1	Fraunhofer accelerator; two technology accelerators; EU incubator with IT focus
R2	Food-specific accelerator	F2	None
R3	Technology accelerator	F3	Start-up competitions; university start-up program; regional founders network
R4	None	F4	Early support by a start-up center, but no formal program
R5	Start-up competitions; university start-up program; two accelerators	F5	None

Note. The results refer to support programs prior or in parallel to the Impact Factory.

Table 11: Ranking of Sponsorship Interventions

Rank	1	2	2	4	4	6	7	8	9
Intervention	Access to external networks	Validation and visibility	Internal mentoring	Peer networking	Education and training	Access to external funding	Co-working space	Milestones and progress tracking	Seed funding
					Start-ups of the fellows program				
F1	1	4	3	5	2	8	7	9	6
F2	2	5	4	1	3	7	8	6	9
F3	1	6	4	7	5	2	3	9	8
F4	2	1	3	4	5	9	8	6	7
F5	1	2	4	3	5	8	9	6	7
Ø	1.4	3.6	3.6	4	4	6.8	7	7.2	7.4
Mechanism	Bridging	Bridging	Buffering	Boosting	Buffering	Bridging	Buffering	Boosting	Buffering
Rank	1	2	3	4	5	6	7	8	9
Intervention	Access to external networks	Education and training	Internal mentoring	Peer networking	Access to external funding	Seed funding	Co-working space	Milestones and progress tracking	Validation and visibility
					Start-ups of the ramp-up program				
R1	3	4	2	6	1	n/a	n/a	7	5
R2	1	2	3	4	6	8	7	5	9
R3	1	2	3	5	4	8	7	9	6
R4	2	4	6	1	5	3	8	7	9
R5	1	4	3	2	7	6	5	8	9
Ø	1.6	3.2	3.4	3.6	4.6	6.25	6.75	7.2	7.6
Mechanism	Bridging	Buffering	Buffering	Boosting	Bridging	Buffering	Buffering	Boosting	Bridging

Note. After the interview, founders were asked to rank nine interventions from 1 (most relevant for social start-ups) to 9 (least relevant for social start-ups). R1 decided not to rate two interventions (given as n/a).

priority on building external relationships, including "connections to funders" and "mentorship" (Global Accelerator Learning Initiative, 2020). The data set by GALI also indicated that "awareness and credibility" and "access to like-minded entrepreneurs" ranked lowest among potential benefits – unlike the views expressed in this study, especially by fellows. A similar discrepancy to Table 11 could be observed in a survey of 14 European impact-oriented incubators: "Revenue strategy," "financial management," and "fundraising" were considered the most important types of support after "strategic support" (Gianoncelli et al., 2020). Finally, a survey of 52 impact-focused accelerators found that the three most significant accelerator benefits were "mentorship," "access to potential investors," and "network" (Lall et al., 2013, p. 118).

Summing up these results, the following trends emerge:

"Networking" and "mentoring" were consistently ranked among the most important services in this study – and in the wider literature. The picture was less clear for "education and training" and "peer networking," which ranked higher in Table 11 compared to other surveys. However, the largest discrepancy was registered in relation to the provision of tangible resources such as "funding" and "coworking." These activities scored surprisingly low in Table 11, despite their prominence in the literature.

4.3.2. Buffering Mechanism

The framework of organizational sponsorship defines the buffering mechanism as the provision of external resources to young organizations that are "understood to lack control over vital resources" (Amezcua et al., 2013, p. 1632). Incubators and accelerators help new organizations build internal resources until they can mobilize them from their environment (Amezcua et al., 2013, p. 1632). The buffering mechanism originally focused on material support, such as office space or consulting services (Breivik-Meyer et al., 2019, p. 8). This study extended the mechanism to include "mentoring," in addition to three services typically used to increase the internal resource base of start-ups: "education and training," "seed funding," and "coworking space." The following section explores how the founders assessed the value of these four services.

Education and Training. The provision of education and training is one of the most prevalent activities of incubators and accelerators. In contrast to mentoring, this activity aims to impart knowledge and skills to start-ups. It can be offered internally, with the help of external experts, and through various formats such as workshops, lectures, and individual training. It can cover various topics, from business skills to legal, operational, and financial issues, either with a standardized curriculum or tailored to the needs of founders.

In the exemplary services portfolio compiled for this study, 25 of 26 studies mentioned activities related to education and training (see Appendix). Moreover, the start-ups in this study mentioned knowledge and training as the most relevant resource need (Table 9 in Section 4.2). Yet, despite this prominence, they did not consider education and training the most important service provided by incubators, as Table 11 shows: Ramp-ups ranked this service in second place, below the access to external networks. Fellows placed it fourth, on the same level as peer networking – not entirely surprising, given their more advanced development stage.

Studying the resource needs in Table 9 leads to two interesting observations regarding education and training: First, the knowledge interests of social start-ups were rather broad, with most topics mentioned only by one or two start-ups. Second, their education needs evolved as the start-ups advanced in their development. How can incubators and accelerators succeed in providing an education and training program that fulfills these diverse requirements?

To answer this question, Table 12 divides the responses of the social start-ups into four categories: what (the desired content), how (the preferred methods), why (the added value), and why not (the limitations). Although all founders were asked about this activity, the latter two categories relied primarily on the views of ramp-ups who, unlike fellows, participated in the formal education program of the Impact Factory.

The topic that was mentioned least frequently was the content of education and training – presumably because the founders had already addressed their knowledge gaps in terms of their resource needs (Table 9). However, half the respondents mentioned one specific area: providing legal advice. Multiple founders said that choosing the legal form is a major decision for a social start-up that can incorporate as a for- or nonprofit entity. Two start-ups (R2 and R3) even changed their legal structure to a "purpose company" on the advice of the Impact Factory.

There was mostly consensus regarding the preferred method of learning in an incubator or accelerator. First, four founders said that the education program needed to be tailored to their specific needs. Second, the education program should be practice-oriented, so replicating or even solving genuine business challenges. Third, outside experts should be included in the education offering. Apart from that, the founders expressed no strong views on whether the education should be provided individually or in groups.

How can an incubator or accelerator then add value through its education and training program? According to the respondents, three activities are key: structuring and speeding up learning processes, solving concrete business challenges, and relating theoretical knowledge, such as accounting practices or financial planning, to the realities of running a start-up. Furthermore, an incubator should ensure that the knowledge it provides represents the state of the art, which even experienced founders such as R1 regarded as helpful.

Simultaneously, all five ramp-ups mentioned limitations. Education and training activities can be useless, if not detrimental, when they are too basic, generic, or theoretical – a logical reversion of the learning preferences expressed in Table 12. They also repeatedly mentioned the time invested as a crucial factor in deciding whether education and training were beneficial.

Summing up these results, three aspects are noteworthy. First, the founders considered education and training helpful overall – in line with the finding that entrepreneurship schooling leads to "significantly higher new venture performance" of accelerator companies (Gonzalez-Uribe & Leatherbee, 2018, p. 1595). Second, human capital – skills, knowledge, and experience – play a key role in how founders appreciate training benefits, as Pandey et al. (2017) have previously shown. In this study, the founders were mostly experienced, with a strong background in business but limited prior knowledge of starting a social-mission venture. Hence, their training needs primarily centered on catalyzing internal learning processes, rather than on receiving business education. Third, these results reflected the views of founders who might have suffered from biases and bounded rationality. For example, it has been found that tailoring accelerator activi-

Table 12: Buffering 1 - Education and Training

Second order code	First order code	Representative quotes	All mentions	
What: content	ecommerce	It would have been easier if we had someone to explain to us at the beginning how a web-shop works and what the best system is that still works as you grow. (F4)	F2, F4	
	business plans	When you start, you go in with a business plan. But the reality is completely different from what you planned. (R2)		R2
	legal advice	Legal is always a big issue. You cannot just read up on it yourself. You don't want to build on a fragile foundation. (R5)	F1, F5	R2, R4, R5
How: methods	customized	Very basic workshops are no longer relevant for us. It's definitely good to have a broad spectrum. The teams are diverse and depending on your needs, you should be able to pick and choose what you need. (F3)	F1, F3	R2, R5
	individual	Exchange is important, but sometimes it adds more value if you really have a specific 1-to-1 conversation.		R5
	collective	It's about doing it directly in dialogue with others, in exchange, in group work. (F1)	F1, F4	
	with outside expertise	The people who led these workshops were really committed and available afterwards. That is really positive. As a start-up, you can't always pay a professional. (R1)	F2, F3, F4	R1
	practical	This is an incubator where you are given a very specific task, which is like building up a new business model. (F1)	F1	R3, R5
Why: added value	learning catalyst	We noticed relatively quickly we had already dealt with all these topics in the past, but the Impact Factory gave us templates of practical use and we were able to organize our thoughts better and improve our processes. (R4)	F2	R4
	problem solver	If you're a founder, the only thing you're looking for are problem solvers for topics that are relevant for you. And if you have one, don't let them go. (R2)		R2, R3, R4
	reality check	I believe that you can learn many basics like accounting on your own if you want to learn that. But to get your product, your offer, your service across to someone else - you don't learn that from a book or an online course. (F1)	F1	R1, R2
	topicality	I studied business administration, but that was a while ago. There are some things that were done differently in the past. And in the Impact Factory they offer workshops that bring you up to date on the latest developments.		R1
Why not: limitations	too basic	The workshops don't always help us one hundred percent because we have already seen a lot. Especially when you do a more general workshop, for example on marketing: we've heard that eight or ten times now. (R5)	F5	R3, R4, R5
	too generic	A standard program, where I already know half of it, is a waste of time for me, which I don't have. Flexibility is extremely important. (R2)		R2, R3, R5
	too theoretical	Of course, we can always get more information. But everything to build up the company, we have to do ourselves. (R3)		R1, R3
	too time-consuming	Maybe these things were offered, but because we are always so busy, maybe they simply slipped through. (F4)	F4	R2

Note. All quotes are taken from the interviews and translated from German by the author.

ties can lead to lower venture performance than standardized offerings (Cohen, Bingham, & Hallen, 2019).

Internal Mentoring. Connecting founders with experienced mentors, often over extended periods, is one of the most significant activities for supporting start-ups (Miller & Stacey, 2014, p. 26). Today, mentoring is mostly associated with accelerators. Pauwels et al. (2016, p. 17) go as far as calling it the service that "most differentiates the accelerator from previous generation incubation models". Indeed, mentoring featured in over 80% of the studies on accelerators and social incubators/accelerators in the exemplary services portfolio (see Appendix). By contrast, none of the studies on incubators mentioned mentoring activities.

Admittedly, the difference between mentoring and related activities, such as coaching and business support, is not clear-cut. In this study, mentoring comprises activities that do not aim to impart technical or business knowledge, but seek to provide "feedback, advice and social support" (Cohen, Fehder, Hochberg, & Murray, 2019, p. 1791). Mentoring is "internally" delivered – formally and regularly – within the accelerator, as opposed to external networking activities.

The social start-ups in this study considered mentoring one of the key accelerator benefits. Seven of the 10 founders identified personal support as a crucial resource need in Table 9. Ramp-ups and fellows ranked internal mentoring as the second and third most important service in Table 11, respectively. Therefore, Table 13 summarizes their views on internal mentoring.

As far as the profile of mentors was concerned, the respondents had clear preferences: Mentors should be experienced, preferably having successfully founded a start-up themselves, and knowledgeable about the industry the start-up planned to enter. R5 explicitly cautioned against mentors for whom "the needs of start-ups are just too far away."

The expectations for the mentoring relationship focused on three aspects. The most frequently mentioned was reflection. Interestingly, the three oldest founders (F4, F5, and R1) primarily expected the mentor to ask the right questions, whereas two younger founders (R3 and R5) would like a mentor to accelerate their growth by setting goals and key milestones. The importance of acceleration is explored further as part of the boosting mechanism, which includes "milestone setting." Besides, mentors should provide emotional support, as R3 mentioned.

Considering that the view of internal mentoring is quite positive both in this study and in the literature (Casasnovas & Bruno, 2013; Cohen, Fehder, et al., 2019; Radojevich-Kelley & Hoffman, 2012), it is worth studying risk factors for mentoring relationships. For instance, R4 said that mentors might not follow the development of a start-up closely enough, while R3 saw a gap between the growing sophistication of their support needs and generalist mentors. More surprising were the views of two fellows who had declined to join a formal program: F4 observed that experts from other sectors rarely added value due to the complexity of F4's activities. F2 signaled a reluctance to be challenged too early by a mentor. In their cases, even a popular service such as internal mentoring failed to add value. However, except F2 and F4, all founders in this study appreciated internal mentoring as highly beneficial.

Coworking Space. Offering coworking spaces at reduced costs to entrepreneurs has been a key feature of business incubators since their emergence in the 1950s (Cohen, Fehder, et al., 2019, p. 1792). Although recent incubator generations have shifted from providing infrastructure to networks, learning processes, services, and capital (Bruneel et al., 2012), subsidized workspaces remain prevalent in entrepreneurial support programs. Accordingly, 90% of studies on incubators and 60% of studies on accelerators and social incubators/accelerators mentioned this service in the exemplary services portfolio (see Appendix).

Despite the prevalence of coworking spaces, the findings on their benefits are mixed at best. Gonzalez-Uribe and Leatherbee (2018, p. 1569) have seen "no evidence that basic accelerator services of cash and coworking space have a treatment effect on fundraising, scale, or survival". Another study has found that workspace provision is "associated with lower performance in terms of maximum valuation" (Cohen, Fehder, et al., 2019, p. 1795). Y Combinator intentionally declined to provide space to avoid unhealthy codependencies and encourage independence (Cohen, Fehder, et al., 2019, p. 1792). However, studies have also highlighted the benefits of shared spaces for peer exchange (Miller & Stacey, 2014, p. 29) and overcoming the loneliness related to entrepreneurship (Duff, 1994, p. 17).

For the social start-ups in this study, the provision of coworking space did not appear to play a significant role. They mentioned it only once as a resource need in Table 9. In the ranking of sponsorship services (Table 11), both ramp-ups and fellows ranked coworking space seventh out of nine. However, it should be noted that the Impact Factory in 2020 had to offer most services virtually due to Covid-19. Before the pandemic, on-site activities were scheduled in bi-weekly blocks to allow start-ups from all over Germany to attend.

As a result, none of the start-ups used the coworking space offered by the Impact Factory, although half signaled an intention to use it if distance or time allowed, as Table 14 shows. Contrary to a study that found rental subsidies to be the main attraction of incubator programs (K. Chan & Lau, 2005), only two founders mentioned cost considerations. The primary motivations for co-locating were immaterial, such as collaboration, creativity, and the shared experience, or as F1 put it: "Being together, suffering together, but also celebrating together."

Coincidentally, the only start-ups that expressed an outright negative view of coworking, F4 and R1, were both founded in 2015, so three years before the other start-ups in this study. This observation suggests that the maturity of a start-up affects its perception of coworking benefits.

Seed Funding. Securing financial resources to launch and grow is one of the most vital tasks for a start-up founder. Incubators and accelerators can help by offering two types of funding: Directly, through seed funding in the form of grants, debt, equity, and hybrid instruments, and indirectly, by pro-

Table 13: Buffering 2 - Internal Mentoring

Second order code	First order code	Representative quotes and further mentions
Desired mentor profile	experienced	For me, mentors are similar to professors. They have done it all before. (F1, also mentioned by R1 and R5)
	knowledgeable	It's important to have a contact person who can guide you with industry knowledge. (F3, also mentioned by R4)
	ask questions	Ask critical questions (F4) / Ask good questions (F5) / Ask uncomfortable questions (R1)
	reflection	Help prioritize and provide outside view (F2) / Serve as reality-check (F1)
Expectations	acceleration	Define KPIs (R3) / Set goals (R5)
	support	Point out risks (R2) Address the emotional level and absorb shocks (R3)
	lack of closeness	Things developed so fast in our start-up and the mentors also have different things on their mind. And you don't meet them every day. (R4)
Risk factors	lack of knowhow	Asking the right questions is not always so easy. (F4)
	increased expectations	Talking about general topics is more important at the beginning. Then the intervals at which you meet become greater because you become better at it. (F3)
	reluctance to be challenged	I ask myself: when will the point come when I need to enter an incubator program? When do I need a sparring partner to challenge my business idea? Otherwise, it can make sense to simply pursue an idea and develop it if it works. (F2)

Note. All quotes are taken from the interviews and translated from German by the author. Some quotes are summarized when the full quote added no necessary context.

Table 14: Buffering 3 – Coworking Space

Second order code	First order code	Representative quotes and further mentions
Potential benefits	collaboration	It's great if you have a space where you can go, where you can work together. (R3, also mentioned by F1 and R5)
	creativity	Just sitting together, brainstorming something new, to challenge ourselves and even get a bloody nose sometimes. (F1, also R2)
	cost savings	It's very difficult to get coworking places. Sometimes they cost a fortune. (R5, also R2)
	shared experience	Start-ups live from being together, from suffering together, but also from celebrating together. (F1, also mentioned by F3)
Reasons against	distance	I was barely there, because Duisburg is two and a half hours from my home. (R2, also R3, R4 and R5)
	time	I was a bit jealous of the start-ups in the Impact Factory. I would love to do that but maybe I don't have the time anymore. (F1)
	no added value	We are often offered shared office space to exchange ideas with, for example, people who make websites. But that never helped us. (F4)
	preference for own facilities	We are simply too well equipped in that respect already. And even in the initial phase, I had my own rooms at home. (R1)

Note. All quotes are taken from the interviews and translated from German by the author.

viding access to external investors such as business angels or VCs. As part of the internal buffering mechanism, this section now focuses on the former, while connecting the start-ups with external investors is explored subsequently as part of the bridging mechanism.

Although the provision of financing has been identified as a primary motivation for joining an incubator (SIM, 2020; Van Weele et al., 2017), it is not that common. Less than a third of impact-oriented accelerators in Europe offer funding in the pre-acceleration phase (Gianoncelli et al., 2020, p. 22). 8% of German incubators and accelerators take an equity stake in their ventures (SIM, 2020, p. 44). In the exemplary services portfolio, 40% of studies on conventional incubators and accelerators mentioned direct funding (see

Appendix).

The low prevalence of seed funding offered by incubators may explain why only four founders named seed financing as a resource need to be addressed by an incubator (Table 9). Moreover, fellows ranked seed financing ninth out of nine and ramp-ups sixth out of nine in the ranking of sponsorship services (Table 11) – in line with the finding by Miller and Bound (2011, p. 26) that "the money that accelerator programs offer is a valuable part of the package. . . but it was rarely rated the most important consideration". Yu (2020) has even argued that the cost of dilution leads founders with the best ideas not to apply to accelerators.

If social entrepreneurs do not consider incubators and accelerators a primary source of funding – how do they finance their seed phase? Table 15 shows that for eight founders in this study, the answer was bootstrapping, i.e., drawing on their savings or reinvesting their first revenues. Only two founders worked with external investors: F5 could rely on business angels from the start and R4 after six months. The other founders used a diverse array of funding sources, including grants, personal loans, donations, and start-up competitions.

Considering that social start-ups "often rely heavily upon a range of funding sources" (Austin et al., 2012, p. 377), these results are not surprising. A survey of social entrepreneurs in Germany has found that 73% tap into their savings, 31% rely on friends and family, and 23% on public funds; only 9% receive funding from business angels and 8% from incubators or accelerators – a far lower share than conventional start-ups (DSEM, 2020, p. 46).

The responses of the founders in this study suggest that the low take-up of external seed funding by social start-ups may not only be related to supply but also to demand. Bearing in mind that the Impact Factory did not provide direct funding to its ventures, Table 16 shows that the founders had a balanced (R4), if not critical (R3) view regarding this source of financing.

More importantly, Table 16 illustrates how the social-mission focus influenced their perspectives on seed funding in general. As discussed in Table 7, most founders signaled a "priority of purpose," meaning that social purpose takes priority over economic goals. Table 16 provides concrete examples of how the founders financed the launch and growth of their start-ups: They worked for low (or no) salary – sometimes for years. The founders had additional jobs, used savings from previous careers, and asked for the help of friends. Together, these views show the resilience and creativity of social entrepreneurs in gathering seed funding – irrespective of any funding support by an incubator or accelerator.

4.3.3. Bridging Mechanism

Incubators and accelerators do not only help start-ups expand their internal resource base through buffering – they also facilitate inter-organizational relationships through bridging, "establishing a conduit through which essential resources can flow more efficiently between external resource providers and new organizations" (Amezcua et al., 2013, p. 1633). From this perspective, the environment is not seen as a threat but as a source of resources and knowledge for improving "the competitive positions and survival chances" of nascent organizations (Amezcua et al., 2013, p. 1633). Sponsors such as incubators and accelerators can serve as connective intermediaries to strengthen a start-up's external relationships – and increase not only its social capital and legitimacy but also its financial resources (Breivik-Meyer, 2020). This section presents three incubator and accelerator benefits aimed at achieving these outcomes: "access to external networks," "validation and visibility," and "access to external funding."

Access to External Networks. The provision of social capital, defined as "the aggregate of resources embedded within, available through, and derived from the network of relationships possessed by an individual or organisation" (Inkpen & Tsang, 2005, pp. 150-151), has been identified as a key benefit of organizational sponsorship since its first formulation (Flynn, 1993b). Relational resources "provide opportunities to exchange information, leverage interpersonal relationships, and realize objectives" (Dacin et al., 2010, p. 50). This, in turn, requires incubation programs to "be linked into the right types of networks" (Miller & Stacey, 2014, p. 28).

The exemplary services portfolio confirmed the significance of networking, since "networking" emerged as the most prevalent bridging intervention (see Appendix). Interestingly, studies on social incubators and accelerators mentioned "networking" more often (90%) than studies on their conventional counterparts (70%) – perhaps an indication that social start-ups, operating in a resource-constrained environment and seeking to achieve systemic change, are more dependent on social capital than other start-ups. Indeed, a study has suggested that social ventures use relational resources differently than conventional entrepreneurs – not "to set up competitive barriers," but in a "cooperative fashion" (Dacin et al., 2010, p. 50).

Among the social start-ups in this study, there was consensus that networking is crucial – and that incubators had a key role in supporting it. Accordingly, "networking" was identified as the single most valuable resource need with seven mentions (Table 9), as well as five mentions of "partnerships." In the ranking of sponsorship services, "access to external networking" topped the list for both ramp-ups and fellows (Table 11), with nine of the 10 start-ups naming it as their first or second priority. But what are social start-ups aiming to gain from this intervention?

The results in Table 17 suggest that the Impact Factory helped start-ups connect with a multitude of external stakeholders – from other companies, as partners or customers for B2B start-ups, to industry experts and service providers, such as marketing agencies. Four respondents also considered the program an effective bridge into the nonprofit sector. The funders of the Impact Factory were themselves seen as crucial networking targets.

Moreover, the respondents named a wide range of benefits of networking. Incubators can support start-ups by helping them enter new markets and industries, build new rela-

Table 15: Buffering 4.1 – Experiences with Seed Funding

	Boot-strapped	Additional funding		Boot-strapped	Additional funding
F1	Yes	Research grant	R1	Yes	Donations
F2	Yes	No	R2	Yes	No
F3	No	Local and national public grants; grant from the Federal Employment Agency	R3	Yes	No; seeking seed investment
F4	Yes	Personal loan from a friend	R4	Yes	Business angels
F5	No	Business angels (personal friends); seeking seed investment from VC	R5	Yes	Small start-up competitions

Note. All information taken from the interviews. "Additional funding" refers to the acquisition of outside funding in the first 12 months of the start-up's operations (personal savings are not included).

Table 16: Buffering 4.2 – Perspectives on Seed Funding

Second order code	First order code	Representative quotes and further mentions
Seed funding (by accelerators)	positive view	You would strengthen the tie between the start-up and the incubator, because the start-up also says: Okay, I'm going to get more involved. (R4)
	negative view	I always find that difficult, because especially at this early stage you might not even want to think about giving away shares. (R3) It is almost too late now. It could be a problem and a reason not to go into the accelerator because we already have a high valuation. (R4)
Seed funding (by investors)	business angels	We have started a convertible loan round with business angels and have started to work with business angels and finance ourselves through them. (R4)
	business angels (friends)	We founded F5 with our own money, but very quickly we took three business angels on board who were all basically our friends. (F5)
Seed funding (other sources)	personal loan	We have no external support, only a personal loan from a friend of ours. (F4)
	private savings	I financed the starting capital myself, through my savings. I do not want any dependence because of capital. (R2, also R1 and R3)
	grants	Foundation grants (R1); public grants (F3); research grant. (F1)
	supplementary income	We have financed it from our own resources. I always work on the side, I do 2 or 3 freelance jobs from time to time to earn a little bit. (R3, also R5)
	voluntary salary waiver	We did it for free for a year and a half, and even back then we said to ourselves: The impact must always come first. If that means that we can't pay out a cent for the first two or three years, then that's perfectly fine. (R5, also R1 and R3)

Note. All quotes are taken from the interviews and translated from German by the author.

tionships, and generate new business. The results indicated that the aim of this intervention is not simply to provide resources, but to empower start-ups to build up their required resources internally, as reflected in the benefits "accelerate new learning" and "provide targeted introductions."

Not surprisingly, given their uniformly positive views of networking, the founders named only two external limitations to this activity: Covid-19 and strong reservations in a particular sector. However, the resource needs presented in Table 9 also suggest that social start-ups become more selective in their networking needs as their development advances – in line with the finding that successful social entrepreneurs differ from struggling ones by "proactively planning for, and being alert to, expanding personal networks" (Katre & Salipante, 2012, p. 980).

Validation and Visibility. Legitimacy, defined as "a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions" (Suchman, 1995, p. 574), is a crucial resource for the survival and growth of new ventures lacking a track-record or established networks (Zimmerman & Zeitz, 2002). Acquiring and increasing social capital and legitimacy are, therefore, central concepts in the organizational sponsorship framework (Breivik-Meyer, 2020).

Consequently, the literature on incubators and accelerators recognizes the importance of strengthening the legitimacy of start-ups for attracting external resources. Bruneel

Table 17: Bridging 1 – Access to External Networks

Second order code	First order code	Representative quotes and further mentions	
With whom: stakeholders	large companies (F1, F5, R3)	B2B customers (F1, R4)	experts (R4, R5)
	incubator funders (F1, F3, F5, R2)	non-profit sector (F1, F3, F5, R2)	service providers (F5, R2, R4)
Why: benefits	accelerate new learning	They put us in touch with an agency who showed us how to set up low-touch marketing. That was very valuable for us. (F5, also mentioned by R3)	
	build new relationships	We can talk to mentors who always give us an intro to someone else. They were super helpful to get ahead. (R3, also mentioned by F3, R5)	
	enter new markets	The Impact Factory convinced us because the Ruhr area is very exciting for us. It is one of the largest urban areas in Germany. (R4)	
	enter new industries	If you operate in a wide range of industries like us with our versatile platform, you need the right contacts in the right places. (F1, also F4)	
	generate new business	Networking has a positive effect in business terms. (F1)	
	provide targeted introductions	Once you have spoken to them, you have a direct contact. But intros to gatekeepers or key persons were very helpful. (R3, also F5 and R5)	
Why not: limitations	Covid-19	Due to the digital offering, networking is being neglected. They tried to say: Hey, stay a little longer after the session. But nobody wants to do that. (R5)	
	sectoral reservations	I still hope the social economy will wake up at some point. The Impact Factory tried to play their contacts. But you always encounter fears. (R1)	

Note. All quotes are taken from the interviews and translated from German by the author. Quotes in the "stakeholders" category have been shortened for clarity.

et al. (2012, p. 112) have described how a "third wave" of incubators since the 1990s focused on "facilitating access to external resources, knowledge and legitimacy". In the exemplary services portfolio (see Appendix), half the studies mentioned incubator and accelerator activities that contributed to the legitimacy of new ventures, such as brand enhancement, media exposure, and public graduations – summarized in this study as the provision of "validation and visibility."

Considering the prominence of "legitimacy" in the literature, including in the resource classification Table 9 was based on (Van Weele et al., 2017), it is notable that the social start-ups in this study did not mention it at all as a resource need (Table 9). In addition, "validation and visibility" showed the largest discrepancy of all nine interventions in the ranking of sponsorship services (Table 11): While fellows ranked it second, it placed last with ramp-ups.

To analyze this result, Table 18 distinguishes two related but distinct concepts – "providing validation," i.e., intangible effects on a start-up's legitimacy, and "creating visibility" through incubator activities, such as pitch training or demo days. In line with the priorities expressed in Table 11, the positive views on validation – such as signaling credibility, reputation, and trust – were mentioned almost exclusively by fellows. Among fellows, F3 appeared particularly receptive to this intervention, as most factors conducive to validation were expressed by F3. In contrast, the negative views, including the difficulty to measure validation effects, the lack of brand recognition, and timing issues, were mostly highlighted by ramp-ups.

To a degree, these views could be attributed to the characteristics of the two subgroups. Whereas ramp-ups participated in a full incubation program, which included mentoring, training, and coaching, fellows deliberately opted for a loser relationship that put a stronger emphasis on validation. The development stage of the start-ups might also play a role. While ramp-ups were still mostly developing their product, fellows were already active on the market. A third explanation might be the novelty of the Impact Factory, which launched in 2019. The reputation of an accelerator influences its ability to increase a start-up's legitimacy (Drori & Wright, 2018, p. 13) – or as R5 said: "The accelerator is not really known – it's something else when you have Y Combinator on there, right?" Being selected by a well-regarded program is a potent validation signal (Miller & Bound, 2011, p. 27).

There are, however, factors that go beyond the reputation of the Impact Factory. The strong mission focus identified in Table 7 could be regarded as a source of legitimacy, which social entrepreneurs can "leverage with internal as well as external constituencies" (Dacin et al., 2010, p. 50). Thus, social entrepreneurs might find garnering access to external resources easier than conventional entrepreneurs. Another factor are the low customer acquisition costs for the digital platform models that characterized most ramp-ups in this study – thereby reducing the need for external resources to reach customers and increase market share (Miller & Bound, 2011, p. 22). In sum, Table 7 shows that the views on "validation and visibility" were mixed – and that its perceived value for

start-ups depends on the reputation and prominence of the organizational sponsor.

Access to External Funding. The opportunity to connect with external investors during demo days is one of the main attractions of accelerator programs (Cohen, 2013, p. 24). Indeed, most accelerators "measure themselves on the proportion of their companies that go on to raise further investment" (Miller & Bound, 2011, p. 27). Studies have shown that accelerator-backed start-ups receive follow-up financing sooner (S. Smith, Hannigan, & Gasiorowski, 2013), raise more external funding post-graduation, and achieve higher valuations (Cohen, Fehder, et al., 2019).

Considering these findings, it may be surprising that the start-ups in this study appeared to put less faith in the incubator to provide them with tangible resources than with intangible ones (Table 9), even though three ramp-ups indicated a greater need for growth capital at the end of the program than at the beginning. In the ranking of sponsorship services (Table 11), the social start-ups also regarded the provision of external funding by an incubator as not highly relevant, with ramp-ups and fellows ranking it in sixth and fifth place out of nine, respectively.

To explore these observations, Table 19 first presents the start-ups' experiences with external funding. As far as their willingness to acquire external funding is concerned, Table 19 presents a divided picture: Five founders showed no interest in taking on external investors – for now (F2, R1, R2, and R5) but also ruling it out categorically (F4). The other five had already received external investments (F1, F5, and R4) or were planning to (F3 and R3). Yet, when it came to the Impact Factory, the opinion of ramp-ups was unanimous: None of them regarded it beneficial for accessing external funding, while fellows expressed no opinion on this activity.

Table 19 shows that the experiences of the social start-ups with external funding varied – similar to their experiences with seed funding previously presented in Table 15. Regardless of their willingness to raise external funding, ramp-ups considered the support of the Impact Factory unhelpful for accessing external funding.

These results can be viewed from different perspectives. Partly, they can be attributed to the nature of the research setting described in earlier sections, such as its novelty. Moreover, the Impact Factory accepted start-ups at an early stage, potentially reducing its value for more advanced ventures. That seven of the 10 start-ups in the sample operated a digital platform model should also be noted, as "cheaper technology costs, easier routes to customer acquisition, and better forms of direct monetization" reduce the initial costs of growing a start-up (Miller & Bound, 2011, p. 21). What all the start-ups had in common, however, was their pursuit of a social mission. How did this affect their approach to fundraising? Table 20 explores the external barriers and internal deliberations social start-ups experience before raising external capital.

Many challenges for social start-ups relate to their hybrid nature: Pursuing both a social and economic mission involves significant trade-offs and higher complexity (Austin et al., 2012; Wilson & Post, 2013). The tensions resulting from competing social and financial goals influence their ability to mobilize external resources. Looking specifically at financial resources, Doherty et al. (2014, p. 424) identified two factors. On the one hand, social enterprises tend to generate less profit since they internalize social costs, making them less attractive to VCs. This was exemplified by R1, who refused to pass on the return expectations of an investor to the vulnerable beneficiaries of his services. On the other hand, funders find it challenging to fit the products and services of social start-ups into their categories. R2 was told by his bank: "What you have in mind does not fit into our spreadsheet." This barrier is particularly pronounced for social tech ventures aiming to pitch to commercial investors (Arena, Bengo, Calderini, & Chiodo, 2018, p. 157). In addition to financial and cultural issues, R3 highlighted legal barriers – in line with a survey by the EU Commission that showed regulations can pose a barrier for social enterprises seeking investments, for example, through limits on profit distribution (Wilkinson, 2015, p. 56).

Even though the social start-ups in this study confirmed the existence of these external barriers, they spoke about their internal deliberations in greater detail. Far from seizing every opportunity to raise external capital, Table 20 indicates that founders prefer "a deliberate and careful approach to both their capital and governance structures," which includes "selecting highly value-aligned investors" to "avoid philosophical or strategic conflicts," as Wilson and Post (2013, p. 729) have phrased it. To safeguard their strong social-mission focus, the respondents remained careful in selecting potential investors, attributing a higher significance to impact commitment, long-term perspectives, and mission alignment than to beneficial financing terms. By aligning their mission and method with their capital and governance structure, social start-ups "aim to avoid the distortion often imposed by the public capital markets, or traditional venture capitalists and private equity investors" (Wilson & Post, 2013, p. 729).

Furthermore, social start-ups carefully weigh the rewards against the risks of taking on external investments. While R2 and F2 acknowledged that an infusion of external capital could accelerate the pursuit of their social mission, there was a stronger emphasis on potential risks, such as internal conflict, loss of control, mission drift, and wrong incentives. After experiencing corporate life for 10 years, F2 refused "to sit at a table with an investor at the end of a day and discuss purchase prices from my manufacturers. I just don't want that, because those are the foundations of our business, and I can't touch them." This statement demonstrates the reluctance of social entrepreneurs to give up control or undermine their mission (Arena et al., 2018) – making it more challenging for an incubator or accelerator to bridge the gap to providers of external capital.

4.3.4. Boosting Mechanism

The services of incubators and accelerators assessed so far relate to buffering and bridging – the two mechanisms

Table 18: Bridging 2 – Validation and Visibility

Second order code	First order code	Representative quotes and further mentions
		Providing validation
Validation: positive views	association effect	We really want to show that we are part of this network. (F5)
	credibility	To show our business model works, incubators are a tremendous help. (F1)
	public reach	We took part in a social media campaign. That was definitely good for our reach. (F3)
	reputation	Our reputation may be enhanced by being connected to the Impact Factory. (F4)
	seal of approval	To have a seal from an accelerator has helped in a conversation with business angels. This was really an advantage. (R4, also F1, F4 and R3)
	trust	It's important for a lot of start-ups: you need things that inspire trust. (F1, also F4)
Validation: negative views	hard to measure	I doubt this really has a measurable effect or will bring a return. (F2, also F5)
	low brand recognition	Maybe this has an effect subconsciously, but the accelerator is not really known. I wouldn't say that there is significant validation. (R5, also R1 and R4)
	wrong timing	The Impact Factory may have been too late - we opened a lot of doors ourselves. (R2)
Validation: conducive	branding	With the Factory, the name speaks for itself. (F3)
	popularity	The better known the incubator, the better for the start-up. (F3)
	reciprocity	We took part in a social media campaign of the Impact Factory. That was definitely good for our reach, but also the other way round. (F3)
	scale effects	In a start-up there is always a certain amount of marketing that you have to do, and the more partner logos you have, the better it is. (R3)
		Creating visibility
Creating visibility: conducive	public events	What helped us the most were public events, where investors attended. (R3, also F1)
	storytelling help	We realized bringing it across is incredibly complicated. We should have worked on that, explaining it simply with storytelling. (F3).
Creating visibility: detrimental	forced rankings	We pitched a lot at the beginning with rankings. I was at the bottom of the list, and that always pulled me down. I don't have any benefit from that. (R2)
	social distancing	We had bad luck, because of Corona we couldn't pitch. (R1)
	superficiality	You should not put much effort into making everything look great, and if someone asks 2 or 3 questions, the whole thing collapses. (F4)

Note. All quotes are taken from the interviews and translated from German by the author.

of the original framework of organizational sponsorship proposed by Amezcua et al. (2013). While buffering aims to insulate new organizations against external threats and market pressures, bridging enhances their chances of survival by facilitating access to external resources, promoting knowledge spillovers, and increasing their legitimacy.

This study, however, follows Breivik-Meyer (2020, p. 181) in extending the original framework with a third mechanism, boosting, to understand how sponsorship can "not only promote survival, but also increase the growth of new firms". Originally proposed by Autio and Rannikko (2016), this mechanism comprises two activities of organizational sponsors aimed at boosting the capacity for growth of new ventures: setting and controlling milestone achievement, and promoting networking among peers.

Milestones and Progress Tracking. The first intervention of the boosting mechanism refers to setting and controlling milestones to accelerate the development of new ventures. Unlike the other interventions discussed so far, this activity has been rarely mentioned in the reviewed literature, at least by this name. The exemplary services portfolio (see Appendix) included only three activities that could be subsumed under this intervention: "pressure and discipline" (Miller & Bound, 2011), "ongoing proof of concept" (Dempwolf et al., 2014), and "counseling services to track progress" (Pauwels et al., 2016).

Admittedly, incubators and accelerators can accelerate their ventures through other activities, like mentoring and coaching. By organizing a demo day at the end of their program, for example, accelerators can set ambitious deadlines. Regardless, it is worth exploring how social start-ups assess this service. Can – and should – organizational sponsors accelerate the development of new start-ups by setting and tracking concrete milestones?

When asked directly whether incubators or accelerators add value by setting milestones, the social start-ups

Table 19: Bridging 3.1 – Experiences with External Funding

Willingness to take on external investment					
F1	Yes	Early research grant from a tech fund	R1	No	No suitable match yet
F2	No	Repeatedly approached, but always refused so far	R2	No	No interest in outside capital yet
F3	Yes	Financing round planned	R3	Yes	First financing round ongoing
F4	No	No interest in outside capital	R4	Yes	Working with business angels
F5	Yes	Business angels from the start, VC financing planned	R5	No	No suitable match yet

Negative views of the incubator's support	
R1	I realized quickly that the Impact Factory can't help me find any investors who are ready to finance the branches at conditions where I can repay them.
R2	A workshop on funding, where all kinds of sources were described, left me disappointed. I need someone who finally listens to what I need.
R3	When I look at the investors we are talking with right now, there was no direct contact through the network of the accelerator. But of course, it helps if you get an intro.
R4	The business angels we are working with did not come through the Impact Factory. We would have liked to focus more on this and a little bit earlier.
R5	With the Impact Factory, the focus is not so clearly on financing.

Note. All quotes and information are taken from the interviews and translated from German by the author.

Table 20: Bridging 3.2 – Perspectives on External Funding

Second order code	First order code	Representative quotes and further mentions
Access barriers	cultural	I talked to banks and I asked for 35,000 euros, which is peanuts. And the risk is nothing. Yet they said: What you have in mind does not fit into our spreadsheet. (R2)
	financial	Investors want a share, eight percent would probably be cheap. But we simply lack this profit motive. We have built the company completely differently. (R1, also R3)
	legal	We are a company in steward ownership. We have committed to reinvesting all profits and donating the rest. This is not yet established in the financing landscape. (R3)
Investor selection criteria	impact commitment	If I would even think about working with a business angel or bring in an external investor, then only if the social impact component is safeguarded. (F2)
	long-term perspective	We are looking for somebody who wants to give patient capital and share in the profits. Not just put money in and bang, away with it. (F3)
	mission alignment	We didn't want to take the classic start-up path of directly taking on investor who is not 100 percent committed to the development and the impact idea. (R5)
Rewards of external funding	accelerate growth	I don't want to depend on external people. But if you spend a little money, it goes faster, and in the end, it may be more efficient. (R2)
	increase resources	Investments are something every founder thinks about. You ask yourself: Does it make sense to raise capital somewhere else, because your resources are always limited. (F2)
Risks of external funding	internal conflict	I would end up sitting at the table with an investor who may not have the same goal as me, which could lead to major conflicts about how to run the company. (F2)
	loss of control	You have to be careful not to let people take too many shares. There are many traps. It can get a bit ugly. (F3)
	mission drift	You could get investors with return expectations, but in the end you don't solve the problem. I don't want to charge 35 euros per hour to look after a poor mother. (R1)
	wrong incentives	They say: If you grow 15x in the first few years and we increase our money fivefold, than that's cool, but the impact is not the main focus. That is a pity. (R5, also F3)

Note. All quotes are taken from the interviews and translated from German by the author.

seemed indifferent. In the ranking of sponsorship activities (Table 11), both ramp-ups and fellows ranked milestone setting eighth out of nine. A more nuanced picture emerged when they discussed the Impact Factory's role in accelerating their growth. Table 21 shows that two founders benefited from regular milestones set by the Impact Factory, whereas the other founders described no effect. Interestingly, this response suggested a clear demographic divide: The two positive views were expressed by very experienced founders (R1 and R2), whereas the four youngest respondents in the sample were also the least convinced of the benefits of milestones set by an accelerator.

Nevertheless, even the skeptical founders mentioned activities that incubators and accelerators can use to accelerate their growth, such as setting tight deadlines, providing focus and structure, monitoring progress, and flagging blind spots. The importance of speed was repeatedly highlighted – in line with recent findings that "time-compressed scaling" is a distinguishing feature of successful start-up acceleration (Shankar & Clausen, 2020, p. 102174). Accelerators also resolve uncertainty faster (Yu, 2020) and shorten learning cycles by providing intensive consultation and rapid feedback (Cohen, Bingham, & Hallen, 2019).

By contrast, there is no consensus in the literature on the value of setting ambitious goals, which R3 argued for. While some founders welcome forced progress (Miller & Bound, 2011, p. 28), others perceive too much intervention by incubators as an interference (Patton, Warren, & Bream, 2009, p. 629). The perception of this activity appears to depend on the motivation and drive of the founders. Two founders (F2 and R3) even suggested that an accelerator should help founders slow down, rather than accelerate, by providing balance and emotional support. In this sense, Table 21 reflects that all start-ups in this study (except R1) had completed the program successfully and developed, or even launched, a product or service. As a highly driven group, their social-mission focus provided a sufficient boost without the need for external milestone setting.

Peer Networking. The value of external networking was addressed as part of the bridging mechanism. Nonetheless, peer networking is discussed separately here for its potential effect on the "rapid organizational growth" of start-ups through the "exchange of experiential insights" (Autio & Rannikko, 2016, p. 44). Organizational sponsors can boost the acceleration of start-ups by serving as connective intermediaries between them (Breivik-Meyer, 2020, p. 181).

The conceptual uniqueness of peer networking as an incubator service was confirmed partially by the literature. 40% of the studies on social incubators and accelerators in the exemplary services portfolio (see Appendix) mentioned networking with like-minded entrepreneurs or peers. Studies on conventional incubators and accelerators named this activity only once (Miller & Bound, 2011). However, this must be qualified to the extent that accelerator studies frequently describe how ventures "enter and exit the programs in groups, known as cohorts or batches" (Cohen, 2013, p. 22) – and the effect these activities have on them.

Accelerators regularly foster collaboration between their start-ups through explicit and implicit activities, such as common working spaces and specialized sessions (Drori & Wright, 2018, p. 11). Encouraging peer support between start-ups can take some of the burden off the accelerator management team, allowing it to "focus on bringing in outside expertise" (Miller & Bound, 2011, p. 10). The Impact Factory, for example, hosted regular sessions for peer exchange.

For the founders in this study, peer networking was a motive for joining the incubator in the first place: Four founders named "peer exchange" as a resource need to be addressed in the program (Table 9), while "networking" received seven mentions. The ranking of sponsorship services indicated similar priorities: Ramp-ups and fellows ranked "peer networking" fourth out of nine, below "access to external networks" (Table 11). This mid-table result was surprising, considering how enthusiastic the founders responded when asked directly about their assessment of peer networking. As Table 22 shows, almost all the founders praised their exchanges with other start-ups in the Impact Factory – even those, like F1, F3, and R3, who ranked it only in fifth or sixth place in the ranking of services.

The positive assessment in Table 22 presents a stark contrast to a survey of 4,000 social entrepreneurs who, in applying for accelerator programs, considered peer networking "unimportant to their venture success" and "the least important benefit" (Pandey et al., 2017, p. 19). Another study of start-ups in a technology incubator in Hong Kong delivered an even bleaker verdict on peer networking, reporting that all firms "shut the door, work alone and never chat on product, market and business-related topics" – with tenants going as far as competing, rather than cooperating (K. Chan & Lau, 2005, p. 1226). The results in Table 22 suggest that the Impact Factory was a rather different environment. The founders highlighted the positive effects of peer networking, including joint learning, exchanging ideas, and mutual support. Four founders emphasized the importance of helping less experienced peers. These views confirmed Cohen (2013, p. 22) observation that the cohort experience "fosters uncommonly strong bonds and communal identity between the founders".

The respondents also named success factors for peer networking. First, cohort selection is key: The founders agreed that peer effects are strongest when peers share values and norms but are not too closely matched (S. Smith, Hannigan, & Gasiorowski, 2015, p. 27). A lack of direct competition fosters transparency and information sharing, which can boost performance (Cohen, Bingham, & Hallen, 2019). Second, building a strong alumni network, which is an asset of prominent accelerator programs (Miller & Bound, 2011). Third, enabling physical co-location, although the views here were mixed. R5 agreed with Duff (1994) that proximity was crucial, while R4 welcomed the remote exchange despite the forced Covid-19 restrictions. In sum, peer networking was mostly seen as beneficial. Interestingly, the emotional value of peer exchange was assessed more positively than its prac-

Table 21: Boosting 1 – Milestones and Progress Tracking

Second order code	First order code	Representative quotes and further mentions
Did growth accelerate?	yes	You set priorities in a relatively short time frame, talk them through and see if they are possible or not. That was actually the good part for me. (R1, also R2)
	maybe	You can always look back and say: maybe we could have gotten further. (F2)
	no	I don't think it had any impact on our development. We are simply very market driven and our competitors pushed the envelope. (R4, also F3, R3 and R5)
How to accelerate	ambitious goals	The Impact Factory is an impact bubble. Other accelerators have more of an economic focus and a tougher approach. This mixture is essential if you have a hybrid form. (R3)
	flag blind spots	That definitely accelerated things, because I simply didn't have certain topics on my radar, for example the question of my legal form. (R2, also R4).
	focus and structure	Gut feeling always plays a role in founding, but structure is also very important. Focus, prioritization and structure are the most important drivers. (F2, also R5).
	regular monitoring	I think the exchange with mentors, and also having a check-up on a regular basis, can help a lot. I think it increases accountability. (R5, also R3)
	tight deadlines	It would have been more helpful if we had done the whole thing in less time. Make it compact at the start and less frequent as we develop. (R4, also R1).
How to slow down	emotional support	Managing a company tests your limits every day. I believe an incubator can help you see the lightness of the whole thing and say: hey, it's not so bad. (F2)
	balance	The founder who talks about his 70 hours a week also needs balance. You often see it portrayed in the media. And that's just wrong. Accelerators can have a big influence and invite people to speak about how you can create balance. (R3)

Note. All quotes are taken from the interviews and translated from German by the author.

tical benefit compared to other incubator interventions.

4.3.5. Impact Acceleration

The three sponsorship mechanisms presented so far, while helpful for understanding the perception of incubator and accelerator support, are not unique to social start-ups but applicable to various types of organizational support. Yet, this study seeks to address the question raised by J. Hausberg and Korreck (2017, p. 13): If social businesses face different challenges, do they also require a different kind of assistance? Therefore, this thesis asks whether – and how – incubators and accelerators need to tailor their services portfolio to social start-ups.

During the coding process, two incubator benefits emerged that particularly suited the needs of social start-ups: "demonstrating social impact" and "delivering the social mission." Together, they can be regarded as a novel support mechanism called impact acceleration, in addition to the buffering, bridging, and boosting mechanisms already discussed.

The Uniqueness of Accelerating Social Start-Ups. Social and conventional start-ups share the aim of developing innovative and market-oriented solutions. However, the pursuit of social value creation also characterizes social start-ups. Table 7 presented six manifestations of this social-mission focus: Three shared with conventional start-ups (independence, sustainability, and long-term thinking), and three that primarily relate to social start-ups (priority of purpose, credible impact, and hybridity). The study of nine incubator and accelerator services in Section 4.3 highlighted where the needs and perceptions of social start-ups differ from conventional start-ups, for instance regarding funding, networking, and training.

Nonetheless, this study has yet to address explicitly what is specific about the acceleration of social start-ups. Do these ventures require support activities beyond the support traditionally offered by incubators and accelerators, i.e., the three mechanisms of the organizational sponsorship framework? Answering this question is not straightforward – and the extant literature provides no consensus view, to the extent it has addressed this question at all.

Even though this study did not include a control group for a systematic comparison, it sought to answer this question according to social start-ups. As Table 23 shows, the respondents identified similarities in support needs, including finding a marketable product or service, solving a real need, and building a successful business model around it. As F5 said, "the initial challenges are the same for all start-ups." Yet, the founders also named differences in support needs – from legal know-how to finding employees and investors with the right motivation for social start-ups. But even if these needs are specific to social start-ups, incubators and accelerators should fulfill them with the mechanisms of organizational sponsorship – for example, by providing legal training or matching start-ups with impact-driven investors.

This is not the case, however, for the two support needs most frequently mentioned in Table 23: "demonstrating social impact" and "delivering the social mission," which were both raised by five founders. In the case of these two activi-

Table 22: Boosting 2 – Peer Networking

Second order code	First order code	Representative quotes and further mentions
Assessment	positive	Super essential. (F1) / Very, very good. (F2) / Simply exciting and fun. (F3) / Always great. (F4) / Awesome. (F5) / Very good overall. (R2) / It worked well. (R3) / Brilliant and very helpful. (R4) / One of the best things about the Impact Factory. (R5)
Positive effects	mixed	Yes, we used that. With one team we might cooperate in future. You just have to see. (R1)
	learning from peers	I think you can learn best from each other. You don't have to make the same mistakes again and again - unless you have an eternity to learn. (F1, also mentioned by F3 and F4)
	exchanging ideas	The exchange was very open, very transparent. There was a lot of trust and the exchange itself worked very well despite the physical distance. (R4, also R3)
	giving back	Everyone asks for help. I have always agreed and said: now's the time to give back a little, even if I don't have the time. It's support for the right people. (R2, also F1, F5 and R1)
	mutual support	Building a start-up is a sinus curve. You're always super motivated and then super depressed. In these phases the exchange with other start-ups is super valuable. (R3, also R5)
Success factors	social contacts	As a founder you tend to stay in your bubble, immersed in your business. The greatest value for me is the contact with other founders. Being a founder is a bit lonely. (F2, also R3, R5)
	alumni community	When a program ends, the contact disappears very quickly, and that's a real pity. How could you somehow create a better alumni management? (R5, also F5 and R3)
	cohort diversity	If you look at the founder profiles, we are all very different. There are people who had a life before, and others are younger. I think the mixture actually works pretty well. (R2, also F3)
	physical colocation	We can see it with accelerators that took place digitally. The networking was practically zero. You don't have a really lasting exchange and that's a great, great pity. (R5)
	value alignment	I don't come from an entrepreneurial family. Suddenly you find like-minded people and realize you're not the only crazy person - there are others who tick like you. (F2, also R4, R5)

Note. All quotes are taken from the interviews and translated from German by the author. Quotes for the "positive" assessment were shortened as the full quote provided no further context.

ties, which emerged in the interviews, social start-ups expect an incubator or accelerator to provide benefits that go beyond the support activities discussed so far. Thus, it is of interest to analyze these two activities in more detail.

Demonstrating Social Impact. Noting that social start-ups expect their social impact to be included in the acceleration process may seem self-explanatory. However, this activity is far less prevalent than might be assumed. A study of impact accelerators has found that frequently "developing social impact models was not a core offering" (King et al., 2015, p. 13). This is reflected by the exemplary services portfolio (see Appendix), as only three out of 10 studies on social incubators and accelerators mentioned activities related to measuring and improving social impact. A quantitative study of 83 Italian incubator managers has found that they "gave little importance to social impact measurement services" (Sansone et al., 2020, p. 7); surprisingly, even among social incubators less than half (44%) had social impact metrics for their tenants (p. 8).

The low prevalence of impact measurement services in the service portfolio of incubators and accelerators stands in contrast to the needs of social start-ups. In a survey of German social entrepreneurs, 71% reported having established impact goals and 23% were planning to do so (DSEM, 2020, p. 40), just as the social start-ups in this study mentioned "demonstrating credible impact" as one of their motives in Table 7. F5 summarized this view: As a social start-up, "you have to have [impact measurement], that's your legitimation for saying I really want to achieve outcome and impact. And I am quite rigorous. You have to be serious about that." However, even when the impact is at the core of a social entrepreneur's mission, as for F5, there are barriers to demonstrating the impact effectively that an incubator or accelerator can help overcome.

Table 24 explores four reasons why incubators and accelerators, in the view of social start-ups, should support them in demonstrating their impact. The first two relate to the ubiquity of the term impact and related concepts such as sustainability. For entrepreneurs who take their social or environmental impact seriously, it is difficult to stand out from the myriad of companies seeking to benefit from this trend. This challenge is exacerbated at an early stage, when a ven-

Table 23: Comparing the Support Needs of Social and Conventional Start-Ups

Second order code	First order code	Representative quotes and further mentions
Similarities	build a successful business model	Normal accelerators show you how to build a business model and quickly validate it on the market. You need the same in an impact accelerator. (R3, also F3, R4)
	find a marketable product	The initial challenges are the same for all start-ups. At the beginning you need a product that is marketable. That core challenge is really the same for everyone. (F5)
	solve a real need	You have to be passionate about your product, you have to see a social or a purely economic need, or maybe a market-driven need, which is the same in the end. (F1)
	storytelling	Maybe marketing and storytelling is also different for impact start-ups. You can communicate the story more forcefully. (R5, also mentioned by F3)
Differences	find the right investors	I think especially for social start-ups, a different group of investors might be interesting, or an additional group of investors. (F3, also R3)
	find the right legal form	But also questions such as the structure of a social business and its legal form. (R3, also F1 and R5)
	hire and retain the right talent	In a classic start-up you have to motivate people with money. But how do you motivate a team if it doesn't necessarily have to be monetary? (R5, also R2)
	deliver the social mission	The difference, in our view, is that no matter whether you talk about profit or not, you always take the idea of impact into account. (R4, also F4, F5, R2 and R3)
	demonstrate social impact	Impact measurement is certainly very relevant for us. I think this is something that is very specific for impact start-ups. (R5, also mentioned by F3, F5, R3 and R4)

Note. All quotes are taken from the interviews and translated from German by the author.

ture's impact is hypothetical or small, and the way to measure it "may change with scale" (Miller & Stacey, 2014, p. 27). At the same time, F4 argued that conventional start-ups tend to benefit from a positive bias for start-ups, even when their social impact is negligible – or even negative. F5 added that larger companies are still reluctant to engage with social start-ups to increase their impact.

Table 24 also shows that incubators or accelerators can support social start-ups in a multitude of ways to address the challenge of demonstrating social impact. First, by helping to develop concrete and measurable impact metrics. Originally associated with international development, social impact can be defined as "significant or lasting changes in people's lives, brought about by a given action or series of actions" (Roche, 1999, p. 21). Beyond this widely accepted definition, the term remains contested, referring alternatively to effects at the individual, organizational, and societal level (Ebrahim & Rangan, 2014; Gupta et al., 2020). The lack of credible indicators and metrics for social impact remains a development barrier for social ventures (Arena et al., 2018, p. 161). A survey by the EU has identified "common mechanisms for measuring and demonstrating impact" as a key factor for the visibility of social enterprises (Wilkinson, 2015, p. 99).

While the standardization of impact metrics is beyond the scope of any single institution, an incubator can still support social start-ups by educating them about established frameworks, such as the Theory of Change or Social Return on Investment. Indeed, a survey of 20,000 social ventures has shown that lack of awareness was a key reason for the low adoption rate (25%) of established impact measurement systems (Global Accelerator Learning Initiative, 2020, p. 9). In contrast, training and peer interaction in an accelerator were strongly related to the adoption of social performance measurement (Lall, 2017, p. 2649).

Additionally, Table 24 suggests incubators can help start-ups to apply impact metrics at the operational level and relate them to frameworks such as the SDGs. Thus, incubators can support social start-ups by focusing on substance, rather than presentation, when it comes to social impact. As F4 put it: "You should not put so much effort into making everything look great on the surface – and if someone asks two or three questions, the whole thing collapses like a house of cards."

Delivering the Social Mission. The results in Table 23 show that incubators and accelerators not only have a role in helping social start-ups to quantify and demonstrate their social impact. They can also support them in achieving it. While this phenomenon has not been studied extensively to date, there are indications that incubators and accelerators can provide added value by improving social start-ups' "ability and readiness to deliver predictable and consistent social impact," as a survey of impact-oriented incubators and accelerators in Europe has suggested (Gianoncelli et al., 2020, p. 25). To explore what the added value of an incubator or accelerator in delivering the social mission might entail, Table 25 categorizes the views of social start-ups into three activities: nurturing a culture of humility, authenticity, and sustainability; selecting an impact-driven community, ideas, and values; and supporting social start-ups to navigate impact and profit, manage hybridity, and safeguard their emotional wellbeing.

The expectations expressed in Table 25 suggest that an incubator's role is not only to provide social start-ups with

Table 24: Impact Acceleration 1 – Demonstrating Social Impact

Second order code	First order code	Representative quotes
Why an incubator should help	sustainability hype	I would say everyone is doing it by now. Even normal start-ups. It is simply very much desired by society. Sustainability is much more relevant than five years ago. (F1)
	green washing	I'm surrounded by advertising now where big companies write: Sustainability is not a luxury. And this is true. But I still have to be able to differentiate myself. (F2)
	positive bias for start-ups	Just because it's a start-up, people are already applauding. Monsanto was a start-up once. Just because you're a start-up doesn't make you a good company. (F4)
	reluctance by large companies	When I tell companies about the impact chain, you can actually apply it to them quite well. But it is still completely foreign to them and nobody expects or demands it. (F5)
How an incubator can help	develop concrete impact metrics	The most important point for me: how do you make impact really understandable and measurable with concrete metrics. This is where incubators can really help. (F2)
	include KPIs at operational level	We want to include impact KPIs in our balanced scorecard, in our monitoring. We haven't done that yet, but it is on our list. (F3)
	connect to global goals	We wanted to do something for people without electricity. We stumbled upon the SDGs and realized there are global goals. How can we link them to what we are doing? (F3)
	be an impact sparring partner	It is very important in such an incubator to check out the motivation. Everyone can say: Hey, we are a social start-up. But what kind of impact do we actually want to have? (F4)

Note. All quotes are taken from the interviews and translated from German by the author.

resources, skills, and contacts to survive in competitive markets – the traditional goals of organizational sponsorship. It should also create an environment of trust and mutual support to sustain social entrepreneurs in fulfilling their social mission.

This finding is backed up by the Schwab Foundation for Social Entrepreneurship, which has found that fear of failure is more pronounced among social than conventional enterprises, since "fail fast" for them is often not an option (Zimmer & Pearson, 2019, July). Also, social entrepreneurs feel a strong affective commitment to the beneficiaries of their causes (Renko, 2013, p. 1047). As a result, "too many founders feel the weight of the world on their shoulders and don't find a way of sharing it around" (Miller & Stacey, 2014, p. 11). Beyond accelerating their growth, incubators and accelerators can play a pivotal role in taking an emotional – and material – load off the shoulders of social entrepreneurs.

4.4. Motives for Entrepreneurial Self-Reliance

The previous Section 4.3 explored how social start-ups perceive the benefits of different sponsorship mechanisms. However, the sponsorship literature has observed that resource munificence is not universally beneficial but contingent on boundary conditions (Amezcua et al., 2013; Flynn, 1993b; Jourdan & Kivleniece, 2017). Therefore, "new ventures ought to evaluate carefully the potential impact of accepting such benefits" (Amezcua et al., 2013, p. 1645).

Nascent entrepreneurs in Germany seem to heed this advice, considering that only 50% of them seek professional assistance, although "professional consultancy for potential entrepreneurs is highly subsidized and inexpensive" (Brixy, Sternberg, & Stüber, 2013, p. 157). This observation also applies to social entrepreneurs, as 51% of them have benefited from a support program (DSEM, 2020, p. 50). Yet, the same survey has also pointed out that of the 49% who did seek support, 37% participated in at least two programs. This is consistent with a global survey of 20,000 social enterprises who applied for an accelerator program, in which around a third had prior accelerator experience (Global Accelerator Learning Initiative, 2020, p. 8). Overall, a dichotomy appears: every second founder prefers to go it alone, while one in three uses multiple programs.

The views of the 10 start-ups in this study were similarly polarized. As Table 10 showed earlier, three founders (F2, F4, and F5) decided not to participate in any program beyond the fellow status of the Impact Factory, which came with minimal participation requirements. By contrast, five teams took part in multiple programs – in two (R2), three (R3 and F3), and more than five (R5 and F1), depending on the definition of a support program.

Understanding the factors driving these decisions would be highly beneficial for the design of entrepreneurial support programs. However, no clear pattern emerged between the decision to join a program and the characteristics of the

Table 25: Impact Acceleration 2 – Delivering the Social Mission

Second order code	First order code	Representative quotes
Nurture	humility	I like the climate, the exchange, the fact that nobody feels superior. There are no people with big egos who are there only to show off or to just put on an act. (R2)
	authenticity	How can you build a company that is not only green on the outside, but also treats employees and resources sensibly internally? (R3)
	sustainability	What we liked very much was, of course, simply this mindset of sustainability, which is taken very seriously there. (R4)
Select	the right community	We could identify with it right away because it was a program specifically designed for social start-ups. The feeling was that there was obviously a community. (F2)
	the right ideas	Incubators are often too business driven, although they have to be. What convinced me about the Impact Factory was that they go for sustainable start-up ideas. (F1)
	the right values	It is important that the values of an investor or incubator match with a social start-up. It is the foundation that someone who comes into this company has the same values. (F2)
Support	to navigate impact and profit	You have to add an additional component. What is the social impact of your start-up, what is the social dimension of your legal form or how you deal with profits? (R3)
	to manage hybridity	You move in a strong field of tension. You have to be heard on the market, otherwise you will go down brutally. But you must not overdo it and stray from your values. (F2)
	to safeguard wellbeing	Funding an impact enterprise is a bit different. You put your heart into it and there is the danger you do too much, because your work is multiplied for the good of others. (R5)

Note. All quotes are taken from the interviews and translated from German by the author.

founders, such as industry, age, and professional background. Hence, this study inquired about the motives for not joining a support program. The motivation that emerged from the interviews, called entrepreneurial self-reliance, did not amount to an outright rejection of external support, at least for most founders. Rather, it indicated a deliberate consideration of the advantages and disadvantages of seeking support.

Before discussing these findings, a note of caution. The decision of a social start-up to join a support program is driven by various factors, ranging from the availability of suitable programs to physical proximity, resource considerations, and venture stage (Chmiliar, 2010). It was beyond the scope of this study to analyze these factors in detail. Instead, this section aims to present common motives that, from the point of view of social start-ups, reduce the subjective value of participating in entrepreneurial support programs. By taking them into consideration, incubators and accelerators can increase their attractiveness for these ventures.

As Table 26 indicates, the motives for entrepreneurial self-reliance expressed by the founders in this study can be summarized in three categories: resource trade-offs, previous experiences, and concerns about a negative impact on a start-up's development.

Among these motives, resource trade-offs were the most prominent, given that seven out of 10 founders regard "time" as their main concern – an indication that participating in a support program comes with significant opportunity costs. In a way, time can be considered a reverse proxy for the overall value of a program. As the previous sections have shown, an incubator or accelerator can also speed up developments and save time when its services are considered beneficial. This trade-off is captured by F1: "Of course, some things repeat themselves. You lose time participating in multiple programs. But if you don't do it, there are lot of things you miss out on."

Financial resources were not mentioned as a factor, probably because the Impact Factory was a subsidized program that did not charge fees or provide funding. The start-ups also failed to address a resource trade-off frequently cited in the sponsorship literature, namely the risk of an incubating environment "artificially inflated with resources" that could make an organization vulnerable in the long run (Flynn, 1993b, p. 56) – a phenomenon also known as the "life-support incubation trap" (Clarysse, Wright, & Van Hove, 2015, p. 19).

The second motive was previous experiences. While the results did not show a clear demographic divide, trends still became visible. The three founders who decided not to join formal support programs (F2, F4, and F5) all had more than 10 years of professional experience (see Table 7). Moreover, F4 and F5 were the only founders who had already launched a social-purpose organization before their current venture – in line with a study that found that serial entrepreneurs were less likely to seek professional assistance (Brixy et al., 2013, p. 158). More experienced founders "tend to view a large part of the educational program as redundant and a waste of time" (Drori & Wright, 2018, p. 8). In contrast, three of

Table 26: Motives for Entrepreneurial Self-Reliance

Second order code	First order code	Representative quotes and further mentions
Concerns	need for autonomy	Attendance lists like at university would be negative, because the needs of each start-up are individual. You have to maintain a certain openness. (F2)
	information overload	There are strong distractors. It's important to believe in your idea, even though others may not see it that way. Incubators can be great sources of knowledge, but it is important to find the balance between what is brought in from outside and what you decide yourself. (F2)
	fear of rejection	When you do something new, you always have to explain and defend yourself, especially in Germany where everybody is critical of what you do. You have to learn this is absolutely normal and take critical feedback on board without taking everything to heart. (R3)
Previous experiences	as founders	Did we consciously decide against it? Yes, I would say so. We were both very experienced founders, both over 50. (F5)
	with incubators	We have gone through five incubator programs with different emphasis. That is perhaps the disadvantage, that they are too similar. (R5, also mentioned by F1)
Resource trade-offs	surrendering equity	I always find that difficult, because especially at this early stage you might not even want to think about giving away shares. (R3, also R4)
	time investment	It was an investment in time, just going to Duisburg at the beginning, that was quite a distance. We are limited in terms of our resources. (R4, also F1, F4, F5, R1, R2, R5)
	travel expenses	Well, we invested our time and paid for our travel expenses. (R1, also R4)

Note. All quotes are taken from the interviews and translated from German by the author.

the four teams who took part in more than three programs (F1, F3, R3, and R5) founded their ventures after finishing university, which indicated a greater willingness to seek external support. However, even younger founders such as R5 became more selective over time, carefully weighing whether the program justified the time invested.

Finally, the broadest motive was concerns about a possible negative impact on a start-up's development. While these concerns might appear counterintuitive, considering the mostly positive assessment of incubator and accelerator services by the founders, they have also been observed in the literature. In particular, the founders voiced three concerns.

First, the need for autonomy highlighted by F2. McAdam and McAdam (2008, p. 288) have also argued that young firms grow reluctant of incubator support as it could be associated with "newness, vulnerability and inexperience". This might be one reason why, in a study of 52 impact-focused accelerators worldwide, 75% worked with ventures at a prototype stage and only 23% at a growth stage (Lall et al., 2013, p. 114).

A second concern was information overload. F2 stressed the importance of following your instincts and having the freedom to experiment – which might be constrained by a rigid program. The importance of flexibility has been highlighted by Shankar and Clausen (2020, p. 102174), who argued that early-stage start-ups are not an ideal target group for an acceleration program that leaves limited room for pivots. However, this concern could also be a sign of cognitive biases such as overconfidence or confirmation bias (Cohen, Bingham, & Hallen, 2019). F2 admitted as much by asking: "You can always look back and ask – would we be further ahead if we had gotten support early on? Maybe yes, but maybe we would have ended up in a very different place."

A third concern, fear of rejection, was named by R3, who described the risk of receiving negative feedback early on when the start-up still rests on a brittle foundation. Miller and Stacey called this effect "mentor whiplash" – when teams "find themselves getting conflicting advice and are confused about which direction to take" (Miller & Stacey, 2014, p. 38). Together, these motives show that social start-ups critically weigh the benefits of a support program before joining it.

5. Discussion

The purpose of the present thesis is to improve the understanding of a novel phenomenon – the acceleration of start-ups aiming to tackle societal or environmental challenges. With the aid of an inductive study of multiple cases, the study explored how social start-ups perceived the activities of incubators and accelerators. The results indicate that the social-mission focus of these ventures led to significant differences in how they perceived incubator benefits as compared to commercial ventures. Consequently, this thesis argues that incubators and accelerators should adapt their services to social start-ups. Moreover, it presents a systematic assessment of the mechanisms of organizational sponsorship from the perspective of social start-ups. As its main theoretical contribution, this study extends the organizational sponsorship framework by proposing a novel support mechanism: impact acceleration. Finally, this study explored how social start-

ups decide between joining a formal support program and depending on entrepreneurial self-reliance.

The following section discusses the relevance of these findings in three stages. After clarifying their theoretical contributions, it addresses their practical implications for designing and running entrepreneurial support programs. It closes by addressing the limitations of the present research and by highlighting promising avenues for future inquiries.

5.1. Theoretical Contributions

Evaluating existing theory with a deductive approach was not the aim of this study. Rather, it sought to identify common patterns and develop theoretical constructs to link qualitative evidence with deductive research (Eisenhardt & Graebner, 2007). In light of sparse prior research, this study drew on three emerging strands of research: social start-ups, incubators and accelerators, and organizational sponsorship. This section discusses the findings' contribution to these three literature streams.

5.1.1. Contributions to the Literature on Social Start-Ups

There is a widely held view that hybridity – the pursuit of financial goals and social purpose – defines social enterprises (Dacin et al., 2010; Doherty et al., 2014; Gupta et al., 2020). However, how the "primacy of the social mission," as Nicholls (2006, p. 20) has called it, shapes their perception of entrepreneurial support remains unclear. To investigate this question, this study adopted the perspective of social start-ups – unlike most research on entrepreneurial support, which has focused on the views of incubators. Thanks to this approach, this study offers two contributions to the literature on social start-ups: 1) how the social-mission focus influences key strategic and operational decisions and 2) how it affects the way that social start-ups acquire external resources.

The Social-Mission Focus in Practice. The first insight is that founders of social start-ups, consistent with the literature, do prioritize the pursuit of a social mission over commercial motives. The interviews highlighted the emphasis on achieving societal or environmental goals – often at the expense of growth or revenue targets. Social entrepreneurs value their autonomy, pay great attention to sustainability, and aim to achieve a credible impact. Yet, they are also aware of the competing and sometimes conflicting demands that come with hybridity, such as avoiding mission drift or having to raise external funding. Founders with a corporate background appeared particularly keen to distance themselves from their previous experiences in business.

The second insight relates to how the social-mission focus "affects how social entrepreneurs perceive and assess opportunities" (Dees, 1998, p. 2). The results of this study confirm that the social mission influences key design and operational decisions of a social start-up (Wilson & Post, 2013) – starting with its resource needs. The respondents regarded knowledge, personal support, and social capital as more pressing needs than the provision of financial or physical capital. A similar picture emerged in the assessment of sponsorship services, as the provision of funding ranked consistently lower than relational benefits, education, and validation. Intending to maximize social value, the founders expressed little interest in material gains and appeared selective when it came to seed funding or external investors. They sacrificed potential returns, sought to anchor their social impact at all operational levels, and in two cases (R2 and R3) chose a legal form that prevented them from distributing profits even though it reduced their appeal for investors.

Additionally, the results illustrate that pursuing a social mission is not a binary decision but manifests itself to varying degrees, as some founders prioritized the mission more than others. In a way, this mirrored their founding motivations, which indicated a broad spectrum of motives – in line with Germak and Robinson (2014, p. 18) observation that the motivations of social entrepreneurs are not one-dimensional but a "unique blend of motivational components". For the founders in this study, tackling societal challenges was an important but not exclusive motivation. Seeking fulfillment and developing innovative solutions were also major motivational drivers. In that respect, the founders showed significant overlap with founders of conventional start-ups.

It would be interesting to determine why the social-mission focus varied in intensity among the founders, as the motivation of social entrepreneurs has frequently been studied in the context of their personality traits or demographic factors (Gupta et al., 2020). However, owing to the research design and the relative homogeneity of the sample, this study could not observe a causal link between founder characteristics and the degree to which founders emphasized the social-mission focus. This focus was equally evident for both younger founders and more experienced entrepreneurs who enjoyed greater material security after successful prior careers.

Acquiring External Resources. Building on the social-mission focus and founding motivations, this study also explored how social start-ups acquire external resources. This is a key step for studying the acceleration process. According to the resource-based view, which underpins the theory of organizational sponsorship, incubators and accelerators provide nascent ventures with tangible and intangible resources to support their survival and growth (Carayannis & Von Zedtwitz, 2005; McAdam & McAdam, 2008; Rothaermel & Thursby, 2005).

The literature on social entrepreneurship features two main arguments regarding the resource needs of social start-ups. First, they require resource inputs broadly similar to those that conventional entrepreneurs need (Austin et al., 2012; Dacin et al., 2010). Second, they face challenges in attracting the resources necessary to scale (Austin et al., 2012; Lall et al., 2013). Hence, the resource-constrained environment in which social start-ups operate should make them receptive to external assistance to support their survival and growth (Pandey et al., 2017, p. 8).

The first argument is only partially confirmed by the results. Like most nascent ventures, the social start-ups in this

study expected an incubator or accelerator to provide knowledge and training, particularly about marketing and sales, as well as social capital and personal support. Surprisingly, however, the founders considered tangible resources such as funding and office space as less important than intangible ones – contrary to studies that identified access to tangible resources as a key benefit of an incubator (McAdam & McAdam, 2008; Van Weele et al., 2017).

To a certain degree, this result is contingent on the individual circumstances of the start-ups, including the maturity of most founders, the Impact Factory's decision not to provide seed funding, and the virtual delivery of most services due to Covid-19. Nonetheless, this finding illustrates how social entrepreneurs deal with resource needs. Far from representing a "fragile breed of startups" (J. Hausberg & Korreck, 2017, p. 2), the founders showed resilience and inventiveness in overcoming resource constraints. Faced with environmental barriers in attracting resources, the founders developed creative mechanisms to circumvent them (Dacin et al., 2010, p. 49).

This finding contradicts the linear way in which the resource-based view envisages the role of incubators or accelerators (Becker & Gassmann, 2006; Rothaermel & Thursby, 2005). Rather than relying on sponsors as primary resource providers, social start-ups leverage their relations with various stakeholders to overcome "institutional voids," as Doherty et al. (2014, p. 424) have observed. Their human capital requirements focus more on versatility and rapid knowledge acquisition to overcome resource barriers than is the case for conventional ventures (Harris & Kor, 2013). The growth process of social start-ups should, therefore, not be viewed as a consistent acquisition of resources, but rather as the development of dynamic capabilities – the ability to react to and capitalize on growth opportunities (Brown & Mawson, 2016, p. 820).

5.1.2. Contributions to the Literature on Incubators and Accelerators

Although the popularity of incubators and accelerators is growing, many researchers still treat them as a black box (Hackett & Dilts, 2008) by focusing on their organizational features or their impact on venture performance. Consequently, there have been recent calls to study the process (Crisan et al., 2019) or mechanisms of acceleration (Shankar & Clausen, 2020). In response to these calls, the present study offers three contributions in the context of social start-ups.

First, it clarified the portfolio of services that accelerators offer (Crisan et al., 2019, p. 2) by compiling an exemplary services portfolio (see Appendix), which describes and systematizes their most prominent activities based on 26 studies. Second, it explored how social start-ups "assess the value-proposition of social accelerators" (Pandey et al., 2017, p. 1) by evaluating these activities from the perspective of social start-ups. Third, it put "accelerators' heterogeneity at the core of the analysis" (Colombo et al., 2018, p. 193) by identifying how the acceleration of social start-ups is different from support for start-ups without a social-mission focus.

The Services Portfolio of Incubators and Accelerators. The services portfolio of incubators and accelerators – or in other words, how these programs support their ventures – has received surprisingly little attention. Existing studies of incubator and accelerator activities are mostly conceptual or specific to limited cases, as recent reviews on incubators (J. P. Hausberg & Korreck, 2021) and accelerators (Crisan et al., 2019) have illustrated. The lack of clarity on the support offerings of incubators and accelerators can be attributed to the novelty of the phenomenon, definitional challenges, and the empirical diversity of these programs.

This study contributes to the understanding of incubator and accelerator activities by compiling an exemplary services portfolio. Although the resulting table, based on 26 studies and reports, did not amount to a systematic review, it still offers an indication of the most prominent activities of incubators and accelerators. The informative value of the services portfolio was increased by using consistent terminology; differentiating between services and interventions (actual activities) and mechanisms (processes transforming activities into outcomes) following Crisan et al. (2019, p. 16); and summarizing the studies in three categories: conventional incubators, conventional accelerators, and social incubators/accelerators.

The exemplary portfolio of nine incubator and accelerator services (see Appendix and summarized as in Table 2 in Section 2.4) leads to the following observations. First, two services – "education and training" and "external networking" – were mentioned in almost all the reviewed studies, which emphasizes their prominence in the literature. Second, the activities of conventional incubators and accelerators feature remarkable overlap – aside from "coworking space" and "internal mentoring," which were most strongly associated with incubators and accelerators, respectively. The prominence of all other activities associated with the bridging or buffering mechanism in Table 2 was equal or within a range of 10%. This finding validates this study's decision, explained in Section 2.2.2, to treat incubators and accelerators as conceptually similar entities. Third, it suggests greater variation between conventional incubators/accelerators and those focused on social start-ups. For example, "peer support" and "access to external funding" seemed more prominent among social incubators/accelerators than among their conventional counterparts.

This last finding, however, also indicates the limitations of the services portfolio. While helpful in bringing clarity to a fragmented research area, it is still based on a small sample of studies and has little value in explaining causality. Moreover, it was compiled not only on statistical grounds but also to reflect potential theoretical salience (e.g., the inclusion of the boosting mechanism). It cannot account for novel support requirements of social start-ups left unaddressed in the reviewed studies. Consequently, the services portfolio primarily served as a research device to explore the acceleration of social start-ups in interviews with their founders.

The Value Proposition of Incubators and Accelerators.

The value of incubator and accelerator support for nascent ventures remains disputed. While some studies have reported lower chances of start-up survival (Schwartz, 2013) and achieving key milestones (Yu, 2020), others have found positive effects on exit financing (S. Smith et al., 2015), venture performance (C. S. R. Chan et al., 2020; Gonzalez-Uribe & Leatherbee, 2018), and funding and employee growth (Hallen et al., 2020). That said, studies claiming positive effects have faced selection problems and issues of limited generalizability (Stokan, Thompson, & Mahu, 2015).

Due to its exploratory research design, this study could not contribute to the debate on the quantifiable impact of incubators or accelerators on venture performance. It could, however, study how the beneficiaries of the acceleration process perceive the value of incubators or accelerators. This approach was motivated by recent calls to study incubators and accelerators from the perspective of start-ups (Colombo et al., 2018; Pauwels et al., 2016). To date, the only peer-reviewed study to explicitly explore what social entrepreneurs value in an accelerator is Pandey et al. (2017). That study reported that founding teams' human capital influences their perceptions of social accelerator benefits. The authors concluded their study of 4,000 social ventures by calling for a "fine-grained qualitative examination" (19), which the present study set out to achieve through the ranking of sponsorship services (Table 11) in combination with an in-depth exploration of nine incubator and accelerator activities.

Together, these findings represent one of the first in-depth assessments of incubator and accelerator benefits by social start-ups. While the results have been thoroughly discussed in Section 4.3, it is worth noting how forming two subgroups allowed the comparison of start-ups that participated in a full support program (ramp-ups) with those that enjoyed a limited interaction (fellows). Ultimately, aside from "validation and visibility" and "education and training," the cross-case analysis showed remarkable similarity in the value perceptions of both groups.

Social Start-Up Acceleration. There is great interest in "examining a wide range of organizational contexts in which accelerators operate" (Crisan et al., 2019, p. 20). However, despite the popularity of studies on accelerators run by corporations or the public sector, support for social enterprises has received less scholarly scrutiny (J. Hausberg & Korreck, 2017). Beneath this research gap lies a question of theoretical and practical importance: Is the acceleration of social start-ups a unique organizational context? Or, phrased from the perspective of the beneficiaries, do social start-ups "require a different kind of assistance" (J. Hausberg & Korreck, 2017, p. 13)?

Recent studies on this issue, referenced in Section 2.2.3, have indicated empirical differences between social incubators and their conventional counterparts (Sansone et al., 2020; SIM, 2020). The exemplary services portfolio (see Appendix) also indicates variation in service offerings. Consistent with these results, the present study provides a strong indication that social start-ups have unique support needs that cannot be met by conventional incubator and accelerator benefits. Consequently, this study proposes a novel support mechanism for start-ups aiming to solve societal or environmental issues: impact acceleration.

Importantly, the suggestion of a novel mechanism does not imply that the original sponsorship mechanisms are ineffective for accelerating social start-ups. On the contrary, there was significant overlap in the support needs of conventional and social start-ups. However, this study identifies two services – demonstrating social impact and delivering the social mission – that are not traditionally offered by incubators and accelerators. According to the interviewed founders, these activities differentiate the acceleration of social start-ups from supporting other ventures.

The benefits of incubator specialization have already been highlighted by multiple studies. Graduating from an incubator with a focus on the SDGs helps social start-ups to communicate their impact (Gianoncelli et al., 2020). Alignment between the service portfolio of incubators and accelerators and their tenant profiles increases the impact of their interventions (Bruneel et al., 2012). The key in increasing customer value for tenants with a specialist stance is differentiation through the selection process and service offerings (Vanderstraeten et al., 2020). These findings reinforce this study's claim that social start-ups are best served by specialized programs.

5.1.3. Contributions to the Literature on Organizational Sponsorship

Following the observation that "knowledge about the form (accelerator) is incomplete without knowledge about the mechanism (acceleration)" (Shankar & Clausen, 2020, p. 2), this study used the framework of organizational sponsorship to explore social start-up acceleration. In doing so, it had to contend with the facts that organizational sponsorship is far from an established theory and that its mechanisms are still "somewhat unclear" and "may differ across sponsorship phenomena" (Breivik-Meyer, 2020, p. 185). However, the novelty of the framework provided fertile ground for exploring social start-ups' perceptions of different sponsorship mechanisms.

This research yielded two main contributions to the literature on organizational sponsorship. First, it clarified how social start-ups perceive the sponsorship mechanisms of buffering, bridging, and boosting. Second, it reflected the trade-off between the benefits of sponsorship and its disadvantages.

The Mechanisms of Organizational Sponsorship. Owing to their novelty, the mechanisms of organizational sponsorship remain undefined. While most authors accept the mechanisms of buffering and bridging proposed by Amezcua et al. (2013), others have suggested that sponsors should engage in bolstering (Pandey et al., 2017), curating (Amezcua et al., 2020), sheltering (Breivik-Meyer et al., 2019), and boosting (Autio & Rannikko, 2016). This theoretical fragmentation sparked two questions about incubators and accelerators: Which activities correspond to these mechanisms, and how do start-ups value them?

This study took two steps to clarify this debate. The first step was to disregard all mechanisms that appeared insufficiently distinct from the two original mechanisms (i.e., all but boosting); the second was to match the nine exemplary interventions to the three remaining mechanisms to assess how social start-ups perceive their value. This process generated two main insights into how social start-ups view the mechanisms of organizational sponsorship.

First, the focus of the early sponsorship literature on increasing venture survival (Amezcua et al., 2013; Flynn, 1993b) does not reflect the reality of social start-ups. Rather than relying on an organizational sponsor to shield them from market pressures – the idea at the core of the buffering mechanism – social start-ups appear to be sufficiently resilient and inventive to build up their resource base on their own. A number of the activities associated with buffering, such as education and training and internal mentoring, were highly valued by the founders – but primarily to accelerate their learning, not to ensure their survival. By contrast, they rated the provision of material resources – buffering in a literal sense – as far less important than networking, a benefit related to bridging. This result is consistent with Pandey et al. (2017, p. 19) early indication that social accelerators "engage in less buffering and more bridging". Because of their strong external orientation, social start-ups also seemed unaffected by the risks of resource munificence often associated with the buffering mechanism, such as missing important early feedback (Cohen, 2013) or facing the "life support incubation trap" (Clarysse et al., 2015, p. 19).

The second insight is that the social start-ups only partially valued the boosting mechanism proposed by Autio and Rannikko (2016). The founders welcomed incubator support in boosting their capacity for growth – but not in controlling milestones. In fact, only two founders expressed a favorable opinion of milestones set by an incubator. Most founders appeared sufficiently driven without the need for external intervention. Some even suggested an incubator could help founders slow down and find stability, rather than bringing even more speed into the process. In that respect, impact acceleration seems to differ from the "time-compressed scaling" that characterizes support from conventional accelerators (Shankar & Clausen, 2020).

The Trade-Off Between Sponsorship and Self-Reliance. One major contribution of Amezcua et al. was to challenge the notion that resource munificence is universally beneficial for nascent ventures (2013). Instead of supporting a one-size-fits-all approach, the authors argued for a contingent approach that accounts for the heterogeneity of ventures and environmental conditions. Similarly, Jourdan and Kivleniece (2017) found that the positive effects of sponsorship can be overshadowed by a loss of incentives and internal discipline beyond a certain level.

Despite their largely positive assessment of incubator support, the founders in this study confirmed that seeking external support is a deliberate trade-off. The founders displayed a strikingly broad range of attitudes towards formal support programs. While some refused to participate in any formal program, others took advantage of five or more support programs.

Yet, in explaining their reasons for joining or not joining an incubator, the founders rarely mentioned the risks of resource munificence that the sponsorship literature has highlighted. Rather, the founders were concerned an incubator might jeopardize their autonomy or be a source of rejection. The motives for entrepreneurial self-reliance confirmed Patton et al. (2009) observation that sharing ownership and control of a new business in an incubator can be a "significant source of anxiety" (p. 628), with a "fine line between interference and support" (p. 633).

Even though founders valued an incubator's help in reducing uncertainty and providing validation (Carayannis & Von Zedtwitz, 2005), they also appeared highly selective in weighing the potential benefits and disadvantages of an incubation program – and the time required. Hence, the following section discusses how funders and managers of support programs can ensure that their offerings provide the most value and relevance for social start-ups.

5.2. Practical Implications

Substantial resources are currently being invested in the incubation of early-stage social ventures. Governments, corporations, and private foundations view social start-ups as a promising solution to tackling societal and environmental challenges. However, there is still considerable uncertainty about the optimal design of entrepreneurial support activities. Nonetheless, precisely because of the ongoing debate about the impact of incubators and accelerators on venture performance, there is "great value in understanding the mechanisms that make sponsorship more effective" (Breivik-Meyer, 2020, p. 185). To realize this ambition, this study describes practical implications for three stakeholder groups that would benefit from maximizing the value of social start-up acceleration: funders, incubator and accelerator managers, and social entrepreneurs.

5.2.1. Funders of Incubators and Accelerators

How accelerators structure and run their programs is "largely determined by the objectives of their key shareholders" (Pauwels et al., 2016, p. 22), the influence of whom stems mostly from their financial contributions. In Germany, much of the revenue of incubators and accelerators depends on grants or subsidies – on average, 50% (SIM, 2020, p. 48). The reliance on external funding is exacerbated by a focus on social start-ups, which promise limited financial returns. As a result, funders have a crucial role in shaping support programs.

The essential takeaway of this study for funders is that the right support can, according to social start-ups themselves, have a positive effect on their development. Provided incubators and accelerators consider the unique characteristics of social enterprises, starting from their social-mission focus and resource needs, they can effectively support these organizations' growth. To put this into practice, funders should consider the following recommendations:

- Fund dedicated programs: The complexity of supporting social start-ups entails higher costs. Social start-ups require specialized knowledge on impact measurement, hybrid legal structures, and double-bottom-line financing that is not as readily available as expertise about conventional business incubation. Regardless, funders would benefit from tailoring their programs, for example, by offering dedicated services for impact acceleration.
- Readjust financial expectations: The hybrid nature of social start-ups means that the venture capital model of many accelerators is inapplicable. Instead of planning for a quick exit, social start-ups think long-term and renounce profits to aid their mission. Even if the impact of social start-ups is harder to quantify, their acceleration holds the promise of generating significant societal and environmental returns. Funders should be patient – and shift their return expectations accordingly.
- Evaluate and adapt continuously: The value that incubators and accelerators create for social start-ups depends on many external factors, from shifts in the political landscape to changes in target markets and founder demographics. Funders should continuously evaluate – and adapt – the programs they offer based on changing circumstances, in the same way that start-ups weigh the advantages and disadvantages of support programs.
- Put values into practice: Incubators and accelerators can be a source of emotional support and stability during a precarious and highly stressful period for founders. However, building a culture of trust and humility requires funders to set the right example and put these values into practice. How funders set goals, recruit staff, and allocate resources determines how social start-ups experience their time within an incubator or accelerator.

5.2.2. Managers of Incubators and Accelerators

The managers of incubators and accelerators play a vital role in the acceleration of start-ups. Experienced incubator managers reduce the risk of start-up failures (Wise & Valliere, 2014). The duration, intensity, and range of their interventions strongly affect the business assistance that start-ups receive (Rice, 2002). To increase the effectiveness of their support for social start-ups, incubator managers could, therefore, consider the following recommendations:

- Select the right teams: Value alignment within cohorts is crucial to leverage peer effects.
- Customize education and training activities: Start-ups appreciate when their practical needs are addressed – especially regarding legal issues and raising external funding.
- Build a wide and relevant network: Social capital is potentially the main incubator benefit. Incubator managers need to invest significant resources in expanding their network.
- Nurture personal relations: Experienced mentors and coaches can be highly effective in helping founders keep their focus, deal with complex challenges, and avoid mission drift.
- Avoid dependencies: Social start-ups need to stand independently – or fail – quickly.
- Invest in the brand: Only well-known incubators provide strong validation signals.
- Ensure the social impact: Measuring, demonstrating, and delivering impact is a key benefit of incubators and accelerators, especially for founders with corporate background.

5.2.3. Founders of Social Start-ups

Far from being passive recipients of incubator services, start-up founders actively shape the incubation process – an effect called coproduction (Rice, 2002) and codevelopment (Vanderstraeten et al., 2020). An accelerator's impact is driven by the accelerator and its applicants "in ways that are difficult to untangle" (Cohen, Fehder, et al., 2019, p. 1783). As a result, founders of social start-ups should heed the following recommendations:

- Allocate sufficient resources: To reap the benefits of a support program, social start-ups first need to critically reflect on the time and resources that active participation requires.
- Determine the right fit: Specialized programs may be more effective, but they also place greater burden on the selection process to maximize fit between incubators and start-ups.
- Stay open-minded: Once they are part of a program, founders need to accept their cognitive biases, recognize their knowledge gaps, and remain open to being challenged.
- Practice self-care: Social entrepreneurs are at risk of burnout and overworking. To solve complex societal challenges, funders must first preserve their personal well-being.
- Give back: A key benefit of an incubator or accelerator is peer networking. Social start-ups should invest sufficient time into learning from other start-ups.
- Focus on the social mission: Pursuing a double-bottom line forces social entrepreneurs to juggle a myriad of demands. An incubator can assist them in focusing on the core mission.

5.3. Limitations and Future Research

As an exploratory case study within the research setting of a single incubator, this study's research design had limitations. In a novel and rapidly evolving field such as the acceleration of social start-ups, these issues also constitute research opportunities. The following section outlines five of these limitations and how they might spark further research.

First, the interview method could have led to possible bias. To improve internal and external validity, a standardized questionnaire was combined with further secondary sources. Regardless, the subjective views of the interviewees, the presence of the interviewer, or a poor recollection of past events may have affected the data collection. Moreover, nascent entrepreneurs are not always aware of their knowledge gaps (Vanderstraeten et al., 2020). These limitations could be addressed by increasing the sample size and interviewing multiple team members. Human capital factors also warrant further examination. How does the background and profile of social entrepreneurs influence their perception of incubator benefits? A quantitative study on the support services identified in this thesis could validate its qualitative results.

The second limitation concerns the units of analysis (i.e., social start-ups). There were valid reasons for adopting the perspective of the beneficiaries of incubator services. However, the robustness and reliability of the results could be strengthened by extending the analysis to managers of incubators and other stakeholders who engage in the acceleration process, such as coaches, mentors, and investors. This approach would reflect the interdependencies of these actors.

The third limitation pertains to the generalizability of the findings due to the lack of comparative data. It would be fruitful to compare social start-ups with a control group of conventional start-ups (i.e., without a social mission), either within the same incubation environment or across different support programs. Doing so could confirm that impact acceleration is a unique sponsorship mechanism not sufficiently provided by traditional sponsorship mechanisms. It would also be interesting to compare the support needs of for-profit ventures with those of nonprofit organizations, which are likely to display different acceleration needs. A longitudinal assessment could help validate the observation that the acceleration needs of social start-ups quickly evolve. Further research could compare incubators or accelerators in other locations to confirm whether observations in a high-income country such as Germany are transferable to other regional contexts.

The fourth limitation relates to the effect that acceleration has on the societal or environmental challenges that start-ups seek to address. Questions remain regarding whether incubator interventions make a measurable difference to achieving a social mission and how this impact could be increased through appropriate support activities. Tracking established impact metrics and economic performance indicators, such as revenue growth, increased headcount, or raised funds, could help evaluate the effect of entrepreneurial support programs on both bottom lines of social start-ups.

The final limitation revolves around the effects of Covid-19, which the World Health Organization declared a global pandemic in March 2020. As a result, the Impact Factory had to move the delivery of its program to an all-virtual setting for a significant part of the analyzed period, which ran from June 2019 to October 2020. Surprisingly, the respondents reported that the pandemic had a limited impact on their program experience, for example in regard to the coworking space and the lack of on-site networking opportunities. Nevertheless, one should assume that the unique circumstances of Covid-19 had a considerable effect on the results of this study. It would be interesting to compare the views expressed in this study with a post-Covid-19 incubator experience, for example in terms of meeting face-to-face for networking or mentoring.

6. Conclusion

Social start-ups differ from other entrepreneurial ventures in how they combine commercial activities and a social mission. Increasing access to electricity in Sub-Saharan Africa, protecting biodiversity in Peru, and reducing waste created by disposable food packaging are just three examples of how the start-ups in this study are tackling societal and environmental problems. Although these ventures have the potential to achieve an impact on a global scale, they also face distinct challenges in acquiring the resources required to grow their business and impact models.

Incubators and accelerators have emerged as one of the most popular support models for early-stage social ventures. By providing tangible and intangible resources, these support programs can help social entrepreneurs to develop and grow their ventures. However, there is still insufficient understanding of how social start-ups perceive the value proposition of incubators and accelerators – and of whether social start-ups have unique support requirements.

This study explored the acceleration of social start-ups through an inductive case study of 10 social start-ups in an impact-oriented incubator. It found that their strong social-mission focus affected their resource needs. Social start-ups are resilient, creative, and independent. They are primarily interested in expanding their networks and acquiring new knowledge, and they are highly selective when it comes to sharing control and raising capital from external investors.

To improve their value proposition for social start-ups, incubators and accelerators should adapt their service offerings accordingly. Relying on the mechanisms of organizational sponsorship – buffering, bridging, and boosting – is insufficient. Therefore, this study proposed the new support mechanism of impact acceleration. By helping social start-ups demonstrate their impact and deliver on their mission, incubator and accelerators can overcome the "pioneer gap" and support high-impact ventures in their most critical phases (Lall et al., 2013).

This study showed that social start-ups are cautious in how they allocate their time – but also appreciative of the right support at the right moment. According to R2, "If

you're a founder, the only thing you're looking for is problem solvers... And if you have one, you don't let them go." A successful incubator or accelerator can be more than the sum of its parts. It can create a community of changemakers who assist, motivate, and inspire each other every day. As F1 said, "Start-ups live from being together, from suffering together, but also from celebrating together."

To understand the support needs of social start-ups, further research is required. Hopefully, the findings of this study will facilitate further study of this emerging – and exciting – phenomenon. Understanding how incubators and accelerators can optimize their support for social start-ups would not only benefit funders and the managers of these organizations, but also ventures seeking to solve some of the world's most urgent problems by putting their mission first.

References

Adkins, D. (2011). What are the new seed or venture accelerators? *NBIA Review*. Retrieved from https://inbia.org/wp-content/uploads/2018/07/Accelerators.pdf?x62369

Aernoudt, R. (2004). Incubators: Tool for entrepreneurship? *Small Business Economics*, *32*(2), 127–135.

Albort-Morant, G., & Ribeiro-Soriano, D. (2016). A bibliometric analysis of international impact of business incubators. *Journal of Business Research*, *69*(5), 1775–1779.

Aldrich, E. R., H. auch Auster. (1986). Even dwarfs started small: Liabilities of age and size and their strategic implications. *Research in Organizational Behavior*, *8*, 165–198.

Alter, S. K. (2003). *Social Enterprise: A Typology of the Field Contextualized in Latin America.* Retrieved from https://publications.iadb.org/publications/english/document/Social-Enterprise-A-Typology-of-the-Field-Contextualized-in-Latin-America.pdf

Amezcua, A., Grimes, M. G., Bradley, S. W., & Wiklund, J. (2013). Organizational sponsorship and founding environments: A contingency view on the survival of business-incubated firms, 1994–2007. *Academy of Management Journal*, *56*(6), 1628–1654.

Amezcua, A., Ratinho, T., Plummer, L. A., & Jayamohan, P. (2020). Organizational sponsorship and the economics of place: How regional urbanization and localization shape incubator outcomes. *Journal of Business Venturing*, *35*(4), 105967.

Arena, M., Bengo, I., Calderini, M., & Chiodo, V. (2018). Unlocking finance for social tech start-ups: Is there a new opportunity space? *Technological Forecasting and Social Change*, *127*, 154–165.

Aspen Network. (2014). *Measuring Value Created by Impact Incubators & Accelerators.* Retrieved from https://www.andeglobal.org/resource/dynamic/blogs/20150609_114557_10161.pdf

Austin, J., Stevenson, H., & Wei-Skillern, J. (2012). Social and commercial entrepreneurship: Same, different, or both? *Revista de Administração*, *47*(3), 370–384.

Autio, E., & Rannikko, H. (2016). Retaining winners: Can policy boost high-growth entrepreneurship? *Research Policy*, *45*(1), 42–55.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, *17*(1), 99–120.

Becker, B., & Gassmann, O. (2006). Gaining leverage effects from knowledge modes within corporate incubators. *R&D Management*, *36*(1), 1–16.

Bergek, A., & Norrman, C. (2008). Incubator best practice: A framework. *Technovation*, *28*(1–2), 20–28.

Bioulac, A., Ditsche, J., & Dujacquier, D. (2019). *Revising the Market for Innovation: How Accelerators and Incubators can Reinvent Themselves.* Retrieved from https://www.rolandberger.com/publications/publication_pdf/roland_berger_revisiting_the_market_for_innovation.pdf

Bleijenbergh, I. (2010). Case selection. In A. Mills, G. Durepos, & E. Wiebe (Eds.), *Encyclopedia of Case Study Research* (Vol. 2, pp. 61–63). SAGE Publications.

Branscomb, L., & Auerswald, P. E. (2002). *Between Invention and Innovation - an Analysis of Funding for Early-Stage Technology Development.* Retrieved from https://www.nist.gov/system/files/documents/2017/05/09/gcr02-841.pdf

Breivik-Meyer, M. (2020). Organizational sponsorship: An overview of the state of knowledge and future research directions. In A. Novotny, E. Rasmussen, T. Clausen, & J. Wiklund (Eds.), *Research Handbook on Start-Up Incubation Ecosystems* (pp. 174–193). Edward Elgar Publishing.

Breivik-Meyer, M., Arntzen-Nordqvist, M., & Alsos, G. A. (2019). The role of incubator support in new firms accumulation of resources and capabilities. *Innovation*, *22*(3), 228–249.

Brixy, U., Sternberg, R., & Stüber, H. (2013). Why some nascent entrepreneurs do not seek professional assistance. *Applied Economics Letters*, *20*(2), 157–161.

Brown, R., & Mawson, S. (2016). Targeted support for high growth firms: Theoretical constraints, unintended consequences and future policy challenges. *Environment and Planning C: Government and Policy*, *34*(5), 816–836.

Bruckschlögl, R. (2019). *Bits & Pretzels 2019: „There are Gretas everywhere".* *Munich Startup.* Retrieved from https://www.munich-startup.de/53289/bits-pretzels-2019-tag-1/

Bruneel, J., Ratinho, T., Clarysse, B., & Groen, A. (2012). The evolution of business incubators: Comparing demand and supply of business incubation services across different incubator generations. *Technovation*, *32*(2), 110–121.

Carayannis, E. G., & Von Zedtwitz, M. (2005). Architecting gloCal (global–local), real-virtual incubator networks (G-RVINs) as catalysts and accelerators of entrepreneurship in transitioning and developing economies: Lessons learned and best practices from current development and business incubation practices. *Technovation*, *25*(2), 95–110.

Casasnovas, G., & Bruno, A. V. (2013). Scaling social ventures. *Journal of Management for Global Sustainability*, *1*(2), 173–197.

Chan, C. S. R., Patel, P. C., & Phan, P. H. (2020). Do differences among accelerators explain differences in the performance of member ventures? evidence from 117 accelerators in 22 countries. *Strategic Entrepreneurship Journal*, *14*(2), 224–239.

Chan, K., & Lau, T. (2005). Assessing technology incubator programs in the science park: The good, the bad and the ugly. *Technovation*, *25*(10), 1215–1228.

Chmiliar, L. (2010). Multiple-case designs. In A. Mills, G. Durepos, & E. Wiebe (Eds.), *Encyclopedia of Case Study Research* (pp. 582–584). SAGE Publications.

Clarysse, B., Wright, M., & Van Hove, J. (2015). *A look inside accelerators.* Retrieved from https://media.nesta.org.uk/documents/a_look_inside_accelerators.pdf

Clayton, P., Feldman, M., & Lowe, N. (2018). Behind the scenes: Intermediary organizations that facilitate science commercialization through entrepreneurship. *Academy of Management Perspectives*, *32*(1), 104–124.

Cohen, S. (2013). What do accelerators do? insights from incubators and angels. *Innovations: Technology, Governance, Globalization*, *8*(3-4), 19–25.

Cohen, S., Bingham, C. B., & Hallen, B. L. (2019). The role of accelerator designs in mitigating bounded rationality in new ventures. *Administrative Science Quarterly*, *64*(4), 810–854.

Cohen, S., Fehder, D. C., Hochberg, Y. V., & Murray, F. (2019). The design of startup accelerators. *Research Policy*, *48*(7), 1781–1797.

Cohen, S., & Hochberg, Y. V. (2014). *Accelerating Startups: The Seed Accelerator Phenomenon.* Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2418000

Colombo, M. G., Rossi-Lamastra, C., & Wright, M. (2018). Chapter 10: Accelerators: insights for a research agenda. In M. Wright & I. Drori (Eds.), *Accelerators - successful venture creation and growth* (pp. 188–204). Edward Elgar Publishing Limited.

Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications.

Crisan, E. L., Salanță, I. I., Beleiu, I. N., Bordean, O. N., & Bunduchi, R. (2019). A systematic literature review on accelerators. *The Journal of Technology Transfer*, *46*(1), 62–89.

Dacin, P. A., Dacin, M. T., & Matear, M. (2010). Social entrepreneurship: Why we don't need a new theory and how we move forward from here. *Academy of Management Perspectives*, *24*(3), 37–57.

Dees, J. G. (1998). *The Meaning of Social Entrepreneurship.* Retrieved from https://centers.fuqua.duke.edu/case/wp-content/uploads/sites/7/2015/03/Article_Dees_MeaningofSocialEntrepreneurship_2001.pdf

Dempwolf, C. S., Auer, J., & D'Ippolito, M. (2014). *Innovation Accelerators: Defining Characteristics Among Startup Assistance Organizations.* Retrieved from https://www.sba.gov/sites/default/files/rs425-Innovation-Accelerators-Report-FINAL.pdf

Doherty, B., Haugh, H., & Lyon, F. (2014). Social enterprises as hybrid organizations: a review and research agenda. *International Journal of Management Reviews*, *16*(4), 417–436.

Dooley, L. M. (2002). Case study research and theory building. *Advances in Developing Human Resources*, *4*(3), 335-354.

Dorado, S., & Ventresca, M. J. (2013). Crescive entrepreneurship in complex social problems: Institutional conditions for entrepreneurial engagement. *Journal of Business Venturing*, *28*(12), 69–82.

Drori, I., & Wright, M. (2018). Chapter 1: Accelerators: Characteristics, trends and the new entrepreneurial ecosystem. In I. Drori & M. Wright (Eds.), *Accelerators - successful venture creation and growth* (pp. 1–20). Edward Elgar Publishing.

DSEM. (2020). *Deutscher Social Entrepreneurship Monitor 2019.* Retrieved from https://www.send-ev.de/wp-content/uploads/2021/03/DSEM2019.pdf

Duff, A. (1994). *Best Practice in Business Incubator Management.* Retrieved from http://docplayer.net/18929817-Best-practice-business-incubator-management-andrew-duff.html

Dutt, N., Hawn, O., Vidal, E., Chatterji, A., McGahan, A., & Mitchell, W. (2016). How open system intermediaries address institutional failures: The case of business incubators in emerging-market countries. *Academy of Management Journal*, *59*(3), 818–840.

Ebrahim, A., & Rangan, V. K. (2014). What impact? a framework for measuring the scale and scope of social performance. *California Management Review*, *56*(3), 118–141.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, *14*(4), 532–550.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of Management Journal*, *50*(1), 25–32.

evers & jung. (2016). *Herausforderungen bei der Gründung und Skalierung von Sozialunternehmen. Welche Rahmenbedingungen benötigen Social Entrepreneurs?* Retrieved from https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/herausforderungen-bei-der-gruendung-und-skalierung-von-sozialunternehmen.pdf?__blob=publicationFile&v=13

Flynn, D. M. (1993a). A critical exploration of sponsorship, infrastructure, and new organizations. *Small Business Economics*, *5*(2), 129–156.

Flynn, D. M. (1993b). Sponsorship and the survival of new organizations. *Journal of Small Business Management*, *31*(1), 51–62.

Galbraith, B., McAdam, R., & Cross, S. E. (2019). The evolution of the incubator: Past, present, and future. *IEEE Transactions on Engineering Management*, *68*(1), 265–271.

Germak, A. J., & Robinson, J. A. (2014). Exploring the motivation of nascent social entrepreneurs. *Journal of Social Entrepreneurship*, *5*(1), 5–21.

Gianoncelli, A., Gaggiotti, G., Miguel, A., & Charro, I. (2020). *Enablers of Impact - The Role of Incubators and Accelerators in Bridging Investment and Solutions.* Retrieved from https://evpa.eu.com/uploads/publications/EVPA_MAZE-Enablers_of_Impact_report_2020.pdf

Global Accelerator Learning Initiative. (2020). *The Entrepreneurship Database Program 2019 Data Summary.* Retrieved from https://www.galidata.org/publications/the-entrepreneurship-database-program-2019-data-summary/

Gonzalez-Uribe, J., & Leatherbee, M. (2018). The effects of business accelerators on venture performance: Evidence from start-up chile. *The Review of Financial Studies*, *31*(4), 1566–1603.

Grant, R. M. (1991). The resource-based theory of competitive advantage: implications for strategy formulation. *California management review*, *33*(3), 114–135.

Gupta, P., Chauhan, S., Paul, J., & Jaiswal, M. P. (2020). Social entrepreneurship research: A review and future research agenda. *Journal of Business Research*, *113*, 209–229.

Hackett, S. M., & Dilts, D. M. (2004). A systematic review of business incubation research. *The Journal of Technology Transfer*, *29*(1), 55–82.

Hackett, S. M., & Dilts, D. M. (2008). Inside the black box of business incubation: Study b—scale assessment, model refinement, and incubation outcomes. *The Journal of Technology Transfer*, *33*(5), 439–471.

Hallen, B. L., Cohen, S. L., & Bingham, C. B. (2020). Do accelerators work? if so, how? *Organization Science*, *31*(2), 378–414.

Harris, D., & Kor, Y. (2013). The role of human capital in scaling social entrepreneurship. *Journal of Management for Global Sustainability*, *1*(2).

Hausberg, J., & Korreck, S. (2017). Mapping the social business incubator landscape. In *Paper to be presented at the 1st iese-luiss conference on responsibility, sustainability and social entrepreneurship, rome.*

Hausberg, J. P., & Korreck, S. (2021). Business incubators and accelerators: a co-citation analysis-based, systematic literature review. *Handbook of Research on Business and Technology Incubation and Acceleration*.

Hochberg, Y. V. (2016). Accelerating entrepreneurs and ecosystems: The seed accelerator model. *Innovation policy and the economy*, *16*(1), 25–51.

Impact Factory. (2020a). *Homepage.* Retrieved from https://impact-factory.de/

Impact Factory. (2020b). *Impact Factory - Internal Presentation 2020.*

Inkpen, A. C., & Tsang, E. W. (2005). Social capital, networks, and knowledge transfer. *Academy of management review*, *30*(1), 146–165.

International Labour Organization. (2020). *Sectors Covered.* Retrieved from https://www.ilo.org/sector/sectors-covered/lang--en/index.htm

Isabelle, D. (2013). Key factors affecting a technology entrepreneur's choice of incubator or accelerator. *Technology innovation management review*, 16–22.

Jourdan, J., & Kivleniece, I. (2017). Too much of a good thing? the dual effect of public sponsorship on organizational performance. *Academy of Management Journal*, *60*(1), 55–77.

Katre, A., & Salipante, P. (2012). Start–up social ventures: Blending fine–grained behaviors from two institutions for entrepreneurial success. *Entrepreneurship Theory and Practice*, *36*(5), 967–994.

King, M. W., Bogusky, A., Krout, T., Schuham, R., Neff, J., Webster, B., & Pintal, C. (2015). *Impact Accelerators: Strategic Options for Development and Implementation.* Retrieved from https://matthewkingphd.com/wp-content/uploads/2016/06/Impact-Accelerators_Strategic_Options_for_Development_and_Implementation_12_30_15.pdf

Lall, S. A. (2017). Measuring to improve versus measuring to prove: Understanding the adoption of social performance measurement practices in nascent social enterprises. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, *28*(6), 2633–2657.

Lall, S. A., Bowles, L., & Baird, R. (2013). Bridging the "pioneer gap": The role of accelerators in launching high-impact enterprises. *Innovations: Technology, Governance, Globalization*, *8*(3-4), 105–137.

Lall, S. A., Chen, L.-W., & Roberts, P. W. (2020). Are we accelerating equity investment into impact-oriented ventures? *World Development*, *131*, 104952.

Leirich, L. (2020). *Das Entscheidungsverhalten von Social Entrepreneurship unterstützenden Organisationen bei der Auswahl von Social Ventures - Eine Untersuchung der Unterstützungslandschaft für Social Entrepreneurship im DACH-Raum (Doctoral Dissertation).* Retrieved from https://ubt.opus.hbz-nrw.de/opus45-ubtr/frontdoor/deliver/index/docId/1466/file/DissertationLilliLeirich.pdf

Mair, J., & Marti, I. (2006). Social entrepreneurship research: A source of explanation, prediction, and delight. *Journal of World Business*, *41*(1), 36–44.

Marmer, M., Herrmann, B. L., Dogrultan, E., Berman, R., Eesley, C., & Blank, S. (2011). *Startup Genome Report Extra: Premature Scaling.* Retrieved from https://integral-entrepreneurship.org/wp-content/uploads/2016/07/Startup-Genome-Premature-Scaling.pdf

McAdam, M., & McAdam, R. (2008). High tech start-ups in University Science Park incubators: The relationship between the start-up's lifecycle progression and use of the incubator's resources. *Technovation*, *28*(5), 277–290.

Mian, S., Lamine, W., & Fayolle, A. (2016). Technology Business Incubation: An overview of the state of knowledge. *Technovation*, *50-51*, 1–12.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. SAGE Publications.

Miller, P., & Bound, K. (2011). *The Startup Factories.* Retrieved from https://www.nesta.org.uk/report/the-startup-factories/

Miller, P., & Stacey, J. (2014). *Good Incubation - The Craft Of Supporting Early–Stage Social Ventures.* Retrieved from https://media.nesta.org.uk/documents/good_incubation_wv.pdf

Nicholls, A. (2006). *Social Entrepreneurship: New Models of Sustainable Social Change*. Oxford University Press.

Nicolopoulou, K., Karataş-Özkan, M., Vas, C., & Nouman, M. (2017). An incubation perspective on social innovation: the London Hub–a social

incubator. *R&D Management*, *47*(3), 368–384.
Pandey, S., Lall, S. A., Pandey, S. K., & Ahlawat, S. (2017). The appeal of social accelerators: What do social entrepreneurs value? *Journal of Social Entrepreneurship*, *8*(1), 88–109.
Patton, D., Warren, L., & Bream, D. (2009). Elements that underpin high-tech business incubation processes. *The Journal of Technology Transfer*, *34*(6), 621–636.
Pauwels, C., Clarysse, B., Wright, M., & Van Hove, J. (2016). Understanding a new generation incubation model: The accelerator. *Technovation*, *50*, 13–24.
Perrini, F., & Vurro, C. (2006). Social entrepreneurship: Innovation and social change across theory and practice. In *Social Entrepreneurship* (pp. 57–85). Springer.
Radojevich-Kelley, N., & Hoffman, D. L. (2012). Analysis of accelerator companies: An exploratory case study of their programs, processes, and early results. *Small Business Institute Journal*, *8*(2), 54–70.
Renko, M. (2013). Early challenges of nascent social entrepreneurs. *Entrepreneurship Theory and Practice*, *37*(5), 1045-1069.
Rice, M. P. (2002). Co-production of business assistance in business incubators: An exploratory study. *Journal of Business Venturing*, *17*(2), 163–187.
Roberts, P. W., & Lall, S. A. (2018). *Observing Acceleration: Uncovering the Effects of Accelerators on Impact-Oriented Entrepreneurs*. Springer.
Roche, C. J. (1999). *Impact Assessment for Development Agencies: Learning to Value Change*. Oxfam.
Rothaermel, F. T., & Thursby, M. (2005). University–incubator firm knowledge flows: Assessing their impact on incubator firm performance. *Research Policy*, *34*(4), 305–320.
Sansone, G., Andreotti, P., Colombelli, A., & Landoni, P. (2020). Are social incubators different from other incubators? Evidence from Italy. *Technological Forecasting and Social Change*, *158*, 120132.
Santos, F. M. (2012). A positive theory of social entrepreneurship. *Journal of Business Ethics*, *111*(3), 335–351.
Schwartz, M. (2013). A control group study of incubators' impact to promote firm survival. *The Journal of Technology Transfer*, *38*(3), 302–331.
Shankar, R. K., & Clausen, T. H. (2020). Scale quickly or fail fast: An inductive study of acceleration. *Technovation*, *98*, 102174.
Shieber, J. (2017). Techstars launches an impact accelerator for companies focused on social and environmental impact. *TechCrunch*. Retrieved from https://techcrunch.com/2017/10/18/techstars-launches-an-impact-accelerator-for-companies-focused-on-social-and-environmental-impact/?guccounter=2
SIM. (2020). *Social Innovation Monitor - Full Report on Incubators and Accelerators in Germany.* Retrieved from https://socialinnovationmonitor.com/en/report-incubators/
Smith, S., Hannigan, T., & Gasiorowski, L. (2013). Accelerators and crowdfunding: Complementarity, competition, or convergence in the earliest stages of financing new ventures? In *Paper presented at the University of Colorado-Kauffman Foundation Crowd-Funding Conference, Boulder, CO.*
Smith, S., Hannigan, T., & Gasiorowski, L. (2015). Peering inside: How do peer effects impact entrepreneurial outcomes in accelerators. In *Paper presented at the Academy of Management Proceedings.*
Smith, W. K., & Lewis, M. W. (2011). Toward a theory of paradox: A dynamic equilibrium model of organizing. *Academy of Management Review*, *36*(2), 381–403.
Sonne, L. (2012). Innovative initiatives supporting inclusive innovation in India: Social business incubation and micro venture capital. *Technological Forecasting and Social Change*, *79*(4), 638–647.
Stinchcombe, A. L. (1965). Organizations and Social Structure. In J. March (Ed.), *Handbook of Organizations* (pp. 153–193). Rand McNally.
Stokan, E., Thompson, L., & Mahu, R. J. (2015). Testing the differential effect of business incubators on firm growth. *Economic Development Quarterly*, *29*(4), 317–327.
Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, *20*(3), 571–610.
Tracy, S. J. (2010). Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qualitative Inquiry*, *16*(10), 837–851.
Triebel, C., Schikora, C., Graske, R., & Sopper, S. (2018). Failure in startup companies: why failure is a part of founding. In *Strategies in Failure Management.* Springer.
Van Weele, M., van Rijnsoever, F. J., & Nauta, F. (2017). You can't always get what you want: How entrepreneur's perceived resource needs affect the incubator's assertiveness. *Technovation*, *59*, 18–33.
Vanderstraeten, J., van Witteloostuijn, A., & Matthyssens, P. (2020). Organizational sponsorship and service co-development: A contingency view on service co-development directiveness of business incubators. *Technovation*, *98*, 102154.
Ventures, P. (2020). *What is Steward-Ownership?* Retrieved from https://purpose-economy.org/en/whats-steward-ownership/
Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, *5*(2), 171–180.
Wilkinson, C. (2015). *A Map of Social Enterprises and Their Eco-Systems in Europe - Synthesis Report.* Retrieved from http://apfse.eu/wp-content/uploads/2019/04/5.Synthesis-report-FINAL.pdf
Wilson, F., & Post, J. E. (2013). Business models for people, planet (& profits): Exploring the phenomena of social business, a market-based approach to social value creation. *Small Business Economics*, *40*(3), 715–737.
Wise, S., & Valliere, D. (2014). The impact on management experience on the performance of start-ups within accelerators. *The Journal of Private Equity*, *18*(1), 9–19.
Yang, S., Kher, R., & Newbert, S. L. (2020). What signals matter for social startups? it depends: The influence of gender role congruity on social impact accelerator selection decisions. *Journal of Business Venturing*, *35*(2), 105932.
Yin, R. K. (2003). *Case Study Research: Design and Methods*. SAGE Publications.
Yu, S. (2020). How do accelerators impact the performance of high-technology ventures? *Management Science*, *66*(2).
Zahra, S. A., Gedajlovic, E., Neubaum, D. O., & Shulman, J. M. (2009). A typology of social entrepreneurs: Motives, search processes and ethical challenges. *Journal of Business Venturing*, *24*(5), 519–532.
Zimmer, K., & Pearson, K. (2019). *This is what is still holding social entrepreneurs back.* Retrieved from https://www.weforum.org/agenda/2019/07/this-is-what-holds-social-entrepreneurs-back/
Zimmerman, M. A., & Zeitz, G. J. (2002). Beyond survival: Achieving new venture growth by building legitimacy. *Academy of Management Review*, *27*(3), 414–431.

Junior Management Science 7(2) (2022) 338-353

Junior Management Science

journal homepage: www.jums.academy

Survival Analysis: An Investigation of Covid-19 Patient Data

Survival Analysis: Eine Untersuchung von Covid-19 Patientendaten

Akira Karimkhani

Freie Universität Berlin

Abstract

The aim of this work is to test the feasibility of a model based on survival analysis for Covid- 19 patients. To investigate the feasibility, a Cox regression (CPH-Model) was constructed and evaluated using introduced diagnostic methods and modified using presented extensions. It is shown that disregarding the model assumptions can lead to biased estimation results. Furthermore, a sample analysis of the current literature in which CPH-Model was used revealed that the underlying model assumptions were comprehensibly tested in 40% of the articles reviewed. The novelty value of this work is based on the data analysis showing that the conventional CPH-Model is inappropriate for the Covid-19 dataset studied. In order to apply CPH-Model, the model had to be extended. It was necessary to adjust the functional form of a variable, remove outliers, include time interactions and stratify the data set. Finally, this allowed the creation of a final model that met all assumptions. However, four of the estimated coefficients appear questionable. Therefore, the adequacy of the extended model is doubtful. This implies that when CPH-Model is applied, the fulfillment of the model assumptions should be checked most carefully, and more robust estimation methods should be used in case of nonfulfillment.

Zusammenfassung

Ziel dieser Arbeit ist es die Realisierbarkeit einer Cox-Regression (CPH-Modell) für Covid-19 Patienten zu prüfen. Dafür wird das konstruierte Modell anhand von eingeführten Diagnostik-Methoden ausgewertet und mittels vorgestellter Erweiterungen modifiziert. Weiterhin wurde eine stichprobenartige Analyse der relevanten Literatur durchgeführt. Die Literaturanalyse hat aufgezeigt, dass die zugrundeliegenden Modell-Annahmen in lediglich 40% der untersuchten Artikel nachvollziehbar geprüft wurden. Der Neuigkeitswert dieser Arbeit begründet sich darin, dass gezeigt werden konnte, dass ein konventionelles CPH-Modell für den untersuchten Covid-19 Datensatz unangemessen ist. Um das CPH-Modell anwenden zu können war es notwendig die funktionale Form einer Variable anzupassen, Ausreißer zu entfernen, Zeitinteraktionsterme aufzunehmen und den Datensatz aufzuteilen. Schließlich konnte so ein finales Modell erstellt werden, welches alle Annahmen erfüllt. Allerdings erscheinen vier der geschätzten Koeffizienten fragwürdig. Daher ist die Angemessenheit des erweiterten Modells zweifelhaft. Dies impliziert, dass bei Anwendung des CPH-Modells auf Covid-19 Datensätzen die Erfüllung der Modell-Annahmen genauesten überprüft und bei Nichterfüllung robustere Schätzmethoden verwendet werden sollten.

Keywords: Covid-19; Cox-Regression; CPH-Modell; Proportional Hazards Model; Survival Analysis.

1. Einleitung

Aufgrund der derzeit herrschenden SARS-CoV- 2 (Covid-19) Pandemie kommt es global zu einer beispiellosen Belastung des Gesundheitswesens. Durch steigende Infektionszahlen und den damit einhergehenden schweren Krankheitsverläufen werden die verfügbaren Betten auf Intensivstationen zunehmend knapp. In zahlreichen Ländern der Welt übersteigt die Anzahl der versorgungsbedürftigen Patienten die vorhandenen Kapazitäten der medizinischen Einrichtungen. Damit Krankenhäuser handlungsfähig bleiben können, ist eine effiziente Allokation von verfügbaren Ressourcen unabdingbar. In Krankenhäusern stehen dabei Ärzte in der Pflicht, Patienten nach Erfolgsaussichten zu priorisieren. Um medizinisches Fachpersonal zu entlasten, könnte ein datenbasiertes,

DOI: https://doi.org/10.5282/jums/v7i2pp338-353

statistisches Modell bei der rationalen Entscheidungsfindung unterstützend wirken.

1.1. Problemumfeld

Um die Beziehung zwischen medizinischen Einflussgrößen und dem Tod eines Patienten zu untersuchen, bieten sich die statistischen Verfahren der Survival-Analyse an. Anhand einer Analyse der Überlebenszeiten von Covid-19 Patienten, können Rückschlüsse auf das Mortalitätsrisiko bestimmter Subgruppen gezogen werden. Der Kaplan-Meier Schätzer und die Cox-Regression eignen sich dabei besonders, um anhand von Patientendaten, Aussagen über die Überlebenswahrscheinlichkeiten zu treffen.

1.2. Zielsetzung

Ziel dieser wissenschaftlichen Arbeit ist es, die Realisierbarkeit eines auf der Survival-Analyse basierenden Modells für Covid-19 Patienten zu prüfen. Dabei soll anhand einer Kaplan-Meier Schätzung und Cox-Regression verschiedene Einflüsse auf die Mortalität von Patienten untersucht werden. Besonderes Gewicht soll bei der Modell-Konstruktion den zugrundeliegenden Annahmen zukommen, um eine Angemessenheit des Modells zu begutachten.

1.3. Aufbau der Arbeit

Konkret wird das Ziel der Arbeit erreicht, in dem zunächst in Kapitel 2 die theoretischen Grundlagen der Survival-Analyse erläutert werden. Unter Kapitel 2.1 werden dabei die Begriffe der Hazardrate und Survival-Funktion eingeführt. In Kapitel 2.2 wird explizit auf die Verfahren der Survival-Analyse eingegangen. Dabei wird zunächst der Kaplan-Meier Schätzer als nicht-parametrisches und die Cox-Regression als semi-parametrisches Verfahren vorgestellt. Nach der Einführung beschäftigt sich Kapitel 2.2.3 mit einigen Erweiterungen und Kapitel 2.2.4 mit Diagnostiken der Cox-Regression. Der theoretische Teil der Arbeit endet mit Kapitel 2.3, in dem der aktuelle Stand der Forschung umrissen wird. Kapitel 3 kennzeichnet den empirischen Teil der Arbeit, welcher in Kapitel 3.1 mit einem Konzept zur Datenanalyse beginnt. In Kapitel 3.2 werden die notwendigen Anpassungen des Datensatzes vor Beginn der Analyse erläutert sowie die verwendeten RStudio Pakete vorgestellt. Kapitel 3.3 bildet den Hauptgegenstand dieser Arbeit. Die zuvor eingeführten theoretischen Konzepte werden hier sukzessive angewendet. Dabei wird ein Modell anhand der im theoretischen Teil erläuterten Methoden konstruiert und evaluiert. In Kapitel 4 werden die geschätzten Koeffizienten des finalen Modells interpretiert und das Modell als Ganzes diskutiert. Abschließend wird der methodische und empirische Teil der Arbeit in Kapitel 5 zusammengefasst und weiterhin ein finaler Ausblick auf weitere Forschungsmöglichkeiten gegeben.

2. Notwendigkeit der Survival-Analyse

Die Survival-Analyse bezeichnet eine Klasse von statistischen Verfahren, welche zeitabhängige Ereignisse untersuchen. Dabei ist das Ereignis von Interesse in den meisten Anwendungen der Todeszeitpunkt eines Individuums. Die Methoden der Survival-Analyse wurden ursprünglich vor allem in klinischen Studien zur Analyse der Überlebenszeit von Patienten verwendet, jedoch finden sie auch in nicht-medizinischen Kontexten Verwendung (Sedlacek, 2018, S. 21 f.). Die Datengrundlage von Survival-Analysen (Survival-Daten) unterscheidet sich dabei von konventionellen Daten, die für statistische Verfahren genutzt werden, dahingehend, dass Survival-Daten für gewöhnlich nicht symmetrisch verteilt sind. Daher wäre es unangemessen, diesen Daten eine Normalverteilung zu unterstellen bzw. Modelle, die auf der Normverteilungs-Annahme basieren, anzuwenden (Collett, 2015, S. 1 f.). Weiterhin unterscheiden sich die Methoden der Survival-Analyse gegenüber herkömmlichen Modellierungen darin, dass die abhängige Variable häufig nur begrenzt beobachtet werden kann. So kann es sein, dass das Ereignis von Interesse bei bestimmten Individuen auch bei unendlicher Beobachtung nicht auftritt. Bei Individuen mit endlicher Beobachtungszeit kann es hingegen sein, dass das Ereignis vor oder auch nach Studienende eintritt. Dadurch kommt eine Zensur der Daten zustande. Das bedeutet, dass vollständige Informationen über die Ereigniszeiten aufgrund des Studiendesigns nicht verfügbar sind (Frees, 2009, S. 383 f.). Darüber hinaus kann es sein, dass bestimmte Individuen, deren Ereignis von Interesse nicht im beobachteten Zeitraum liegt, nicht repräsentiert werden. In diesem Fall spricht man von einer Verkürzung der Daten (Klein & Moeschberger, 2006, S. 64). Die Nicht-Berücksichtigung von Zensur und Verkürzung durch schlichtes Auslassen der entsprechenden Beobachtungen kann zu stark verzerrten Schätzern führen (Sedlacek, 2018, S. 26).

Dadurch, dass Survival-Daten Überlebenszeiten von Individuen abbilden, kommt es sowohl bei der Erhebung als auch der Verteilung zu kennzeichnenden Charakteristiken. Eine Analyse von Survival-Daten bedarf daher der Einführung von theoretischen Grundkonzepten, auf die sich die Anwendung im späteren Teil der Arbeit beziehen wird.

2.1. Grundlegende theoretische Konzepte

Es sei T die Zeitdauer bis zum Eintritt eines Ereignisses von Interesse, so ist T eine nicht-negative, stetige Zufallsvariable einer homogenen Population. Es charakterisieren drei Funktionen die Verteilung von T, nämlich die Survival-Funktion, die Hazardrate bzw. Hazardfunktion und die Wahrscheinlichkeits-Dichtefunktion (Dichtefunktion). Sobald eine der drei Funktionen bekannt ist, können die anderen eindeutig bestimmt werden. In der Praxis werden diese drei Funktionen gemeinsam mit einer weiteren, der kumulativen Hazardrate verwendet, um die verschiedenen Aspekte der Verteilung von T zu veranschaulichen (Klein & Moeschberger, 2006, S. 21 f.).

2.1.1. Survival-Funktion

Es sei T die Zeitdauer bis zum Eintritt eines Ereignisses von Interesse, so ist T eine nicht-negative, stetige Zufallsvariable mit der Dichtefunktion $f(t)$ und Verteilungsfunktion

F(t). Die Verteilungsfunktion *F(t)* ist dabei gegeben als

$$F(t) = P(T < t) = \int_0^t f(u)\mathrm{d}u \tag{1}$$

und gibt die Wahrscheinlichkeit an, dass die Überlebenszeit eines Individuums geringer als ein Wert t ist. Danach ist

$$S(t) := P(T \geq t) = 1 - F(t) \tag{2}$$

definiert als die Überlebens- bzw. Survival-Funktion. Sie gibt die Wahrscheinlichkeit an, dass die Überlebenszeit eines Individuums größer oder gleich t ist. Die Survival-Funktion kann daher verwendet werden um die Wahrscheinlichkeit anzugeben, dass ein Individuum bis zu einem Zeitpunkt, welcher nach t liegt, überlebt hat (Collett, 2015, S. 10).
Da das Ereignis von Interesse hier der Tod eines Individuums ist, gibt die Survival-Funktion an, dass ein Individuum den Zeitpunkt t „erlebt" hat. Für $S(t)$ gilt $S(0) = 1$, sowie ein monoton fallender Verlauf (Sedlacek, 2018, S. 29).

2.1.2. Hazardrate

Neben der Survival-Funktion ist die Hazardrate eine zentrale Größe in der Survival-Analyse. Die Hazardrate gibt das augenblickliche Risiko zum Zeitpunkt $T = t$ an, dass das Ereignis von Interesse eintritt, unter der Bedingung, dass es bisher nicht eingetreten ist (Sedlacek, 2018, S. 29). In anderen Worten gibt die Hazardrate die Wahrscheinlichkeit an, dass ein Individuum an Zeitpunkt t verstirbt unter der Bedingung, dass es bis zu diesem Zeitpunkt überlebt hat (Collett, 2015, S. 10). Definiert wird die Hazardrate als

$$h(t) := \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}. \tag{3}$$

Aus Formel (3) lassen sich dabei nützliche Beziehungen zwischen der Survival-Funktion und der Hazardrate herleiten. Unter Berücksichtigung einer der Grunderkenntnisse der Wahrscheinlichkeitstheorie nach Bayes und Price (1763) gilt

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \tag{4}$$

wobei $P(A|B)$ die Wahrscheinlichkeit des Auftretens des Ereignisses A unter der Bedingung von B , $P(A \cap B)$ die Wahrscheinlichkeit für das gemeinsame Auftreten der Ereignisse A und B und $P(B)$ die Wahrscheinlichkeit für das alleinige Auftreten von B ist. Unter der Verwendung von (4) lässt sich der Zähler in (3) umformen zu

$$P(t \leq T < t + \Delta t | T \geq t) = \frac{P(t \leq T < t + \Delta t)}{P(T \geq t)}. \tag{5}$$

Nenner und Zähler aus (5) können unter Verwendung von (1) und (2) ausgedrückt werden als

$$\frac{P(t \leq T < t + \Delta t)}{P(T \geq t)} = \frac{F(t + \Delta t) - F(t)}{S(t)}.$$

Für $h(t)$ gilt daher

$$h(t) = \lim_{\Delta t \to 0} \Big[\frac{F(t + \Delta t) - F(t)}{\Delta t} \Big] \frac{1}{S(t)}, \tag{6}$$

wobei der erste Term dem Differentialquotient der Verteilungsfunktion $F(t)$ entspricht. Daher gilt

$$h(t) = \frac{f(t)}{S(t)}. \tag{7}$$

Aus (7) folgt unmittelbar $h(t) = -\big[\log[S(t)]\big]'$, da

$$\begin{aligned}
-\big[\log[S(t)]\big]' &= -\frac{1}{S(t)} S'(t) \\
-\big[\log[S(t)]\big]' &= -\frac{1}{S(t)} [1 - F(t)]' \\
-\big[\log[S(t)]\big]' &= -\frac{1}{S(t)} [-f(t)] \\
-\big[\log[S(t)]\big]' &= \frac{f(t)}{S(t)}.
\end{aligned}$$

Stellt man nun $h(t) = -\big[\log[S(t)]\big]'$ nach $S(t)$ um erhält man

$$S(t) = \exp[-H(t)], \tag{8}$$

wobei $H(t) = \int_0^t h(u)\, \mathrm{d}u$ ist. Die Funktion H(t) ist definiert als die kumulative Hazardrate und kann durch die Survival-Funktion erhalten werden, da $H(t) = -\log[S(t)]$ gilt (Collett, 2015, S. 12).

2.2. Statistische Verfahren der Survival-Analyse

Als erster Schritt in der Analyse von Survival-Daten wird für gewöhnlich ein grafischer und numerischer Überblick über die Überlebenszeiten von Individuen gegeben. Typischerweise umfasst der Überblick die Schätzung der Survival-Funktion und der Hazardrate mittels der Verteilung angemessener Verfahren. Die Hazardrate und die Survival-Funktion werden dabei aus den beobachteten Überlebenszeiten geschätzt (Collett, 2015, S. 13 f.). Methoden, die zur Schätzung keine Annahmen bezüglich der Form der Verteilung machen, zählen zu den nicht-parametrischen Verfahren der Survival-Analyse. Methoden, die spezifische Annahmen treffen, zählen zu den (semi-)parametrischen Verfahren.

2.2.1. Kaplan-Meier Schätzer

Die in der Anwendung am weitesten verbreitete nicht-parametrischen Verfahren bilden dabei die Sterbetafelmethode (Lifetable Schätzer) und der von Kaplan und Meier (1958) vorgestellte Produkt-Grenzwert Schätzer (Kaplan-Meier Schätzer). Dabei verwendet man die Sterbetafelmethode bei diskreter und den Kaplan-Meier Schätzer bei stetiger Zeitmessung (Windzio, 2013, S. 90). Da im empirischen Teil dieser Arbeit die Überlebenszeit als stetige Variable erfasst wird, wird im folgenden nur auf den Kaplan-Meier Schätzer als nicht-parametrisches Verfahren eingegangen.

Der Kaplan-Meier Schätzer ist ein nicht-parametrischer Schätzer der Survival-Funktion $S(t)$. Er ist für alle Werte von t, die beobachtet werden können, definiert als

$$\hat{S}(t) = \begin{cases} 1 & \text{wenn } t < t_1, \\ \prod_{t_i \leq t}[1 - \frac{d_i}{Yi}], & \text{wenn } t_1 \leq t. \end{cases}$$

Wobei Y_i die Anzahl der noch nicht verstorbenen Individuen und d_i die Anzahl der Todesfälle zum Zeitpunkt t_i ist. Der Kaplan-Meier Schätzer ist damit eine Treppenfunktion, welche Sprünge bei den beobachteten Todesfällen aufweist (Klein & Moeschberger, 2006, S. 84).

Mit dem Kaplan-Meier Schätzer können die Survival-Funktionen verschiedene Subgruppen geschätzt und anschließend verglichen werden. Für einen deskriptiven Vergleich bietet es sich an die Kurven-Verläufe unterschiedlicher Gruppen in einer Grafik abzubilden und zu vergleichen. Für einen statistischen Vergleich durch Hypothesentests bietet sich der Log-Rank Test an (Zwiener, Blettner & Hommel, 2011, S. 166). Der Log-Rank wird dabei verwendet um die Null-Hypothese zu testen, ob zwischen Populationen die gleiche Wahrscheinlichkeit für den Eintritt eines Todesfalls zu beliebigen Zeitpunkten besteht (Bland & Altman, 2004, S. 1073). Die Ablehnung der Null-Hypothese impliziert dabei, dass sich die Hazardraten der Populationen während einem oder mehreren Zeitpunkten unterscheiden (Gad & Rousseaux, 2002, S. 384). Für weiterführende Informationen bezüglich des Log-Rank Test siehe Mantel (1966).

Der Kaplan-Meier Schätzer ist ein geeignetes Verfahren zur Schätzung der Survival-Funktion bei simplen Datensätzen, in denen lediglich die Überlebenszeiten von Individuen erfasst werden. Ein häufig auftretendes Problem bei der Analyse von Survival-Daten ist die Berücksichtigung begleitender Informationen durch Anpassung der Survival-Funktion (Klein & Moeschberger, 2006, S. 45). Um Zusammenhänge zwischen der Überlebenszeit eines Individuums und erklärenden Variablen zu untersuchen, kann ein auf der Regressions-Analyse basierender Ansatz verwendet werden (Collett, 2015, S. 53). Das wohl prominenteste Modell ist hierfür das von Cox (1972) präsentierte proportionale Hazard Modell (CPH-Modell) bzw. Cox-Regression als semi-parametrisches Verfahren.

2.2.2. Cox-Regression

Neben den Überlebenszeiten werden in medizinischen Anwendungen typischerweise weitere Variablen wie etwa das Alter, Geschlecht und Gewicht erhoben. Ähnlich wie in der linearen Regression lässt sich mit der Cox-Regression der Einfluss von erklärenden Variablen auf die abhängige Variable quantitativ ermitteln. Die Cox-Regression basiert dabei auf der Annahme von proportionalen Hazards, jedoch wird den Überlebenszeiten keine bestimmte Wahrscheinlichkeitsverteilung unterstellt. Daher zählt das Modell zu den semi-parametrischen Verfahren (Collett, 2015, S. 54).

Sei der Vektor $\mathbf{x}$ das Tupel der erklärenden Variablen x, so dass $\mathbf{x} = (x_1, x_2, ..., x_p)^T$ und sei $h_0(t)$ die Hazardfunktion für ein Individuum, dessen Wert aller Kovariablen des Vektors $\mathbf{x}$ Null sind. So heißt die unspezifizierte Funktion $h_0(t)$ die Baseline Hazardfunktion. Die Hazardfunktion für ein Individuum i kann definiert werden als

$$h_i(t) = \varphi(\mathbf{x}_i)h_0(t), \tag{9}$$

wobei $\varphi(\mathbf{x}_i)$ eine Funktion der erklärenden Variablen für das Individuum i ist. Die Funktion $\varphi(\mathbf{x}_i)$ kann dabei als Hazardrate für ein Individuum zum Zeitpunkt t mit den erklärenden Variablen $\mathbf{x}_i$, relativ zu einem Individuum bei dem $\mathbf{x} = \mathbf{0}$ gilt, aufgefasst werden, da aus (9)

$$\frac{h_i(t)}{h_0(t)} = \varphi(\mathbf{x}_i) \tag{10}$$

folgt. Die relative Hazardrate $\varphi(\mathbf{x}_i)$ kann nicht negativ werden, daher ist es zweckmäßig sie als $\exp[\eta_i]$ auszudrücken. Dabei ist η_i eine Linearkombination der p erklärenden Variablen in $\mathbf{x}_i$. So dass

$$\eta_i = \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_p x_{pi}$$

gilt, bzw. in Matrix-Notation

$$\eta_i = \beta^T \mathbf{x}_i.$$

Das proportionale Hazard Modell wird somit zu

$$h_i(t) = \exp[\beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_p x_{pi}]h_0(t) \text{ bzw.} \tag{11}$$

$$\log\Big[\frac{h_i(t)}{h_0(t)}\Big] = \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_p x_{pi} \tag{12}$$

(Collett, 2015, S. 55 ff.).

Um das in (11) erhaltene Modell auf einen Datensatz anwenden zu können, müssen die unbekannten Parameter $\beta_1, \beta_2, ..., \beta_p$ und die Baseline Hazardfunktion $h_0(t)$ bestimmt werden. Dabei können beide Komponenten des Modells separat voneinander geschätzt werden. Zunächst werden die β-Koeffizienten durch eine Maximum-Likelihood Schätzung bestimmt. In einem zweiten Schritt werden die geschätzten β-Koeffizienten dazu verwendet, eine Schätzung der Baseline Hazardfunktion zu konstruieren (Collett, 2015, S. 61 f.). Da bei den meisten statistischen Programmen die Konstruktion und Maximierung der Likelihood-Funktion unproblematisch ist, wird dies im Weiteren nicht vertieft. Für weiterführende Information zur Form der Likelihood-Funktion des proportionalen Hazard Modells siehe Cox (1972). Die Maximierung der Likelihood-Funktion erfolgt für gewöhnlich durch das Newton-Raphson-Verfahren. Für weiterführende Informationen zum Newton-Raphson-Verfahren siehe Ben-Israel (1966) und Galántai (2000).

Voraussetzungen der Cox-Regression

Um die bisher vorgestellten Verfahren der Survival-Analyse

anwenden zu können, müssen bestimmte Voraussetzungen erfüllt sein. So darf sowohl beim Kaplan-Meier Schätzer als auch bei der Cox-Regression keine informative Zensur vorliegen. Das heißt, dass wenn eine Zensur der Daten vorliegt, diese zufällig erfolgen muss. Eine nicht-informative Zensur stellt sicher, dass die Überlebenszeiten zensierter Individuen so repräsentativ wie die der nicht zensierten sind (Collett, 2015, S. 4). Weiterhin müssen die beobachteten Überlebenszeiten in beiden Modellen unabhängig voneinander sein. Neben diesen zwei Anforderungen verlangt das CPH-Modell proportionale Hazards. Die Annahme von proportionalen Hazards (PH-Annahme) sagt aus, dass das Verhältnis der Hazardraten verschiedener Beobachtungen im Zeitverlauf konstant ist (Kleinbaum & Klein, 2011, S. 123). Zur Veranschaulichung sei β ein p-dimensionaler Koeffizienten-Vektor und $\mathbf{x}_i$ sowie $\mathbf{x}_j$ p-dimensionale Kovariablen-Vektoren zweier Beobachtungen. So lässt sich das Verhältnis der Hazards (Hazardratio bzw. HR) darstellen als

$$HR = \frac{h_i(t)}{h_j(t)} = \frac{h_0(t)\exp[\beta^T \mathbf{x}_i]}{h_0(t)\exp[\beta^T \mathbf{x}_j]} = \exp[\beta^T(\mathbf{x}_i - \mathbf{x}_j)] = \theta, \tag{13}$$

wobei θ eine Konstante sei (Xue & Schifano, 2017. S. 584). Darüber hinaus verlangt das CPH-Modell, so wie in der linearen Regression, einen linearen Zusammenhang zwischen den kontinuierlichen, erklärenden Variablen und der abhängigen Variable. Formel (12) veranschaulicht diese Annahme. Des Weiteren nimmt das CPH-Modell an, dass die Werte der Kovariablen im Zeitverlauf konstant sind. Ersichtlich ist diese Annahme in Formel (9), da die Funktion $\varphi(\mathbf{x}_i)$ wohl für Individuen i variiert, jedoch nicht für Zeitpunkte t (im Gegensatz zu der Baseline Hazardfunktion $h_0(t)$).

2.2.3. Erweiterungen der Cox-Regression

Sind die im vorherigen Abschnitt genannten Voraussetzungen der Cox-Regression nicht erfüllt, so ist es nicht zwingend notwendig, das Modell zu verwerfen. Stattdessen bietet es sich an, das Grundmodell mittels Modifikationen zu erweitern. Im Folgenden werden einige Erweiterungen des CPH-Modell kurz erläutert, welche im empirischen Teil Anwendung finden.

Nicht-Linearität der kontinuierlichen Kovariablen
Liegt bei einer kontinuierlichen Kovariable ein nicht-linearer Zusammenhang vor, kann durch Transformation der funktionalen Form der Kovariable der Zusammenhang linear ausgedrückt werden. Dabei ist es wie in der linearen Regression erforderlich, dass der Koeffizient der jeweiligen Kovariable linear verbleibt (Casson & Farmer, 2014, S. 592). Für weiterführende Informationen bezüglich der Annahmen der linearen Regression und der Überprüfung dieser siehe Casson und Farmer (2014).

Die stratifizierte Cox-Regression
In manchen Fällen kann es sein, dass die PH-Annahme nicht für die gesamten Daten erfüllt ist. Es kann beispielsweise sein, dass sich die Hazardrate von Patienten, die in einer Studie teilnehmen, in der Placebos eingesetzt werden, unverhältnismäßig zu der Hazardrate von Patienten, die einen Wirkstoff erhalten, verhält. In einer solchen Situation kann es sinnvoll sein anzunehmen, dass die Baseline Hazardfunktion $h_0(t)$ zwischen Subgruppen variiert, während die erklärenden Variablen die PH-Annahme befriedigen. In so einem Fall leistet eine Aufteilung der Beobachtungen in Schichten (Strata) Abhilfe. Es sei $h_{0j}(t)$, die Baseline Hazardfunktion für Schicht (Stratum) j, mit $j = 1, 2, ..., g$ wobei g die Anzahl der verschiedenen Strata ist. So kann das CPH-Modell als $h_{ij}(t)$ für das i-te Individuum in dem j-Stratum dargestellt werden, wobei $i = 1, 2, ..., n_j$ die Anzahl der Beobachtungen in dem j-Stratum ist. So ergibt sich das geschichtete (stratifizierte) CPH-Modell als Erweiterung von Formel (11) als

$$h_{ij}(t) = \exp[\beta^T \mathbf{x}_{ij}]h_{0j}(t), \tag{14}$$

wobei $\mathbf{x}_{ij}$ ein p-Dimensionaler Vektor der erklärenden Variablen $x_1, x_2, ..., x_p$ der Individueen i des j-Stratums ist (Collett, 2015, S. 269 f.).

Cox-Regression mit zeitabhängigen Kovariablen
Wie in Abschnitt 2.2.2 erläutert, nimmt das CPH-Modell an, dass die erklärenden Variablen im Zeitverlauf konstant sind. Ist diese Annahme nicht erfüllt, kann die Cox-Regression um zeitabhängige Kovariablen (TDC) erweitert werden. Es sei $\mathbf{x}_i(t)$ der Kovariablen-Vektor zeitabhängiger erklärenden Variablen $x_i(t)$ zum Zeitpunkt t der i Individuen. So kann das Modell aus Formel (11) für zeitabhängige Kovariablen definiert werden als

$$h_i(t) = \exp[\beta^T \mathbf{x}_i(t)]h_0(t). \tag{15}$$

Dadurch, dass die $x_i(t)$ zeitabhängig sind, ist das Verhältnis der Hazards $\frac{h_i(t)}{h_0(t)}$ nicht mehr konstant (Collett, 2015, S.). Um zeitabhängige Kovariablen in einem CPH-Modell zu berücksichtigen, ist in den meisten Fällen eine Anpassung der Datengrundlage notwendig. Dies rührt daher, dass die TDC ihre Werte schrittweise über die Zeit ändern. Zu jedem Zeitpunkt, indem eine TDC ihren Wert ändert, ist daher ein Intervall erforderlich, in dem die TDC konstant bleibt. Anhand einer Zusammenfassung der Daten in Form eines Zählprozesses kann dem gerecht werden (Box-Steffensmeier, Box-Steffensmeier & Jones, 2004, S. 97 f.).
Da die technische Implementierung eines Zählprozesses unproblematisch ist, wird eine detaillierte Ausführung ausgelassen. Für weiterführende Informationen zur Formulierung von Zählprozessen für die Cox-Regression siehe Andersen und Gill (1982).

2.2.4. Modell Diagnostik

Nach der Konstruktion eines CPH-Modells ist es notwendig, die Modellgüte zu beurteilen. Zum einen darf keine der zuvor genannten Modellannahmen verletzt werden, zum anderen sollte die Anpassung des Modells stimmig sein. Für eine Beurteilung der Modellgüte gibt es unterschiedliche An-

sätze. Eine weit verbreitete Methode ist die graphische Analyse von Residuen. In dem CPH-Modell sind die Cox-Snell- (Cox & Snell, 1968), Deviance- (Therneau, Grambsch & Fleming, 1990), Martingale- (Lagakos, 1981), Schoenfeld- (Schoenfeld, 1982) und Score-Residuen (Cain & Lange, 1984) bzw. Df-Betas besonders relevant. Da im empirischen Teil der Arbeit jedes der genannten Residuen Anwendung findet, werden sie im Folgenden kurz aufgegriffen.

Cox-Snell-Residuen
Die Cox-Snell-Residuen sollten besonders hervorgehoben werden, da durch simple Transformation dieser die Deviance- und Martingale-Residuen hergeleitet werden können. Die Cox-Snell-Residuen sind, für ein Individuum i mit $i = 1, 2, ..., n$ definiert als

$$r_{Ci} = \exp[\hat{\beta}^T \mathbf{x}_i]\, \hat{H}_0(t_i) = \hat{H}_i(t_i), \tag{16}$$

wobei $\hat{H}_0(t_i)$ und $\hat{H}_i(t_i)$ die geschätzten kummulativen Hazard- und Baseline Hazardfunktionen zum Zeitpunkt t für ein Individuum i sind. Aus Formel (8) geht für $\hat{H}_i(t_i)$ hervor, dass

$$r_{Ci} = \hat{H}_i(t_i) = -\log[\hat{S}_i(t_i)] \tag{17}$$

ist, wobei $\hat{S}_i(t_i)$ die geschätzte Survival-Funktion des i-ten Individuums zum Zeitpunkt t ist (Collett, 2015, S. 150 f.).

Die Cox-Snell-Residuen eignen sich zur Überprüfung der allgemeinen Anpassung des Modells. Bei einer korrekten Anpassung des Modells folgen die Cox-Snell-Residuen approximativ einer Exponential-Verteilung im Einheitsintervall. Um zu überprüfen, ob die Cox-Snell-Residuen dieser Verteilung folgen, kann ein Residuenplot der geschätzten kumulativen Hazardfunktion, basierend auf den Cox-Snell Residuen $\hat{H}_r(r_{Ci})$ gegen die Cox-Snell-Residuen r_{Ci}, erstellt werden. Bei einer korrekten Anpassung ergibt sich im Graphen eine Gerade durch den Ursprung mit einer Steigung von eins (Box-Steffensmeier et al., 2004, S. 120). Für eine Herleitung der approximativen Verteilung der Cox-Snell-Residuen siehe Crowley und Hu (1977).

Martingale-Residuen
Durch Subtraktion der Cox-Snell-Residuen von einer Indikatorvariable δ_i, welche 0 bei Zensierung und 1 bei dem Tod eines Individuums annimmt, erhält man die Martingale-Residuen. Die Martingale-Residuen werden folglich definiert als

$$r_{Mi} = \delta_i - r_{Ci}. \tag{18}$$

Die r_{Mi} nehmen dabei Werte zwischen $(-\infty, 1]$ an und haben bei großem n einen Erwartungswert von Null. Allerdings sind sie nicht symmetrisch um Null verteilt (Collett, 2015, S. 153).

Die Martingale-Residuen sind hilfreich, um die Modell-Annahme des linearen Zusammenhangs zwischen dem logarithmierten Hazard und den kontinuierlichen Kovariablen zu untersuchen (Wollschläger, 2010, S. 369). Da die Martingale-Residuen bei großem n einen Erwartungswert von Null haben, können systematische Abweichungen von Null auf eine falsche funktionale Form hinweisen (Box-Steffensmeier et al., 2004, S. 126). Für weiterführende Informationen und einer allgemeineren Herleitung der Martingale-Residuen siehe Lagakos (1981) und Barlow und Prentice (1988).

Deviance-Residuen
Die Deviance-Residuen sind skalierte Martingale-Residuen, welche im Gegensatz zu den Martingale-Residuen approximativ symmetrisch um Null verteilt sind. Die Deviance-Residuen können dabei als eine Verallgemeinerung der kleinsten-Quadrat-Residuen aus der linearen Regression betrachtet werden. Besonders große Deviance-Residuen entsprechen dabei Beobachtungen, welche durch das Modell nicht gut angepasst sind (Collett, 2015, S. 153). Dabei sagt ein negatives Deviance-Residuum aus, dass die vorhergesagte Überlebenszeit geringer ist als die tatsächliche. Das bedeutet, dass für diese Beobachtungen die Hazardrate überschätzt wird. In anderen Worten überlebt die Beobachtung länger, als sie nach dem CPH-Modell sollte. Bei Beobachtungen mit positivem Deviance-Residuum ergibt sich die umgekehrte Logik (Box-Steffensmeier et al., 2004, S. 130 f.).
Die Herleitung der Deviance-Residuen geht über den Rahmen dieser Arbeit hinaus und wird aus diesem Grund ausgelassen. Für eine formale Definition und Herleitung der Deviance-Residuen siehe Therneau et al. (1990).

2.2.5. Score-Residuen/Df-Betas

Um ein CPH-Modell auf einflussreiche Beobachtungen zu untersuchen, können ähnlich wie in der linearen Regression Df-Betas verwendet werden (Wollschläger, 2010, S. 369). Die Df-Betas liefern für jede Beobachtung ein standardisiertes Maß, wie stark sich die geschätzten Parameter ändern, wenn die jeweilige Beobachtung aus den Daten ausgeschlossen wird (Wollschläger, 2010, S. 215). Man erhält die Df-Betas für ein CPH-Modell, indem man eine Matrix von Score-Residuen erstellt und durch Multiplikation der Varianz-Kovarianz-Matrix skaliert. Durch die Skalierung geben die Df-Betas die verursachte Änderung der einflussreichen Beobachtungen in den geschätzten Parametern als Änderung der Standardabweichung an (Box-Steffensmeier et al., 2004, S. 123 ff.). Die Herleitung der Score-Residuen geht über den Rahmen dieser Arbeit hinaus und wird daher ausgelassen. Für eine formale Definition und Herleitung der Score-Residuen siehe Cain und Lange (1984).

Schoenfeld-Residuen
Die Schoenfeld-Residuen können im Wesentlichen als die erwarteten, subtrahiert von den beobachteten Werten der Kovariablen des CPH-Modells betrachtet werden. Anhand einer grafischen Darstellung der i-ten Schoenfeld-Residuen gegen die t_i-ten Überlebenszeiten lassen sich Verletzungen der PH-Annahme aufzeigen. Dabei deuten Veränderungen der Residuen im Zeitverlauf auf eine Zeitabhängigkeit hin (Box-Steffensmeier et al., 2004, S. 121). Darüber hinaus ist

der in Abschnitt 2.2.1 angesprochene Log-Rank Test geeignet um die PH-Annahme anhand der Schoenfeld-Residuen statistisch zu untersuchen. Anhand der Korrelation zwischen Überlebenszeiten und den Schoenfeld-Residuen testet der Log-Rank Test mit einem Chi-Quadrat-Test (Chi-Sq) die Null-Hypothese, ob die PH-Annahme erfüllt ist (Wollschläger, 2010, S. 369). Daher weist ein signifikantes Ergebnis die PH-Annahme zurück, während ein nicht-signifikantes die PH-Annahme bestärkt. Die Herleitung der Schoenfeld-Residuen geht über den Rahmen dieser Arbeit hinaus und wird daher im Folgenden ausgelassen. Für eine formale Definition und Herleitung der Schoenfeld-Residuen siehe Schoenfeld (1982).

2.2.6. Stand der Forschung

Nach der Einführung von Cox (1972) wurde das CPH-Modell stetig um Modifikationen erweitert. Andersen und Gill (1982) haben die Verwendung der Zählprozess- Formulierung von Daten zur Verwendung im CPH-Modell vorgestellt, um die Aufnahme von zeitabhängigen Kovariablen zu realisieren. Die Integration von zeitabhängigen Kovariablen im CPH-Modell wurde dabei erstmals von Zucker und Karr (1990) sowie Gamerman (1991) vorgestellt. Die Aufnahme von Kovariablen, welche mit der Überlebenszeit interagieren, ergibt meist dann Sinn, wenn die PH-Annahme im regulären Modell verletzt ist und eine Zeitabhängigkeit plausibel erscheint.

Um die PH-Annahme grafisch zu überprüfen, hat Schoenfeld (1982) eine verallgemeinerte Form von Residuen vorgestellt, die zeitunabhängig sind. Ergänzend zu den grafischen Verfahren haben Grambsch und Therneau (1994) einen statistischen Test, welcher auf skalierten Schoenfeld-Residuen beruht, zur Überprüfung der PH-Annahme vorgestellt. Liegt eine Verletzung der PH-Annahme vor, ist zu klären, ob und in welcher Form Zeitabhängigkeit auftritt. Fisher und Lin (1999) diskutieren dafür das CPH-Modell mit zeitabhängigen Kovariablen. Dabei bemerken sie, dass die technische Implementierung problematisch sein kann und Potential für Verzerrungen bietet. Insbesondere unterstreichen sie die Bedeutsamkeit, bei der Aufnahme von zeitabhängigen Kovariablen die funktionale Form der Zeitabhängigkeit korrekt zu bestimmen. Fisher und Lin (1999) fassen zusammen, dass die Form der Zeitabhängigkeit nicht unbedingt offensichtlich ist, jedoch durch inhaltliches Verständnis erfasst werden kann. Box-Steffensmeier et al. (2004) teilen die Auffassung von Fisher und Lin (1999) hinsichtlich dem Verzerrungs-Potenzial in der funktionalen Form von Zeitabhängigkeiten. In diesem Zusammenhang nennen Box-Steffensmeier et al. (2004) $\log(t)$, als die in der Anwendung favorisierte funktionale Form, um den Interaktionseffekt zwischen einer Kovariable und der Überlebenszeit zu modellieren.

Keele (2010) demonstriert, dass Nicht-Linearität von metrischen Kovariablen ebenfalls zu der Verletzung der PH-Annahme führen kann. Daher wird die Korrektur der funktionalen Form von Nicht-Linearität als ein wichtiger Aspekt zur Verbesserung der Modell-Spezifikation betrachtet. Bereits Therneau et al. (1990) haben dafür ein auf den Martingale-Residuen basierendes grafisches Verfahren zur Überprüfung der funktionalen Form von kontinuierlichen Kovariablen vorgestellt. Farcomeni und Viviani (2011) weisen ergänzend darauf hin, dass Ausreißer in den Daten ebenfalls zur Verletzung der PH-Annahme führen können.

Bei der Konstruktion eines CPH-Modell sollten die angesprochenen Risiken berücksichtigt werden, um eine Verletzung der PH-Annahme zu vermeiden. Es sei anzunehmen, dass in praktischen Anwendungen, in denen das CPH-Modell Verwendung findet, zumindest die Bekräftigung der PH-Annahme ausgesprochen wird. Nach einer Untersuchung von Altman, De Stavola, Love und Stepniewska (1995) wurden jedoch bei lediglich 5% der Veröffentlichungen in medizinischen Fachzeitschriften die zugrunde liegenden Annahmen des CPH-Modells geprüft. Altman et al. (1995) schlagen daher Richtlinien zur Veröffentlichung von Survival-Analysen vor. Rulli et al. (2018) haben eine ähnliche Untersuchung angestellt. Dabei führen sie eine systematische Analyse klinischer Studien mit dem Ziel durch, die Angemessenheit der dort verwendeten Cox-Regressionen basierend auf der PH-Annahme zu beurteilen. Rulli et al. (2018) stellen darin fest, dass in lediglich 4 von 115 Artikeln, in denen die Cox-Regression verwendet wurde, die Bestätigung der PH-Annahme erwähnt wurde.

Aktuell wird die Cox-Regression in einer Vielzahl von Studien verwendet, um Forschungsfragen bezüglich der Covid-19 Pandemie nachzugehen. Stand Dezember 2020 gibt es auf der medizinischen Datenbank PubMed (*https://pubmed.ncbi.nlm.nih.gov/*) auf die Suchanfrage „cox regression proportional hazard covid 19“ 125 veröffentlichte Artikel. Zu den 10 relevantesten Artikel zählen Veröffentlichungen von Bhandari et al. (2020), Chen et al. (2020), Cheng et al. (2020), Cummings et al. (2020), Grasselli et al. (2020), Ji et al. (2020), Li et al. (2020), Rosenberg et al. (2020), Shi et al. (2020) und Zha et al. (2020). Dabei wird die Relevanz anhand der „Best Match“-Suchoption gemessen. Diese berücksichtigt die frühere Verwendung, das Publikationsdatum, die Relevanz-Bewertung und die Art des Artikels. In jedem der veröffentlichten Artikel wird ein (erweitertes) CPH-Modell verwendet, was nahelegen sollte, dass zumindest die Überprüfung der PH-Annahme eine Erwähnung findet. Bei lediglich 40% (Rosenberg et al., 2020 , Cummings et al. (2020), Cheng et al. (2020) und Grasselli et al. (2020)) werden die Modell-Annahmen jedoch nachvollziehbar im Sinne der Richtlinien von Altman et al. (1995) geprüft. Bezüglich der Annahmen des CPH-Modells heißt das, dass die Autoren zumindest eine Bekräftigung der PH-Annahme aussprechen bzw. deren Überprüfung erläutern. Tabelle 7 im Appendix fasst die zehn Studien hinsichtlich der Überprüfung der PH-Annahme und deren Nachvollziehbarkeit zusammen. Dabei sind neben den Autoren zusätzlich relevante Zitate, welche die Nachvollziehbarkeit der Untersuchung unterstützen, abgebildet. Da 60% der Veröffentlichungen aus Tabelle 7 im Appendix die PH-Annahme nicht nachvollziehbar prüfen, ist es fraglich, ob die Anwendung des CPH-Modells für Covid-19 Patienten in diesen Fällen

angemessen ist. Im Folgenden soll daher ein (erweitertes) CPH-Modell für einen beispielhaften Covid-19 Patienten-Datensatz konstruiert werden. Insbesondere sollen dabei die Modell-Annahmen im Einzelnen geprüft werden, um die Angemessenheit des Modells einschätzen zu können.

3. Survival-Analyse an Covid-19 Patientendaten

Im Folgenden werden die zuvor eingeführten Methoden der Survival-Analyse an einem beispielhaften Datensatz angewendet. Bei dem betrachteten Datensatz handelt es sich um offizielle, mit Covid-19 assoziierte Patientendaten, die von der mexikanischen Regierung erhoben worden sind. Dabei stammt der Datensatz nicht unmittelbar von der Website des mexikanischen Gesundheitsamtes (*www.gob.mx/salud/*). Stattdessen wurden die Daten von der Online-Plattform Kaggle (*www.kaggle.com*) bezogen. Dort wurde eine ins englische übersetzte Version des Datensatzes zur Verfügung gestellt, auf den sich die folgende Analyse bezieht. Der Datensatz enthält insgesamt 566.602 Beobachtungen bzw. Patienten, welche im Zeitraum vom 01.01.2020 - 29.06.2020 in mexikanische Krankenhäuser eingeliefert worden sind. Zu jedem Patient wurden 23 verschiedene Kovariablen erfasst, auf die in Kapitel 3.2 genauer eingegangen wird. Voranzustellen ist allerdings, dass für jeden Patienten ein Einlieferungs- und falls eingetreten Sterbedatum, vorliegend ist. Aus technischer Sicht eignet sich der Datensatz daher für eine Survival-Analyse.

3.1. Konzept

Im Weiteren werden zunächst einige Anpassungen des Datensatzes aus Gründen der Praktikabilität vorgenommen. Anschließend wird der in Abschnitt 2.2.1 eingeführte Kaplan-Meier Schätzer konstruiert und ausgewertet. Für eine tiefgreifendere Analyse wird ein CPH-Modell im Sinne des Abschnitt 2.2.2 erstellt. Um die Anwendbarkeit des CPH-Modells zu gewährleisten, müssen die zugrunde liegenden Annahmen der Cox-Regression (siehe Abschnitt 2.2.2) und die Angemessenheit des Modells (siehe Abschnitt 2.2.4) bestätigt werden. Daher wird nach Aufnahme der Kovariablen die allgemeine Angemessenheit des Modells mittels Cox-Snell-Residuen überprüft. Daraufhin wird die funktionale Form der kontinuierlichen Kovariablen mittels Martingale-Residuen untersucht. Einflussreiche Beobachtungen werden anhand der Df-Betas und Ausreißer mittels der Deviance-Residuen identifiziert. Schließlich wird die für das CPH-Modell entscheidende PH-Annahme anhand der Schoenfeld-Residuen überprüft. Bei Unzulänglichkeiten soll das Modell anhand der in Abschnitt 2.2.3 eingeführten Erweiterungen modifiziert werden. Abschließend werden die Ergebnisse des finalen Modells kritisch gewürdigt und interpretiert.

3.2. Vorbereitungen

Um das in Abschnitt 3.1 beschriebene Konzept durchführen zu können, müssen einige Anpassungen am Datensatz vorgenommen werden. Neben der Datenaufbereitung werden im Folgenden die verwendeten RStudio-Pakete aufgegriffen. Verweise auf den beigefügten RStudio-Code werden durch Fettdruck kenntlich gemacht.

3.2.1. R-Pakete

In dem empirischen Teil der Arbeit wird RStudio Version 1.2.5001 verwendet. In der Analyse finden dabei die folgenden Pakete Anwendung. Das Paket **survival** (Therneau, 2020) liefert den Großteil der benötigten Funktionen für die Survival-Analyse. Das Paket **ggplot2** (Wickham, 2016) bzw. ergänzend **survminer** (Kassambara & Kosinski, 2020) dient der Erstellung von Visualisierungen. Das Paket **MASS** (Venables & Ripley, 2002) wird für eine schrittweise Variablen-Selektion für das optimale Modell nach dem Akaike-Informationskriterium (AIC) benötigt. Schließlich wird das Paket **Greg** (Gordon & Seifert, 2020) benötigt, um die Überlebenszeiten von den Beobachtungen in Teilintervalle zu zerlegen, falls Zeitabhängigkeit vorliegt. Bevor die genannten RStudio-Pakete auf den Datensatz angewendet werden können, müssen die zugrunde liegenden Daten für die Survival-Analyse aufbereitet werden. Im Folgenden wird das genaue Vorgehen als unterstützende Erläuterung des Codes beschrieben.

3.2.2. Datenaufbereitung

In dem Datensatz **covid.csv** sind ursprünglich 566.602 Beobachtungen enthalten. Um zu gewährleisten, dass in der Analyse ausschließlich Covid-19 Erkrankte berücksichtigt werden, wird der Datensatz auf Beobachtungen mit positivem Covid-19 Test reduziert (**covid_res** $= 1$). Daraufhin verkürzt sich der Datensatz auf 220.657 Beobachtungen. Um für die nachfolgende Analyse Überlebenszeiten zu konstruieren werden die Variablen Einlieferungs- (**entry_date**) und Sterbedatum (**date_died**) als Datumstyp formatiert. Im Datensatz wäre alternativ zu dem Einlieferungsdatum ein Symptom-Eintrittsdatum vorhanden. In vielen Fällen entspricht dieses jedoch dem Einlieferungsdatum, was darauf hindeutet, dass für jede Beobachtung, für die kein Symptom-Eintrittsdatum festgestellt werden konnte, schlicht das Einlieferungsdatum gewählt worden ist. Aus diesem Grund fällt die Entscheidung bei der Konstruktion der Überlebenszeiten auf das Einlieferungsdatum. Um die Überlebenszeiten zu konstruieren, wird die Variable **dfentrydeath** erstellt, welche die Differenz zwischen dem Einlieferungs- und dem Sterbedatum in Tagen abbildet. Da nicht für alle Beobachtungen ein Tod im Erhebungszeitraum festgestellt worden ist, enthält die Variable **dfentrydeath** für alle zensierten Einträge den Wert „NA“. Da Berechnungen mit dem Wert „NA“ nicht möglich sind, werden die Variablen **time** und **death** erstellt. Die Variable **time** nimmt bei zensierten Überlebenszeiten den Wert 0 und für die beobachteten Todeszeitpunkte den Wert von **dfentrydeath** an. Die Variable **death** stellt eine Indikatorvariable dar, welche den Wert 0 bei Zensur und 1 für einen Todesfall annimmt. Anschließend werden 60 Beobachtungen, welche negative Überlebenszeiten aufweisen (**time** < 0) aus dem Datensatz entfernt, da diese unplausibel

erscheinen und auf Fehler in der Erhebung hindeuten. Weiterhin werden 2032 Beobachtungen, welche bei Einlieferung einem sofortigen Tod erliegen (Überlebenszeit $= 0$) aus der Analyse ausgeschlossen, da für sie $S(0) = 1$ nach Formel (1) nicht gilt. Anschließend werden 454 Beobachtungen mit **age** kleiner 1 und größer 100 entfernt, um unplausible Altersangaben auszuschließen. Weiterhin werden die Ausprägungen von den nominal-skalierten Variablen aufgrund geringen Aussagegehalts von *„1 = Ja , 2 = Nein, 97 = Nicht Anwendbar, 98 = Ignoriert, 99 = Unbestimmt“* zu *„1 = Ja und 0 = Nein (inklusive 97, 98 und 99)“* zusammengefasst. Abschließend werden die nominal-skalierten Variablen faktorisiert und die nicht benötigten Variablen vom Datensatz entfernt. Zu den entfernten Variablen zählen **outpatient**, **other_disease** und **other_contact_covid**. Außerdem wurden die zuvor erwähnten, zur Konstruktion der Überlebenszeiten verwendeten Variablen entfernt. Die Variable **outpatient** gibt dabei an, ob es sich um ambulante oder stationär behandelte Patienten handelt. Die Variable **other_disease** gibt an, ob bei dem Patienten weitere Krankheiten vorliegen, welche nicht durch die restlichen Variablen erfasst worden sind. Die Variable **other_contact_covid** gibt an, ob ein Kontakt zwischen dem Patienten und einer Covid-19 infizierten Person bestand. Da sich die Survival-Analyse auf spezifische, medizinische Einflussfaktoren beschränken soll, sind die genannten Variablen für die weitere Untersuchung irrelevant. Für die anschließende Analyse enthält der Datensatz schließlich noch 218.111 Beobachtungen mit den in Tabelle 1 aufgezählten 17 Variablen.

3.3. Survival-Analyse von Covid-19 Patientendaten

Die angepassten Daten erfüllen nach der Aufbereitung in Abschnitt 3.2.2 aus technischer Sicht die Voraussetzung für die anschließende Survival-Analyse.

3.3.1. Kaplan-Meier Schätzung

Um eine Kaplan-Meier Schätzung zu konstruieren, wird zunächst eine **Surv()**-Funktion erstellt, welche die Indikatorvariable **death** und die Überlebenszeit-Variable **time** in einem Survival-Objekt zusammenfasst. Nach der Erstellung der **Surv()**-Funktion wird mittels der Funktion **survfit()** die Survival-Funktion per Kaplan-Meier Methode geschätzt. Beide Funktionen sind in dem **survival**-Paket enthalten. Tabelle 2 sind die geschätzten Survival-Funktion $\hat{S}(t)$ für $t = 0 - 5, 10, 15, 20$ und 25 Tagen, sowie die jeweiligen 0.95-Konfidenzintervalle (0.95-KI) zu entnehmen. Die Notation in der Kopfzeile erfolgt dabei nach Abschnitt 2.2.1. Neben dem numerischen Überblick in Tabelle 2 wird in Abbildung 1 im Appendix eine Visualisierung von $\hat{S}(t)$ für die Überlebenszeit t dargestellt. Aus Abbildung 1 im Appendix wird ersichtlich, dass die mediane Überlebenszeit der Beobachtungen bei 6 Tagen liegt. Die mediane Überlebenszeit sagt hier aus, dass 50% der Covid-19 Patienten nach 6 Tagen Krankenhausaufenthalt verstorben sind. Die Kaplan-Meier Schätzung gibt einen ersten deskriptiven Einblick über das Mortalitätsrisiko von Covid-19 Patienten. Um die in Tabelle 1 erhobenen Variablen als Einflussfaktoren auf die Überlebenszeiten zu betrachten, wird im Folgenden eine Cox-Regression konstruiert.

3.3.2. Konstruktion der Cox-Regression

Die Konstruktion des CPH-Modells erfolgt in der Praxis nach einer ähnlichen Vorgehensweise, wie in der konventionellen Regressions-Analyse üblich. Die erklärenden Variablen werden auf die abhängige Variable regressiert, wobei die abhängige Variable hier das zuvor erwähnte Survival-Objekt ist. Anhand einer beidseitigen Variablenselektion nach dem AIC findet dabei die Modellauswahl für ein Step-Modell (**CoxStep**) statt. In Tabelle 1 im Appendix sind die Kovariablen des **CoxStep**-Modells, deren Koeffizienten im Exponenten der Exponentialfunktion, das 0.95-KI der Koeffizienten und der probabilitas-Wert (p-Wert) dieser zu entnehmen. Aus Tabelle 1 im Appendix geht hervor, dass alle Koeffizienten abgesehen von **copd** auf einem Alpha-Niveau von 5% einen signifikanten Einfluss auf die Hazardraten haben. Des Weiteren ist das Gesamt-Modell nach Likelihood-Quotienten- (LQ-Test), Wald- (W-Test) und Log-Rank-Test (LR-Test) signifikant und weist einen AIC von 456261.5 auf. Allerdings sind die Koeffizienten an dieser Stelle noch nicht interpretierbar, da zunächst die Modellannahmen geprüft werden müssen.

3.3.3. Überprüfung der Annahmen der Cox-Regression

Um valide Schätzergebnisse des CPH-Modells zu gewährleisten, müssen die Modellannahmen und eine allgemeine Angemessenheit der Cox-Regression bestätigt werden. Im Folgenden werden die in Abschnitt 2.2.4 eingeführten Residuen zur Modell-Diagnostik herangezogen.

Allgemeine Anpassung des Modells

Um in einem ersten Schritt die allgemeine Anpassung des Modells zu untersuchen können die Cox-Snell-Residuen verwendet werden. In RStudio existiert derzeit keine simple Funktion zur Ausgabe der Cox-Snell-Residuen. Diese lassen sich allerdings unproblematisch durch Umstellung der Martingale-Residuen in Ausdruck (18) herleiten. Dafür werden zunächst die Martingale-Residuen für jede Beobachtung mit der Funktion **residuals(coxstep, type = “martingale“)** berechnet und anschließend von der Indikatorvariable **death** subtrahiert. Nach Erstellung der Cox-Snell-Residuen wird eine Survival-Funktion vom Kaplan-Meier Typ geschätzt (siehe Abschnitt 3.3.1), welche auf den zuvor erstellten Cox-Snell-Residuen basiert. Anschließend wird der negative Logarithmus der Survival-Funktion berechnet, um wie nach Formel (8) die kumulative Hazardrate zu erhalten. Schließlich wird die kumulative Hazardrate gegen die Cox-Snell-Residuen in einem Graph dargestellt. Dies veranschaulicht Abbildung 2 im Appendix. Im Graphen zeigt sich eine Gerade mit Ursprung in Null und einer konstanten Steigung von circa Eins. Nach Abschnitt 2.2.4 deutet dies auf eine allgemeine Angemessenheit der Modell-Anpassung hin. Da die grafische Darstellung der Cox-Snell-Residuen nur ein ungefähres Bild über die Eignung des Modells gibt, sind weitere Untersuchungen der Modellannahmen notwendig.

Tabelle 1: Übersicht der in der Analyse verwendeten Variablen (eigene Darstellung)

Variable	Bedeutung	Skalenniveau	Ausprägung
age	Alter	metrisch	Alter in Jahren
asthma	Asthma-Erkrankung	nominal	ja = 1, nein = 0
cardiovascular	Herz-Kreislauf-Erkrankung	nominal	ja = 1, nein = 0
copd	Lungenerkrankung (COPD)	nominal	ja = 1, nein = 0
death	Todesfall	nominal	ja = 1, nein = 0
diabetes	Diabetes-Erkrankung	nominal	ja = 1, nein = 0
hypertension	Bluthochdruck	nominal	ja = 1, nein = 0
icu	Aufnahme in die Intensivstation	nominal	ja = 1, nein = 0
inmsupr	Autoimmunerkrankung	nominal	ja = 1, nein = 0
intubed	künstliche Beatmung	nominal	ja = 1, nein = 0
obesity	Übergewicht	nominal	ja = 1, nein = 0
pneumonia	Lungenentzündung	nominal	ja = 1, nein = 0
pregnancy	Schwangerschaft	nominal	ja = 1, nein = 0
renal_chronic	chronisches Nierenversagen	nominal	ja = 1, nein = 0
time	Überlebenszeit	metrisch	überlebte Tage
tobacco	Tabakkonsum	nominal	ja = 1, nein = 0
sex	Geschlecht	nominal	w = 1, m = 2

Tabelle 2: Kaplan-Meier Schätzung von $\hat{S}(t)$ (eigene Darstellung)

t	Y_i	d_i	$\hat{S}(t)$	0.95-KI
0	218111	0	1	(1.0; 1.0)
1	25010	2745	0.8902	(0.8864; 0.8941)
2	22265	2525	0.7893	(0.7842; 0.7944)
3	19740	2326	0.6963	(0.6906; 0.7020)
4	17414	2082	0.6130	(0.6070; 0.6191)
5	15332	8817	0.5377	(0.5316; 0.5439)
10	7567	6942	0.2601	(0.2548; 0.2656)
15	3485	3524	0.1192	(0.1153, 0.1233)
20	1591	1641	0.0536	(0.0509; 0.0565)
25	718	725	0.0246	(0.0228; 0.0266)

Linearität der kontinuierlichen Kovariablen

Wie in Abschnitt 2.2.6, angesprochen haben Therneau et al. (1990) vorgeschlagen, die Linearität von kontinuierlichen Kovariablen anhand der Martingale-Residuen zu überprüfen. Demnach wird zunächst eine Schätzung des **Cox-Step**-Modells ohne die Kovariable **age** in einem Test-Modell (**testFitAge**) ermittelt. Anschließend werden die Martingale-Residuen von **testFitAge** berechnet. Schließlich werden die Martingale-Residuen von **testFitAge** gegen die ausgelassene Kovariable **age** dargestellt. Abbildung 3 im Appendix visualisiert dieses Vorgehen. Anhand der Grafik sind systematische Abweichungen von Null erkennbar, was einen Hinweis dafür liefert, dass für die Kovariable **age** Nicht-Linearität vorliegt. Um sich der funktionalen Form von **age** anzunähern ist die Funktion **ggcoxfunctional()** aus dem Paket **survminer** geeignet. Die Funktion **ggcoxfunctional()** erstellt dabei Graphen für verschiedene Transformationen der kontinuierlichen Kovariablen, die gegen die Martingale-Residuen dargestellt werden sollen. Denkbar sind hier Transformationen von **age**, wie etwa $\mathbf{age}^2$, $\sqrt{\mathbf{age}}$ oder log[**age**]. Da die Martingale-Residuen einen Erwartungswert von Null haben, werden Abweichungen von Null mit den vorgeschlagenen Transformationen verglichen. Abbildung 4 im Appendix veranschaulicht dieses Vorgehen. Der Grafik ist zu entnehmen, dass die Transformation von **age** in Form von $\sqrt{\mathbf{age}}$ und log[**age**] am ehesten die funktionale Form der Martingale-Residuen abbildet. Folglich werden beide Transformationen für die weitere Modell-Konstruktion berücksichtigt.

Einflussreiche Beobachtungen

Wie in Abschnitt 2.2.5 erläutert, eignen sich die Df-Betas zur Identifizierung von einflussreichen Beobachtungen. Dabei sollten im Datensatz keine zu einflussreichen Beobachtungen enthalten sein, um zu gewährleisten, dass das Modell keine verzerrten Ergebnisse liefert. Die Konstruktion der Df-Betas erfolgt ähnlich wie die der Martingale-Residuen. Anstelle der Martingale-Residuen wird in der **residuals()**-

Funktion als Typ „dfbetas“ eingetragen. Nach der Berechnung werden die Df-Betas für jede Kovariable visualisiert. In Abbildung 5 und 5 im Appendix wird ersichtlich, dass es einige Beobachtungen gibt, welche im Vergleich zu den restlichen Beobachtungen einen starken Ausschlag in den Df-Betas haben. Dabei messen die auf der Y-Achse aufgetragenen Df-Betas, nach Abschnitt 2.2.5, die Veränderung der Koeffizienten durch Auslassung der einzelnen Beobachtungen in Höhe der Standardabweichung. Da in Abbildung 5 und 6 im Appendix keine Beobachtung einen Ausschlag von annähernd größer als Eins besitzt, deutet dies daraufhin, dass keine unverhältnismäßig einflussreichen Beobachtungen im Datensatz vorliegen.

Proportionales Hazard

Zur Überprüfung der für die Cox-Regression entscheidende PH-Annahme werden die Schoenfeld-Residuen verwendet. Dabei werden, wie in Abschnitt 2.2.5 erläutert, die Schoenfeld-Residuen gegen den Zeitverlauf abgebildet, um systematische Abweichungen von Null aufzuzeigen. Systematische Abweichungen von Null im Graphen sind dabei ein Hinweis auf eine Zeitabhängigkeit der Kovariablen. In Abbildung 7 und 8 im Appendix werden die Schoenfeld-Residuen für jede Kovariable gegen die Überlebenszeiten abgebildet. Aus den Graphen geht hervor, dass für alle Kovariablen bis auf **icu** eine Zeitabhängigkeit denkbar ist. Insbesondere scheinen hohe Überlebenszeiten einen systematischen Trend auszulösen.

Neben der grafischen Überprüfung der PH-Annahme ist ergänzend ein Log-Rank Test der Schoenfeld-Residuen auf Zeitabhängigkeit sinnvoll. Hierfür ist die **cox.zph()**-Funktion als Test auf Zeitabhängigkeit geeignet, da sie einen Log-Rank Test für jede Kovariable und das gesamte Modell vornimmt. In Tabelle 3 ist das Ergebnis eines Log-Rank Test für das **StepCox**-Modell zu entnehmen. Die Interpretation des p-Wert des Log-Rank Test auf Zeitabhängigkeit folgt dabei der in Abschnitt 2.2.5 erläuterten Logik. Von **asthma**, **copd** und **icu** abgesehen kann für alle Kovariablen eine Verletzung der PH-Annahme nicht verworfen werden. Des Weiteren wird ein Log-Rank Test für die transformierten Kovariablen $\sqrt{\textbf{age}}$ und log[**age**] durchgeführt, um die Erfüllung der PH-Annahme zu überprüfen. Dafür wird ein Test-Modell (**FitAgeTransform**) erstellt, welches lediglich die Kovariablen $\sqrt{\textbf{age}}$ und log[**age**] enthält. Aus Tabelle 2 im Appendix wird ersichtlich, dass die PH-Annahme auf einem Alpha-Niveau von 5% für log[**age**] erfüllt und für $\sqrt{\textbf{age}}$ verletzt ist. Die Modell-Konstruktion wird daher mit log[**age**] fortgeführt.

Ausreißer

Ausreißer im Datensatz werden mit den Deviance-Residuen identifiziert. Dabei wird erneut die **residuals()**-Funktion verwendet, jedoch wird als Typ „deviance“ eingetragen. Wie in Abschnitt 2.2.4 erläutert entsprechen Deviance-Residuen mit hohen absoluten Werten Beobachtungen, welche nicht gut durch das Modell angepasst sind. In Abbildung 9 im Appendix sind die Deviance-Residuen des **CoxStep**-Modells gegen die Überlebenszeiten abgebildet. Aus der Grafik geht hervor, dass das Modell Beobachtungen mit besonders hohen Überlebenszeiten schlecht anpasst. Dieses Ergebnis ist insoweit nachvollziehbar, da es intuitiv erscheint, dass einige wenige Patienten mit sehr hohen oder niedrigen Überlebenszeiten die Schätzergebnisse verzerren. Aus Abbildung 9 geht weiterhin hervor, dass der Großteil der Beobachtungen im Intervall [-2;2] liegt und mit zunehmender Überlebenszeit die Anpassung des Modells nachlässt. Daher wird im Datensatz eine neue Variable **Ausreißer** erstellt. Die Variable **Ausreißer** nimmt dabei für jede Beobachtung 1 an, wenn das jeweilige Deviance-Residuum Werte kleiner -2 und 0, falls es Werte größer -2 annimmt. Anschließend werden alle Beobachtungen mit **Ausreißer** $= 1$ aus dem Datensatz entfernt. Insgesamt werden dadurch 285 Ausreißer entfernt, was 0,13% aller Beobachtungen oder 1,14% der Beobachtungen mit Todesfall ausmacht.

Nachdem die Ausreißer entfernt wurden, wird ein neues Step-Modell (**CoxOptimal**) mit Transformation der kontinuierlichen Variable zu log[**age**] konstruiert. In Tabelle 3 im Appendix wird ersichtlich, dass für das **CoxOptimal**-Modell die gleichen Variablen für das Modell ausgewählt werden wie in der vorherigen Variablenselektion. Außerdem weist das **CoxOptimal**-Modell gegenüber dem **CoxStep**-Modell einen niedrigeren AIC (450342.7) und ausschließlich signifikante Koeffizienten bei einem Alpha-Niveau von 5% auf. Nach der Konstruktion erfolgt ein erneuter Log-Rank Test um die PH-Annahme des **CoxOptimal**-Modells zu überprüfen. Nach Tabelle 4 erfüllt das **CoxOptimal**-Modell global die PH-Annahme nicht, allerdings verletzen nur noch die Kovariablen **icu**, **pneumonia** und **renal_chronic** die PH-Annahme individuell. Da die Aufnahme in eine Intensivstation und die Entwicklung einer Lungenentzündung bzw. chronischen Nierenversagens im Krankheitsverlauf eines Covid-19 Patienten auftreten können, ist bei den genannten Kovariablen eine Zeitabhängigkeit denkbar. Folglich sollte das **CoxOptimal**-Modell um die in Abschnitt 2.2.3 angesprochenen Modifikationen erweitert werden.

3.3.4. Erweiterung des Cox-Regression

Im vorherigen Abschnitt wurde dargelegt, dass trotz der Entfernung von Ausreißern und einer Korrektur der funktionalen Form von **age** die PH-Annahme für die Kovariablen **icu**, **pneumonia** und **renal_chronic** verletzt ist. Da bei den drei genannten Kovariablen eine Zeitabhängigkeit denkbar ist, wird eine Interaktion der Kovariablen mit den Überlebenszeiten im Sinne von Abschnitt 2.2.3 implementiert. Dafür müssen im Datensatz zunächst die Beobachtungsintervalle in Teilintervalle aufgeteilt werden. Die Aufteilung der Beobachtungsintervalle erfolgt dabei mit der **timeSplitter()**-Funktion des **Greg**-Pakets. Im Wesentlichen wird durch die **timeSplitter()**-Funktion die Variable **time** in 0,5-Intervalle unterteilt, wobei die neuen Variablen **Start_time** und **Stop_time** die Intervall-Grenzen erfassen. Hervorzuheben ist hierbei, dass obwohl ein neues Survival-Objekt (**Surv_intervall**) erstellt wird, sich das darauf basierende Intervall-Modell nicht von dem vorherigen **CoxOptimal**-Modell unterscheidet. Nach der Aufteilung der Beobachtungsintervalle sind die technischen

Tabelle 3: Log-Rank Test der Kovariablen des **CoxStep**-Modells auf PH-Annahme (eigene Darstellung)

Kovariable	Chi-Sq	Freiheitsgrade	p-Wert
age	10.907	1	0.00096
asthma	3.647	1	0.05617
copd	3.369	1	0.06645
diabetes	8.981	1	0.00273
icu	0.155	1	0.69351
intubed	5.919	1	0.01498
obesity	4.933	1	0.02634
pneumonia	7.231	1	0.00716
renal_chronic	18.459	1	< 0.0001
GLOBAL	55.666	9	< 0.0001

Tabelle 4: Log-Rank Test der Kovariablen des **CoxOptimal**-Modells auf PH-Annahme (eigene Darstellung)

Kovariable	Chi-Sq	Freiheitsgrade	p-Wert
$\log[age]$	0.73473	1	0.39136
asthma	2.45537	1	0.11712
copd	0.36076	1	0.54808
diabetes	0.00217	1	0.96283
icu	14.10517	1	**0.00017**
intubed	0.34892	1	0.55472
obesity	0.20561	1	0.65023
pneumonia	3.41762	1	**0.06450**
renal_chronic	4.95804	1	**0.02597**
GLOBAL	32.75795	9	0.00015

Voraussetzungen für eine Zeitinteraktion der Kovariablen gegeben. Zunächst wird das zuvor konstruierte Modell um eine Zeitinteraktion mit **Start_time** bei den Kovariablen **icu**, **pneumonia** und **renal_chronic** erweitert (**CoxTDC**-Modell). Anschließend wird ein Log-Rank Test durchgeführt. Tabelle 4 im Appendix ist der Log-Rank Test des **CoxTDC**-Modell zu entnehmen. Aus diesem geht hervor, dass die PH-Annahme für die Kovariablen **intubed** und **renal_chronic** sowie das gesamte Modell verletzt ist. Wie in Abschnitt 2.2.6 erwähnt ist der Logarithmus in der Literatur eine präferierte Transformation für die Zeitinteraktion. Daher wird die Variable **logtime** erstellt, welche den Wert $\log[\textbf{Start\_Time} + \textbf{1}]$ für jede Beobachtung annimmt. Anschließend wird ein CPH-Modell, welches eine Interaktion der Kovariablen mit **logtime** besitzt, konstruiert (**CoxTDC_logtime**). Im Anschluss wird ein ein Log-Rank Test des **CoxTDC_logtime**-Modells durchgeführt. Aus Tabelle 5 im Appendix geht hervor, dass das **CoxTDC_logtime**-Modell die PH-Annahme global verletzt, allerdings ist **renal_chronic** die einzige Kovariable, bei der eine individuelle Verletzung vorliegt.

Da im **CoxTDC_logtime**-Modell bei der Kovariable **renal_chronic** durch die Aufnahme einer Zeitinteraktion keine Befriedigung der PH-Annahme zustande kommt, ist es denkbar, dass eine Stratifizierung im Sinne von Abschnitt 2.2.3 Abhilfe verschaffen kann. Durch die Aufnahme des Ausdrucks **strata(renal_chronic)** in der Formel des **CoxTDC_logtime**-Modells wird das Modell in zwei Schichten (**renal_chronic** $= 1$ und **renal_chronic** $= 0$) aufgeteilt. Dabei sollte hervorgehoben werden, dass durch die Stratifizierung für die Kovariable **renal_chronic** kein β-Koeffizient geschätzt werden kann, wodurch ein gewisser Informationsverlust entsteht. Tabelle 6 im Appendix veranschaulicht den Log-Rank Test des stratifizierten CPH-Modells mit zeitabhängigen Kovariablen (**CoxTDC_Strat**), aus dem hervorgeht, dass das **CoxTDC_Strat**-Modell für alle Kovariablen die PH-Annahme erfüllt.

4. Ergebnisse

Da das **CoxTDC_Strat**-Modell schließlich alle Modellannahmen erfüllt, ist eine Interpretation der Koeffizienten möglich. Die Koeffizienten des finalen Modells sind Tabelle 5 zu entnehmen.

4.1. Interpretation der Koeffizienten

Laut Tabelle 5 leisten alle im Modell enthaltenen Koeffizienten nach einem Wald-Test auf einem Alpha-Niveau von 5% einen signifikanten Einfluss auf die Hazardratios von Covid-19 Patienten. Die folgende Interpretation des Einflusses der einzelnen Kovariablen erfolgt dabei anhand der $\exp[\hat{\beta}]$, da diese nach Formel (14) die Hazardratios ausdrücken.

Tabelle 5: Übersicht der $\hat{\beta}$-Koeffizienten des **CoxTDC_Strat**-Modells (eigene Darstellung)

Kovariable	exp[**Koeffizient**]	**0.95-KI**	**p-Wert**
$\log[age]$	1.0779	[1.0268; 1.1316]	0.00249
asthma	0.8807	[0.8061; 0.9623]	0.00495
copd	1.0771	[1.0159; 1.1420]	0.01289
diabetes	1.1233	[1.0939; 1.1535]	$<$ 0.00001
icu $*logtime$	0.8849	[0.8649; 0.9055]	$<$ 0.00001
intubed	0.9019	[0.8663; 0.9391]	$<$ 0.00001
obesity	1.0886	[1.0573; 1.1208]	$<$ 0.00001
pneumonia $*logtime$	1.0559	[1.0397; 1.0724]	$<$ 0.00001
$n = 372518$	$d_i = 24725$	AIC = 438285.2	
LQ-Test: p $<$ 0.0001	W-Test:p $<$ 0.0001	LR-Test:p $<$ 0.0001	

Die Interpretation der $\hat{\beta}$-Koeffizienten lautet wie folgt.
log[**age**]: Während alle anderen Kovariablen konstant gehalten werden (c.p.), erhöht jedes zusätzliche logarithmierte Lebensjahr die Hazardrate, also das Risiko eines Patienten, zu einem gegebenen Zeitpunkt zu sterben, durchschnittlich um 7.79% (um den Faktor 1.0779).
asthma: Zu einem gegebenen Zeitpunkt ist es für einen Patienten, bei dem Asthma diagnostiziert wurde, durchschnittlich 0.8807 mal so wahrscheinlich (11,93% weniger wahrscheinlich) zu sterben, als ein Patient bei dem kein Asthma diagnostiziert wurde (c.p.).
copd: Zu einem gegebenen Zeitpunkt ist es für einen Patienten, bei dem eine chronische Lungenerkrankung diagnostiziert wurde, durchschnittlich 1.0771 mal so wahrscheinlich (7,71% wahrscheinlicher) zu sterben, als ein Patient bei dem keine chronische Lungenerkrankung diagnostiziert wurde (c.p.).
diabetes: Zu einem gegebenen Zeitpunkt ist es für einen Patienten, bei dem Diabetes diagnostiziert wurde, durchschnittlich 1.1233 mal so wahrscheinlich (12,33% wahrscheinlicher) zu sterben, als ein Patient bei dem kein Diabetes diagnostiziert wurde (c.p.).
intubed: Zu einem gegebenen Zeitpunkt ist es für einen Patienten, der künstlich beatmet werden muss, durchschnittlich 0.9019 mal so wahrscheinlich (9.81% weniger wahrscheinlich) zu sterben, als ein Patient welcher nicht künstlich beatmet werden muss (c.p.).
obesity: Zu einem gegebenen Zeitpunkt ist es für einen Patienten, bei dem Übergewicht diagnostiziert wurde, durchschnittlich 1.0886 mal so wahrscheinlich (8,86% wahrscheinlicher) zu sterben, als ein Patient bei dem kein Übergewicht diagnostiziert wurde (c.p.).

Die Ergebnisse für die Kovariablen log[**age**], **copd**, **diabetes** und **obesity** erscheinen plausibel. Es ist nachvollziehbar, dass sich bei einer Diagnose mit einer der genannten Krankheiten die Wahrscheinlichkeit zu sterben erhöht. Gleiches gilt für die Zunahme des Alters eines Patienten. Die Wirkungsrichtung des Koeffizienten von **intubed** ist dagegen diskussionswürdig. Zum einen ist anzunehmen, dass bei der künstlichen Beatmung eines Patienten ein schwerer Krankheitsverlauf vorliegt, welcher ein höheres Sterberisiko impliziert. Zum anderen ist jedoch auch anzunehmen, dass sofern ein schwerer Krankheitsverlauf vorliegt, eine künstliche Beatmung die Überlebenschancen des Patienten positiv beeinflusst. Der Koeffizient von **intubed** erscheint nach dieser Überlegung plausibel. Widersprüchlich ist hingegen der Koeffizient von **asthma**. Aus medizinischer Sicht erscheint es unplausibel, dass eine Asthma-Erkrankung die Sterbewahrscheinlichkeit vermindert, da Asthma eine Atemwegserkrankung ist. Weitere Untersuchungen oder inhaltliche Vergleiche mit anderen Studien könnten hier neue Erkenntnisse liefern.

Die Interpretation der $\hat{\beta}$-Koeffizienten von **icu** und **pneumonia** ist durch die Aufnahme der Zeitinteraktion mit **logtime** komplexer als die der zuvor beschriebenen Kovariablen. Da die $x_i(t)$ im Zeitverlauf variieren, ist die Hazardratio nach Formel (15) abhängig von den Überlebenszeiten und muss bei einer Interpretation der Koeffizienten berücksichtigt werden.

icu $*logtime$: Die Hazardrate eines Patienten, welcher auf die Intensivstation aufgenommen wird, ist durchschnittlich um den Faktor $\mathbf{0.8849} * \log[\mathbf{Start\_Time} + \mathbf{1}]$ höher/niedriger als die eines Patienten, welcher nicht in die Intensivstation aufgenommen wird. Für Überlebenszeiten von 1, 2, 3 und 4 Tagen ergeben sich Hazardratios in Höhe von 0.6134, 0.9722, 1.2267 und 1.4242. Inhaltlich sagen die Hazardratios aus, dass ein Patient, welcher seit 1,2,3 und 4 Tagen künstlich beatmet wird, durchschnittlich um den Faktor 0.6134, 0.9722, 1.2267 und 1.4242 so wahrscheinlich verstirbt wie ein Patient, welcher nicht künstlich beatmet wird (c.p.). Insgesamt scheint im Zeitverlauf das Sterberisiko eines Patienten, welcher in eine Intensivstation eingewiesen wird, gegenüber einem, der nicht eingewiesen wird, stark anzusteigen. Inhaltlich kann dieses Ergebnis wie folgt interpretiert werden. Wird ein Patient in die Intensivstation eingeliefert, ist dessen Sterbewahrscheinlichkeit gegenüber einem Patienten, welcher nicht eingeliefert wird, zunächst geringer. Mit zunehmender Zeit auf der Intensivstation steigt die Wahrscheinlichkeit zu sterben gegenüber einem Patienten in gewöhnlicher Behandlung an. Es ist anzunehmen, dass mit einer Verlegung auf die Intensivstation ein schwerwiegenderer Krankheitsverlauf einhergeht, was eine geringere Überlebenswahrscheinlichkeit impliziert. Allerdings ist auch

anzunehmen, dass auf einer Intensivstation eine bessere medizinische Versorgung zugänglich ist, was eine Verminderung der Sterbewahrscheinlichkeit in den ersten Tagen erklärt. Bei Patienten mit voranschreitender Aufenthaltszeit auf der Intensivstation ist anzunehmen, dass sich keine Verbesserung des Gesundheitszustands einstellt, was die ansteigende Hazardratio aufzeigen könnte. Der Koeffizient von **icu** $*logtime$ erscheint nach genauerer Betrachtung plausibel.

pneumonia $*logtime$: Für den $\hat{\beta}$-Koeffizienten von **pneumonia** ergibt sich eine ähnliche Systematik wie für den von **icu**. Die Hazardrate eines Patienten mit einer Lungenentzündung und Überlebenszeiten von 1,2,3 oder 4 Tagen ist durchschnittlich um den Faktor 0.7319, 1.160, 1.4638 und 1.6994 mal so hoch wie bei einem Patienten ohne Lungenentzündung (c.p.). Während der Koeffizient für eine Überlebenszeit von einem Tag unplausibel erscheint, lässt sich die Zunahme der Hazardratio im weiteren Zeitverlauf nachvollziehen. Es ist anzunehmen, dass Patienten, bei denen eine Lungenentzündung diagnostiziert wird, unter besonderer Beobachtung stehen. Bei einem medizinischem Notfall könnten diese vermutlich schneller verpflegt werden als Patienten, bei denen keine Lungenentzündung diagnostiziert wurde. Stellt sich kurzfristig keine Verbesserung des Gesundheitszustandes ein, steigt die Hazardratio. Anhand dieser Überlegung erscheint das zunächst verminderte Sterberisiko eines Patienten mit Lungenentzündung plausibel.

4.2. Diskussion des Modells

Um valide Schätzergebnisse eines CPH-Modells zu erhalten, ist es notwendig, die Erfüllung der Modell-Annahmen genauestens zu überprüfen. In Abschnitt 3.3.3 konnte gezeigt werden, dass eine Beurteilung der allgemeinen Anpassung anhand der Cox-Snell-Residuen nicht ausreicht. Dies wird daran deutlich, dass eine Visualisierung der Cox-Snell-Residuen in Abbildung 2 im Appendix zunächst eine gute Modell-Anpassung vermuten lässt. Im weiteren Verlauf der Analyse stellt sich jedoch heraus, dass mehrere Annahmen verletzt sind. Insbesondere ist, wie in Abschnitt 3.3.3 dargelegt, die Annahme eines linearen Zusammenhangs zwischen der abhängigen und der kontinuierlichen Kovariable (**age**) verletzt. Eine Transformation der Kovariable in Form von log[**age**] gewährleistet den benötigten linearen Zusammenhang. Die funktionale Form ist hier fragwürdig, da die Kovariable für Säuglinge (**age** $= 0$) nicht definiert ist. Dies ist hier unproblematisch, da der bereinigte Datensatz keine Beobachtungen mit **age** < 1 enthält. Die in Abschnitt 3.3.3 angesprochene Entfernung von Ausreißern ist ebenfalls diskussionswürdig. Im Allgemeinen ist es nicht abwegig unplausible Beobachtungen zu entfernen, wenn diese etwa auf einen Erhebungsfehler hindeuten. Allerdings geht mit der Reduzierung des Datensatzes ein gewisser Informationsverlust einher. Nachdem die funktionale Form von **age** angepasst und die identifizierten Ausreißer aus dem Datensatz entfernt wurden, konnte die PH-Annahme für die Kovariablen **icu**, **pneumonia** und **renal_chronic** nicht bestätigt werden. Das CPH-Modell wurde daher in Abschnitt 3.3.4 erweitert. Während die Aufnahme eines Zeitinteraktions-Terms für die Kovariablen **icu** und **pneumonia** Abhilfe verschaffte, konnte für **renal_chronic** auf diese Weise keine Problemlösung erzielt werden. Daher wird für **renal_chronic** stratifiziert, was zur Folge hat, dass kein $\hat{\beta}$-Koeffizient für **renal_chronic** geschätzt werden kann. Die Interpretation der $\hat{\beta}$-Koeffizienten für **icu** und **pneumonia** wird nach Aufnahme der Zeitinteraktion komplexer als die der restlichen Kovariablen des Modells. Aussagen über die Hazardratios lassen sich für **icu** und **pneumonia** nur in Abhängigkeit von spezifischen Überlebenszeiten treffen, was eine allgemeine Interpretation erschwert. Darüber hinaus ist die Wirkungsrichtung der $\hat{\beta}$-Koeffizienten von **asthma**, **intubed**, **pneumonia** und **icu** fragwürdig. Im finalen Modell **CoxTDC_Strat** erfüllen alle Kovariablen die PH-Annahme. Allerdings sollte hervorgehoben werden, dass der LR-Test die Null-Hypothese von zeitunabhängigen Kovariablen bei **asthma** und **intubed** nur auf einem Alpha-Niveau von 10% nicht verwerfen kann. Inhaltlich deutet dies an, dass das Modell formal die PH-Annahme erfüllt, jedoch nicht mit hoher Wahrscheinlichkeit. Es ist daher also fraglich, ob die Aufnahme von zeitabhängigen Kovariablen das Vorhandensein von nicht proportionalen Hazards optimal handhabt.

5. Fazit und Ausblick

Ziel dieser wissenschaftlichen Arbeit war es, die Realisierbarkeit eines auf der Survival-Analyse basierenden Modells für Covid-19 Patienten zu prüfen. Dabei wurden anhand eines erweiterten CPH-Modells verschiedene Einflüsse auf die Mortalität untersucht. Es konnte gezeigt werden, dass die Überprüfung der Linearität der kontinuierlichen Kovariablen, die Untersuchung der PH-Annahme und die Identifikation von Ausreißern von größter Relevanz für die Modell-Konstruktion sind. Werden die zugrunde liegenden Voraussetzungen beachtet, steht die Realisierbarkeit des Modells dennoch offen. In dem untersuchten Datensatz war die entscheidende Annahme von proportionalen Hazards an mehreren Stellen verletzt. Anhand einer Visualisierung der Martingale-Residuen konnte gezeigt werden, dass zwischen der abhängigen und der kontinuierlichen Kovariable (**age**) ein nicht linearer Zusammenhang besteht. Durch Transformation zu log[**age**] konnte ein linearer Zusammenhang geschaffen werden. Mittels Deviance-Residuen konnten Ausreißer im Datensatz identifiziert werden. Dabei wurden Beobachtungen mit besonders hohen Überlebenszeiten entfernt, da sie durch das Modell schlecht angepasst werden. Um die konstruierten Modelle auf die PH-Annahme zu überprüfen wurden die Schoenfeld-Residuen und der auf ihnen basierende Log-Rank-Test herangezogen. Nachdem die funktionale Form von **age** korrigiert und die Ausreißer entfernt wurden, erfüllten alle Kovariablen abgesehen von **icu**, **pneumonia** und **renal_chronic** die PH-Annahme. Durch die Aufnahme eines Interaktionstermes mit der logarithmierten Überlebenszeit konnte die PH-Annahme für die Kovariablen **icu** und **pneumonia** befriedigt werden. Durch Stratifizierung des Datensatzes anhand der Kovariable **renal_chronic** wurde schließlich ein finales Modell konstruiert, welches die

PH-Annahme für alle Kovariablen erfüllt. Die $\hat{\beta}$-Koeffizienten des finalen Modells von log[**age**], **copd**, **diabetes** und **obesity** lassen sich plausibel interpretieren. Die $\hat{\beta}$-Koeffizienten von **intubed**, **pneumonia** und **icu** sind dagegen diskussionswürdig. Der $\hat{\beta}$-Koeffizient **asthma** erscheint dagegen unplausibel. Darüber hinaus kann das finale Modell die Null-Hypothese der PH-Annahme für die Kovariablen **asthma** und **intubed** nur auf einem Alpha-Niveau von 10% nicht verwerfen. Es ist daher also fraglich, ob die zugrunde liegende Annahme von proportionalen Hazards sinnvoll und damit das Modell angemessen ist.

Zusammenfassend lässt sich sagen, dass sich die untersuchten Covid-19 Patientendaten aus technischer Sicht für eine Survival-Analyse mittels Cox-Regression eignen. Um die Cox-Regression auf die untersuchten Covid-19 Patientendaten anwenden zu können, musste das CPH-Modell jedoch erweitert werden. Die Modifikation des Modells bringt nicht nur technischen Aufwand mit sich, sondern führt auch zu einer erschwerten Interpretation der $\hat{\beta}$-Koeffizienten. Wie in der Datenanalyse gezeigt wurde, ist die Interpretation bei zeitabhängigen Kovariablen komplexer, da nur Aussagen zu spezifischen Überlebenszeiten getroffen werden können. Es lässt sich summieren, dass das CPH-Modell durch die Aufnahme von stratifizierten und zeitabhängigen Kovariablen sowie Transformationen der kontinuierlichen Kovariablen und der Entfernung von Ausreißern an die Grenzen seiner Praktikabilität kommt. Es sollte schließlich anhand von den vorgestellten Diagnostik-Methoden abgewogen werden, ob notwendige Anpassungen in einem Verhältnis mit dem einhergehenden Informationsverlust stehen oder ob für die Analyse ein besser geeignetes Modell herangezogen werden sollte.

Zukünftige Forschung könnte mit einer Untersuchung von alternativen Schätzmethoden des CPH-Modells anknüpfen. Wie aus der Datenanalyse hervorging, traten vor allem Probleme bei der Einbindung von Ausreißern und der Verletzung der PH-Annahme auf. Als alternativen Ansatz zur Handhabung von Ausreißern und Verletzungen der PH-Annahme ist eine robuste, gewichtete Schätzung des CPH-Modells denkbar. Bei einem gewichteten CPH-Modell wird die zu maximierende Likelihood-Funktion beispielsweise so gewichtet, dass früheren Todesfällen höheres und späteren Todesfällen geringeres relatives Gewicht in der Schätzung zukommt. In der Anwendung sind bei Covid-19 Patientendaten bisher vor allem konventionelle CPH-Modelle verwendet worden. Weiterführende Forschung könnten daher die Anwendung von robusten, gewichteten CPH-Modellen für Covid-19 Patientendaten untersuchen.

Literatur

Altman, D. G., De Stavola, B. L., Love, S. B. & Stepniewska, K. A. (1995). Review of survival analyses published in cancer journals. *British Journal of Cancer*, *72* (2), 511–518.

Andersen, P. K. & Gill, R. D. (1982, 12). Cox's regression model for counting processes: A large sample study. *Ann. Statist.*, *10* (4), 1100–1120.

Barlow, W. E. & Prentice, R. L. (1988, 03). Residuals for relative risk regression. *Biometrika*, *75* (1), 65-74.

Bayes, M. & Price, M. (1763). An essay towards solving a problem in the doctrine of chances. by the late rev. mr. bayes, f. r. s. communicated by mr. price, in a letter to john canton, a. m. f. r. s. *Philosophical Transactions (1683-1775)*, *53*, 370–418.

Ben-Israel, A. (1966). A newton-raphson method for the solution of systems of equations. *Journal of Mathematical Analysis and Applications*, *15* (2), 243 - 252.

Bhandari, S., Tak, A., Singhal, S., Shukla, J., Shaktawat, A. S., Gupta, J., ... others (2020). Patient flow dynamics in hospital systems during times of covid-19: Cox proportional hazard regression analysis. *Frontiers in public health*, *8*.

Bland, J. M. & Altman, D. G. (2004). The logrank test. *Bmj*, *328* (7447), 1073.

Box-Steffensmeier, J. M., Box-Steffensmeier, J. M. & Jones, B. S. (2004). *Event history modeling: A guide for social scientists*. Cambridge University Press.

Cain, K. C. & Lange, N. T. (1984). Approximate case influence for the proportional hazards regression model with censored data. *Biometrics*, 493–499.

Casson, R. J. & Farmer, L. D. (2014). Understanding and checking the assumptions of linear regression: a primer for medical researchers. *Clinical & experimental ophthalmology*, *42* (6), 590–596.

Chen, R., Liang, W., Jiang, M., Guan, W., Zhan, C., Wang, T., ... others (2020). Risk factors of fatal outcome in hospitalized subjects with coronavirus disease 2019 from a nationwide analysis in china. *Chest*, *158* (1), 97–105.

Cheng, Y., Luo, R., Wang, K., Zhang, M., Wang, Z., Dong, L., ... Xu, G. (2020). Kidney disease is associated with in-hospital death of patients with covid-19. *Kidney international*, *97* (5), 829–838.

Collett, D. (2015). *Modelling survival data in medical research*. CRC press.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34* (2), 187–220.

Cox, D. R. & Snell, E. J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society. Series B (Methodological)*, *30* (2), 248–275.

Crowley, J. & Hu, M. (1977). Covariance analysis of heart transplant survival data. *Journal of the American Statistical Association*, *72* (357), 27–36.

Cummings, M. J., Baldwin, M. R., Abrams, D., Jacobson, S. D., Meyer, B. J., Balough, E. M., ... others (2020). Epidemiology, clinical course, and outcomes of critically ill adults with covid-19 in new york city: a prospective cohort study. *The Lancet*, *395* (10239), 1763–1770.

Farcomeni, A. & Viviani, S. (2011). Robust estimation for the cox regression model based on trimming. *Biometrical Journal*, *53* (6), 956–973.

Fisher, L. D. & Lin, D. Y. (1999). Time-dependent covariates in the cox proportional-hazards regression model. *Annual review of public health*, *20* (1), 145–157.

Frees, E. W. (2009). *Regression modeling with actuarial and financial applications*. Cambridge University Press.

Gad, S. C. & Rousseaux, C. G. (2002). 15 - use and misuse of statistics in the design and interpretation of studies. In W. M. HASCHEK, C. G. ROUSSEAUX & M. A. WALLIG (Hrsg.), *Handbook of toxicologic pathology (second edition)* (Second Edition Aufl., S. 327 - 418). San Diego: Academic Press.

Galántai, A. (2000). The theory of newton's method. *Journal of Computational and Applied Mathematics*, *124* (1), 25 - 44. (Numerical Analysis 2000. Vol. IV: Optimization and Nonlinear Equations)

Gamerman, D. (1991). Dynamic bayesian models for survival data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *40* (1), 63–79.

Gordon, M. & Seifert, R. (2020). Greg: Regression helper functions [Software-Handbuch]. Zugriff auf https://cran.r-project.org/web/packages/Greg/index.html

Grambsch, P. M. & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, *81* (3), 515–526.

Grasselli, G., Greco, M., Zanella, A., Albano, G., Antonelli, M., Bellani, G., ... others (2020). Risk factors associated with mortality among patients with covid-19 in intensive care units in lombardy, italy. *JAMA internal medicine*, *180* (10), 1345–1355.

Ji, D., Zhang, D., Xu, J., Chen, Z., Yang, T., Zhao, P., ... others (2020). Prediction for progression risk in patients with covid-19 pneumonia: the call score. *Clinical Infectious Diseases*, *71* (6), 1393–1399.

Kaplan, E. L. & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, *53* (282), 457–481.

Kassambara, A. & Kosinski, M. (2020). survminer: Drawing survival curves using "ggplot2" [Software-Handbuch]. Zugriff auf https://cran.r-project.org/web/packages/survminer/

Keele, L. (2010). Proportionally difficult: Testing for nonproportional hazards in cox models. *Political Analysis*, *18* (2), 189–205.

Klein, J. P. & Moeschberger, M. L. (2006). *Survival analysis: techniques for censored and truncated data*. Springer Science & Business Media.

Kleinbaum, D. G. & Klein, M. (2011). The cox proportional hazards model and its characteristics. In *Survival analysis* (S. 97–159). New York, NY: Springer New York.

Lagakos, S. W. (1981). The graphical evaluation of explanatory variables in proportional hazard regression models. *Biometrika*, *68* (1), 93–98.

Li, X., Xu, S., Yu, M., Wang, K., Tao, Y., Zhou, Y., ... others (2020). Risk factors for severity and mortality in adult covid-19 inpatients in wuhan. *Journal of Allergy and Clinical Immunology*, *146* (1), 110–118.

Mantel, N. (1966). Models for complex contingency tables and polychotomous dosage response curves. *Biometrics*, *22* (1), 83–95.

Rosenberg, E. S., Dufort, E. M., Udo, T., Wilberschied, L. A., Kumar, J., Tesoriero, J., ... others (2020). Association of treatment with hydroxychloroquine or azithromycin with in-hospital mortality in patients with covid-19 in new york state. *Jama*, *323* (24), 2493–2502.

Rulli, E., Ghilotti, F., Biagioli, E., Porcu, L., Marabese, M., D'Incalci, M., ... Torri, V. (2018). Assessment of proportional hazard assumption in aggregate data: a systematic review on statistical methodology in clinical trials using time-to-event endpoint. *British journal of cancer*, *119* (12), 1456–1463.

Schoenfeld, D. (1982). Partial residuals for the proportional hazards regression model. *Biometrika*, *69* (1), 239–241.

Sedlacek, G. (2018). *Analyse der studiendauer und des studienabbruchrisikos: unter verwendung der statistischen methoden der ereignisanalyse*. Peter Lang International Academic Publishers.

Shi, Q., Zhang, X., Jiang, F., Zhang, X., Hu, N., Bimu, C., ... others (2020). Clinical characteristics and risk factors for mortality of covid-19 patients with diabetes in wuhan, china: a two-center, retrospective study. *Diabetes care*, *43* (7), 1382–1391.

Therneau, T. M. (2020). A package for survival analysis in r [Software-Handbuch]. Zugriff auf https://CRAN.R-project.org/package=survival

Therneau, T. M., Grambsch, P. M. & Fleming, T. R. (1990). Martingale-based residuals for survival models. *Biometrika*, *77* (1), 147–160.

Venables, W. N. & Ripley, B. D. (2002). *Modern applied statistics with s* (Fourth Aufl.). New York: Springer. (ISBN 0-387-95457-0)

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Zugriff auf https://ggplot2.tidyverse.org

Windzio, M. (2013). *Regressionsmodelle für zustände und ereignisse: Eine einführung* (2013. Aufl.). Wiesbaden: Springer Fachmedien.

Wollschläger, D. (2010). *Grundlagen der datenanalyse mit r*. Springer.

Xue, Y. & Schifano, E. D. (2017). Diagnostics for the cox model. *Communications for statistical Applications and Methods*, *24* (6), 583–604.

Zha, L., Li, S., Pan, L., Tefsen, B., Li, Y., French, N., ... Villanueva, E. V. (2020). Corticosteroid treatment of patients with coronavirus disease 2019 (covid-19). *Medical Journal of Australia*, *212* (9), 416–420.

Zucker, D. M. & Karr, A. F. (1990). Nonparametric survival analysis with time-dependent covariate effects: A penalized partial likelihood approach. *The Annals of Statistics*, *18* (1), 329–353.

Zwiener, I., Blettner, M. & Hommel, G. (2011). Überlebenszeitanalyse. *Dtsch Arztebl International*, *108* (10), 163-169.

Junior Management Science 7(2) (2022) 354-389

Junior Management Science

journal homepage: www.jums.academy

Unraveling the Process of Knowledge Integration in Agile Product Development Teams

Julia Haselsteiner

Johannes Kepler Universität Linz

Abstract

Agile product development seems to be the solution for many companies to drive innovation and shorten time-to-market, but what mechanisms lie behind the promises of faster development times and more innovative products? Defined as locus of innovation and driver of dynamic performance, the concept of cross-functional knowledge integration and the organizational learning literature have the potential to provide answers here. Recent empirical studies imply that knowledge integration happens on multiple levels that influence each other, and that environmental uncertainty leads to changes in the knowledge integration process. However, the interplay of individual-level and group-level knowledge integration has not yet received adequate attention and prior studies do not show how knowledge integration changes over shorter periods of time. This paper takes a grounded theory approach to explore the knowledge integrating mechanisms in two agile product development teams. The resulting iterative process model shows how agile teams integrate diverse contributions of individual team members into a new product, how internal and external factors trigger alterations in knowledge integration practices, and how agile teams adapt to changes in coordination and collaboration demands.

Keywords: Agile product development; knowledge integration; cross-functional teams; organizational learning.

1. Introduction

With the growing need for flexibility and reactivity of companies to successfully compete in an increasingly complex and ambiguous environment, practitioners and scholars have begun to search for organizational designs and management approaches that enhance innovative performance and fit a dynamic world. Over the last few years the buzzword "agile" has been appearing in numerous popular scientific papers and books on management providing guidance for organizations to becoming "agile", which is often referred to as being flexible and adaptive (Hasenzagl & Link, 2017, p. 47). Agile innovation methods, like Scrum or Design Thinking, have been contributing to software development success for 30 years, being proposed to improve quality and time-to-market of the products as well as motivation and productivity in the product development teams (Rigby, Sutherland, & Takeuchi, 2016, p. 41). More recently, the benefits and enhancements that the implementation of agile values, principles, and practices promise, have induced companies of other industries to adopt the agile management approach that radically differs from traditional command-and-control-styles, for enhancing product development, marketing, HR and strategy making (Rigby et al., 2016, p. 42). An integral part of the agile principles, established for software development, is the implementation of self-organizing teams composed of motivated individuals from diverse functions or areas of expertise that shall receive the support they need and be trusted to achieve the successful completion of a project (Beedle et al., 2001). This central requirement coincides with the general trend inherent in organizations to structure around teams rather than individuals. However, research on agile teams outside the software development context remains scarce, raising questions of whether the positive effects that the agile methodology promises can also be achieved in other settings and if so, what the mechanisms behind increased quality and time-to-market are.

The literature on new product development has long been dealing with the factors that determine new product success. An aspect that has been receiving considerable attention is the role of cross-functional teams in achieving innovative outcomes. For example, in the rugby approach to new product development, that serves as a theoretical foundation for Scrum, multidisciplinary teams represent a central feature (Takeuchi & Nonaka, 1986, p. 138). Cross-functional team

DOI: https://doi.org/10.5282/jums/v7i2pp354-389

compositions, entailing not only different functional specializations, but also diverse thought processes and behavioral patterns, are thereby proposed to foster new ideas and concepts (Takeuchi & Nonaka, 1986, p. 140). Thus, the proposed value of the team approach is based on the incorporation of different perspectives and expertise from different domains to collective task completion (Edmondson & Nembhard, 2009, p. 125). In general, there is consensus in the literature on new product development that effective cross-functional teamwork is critical to team performance and new product success (S. Brown & Eisenhardt, 1995). For example, in their study of software development teams, Faraj and Sproull (2000, p. 1564) found a strong relationship between expertise coordination, i.e. the management of skill and knowledge dependencies, and team performance. Similarly, Hoegl and Gemuenden (2001, p. 437) showed that coordination, defined as "the degree of common understanding regarding the interrelatedness and current status of individual contributions", improved team work quality and subsequently team performance in 145 German software development projects. In providing a structure for the integration of diverse skill sets, perspectives, and other specialist knowledge, the cross-functional team enables the timely exposure of interdependencies across functions early in the new product development process, where corrections are rather easy and inexpensive (Edmondson & Nembhard, 2009, p. 126).

Likewise, the literature on organizational learning acknowledges the value of the cross-functional team for creating new knowledge, e.g. in the form of product or process innovations, through integrating diverse knowledge bases. For example, Leonard-Barton (1995, p. 64) suggests the locus of innovation to lie at the boundaries between specializations and their respective mindsets. Similarly, Nonaka (1994, p. 24) describes the cross-functional team "as the basic building block for structuring the organizational knowledge creation process". However, innovating through integrating knowledge from diverse sources has been representing a core challenge for organizations ever since (Carlile & Rebentisch, 2003, p. 1181). Knowledge differences have been suggested to hinder the performance of cross-functional teams by complicating knowledge integration, especially under novel or uncertain circumstances (Majchrzak, More, & Faraj, 2012, p. 951). The organizational learning literature offers two partly contradicting perspectives on how the specialized knowledge of individuals can be successfully integrated into organizational knowledge (Grunwald, 2003, p. 3): a "cross-learning" perspective suggesting that knowledge needs to be transferred between specialists and combined in their individual brains to enable organizational learning (Kieser & Koch, 2008, p. 331) and a specialization perspective promoting the exertion of structural mechanisms to reduce knowledge transfer, thus accounting for individuals' limited cognitive capacities and economizing on specialization (Grant, 1996, p. 114; Kieser, Beck, & Tainio, 2001, p. 600).

Thus far, contemporary empirical research seems to agree that knowledge integration may involve merely mentioning and demonstrating knowledge, rather than a "deep" sharing of interpretations and interests (Mengis, Nicolini, & Swan, 2018, p. 597). However, extant research also suggests that the intensity of knowledge sharing in teams depends on the novelty and uncertainty they face. For example, Schmickl and Kieser (2008, p. 485) found that the more radical the innovation, the more knowledge was exchanged between project members (Schmickl & Kieser, 2008, p. 485). Furthermore, research that frames knowledge integration as a process rather than an outcome implies that the process of integrating knowledge may change over time (Mengis et al., 2018, p. 598). Studies of this kind especially highlight that knowledge integration progresses between phases of "working together alone" (Bruns, 2013; Enberg, Lindkvist, & Tell, 2006) and phases of closer interaction. While most of the studies on cross-functional teams in new product (e.g. Schmickl & Kieser, 2008) or new process development (e.g. Majchrzak et al., 2012) consider the overall level of uncertainty and its impact on knowledge integration, they do not show how knowledge integration mechanisms and practices change over time within the course of a project. However, unexpected problems or changes in customer demand that increase task uncertainty and novelty may prompt the reliance on coordination mechanisms other than those currently in place (Grant, 1996, p. 113). Finally, despite prior research evidenced the mutual influence of individual and collective knowledge integrating practices (Enberg et al., 2006), thereby implying the need for of a multi-level perspective on knowledge integration processes, the interplay of individual-level and group-level knowledge integration practices has not yet received adequate attention in empirical research on knowledge integration in new product development teams.

To shed light on the success factors of agile teamwork, i.e. the mechanisms behind increased product quality and time-to-market, this master's thesis adopts a process-based view of knowledge integration and deals with changes in knowledge integration practices and mechanisms over the course of agile product development projects and how agile teams cope with such deviations. Agile product development is characterized by its great flexibility in dealing with changes in market demand or technology and in shortening the time-to-market of new products or services. However, the promises of speed and flexibility should not be misleading, because the agile methodology goes hand in hand with a tightly structured process that requires precise planning and includes regular meetings, recurring ceremonies, and other formal coordination mechanisms. Therefore, this product development method is ideally suited to observe the effect of changes in coordination and collaboration requirements, triggered by unexpected changes or emerging uncertainties, on formal and informal mechanisms and practices within the course of a project.

The research method used is consistently qualitative. In taking a grounded theory approach according to the guidelines of the Gioia methodology (Gioia, Corley, & Hamilton, 2013), the knowledge integrating mechanisms and practices in two different agile product development teams of an Aus-

trian manufacturing company specialized in the development and production of special machinery and associated spare parts were explored. Data were predominantly collected through semi-structured interviews with 14 individuals from the teams, including all three roles in agile product development. The fully transcribed interviews and the associated research notes were analyzed by being subjected to a first-order and a second-order analysis as suggested by Gioia et al. (2013). Ultimately, an iterative process model of knowledge integration in agile product development projects in the manufacturing industry that accounts for changes and adaptions in knowledge integration mechanisms and practices was developed. The proposed model shows that the process of knowledge integration in agile new product development teams iterates between group-level practices, in which coordination-related and problem-centered knowledge exchange take place, and individual-level practices, in which team members individually process the information they receive and integrate their cognition in their individual contributions to the project. Knowledge integration in this model is aided by support mechanisms such as a shared information base, physical and psychological proximity, transparency of responsibilities, and project leadership. These mechanisms, individually and in combination, reduce the need for extensive coordination-related knowledge sharing among team members. However, as will be shown, internal conflicts potentially disrupt the effectivity of these mechanisms in reducing knowledge exchange by making cross-functional communication difficult and forcing an increase in conflict resolution measures. Furthermore, uncertainty and novelty temporarily increase the need for both coordination-related and problem-centered knowledge exchange, hampering the routine knowledge integration process, which gets by without an intensive exchange of knowledge. To cope with these disruptive factors, team members increasingly engage in dialogical knowledge exchange, in the form of group-problem-solving and decision-making as well as bilateral exchanges outside the planned meetings, for a short time, leading to structural adaptions in some cases. The model that will be developed in the results section thus represents an iterative process model of knowledge integration that shows (a) how agile teams integrate the diverse contributions of the individual team members into a new product, (b) how internal and external factors trigger alterations in knowledge integration practices, and (c) how agile teams adapt to the resulting changes in coordination and collaboration demands.

The present study contributes to the organizational learning literature on knowledge integration in three ways. First, it provides an understanding about the contingencies that determine the depth of knowledge exchange in project teams and the effectiveness of structural support mechanisms in reducing the need for extensive knowledge sharing. In doing so, the consistent substitutive effect on knowledge exchange that is attributed to structural support mechanisms (e.g. Schmickl & Kieser, 2008) is called into question. It rather appears that the effectiveness of support mechanisms varies over the course of a project, depending on the project phase, the project context, and the quality of teamwork. Furthermore, while prior studies have discovered that the degree of innovation of a project determines the depth of knowledge integration (Majchrzak et al., 2012, p. 958; Schmickl & Kieser, 2008, p. 485), the present study implies that disruptions trigger changes in the depth of knowledge integration within the course of a project, temporarily increasing the need for deep-level knowledge exchange among experts. Second, empirical insights on the locus of knowledge integration are added. While prior studies mostly examined individual, group, and organizational-level knowledge integration in isolation, this master's thesis adopts a multi-level perspective on knowledge integration (Enberg et al., 2006) and shows how individual-level and group-level knowledge integration practices interact and how the locus of knowledge integration changes as a reaction to uncertainty. Third, the model developed in this master's thesis enriches our knowledge about the factors that influence knowledge integration practices with a team internal perspective that prior research in the field has neglected. While breakdowns (Lok & Rond, 2013, p. 186) in the knowledge integration process are proposed to be triggered by epistemic uncertainty and lead to temporary intensifications of collaboration in integrating knowledge (Mengis et al., 2018), the present thesis indicates that breakdowns in knowledge integration may also be triggered by smaller events, like conflicts, albeit having different consequences. In this regard, the present thesis suggests that disruptions may lead to temporary changes not only in the nature of collaboration but also in the characteristics of coordination. The central practical implication that arises from this master's thesis is that effectiveness and efficiency of the knowledge integration process can be influenced by actively managing the internal factors that tend to erode the effects of support mechanisms. However, the data also suggest that the agile product development methodology may only be appropriate in uncertain and novel contexts, in concept phases, and other phases were close coordination and collaboration are necessary. In rather routine projects of low innovativeness, in which interdependencies are low, the agile project structure could tie up too many resources.

2. Theoretical background

2.1. Organizational learning and knowledge integration

In contemporary literature on organizational learning, the change or increase in organizational knowledge occupies a central position (Schreyögg & Eberl, 1998, p. 519). An organization learns when the knowledge of its individual members is integrated into new organizational knowledge or through the recombination of existing knowledge (Kieser & Koch, 2008, p. 329). Especially, innovation and development projects demand the combination of a broad range of technical and functional knowledge to create new organizational knowledge in the form of new products or processes (Grant, 1996, p. 378). Emphasizing the significance of knowledge integration, Iansiti and Clark (1994, p. 557) propose that

> "[t]he capacity to integrate diverse knowledge bases is the foundation of knowledge building in an organization, and is therefore a critical driver of dynamic performance."

Leonard-Barton (1995, p. 64) suggests the locus of new knowledge to lie at the boundaries between specializations and their respective mindsets, implying the significant role of effective cross-boundary or cross-functional coordination for enabling the creation of competitive advantage (Carlile, 2004, p. 555). However, the creation of new organizational knowledge through integrating knowledge from diverse sources has been representing a core challenge for organizations ever since (Carlile & Rebentisch, 2003, p. 1181). Thus, the integration of individual-level knowledge into organizational knowledge represents one of the key problems dealt with in organizational learning research.

Knowledge comprises information on the one hand, and know-how on the other (Kogut & Zander, 1992, p. 386). Information, i.e. knowing what something means, includes facts, basic statements and symbols that are easy to transmit (J. S. Brown & Duguid, 1998, p. 91; Kogut & Zander, 1992, p. 386). Know-how is defined as "the accumulated skill or expertise which allows one to do something smoothly and efficiently" (Hippel, 1988; cf. Kogut & Zander, 1992, p. 386). In other words, know-how is the ability to convert information into required actions (J. S. Brown & Duguid, 1998, p. 95). Organizational knowledge may then be seen as "the sum of individual knowledge used in the value creation process and the knowledge embedded in collective action" (Schüppel, Müller-Stewens, & Gomez, 1998, p. 227). In the organizational learning literature there are various, sometimes contradicting ideas about how the specialized knowledge of individuals can be successfully integrated into organizational knowledge. The different concepts can be assigned to two perspectives, the "cross-learning" perspective and the specialization perspective (Grunwald, 2003, p. 3). The two views differ essentially in the extent to which they demand that knowledge has to be exchanged between organizational members to create a basis for integrating the knowledge of individual organizational members.

2.1.1. The cross-learning perspective

The cross-learning approach to integrating knowledge across domains of expertise suggests that for organizational learning to occur, knowledge needs to be transferred between specialists and combined in their individual brains (Kieser & Koch, 2008, p. 331). Among the first to propose this need for intensive cross-learning between specialists were Argyris and Schön (1978). The authors conceptualize organizational learning processes from a cognitive-theoretical perspective and assume that individual members of an organization develop to a certain extent common basic assumptions that manifest themselves in mostly unconscious organizational theories of action and are expressed in the actions of the organization. According to the authors, organizational learning means a change in collective theories of action. Argyris and Schön (1978, p. 17) argue that the pre-condition for successful knowledge recombination and creation is the development of shared "organizational maps", which is achieved through extensive knowledge transfer between individuals. Similarly, in their theory of organizational knowledge creation, Nonaka and Takeuchi (1995, p. 58) draw attention to the differences between the syntactic and semantic aspects (i.e. volume vs. meaning) of information. The authors emphasize that knowledge, like information, is about meaning, highlighting its context-specific and relational nature (Nonaka & Takeuchi, 1995, p. 58). Nonaka (1994, p. 24) emphasizes that through processes of socialization, teams need to form a "common base of understanding" or a "shared implicit perspective" to enable the continuous, time-consuming dialogues necessary for externalization, the critical learning process that involves transforming tacit into explicit knowledge that is accessible for the other members of an organization. In another approach, Kim (1993) proposes that individual and organizational learning are linked through mental models, defined as the thought constructs that influence individual and organizational actions. The integration of individual knowledge in organizational knowledge in this perspective is aided by shared mental models. According to Kim (1993), the vast majority of organizational knowledge (i.e. know-how and know-why) is stored in these shared mental models that are in turn deposited in the heads of the individual members of the organization (Kim, 1993, p. 44).

Research in the cross-learning paradigm concentrating on innovation in organizations especially points out the importance of a common base of understanding in product and process development, indicated by the holistic organizational learning concepts described above. Scholars in this field argue, that difficulties in knowledge transfer arise due to differences in the "thought worlds" of specialists (Dougherty, 1992, p. 182) and a lack of "common ground" among individuals of different departments (Bechky, 2003, p. 326). Research dealing with communities-of-practice points to the strong influence of a community's culture on knowledge sharing and underlines the connection of knowledge to the context in which it is learned (J. S. Brown & Duguid, 1991, p. 48; Lave & Wenger, 1991). It follows from this view that knowledge is assumed to be "largely tacit, situated, and experiential, and not easily articulated or codified" (Kellogg, Orlikowski, & Yates, 2006, p. 24). Dougherty (1992, p. 182) suggests that members of cross-functional teams representing different departments understand and interpret tasks or problems to be solved in different ways, due to their differing "thought worlds". Differences in the "fund of knowledge" (i.e. what is known) and the "system of meaning" (i.e. how do people know) of these thought worlds, lead to difficulties in sharing knowledge across departmental boundaries (Dougherty, 1992, p. 182). Similarly, Cronin and Weingart (2007, p. 763) argue that information sharing in functionally diverse teams might be impeded because of team members' differing problem representations stemming from the differences in the knowledge they hold. In creating di-

verging perceptions about how a problem should be tackled, these representational gaps complicate coordination between teammates (Cronin & Weingart, 2007, p. 762). Building on the work of Dougherty (1992), Boland and Tenkasi (1995, p. 358) argue that in order to integrate knowledge, organizational members need to engage in a process of perspective taking in which "the unique thought worlds of different communities of knowing are made visible and accessible to others". To overcome barriers to knowledge integration, cross-learning proponents basically suggest that individuals need to "identify, elaborate, and then explicitly confront the differences and dependencies across the knowledge boundaries" (Majchrzak et al., 2012, p. 951). Carlile (2002, 2004) distinguishes among three progressively complex types of knowledge boundaries that may occur between functionally diverse departments, presenting barriers to effective knowledge transfer: a syntactic or information-processing boundary that arises in the absence of a common lexicon, a semantic or interpretive boundary caused by inconsistencies in meanings, and a pragmatic or political boundary resulting from diverging interests. Carlile (2004, p. 556) argues that with increasing levels of difference, dependence and novelty of knowledge in a development project, the complexity of knowledge boundaries rises, posing new demands on the effective management of knowledge across a boundary.

A syntactic or information-processing boundary may cause a breakdown in knowledge transfer triggered by the absence of a shared syntax, i.e. a common lexicon, and the resulting mismatches in codes, routines, protocols, and other means of expression (Carlile, 2002, p. 443, Carlile, 2004, p. 558; Kellogg et al., 2006, p. 23). This boundary is theoretically grounded in the mathematical theory of communication (Shannon and Weaver (1949)) and the information-processing perspective on boundaries in organization theory (Galbraith, 1973; Lawrence & Lorsch, 1967), which basically assumes that knowledge is "external, explicit, and capable of being codified, captured, stored, retrieved, and transferred across people and contexts" (Kellogg et al., 2006, p. 24). This perspective holds that as soon as a syntax is shared among different parties, information can be processed, and knowledge can be transferred (Carlile, 2004, p. 558). Knowledge transfer in this view may be enabled by the development of information artifacts, such as standards, repositories, and specifications, that assist in communication across specialties under uncertainty (Kellogg et al., 2006, p. 23). A shared and stable syntax is suggested to be efficient, as specifications and agreements about knowledge differences and dependencies have been made in advance (Carlile, 2002, p. 453). Cross-specialty knowledge sharing in development projects may also be facilitated through actively involving individuals from multiple disciplines from start to finish (Lawrence & Lorsch, 1969; cf. Kellogg et al., 2006, p. 23)). Further, organizational members may take the roles of "technological gatekeepers" (Allen, 1970, p. 16) or "liaison-engineers" (K. B. Clark & Fujimoto, 1991, p. 103) installed to link different departments or communities with one another, thereby facilitating the flow of information and knowledge across boundaries.

A fundamental proposition of the information-processing view is that increasing levels of task uncertainty demand decision makers to process a higher amount of information. To be able to cope with the increasing information-processing demands, organizations need to adopt integrating mechanisms, like rules and programs, hierarchy, or goals, to increase their information-processing capabilities (Galbraith, 1974, pp. 28–29). Product development research in this tradition suggests the use of team-based structures, intensive communication, and information sharing true to the motto "more is better" to respond to uncertainty (Carlile, 2002, p. 444). However, some scholars critically note that information-processing or knowledge transfer frameworks may only offer satisfactory explanations in stable conditions that enable the development of a common lexicon or shared language between groups (e.g. Bechky, 2003, p. 313; Carlile & Rebentisch, 2003, p. 1182). For example, Carlile and Rebentisch (2003, p. 1182) argue that this simple view on knowledge transfer is not adequate to handle the complexity and ambiguity of knowledge integration activities inherent in contemporary organizations. Further, Bechky (2003, p. 313) claims that the universality of meaning and the homogeneity of context implied by simple knowledge transfer, do not mirror reality. Thus, Carlile (2004, p. 558) proposes that a syntactical boundary becomes a semantic boundary when novelty arises and the existing lexicon is no longer sufficient to clarify the newly emerging differences and dependencies.

Semantic or interpretive boundaries arise as a result of ambiguity in meaning (Carlile, 2004, p. 558). The interpretive or semantic approach to boundary spanning highlights that despite the presence of a common language, individual interpretations often differ, making communication and teamwork problematic (Carlile, 2002, p. 444). According to the interpretive perspective, just processing information is insufficient to overcome differences in meaning. Rather individuals need to learn about and understand the differences, dependencies and boundaries between each other's knowledge to enable knowledge sharing across a semantic barrier (Carlile, 2002, p. 444; Majchrzak et al., 2012, p. 952). Thus, scholars argue that overcoming semantic boundaries and enabling knowledge sharing necessitates the development of a "common meaning" (Carlile, 2004, p. 555).

To address the interpretive differences across boundaries and to arrive at a common meaning, Carlile (2004, p. 558) suggests knowledge translation as central practice. To do so, scholars stress the importance of cross-functional teams, colocation, and the utilization of shared practices (Carlile, 2004, p. 558). Further, insights from the communities-of-practice literature (J. S. Brown & Duguid, 1991; Lave & Wenger, 1991) imply that shared meanings can be developed through participation in similar activities (Carlile, 2004, p. 558). J. S. Brown and Duguid (1998, p. 103) also highlight the role of individuals as "organizational translators" and "knowledge brokers" in encouraging the movement of knowledge between different communities-of-practice. While translators are in place to "frame the interests of

one community in terms of another community's perspective" (p. 103) like mediators, brokers truly participate in several communities-of-practice allowing them to "broker" knowledge between the domains (J. S. Brown & Duguid, 1998, p. 103). For example, in an ethnographic study of the product design firm IDEO, Hargadon and Sutton (1997, p. 716) showed how the designers at the firm acted as technology brokers by connecting current design problems with existing solutions from other domains or industries to create new products. Boland and Tenkasi (1995, p. 362) emphasize the fundamental role of boundary objects for the process of perspective taking, as they enable individuals to bring their distinctive viewpoints into dialogue. Boundary objects can be physical objects like documents, protocols, concepts and prototypes, but also technologies or procedures that are jointly used by the communities but may be perceived of in distinct ways by the parties involved. These objects assist in clarifying commonalities and differences in the practices, attitudes, and world views of the differing knowledge domains and facilitate the development of a common understanding (J. S. Brown & Duguid, 1998, p. 104). However, in some cases novelty might generate differences in the interests of different actors, impeding their ability to share and assess knowledge (Carlile, 2004, p. 560). Under these circumstances, translating different meanings will not prove sufficient. Instead negotiating interests and transforming knowledge between actors become central processes to enable knowledge transfer, prompting the transition of a semantic to a pragmatic boundary (Carlile, 2004, p. 559).

As already implied, a pragmatic or political boundary concerns differences in interests, goals, practices, and other commonly held community-specific aspects of diverse knowledge domains (Carlile, 2004, p. 559). Research dealing with these pragmatic differences stresses the importance of understanding the consequences that differences and dependencies in knowledge across boundaries bring about (Carlile, 2002, p. 445). In the pragmatic view, knowledge is assumed to be localized, embedded, and invested in practice, recognizing the worth of knowledge and that it is "at stake" for those who created it (Carlile, 2004, p. 559). Carlile (2002, p. 445) emphasizes the complexity of overcoming a pragmatic boundary in outlining that

> "[t]he cross-boundary challenge is not just that communication is hard, but that to resolve the negative consequences by the individuals from each function they have to be willing to alter their own knowledge, but also be capable of influencing or transforming the knowledge used by the other function."

Thus, reducing differences in interests necessitates the joint transformation of existing domain-specific as well as common knowledge into new knowledge (Carlile, 2004, p. 559). However, the process of knowledge transformation is challenging, as it demands the actors involved to invest in relationship building and make trade-offs in practices, interests, and jurisdictions (Kellogg et al., 2006, p. 24). To support the negotiation of interests and the transformation of knowledge, scholars stress the importance of team structures and shared artefacts, methods, and practices that provide a common ground for sharing and assessing knowledge (Bechky, 2003, p. 326). While boundary objects may enable knowledge transfer across all three types of boundaries, they have been found to be particularly useful in overcoming pragmatic barriers in product development settings (Carlile, 2004, p. 559). According to Carlile (2002, p. 453) effective boundary objects need to: (1) establish a shared language, (2) provide individuals with a way to learn about the differences and dependencies of knowledge across boundaries, (3) enable processes of knowledge transformation in which individuals can alter their domain-specific knowledge into cooperatively created "common knowledge" (Carlile, 2004, p. 559). In studying misunderstandings between engineers, technicians, and assemblers on a production floor, Bechky (2003, p. 312) demonstrated how the cocreation of such "common ground" transformed individuals' perception of the product and the production process leading to a deeper understanding of the product and its problem areas.

However, the central assumption in cross-learning approaches that intensive knowledge exchange between actors is the prerequisite for integrating the diverse knowledge bases of specialists has been criticized by some authors. These scholars argue that the idea that specialists need to acquire the knowledge of other specialists to achieve integration would pose extraordinary challenges for the members of the organization in light of their limited cognitive capacities and therefore seems impractical (Kieser, 2001, p. 244). Thus, they cast doubt on the effectiveness of cross-learning as a mechanism to integrate knowledge, even if transfer is supported by boundary objects or boundary spanners (Kieser & Koch, 2008, p. 332). For example, Demsetz (1991) argues that the need for intensive learning between specialists in an organization would undermine the advantages of the division of labor.

> "Although knowledge can be learned more effectively in a specialized fashion, its use to achieve high living standards requires that a specialist somehow uses the knowledge of other specialists. This cannot be done only by learning what others know, for that would undermine gains from specialized learning." (Demsetz, 1991, p. 172)

Moreover, intensive cross-learning between specialists is associated with significantly high resource and time expenses and assumed to be an ineffective means of integrating diverse knowledge bases (Enberg et al., 2006, p. 145; Majchrzak et al., 2012, p. 951). Thus, scholars argue that there must be mechanisms "that bring about the recombination of knowledge but do not strongly depend on cross-learning and human cognitive abilities" (Kieser & Koch, 2008, p. 332). The approaches turned to next, show how such mechanisms may enable learning processes while maintaining specialization and do not require an in-depth knowledge transfer. Among

them are experience-based learning concepts in the tradition of the behavioral theory of the firm (March & Olsen, 1975) and the knowledge-based view (Grant, 1996).

2.1.2. The specialization perspective

The organizational learning concepts rooted in the behavioral theory of the firm (Cyert & March, 1963), assume that rules and routines are the starting point for and the outcome of learning processes (Kieser et al., 2001, p. 599). The behavioral theory of the firm (Cyert & March, 1963) portrays organizations as goal-oriented and rule-based systems that learn from experience. In this view, organizations learn by drawing conclusions from experience and appropriately adapting their standard operating rules to the changing environment (Cyert & March, 1992, p. 120). Grounded in these assumptions, March and Olsen (1975, pp. 148–150) propose a learning cycle with four ideal phases describing how experience-based learning transforms individual actions into organizational actions: (1) individual actions are based on individual perception and preferences, (2) these individual actions lead to organizational actions, (3) the organizational actions trigger certain environmental reactions, (4) which in turn influence individual perceptions and preferences. However, according to March and Olsen (1975, p. 158) certain barriers to learning may disrupt this cycle, thus making it incomplete. These are: (1) role-constrained experiential learning, in which individual learning is hampered by restrictive role-definitions and detailed standard operating procedures, (2) audience experiential learning, pointing to the limited influence of individuals in changing organizational actions, (3) superstitious experiential learning, which arises when organizational members misleadingly attribute a rule-based change in certain behaviors to a change in the environment, and (4) experiential learning under ambiguity, in which organizational members cannot draw clear conclusions from environmental reactions due to ambiguity.

A basic assumption in concepts rooted in the behavioral theory of the firm, is that organizations learn through experience and that learning changes their behavior. As opposed to the cross-learning approach, this view postulates that the outcomes of organizational learning are stored in artefacts, i.e. standard operating rules, and not in individual employees (Kieser et al., 2001, p. 599). In informing organizational members how to handle and process information, make decisions, and evaluate the results of decisions in particular situations, standard operating procedures serve as the organization's primary memory, enabling the transfer of past learning (Cyert & March, 1992, pp. 123–127). Accordingly, rules and standards incorporate solutions to organizational problems, thereby complementing specialization as a mechanism that enables cooperation of organizational members despite their limited rationality and cognitive capacity (Kieser et al., 2001, p. 600; Kieser & Koch, 2008, p. 331).

The knowledge-based view (KBV), proposes that the primary role of the firm is the integration of individuals' specialist knowledge to create organizational capabilities (Grant, 1996, p. 375). Knowledge is supposed to be heterogeneously distributed among individual specialists and functional departments in an organization, due to individuals' bounded rationality and limited cognitive capacity that restricts the infinite absorption of knowledge. A fundamental task is therefore the coordination of diverse specialists' efforts to economize on specialization and achieve organizational goals (Grant, 1996, p. 113). From the perspective of the knowledge-based view, the coordination and integration of the knowledge of different domains does not require extensive knowledge sharing (Grant, 1996, p. 114). Quite contrarily, Grant (1996, p. 114) claims that

> "[t]ransferring is not an efficient approach to integrating knowledge. If production requires the integration of many people's specialist knowledge, the key to efficiency is to achieve effective integration while minimizing knowledge transfer through cross-learning by organizational members."

In doing so, Grant (1996) formulates quite clear requirements for the mechanisms of knowledge integration, rejecting the notion of intensive cross-learning and instead emphasizing the reduction of knowledge transfer. According to Grant (1996, p. 114), there are four mechanisms that facilitate the integration of specialized knowledge into organizational knowledge: (1) rules and directives, (2) sequencing, (3) routines and (4) problem-solving and decision-making in groups. Rules and directives, involving plans, lists, forecasts, guidelines, and procedures, enable the conversion of implicit knowledge into explicit knowledge. They promote knowledge integration by coordinating the interactions between specialists through the use of common standards and significantly reduce the need for shared knowledge (Grant, 1996, p. 114). Sequencing encompasses the arrangement of activities in a chronological sequence, so that various specialists may complete their diverse subtasks independently (Grant, 1996, p. 115). The modularization of tasks and the accompanied reduction in communication and coordination efforts are particularly important for complex projects that demand for broad-scope knowledge integration (Grant, 1996, p. 381). Organizational routines also provide a mechanism for integrating knowledge. In generating routines, specialists develop sequential patterns of interaction, which allow for the integration of specialist knowledge in absence of verbal communication (Grant, 1996, p. 379). Group problem-solving and decision-making are more personal and communication-intensive forms of integration that are used for unusual, complex and important tasks, in situations in which rules, instructions and routines reach their limits (Grant, 1996, p. 115). According to Grant (1996, p. 115) knowledge integration should be carried out using the first three mechanisms, if possible, to avoid extensive knowledge transfer and the associated costs.

While Grant (1996, p. 114) argues, that knowledge transfer through cross-learning in organizations should be minimized, he also points out that for knowledge to be integrated, members of an organization need to dispose of a

certain level of common knowledge, representing the intersection of their individual knowledge bases (Grant, 1996, p. 115). It follows that, in innovation projects, organization members need at least a rough understanding of the overall process or product in order to better coordinate their work with others. Grant (1996, p. 116) outlines five types of common knowledge: (1) a common language, which allows an unambiguous understanding between departments and specialists at the organizational and project level, (2) other forms of symbolic communication that complement the verbal aspects of language, (3) a basic stock of common specialist knowledge, (4) shared interpretations that facilitate the interdisciplinary transfer of tacit knowledge, and (5) recognition of individual knowledge domains, involving knowledge about who knows what in the organization or project. However, how exactly individual knowledge integration comes about in the knowledge-based view proposed by Grant (1996) remains unclear, as his statements are rather abstract.

2.2. State of the art in empirical research

A growing body of empirical work on knowledge integration in general and in innovation projects in particular indicates that deep-level knowledge sharing as demanded by cross-learning approaches is not always necessary to enable knowledge integration (Majchrzak et al., 2012, p. 954). Already, Donnellon, Gray, and Bougon (1986, p. 43) showed that organized action among organizational members can happen in absence of shared interpretations. The authors observed that the creation of "equifinal meaning" through communication enabled coordinated action. Further evidence comes from Galison (2000, p. 46), who found that physicists from different subcultures (i.e. theorists, experimentalists, and engineers) advancing divergent viewpoints, were able to align their activities without the development of shared interpretations, identities and interests. Galison (2000, p. 46) uses the metaphor of a "trading zone" to describe an "intermediate domain" in which activities may be coordinated locally, even if the broader meanings of different subcultures collide.

Similarly, studies on R&D projects and product development imply that extensive cross-learning is not a precondition for successful knowledge integration and innovative performance (e.g. Enberg et al., 2006; Faraj & Xiao, 2006; Majchrzak et al., 2012; Schmickl & Kieser, 2008). For example, in their study of an interactive marketing organization, Kellogg et al. (2006, p. 40) found that coordination within projects can be achieved without building "deep commitments to shared meanings or transformed knowledge". Instead, project team members enacted "trading zones" (Galison, 2000) by agreeing on the general procedures of knowledge exchange, i.e. the technology-based coordination practices, which allowed them to interact across boundaries even if their local understandings of a task diverged (Kellogg et al., 2006, p. 42). Members of different communities in the project integrated their diverse knowledge by sharing their contributions in a "common digital space", in the form of e-mails, PowerPoint or Word documents, thereby making their work mutually visible and legible, which allowed for ongoing adjustments and alignment of the diverse contributions (Kellogg et al., 2006, p. 40). Similarly, Enberg (2012) found that in coopetitive R&D projects knowledge integration was built on a shared understanding of the process and not the content of project work. Planning and process specifications, and presentation genres were found to enable knowledge integration. While these mechanisms aided the processes for integrating knowledge, they simultaneously constrained the breadth of knowledge to be exchanged and provided a structure for collaboration and face-to-face discussions and thus reduced knowledge transfer among specialists (Enberg (2012, p. 771)). In studying successful innovation teams in an electrotechnical company, Schmickl and Kieser (2008, p. 487) also only found limited evidence for deep-knowledge sharing. Rather, team members exchanged "broad" and "rough" knowledge to integrate their diverse viewpoints. The exchange of "narrow" and "detailed" knowledge happened selectively on occasion through a kind of question and answer game (Schmickl & Kieser, 2008, p. 487). Schmickl and Kieser (2008, p. 488) identified modularization, prototyping and transactive memory as integration mechanisms that, in combination, substitute for cross-learning in product development. Modularization refers to breaking down products or processes into simpler components and may reduce knowledge transfer by allowing specialists to work independently on their components to a certain extent (Schmickl & Kieser, 2008, p. 476). Prototyping may be defined as iterative trial and error process, in which team members engage to align their diverse efforts. Knowledge transfer may be reduced, because prototyping directs attention to identifying and explaining problems between components, thus narrowing the scope of knowledge that needs to be shared. Transactive memory, i.e. knowledge about who knows what, helps team members to readily localize specialists with relevant knowledge, thus curtailing search processes and reducing the need for knowledge transfer (Schmickl & Kieser, 2008, p. 477). Likewise, Enberg et al. (2006) showed how the reliance on certain mechanisms and practices allowed team members in a development project in the manufacturing industry to work on their tasks in isolation "without a lot of effort devoted to knowledge sharing" (Enberg et al., 2006, p. 158). Knowledge integration in the stacker project under study was achieved through ad hoc problem-solving, experience rather than knowledge sharing, and "individuals' idiosyncratic representations of the stacker artefact" (Enberg et al., 2006, p. 158). In this regard, Enberg et al. (2006, p. 157) highlight the complementary role of formal meetings in achieving coordination among functionally diverse team members. In the project they studied, regular meetings strengthened the commitment of team members to integrate their diverse knowledge bases by aiding the development of a common goal and a general sense of being a team. Furthermore, the focus on experience sharing, rather than knowledge exchange, in the time-limited meetings aligned team members' mutual expec-

tations of task completion, thereby facilitating coordinated action (Enberg et al., 2006, p. 157). Similarly, Majchrzak et al. (2012) showed how brainstorming workshops, strategic planning meetings, and other face-to-face meetings provided an arena for the members of cross-functional teams to integrate their diverse knowledge bases. The team members engaged in five dialogic practices that allowed them to cocreate a solution without emphasizing the differences among their knowledge domains despite task novelty (Majchrzak et al., 2012, p. 958).

Thus far, contemporary empirical research seems to agree that knowledge integration may involve merely mentioning and demonstrating knowledge, rather than a "deep" sharing of interpretations and interests (Mengis et al., 2018, p. 597). However, extant research also suggests that the intensity of knowledge sharing in teams depends on the novelty and uncertainty they face. For example, Schmickl and Kieser (2008, p. 485) found that the more radical the innovation, the more knowledge was exchanged between project members (Schmickl & Kieser, 2008, p. 485). Furthermore, research that frames knowledge integration as a process rather than an outcome implies that the process of integrating knowledge may change over time (Mengis et al., 2018, p. 598). For example, Bruns (2013, p. 62) discovered that coordination across domains of expertise in a group of scientists was achieved through switching between the collaborative practices of counterprojection and alignment when the scientists were working alone, and the more dialogic coordination practices of joint assessment and consultation when they worked together. In the trauma care context, Faraj and Xiao (2006, pp. 1164–1165) observed that as unexpected events occurred, trauma teams sometimes had to abandon established treatment protocols and relied more on dialogic coordination, involving joint sensemaking, to arrive at a new collective understanding of the patient. Similarly, Mengis et al. (2018) suggest that knowledge integration is not a consistent process, but rather requires alternations between different coordination practices over time. In their study of collaboration efforts of scientists involved in the development of a highly novel bioreactor, they found that particularly in cases of "epistemic breakdowns", i.e. disruptive events that erode extant understandings of a problem, scientists changed their coordination practices. When such breakdowns occur, the scientists changed from a coordination mode of "working together alone" (Bruns, 2013; Enberg et al., 2006) to a more intensely collaborative one that focused on dialogically drawing distinctions to expand collective knowledge. As soon as the scientists had drawn new distinctions they established a revised division of labor that again shifted their focus towards coordinating their work, albeit in a more political fashion centered on the alignment of expectations and obligations (Mengis et al., 2018, p. 607).

Taken together, the presented findings imply that extensive deep-level knowledge sharing is not key to successful knowledge integration, thereby contradicting the central conjecture of the cross-learning approach. Rather, extant empirical work on knowledge integration seems to comply with the specialization perspective, in particular the knowledge-based view, suggesting that certain practices and structural mechanisms reduce the need for knowledge sharing. However, recent studies on product development projects suggest that the need for and the engaging in knowledge sharing are dependent on the level of task novelty or innovativeness of the product or project (e.g. Schmickl & Kieser, 2008). Moreover, studies dealing with the collaboration of diversely specialized scientists in highly novel and uncertain settings, point to the processual nature of knowledge integration that implies the variation of coordination and collaboration practices over time. While most of the studies on cross-functional teams in new product (e.g. Schmickl & Kieser, 2008) or new process development (e.g. Majchrzak et al., 2012) consider the overall level of uncertainty and its impact on knowledge integration, they do not show how knowledge integration mechanisms and practices change over time within the course of a project. For example, the concept phase in new product development may raise other coordination demands than the implementation phase in both radical and incremental innovations. Furthermore, unexpected problems or changes in customer demand may trigger the emergence of coordination mechanisms other than those currently in place (Grant, 1996, p. 113). Finally, despite prior research implied the need for of a multi-level perspective on knowledge integration (Enberg et al., 2006), the interplay of individual-level and group-level knowledge integration practices as well as the changes in the primary locus of knowledge integration within the course of a project have not yet received adequate attention in empirical research on knowledge integration in new product development teams.

3. Methods

3.1. Research design

The starting point for this master's thesis was my curiosity about the ubiquitous organizational agility hype that has been spreading across various industries and affecting project work in numerous areas of application. An initial search of the EBSCOhost database "Business Source Premier" revealed a lack of scientific research on agile methods in the leading management journals, and a general scarcity of studies outside the software development context. As mentioned in the introduction, popular scientific papers and books often associate the agile methodology with increased team performance and velocity and attribute a superiority over classic project management methods to the agile way of working. However, empirical support for these claims is missing, which was the decisive factor to focus my study on the phenomenon of (successful) agile teamwork in product development outside of the software industry. A suitable manufacturing company, which allowed the investigation of two of their agile development teams, was found through my personal network. A first meeting with CEO, Head of Innovation and Head of HR was held in July 2019 to clarify mutual expectations and to agree on the anticipated result of the master's thesis. As

a result, the initial phase of the research project was guided by the following question: "What are the success factors of agile teamwork in product development projects in the manufacturing industry?".

Collaboration and coordination in cross-functional or agile teams are complex phenomena. Therefore, researchers need to undertake a deep dive into individuals' attitudes, motives, and perceptions to be able to gain an understanding about the processes lying beneath individual and group behavior that are associated with the creation of new products and services. Qualitative methods offer the opportunity to explore the complexity of human behavior by allowing respondents to share their experiences, perceptions, and feelings about a certain topic with the researcher in their own words (Berg, 2001, p. 7). Furthermore, the lack of research on the success factors of agile teamwork calls for an inductive, interpretive approach that enables the elucidation of factors that affect the realization of agile product development projects and that have not been previously discovered. Thus, a grounded theory methodology (Glaser & Strauss, 1967) was applied. More precisely, the research design was set up adhering to the principles of the Gioia Methodology (Gioia et al., 2013).

The grounded theory method was introduced by Glaser and Strauss (1967) in their book "The Discovery of Grounded Theory". The methodology encompasses an inductive and comparative approach to research with the ultimate aim of building theory embedded in empirical data. Within this approach, data collection and analysis happen simultaneously rather than sequentially, thereby mutually influencing each other. Throughout the project, researchers iteratively process back and forth between empirical data and the emerging analysis (Bryant & Charmaz, 2007, p. 1). As opposed to hypothesis-based methods, the starting point for grounded theory research resides in empirical data and not in the extant literature. Thus, researchers should not start by formulating hypotheses to be tested, but rather by gathering data relevant to a phenomenon of interest that form the foundation for developing the framework of the research project (Bryant & Charmaz, 2007, p. 126). By applying theoretical sampling, i.e. the selection of a case or a sample against the background of theoretical considerations, data collection gets refined and more focused over time, which allows for an increasingly focused theoretical analysis (Bryant & Charmaz, 2007, p. 1; Flick, 2009, p. 432). Data is collected and analyzed until theoretical saturation is reached, which means that no new insights can be obtained with additional data (Flick, 2009, p. 436). Features like the proximity to the data or the empirical referents, the structure provided for analyzing complex social phenomena, and the ability of developing meaningful practical theories have been contributing to the popularity of the grounded theory methodology among researchers (Bryman, 2000, p. 84). The approach has been offering profound contextual theoretical accounts of organizational phenomena (Gioia et al., 2013, p. 16). However, despite its advantages the approach is not free of criticism. Some scholars claim that inductive methods are inappropriate to achieve the high criteria that apply to validating scientific progress (Gioia et al., 2013, p. 17). Accordingly, qualitative researchers are regularly accused of reluctance to incorporate theoretical aspects into their studies (Bryman, 2000, p. 85). Furthermore, dealing with the flood of data associated with grounded theory methods, disregarding extant theories relating to the topic of interest, and the selection of a research site constitute severe practical challenges for this kind of investigation (Bryman, 2000, p. 87).

To counter the allegations of inappropriateness and to increase practicability of grounded theory, Gioia and his colleagues (S. M. Clark, Gioia, Ketchen, & Thomas, 2010; Corley & Gioia, 2004; Gioia & Chittipeddi, 1991; Gioia, Price, Hamilton, & Thomas, 2010; Harrison & Corley, 2011; Nag, Corley, & Gioia, 2007; Nag & Gioia, 2012) have been developing and successively refining a holistic approach to inductive concept development, which is intended to ensure highly qualitative scientific rigor, and which is becoming increasingly popular among organizational researchers. The methodology is centered around two basic assumptions: (1) the organizational world is socially constructed, and (2) organization members are "knowledgeable agents", meaning that they have an idea of what is going on around them and can provide qualified information about their thoughts, intentions and actions. Accordingly, this approach attaches great importance to the interpretations of the informants right from the start of data collection and analysis, creating fruitful opportunities for concept discovery rather than confirmation. Furthermore, the method's developers assume that scientists are also well-informed and therefore able to recognize patterns, relationships, and emerging concepts in collected data and place them in a theoretical context (Gioia et al., 2013, p. 17). Gioia et al. (2013, p. 16) emphasize that, in their view, "concepts are precursors to constructs in making sense of organizational worlds", by which they mean that for the purpose of theory building, researchers first of all need to discover relevant concepts that can subsequently direct the formation and validation of theoretical constructs (Gioia et al., 2013, p. 16). While research design and data collection within the Gioia Methodology are similar to traditional grounded theory approaches, data analysis and the articulation of grounded theory are carried out distinctively (Gioia et al., 2013, p. 26). In the following, I will first shortly describe the research setting before providing detailed accounts of my approach to data collection and analysis.

3.2. The case of A-Machining

As already implied, for my research I got access to two agile product development teams of an Austrian company, which produces and sells special machinery and associated spare parts. For the remainder of this thesis, I will call the company A-Machining. In the fiscal year 2018/19 the company had sales of 382 million Euro, 90 percent of which were from abroad. Currently the company employs approximately 1,900 people worldwide. Founded in the 19th century, the family-owned company can look back on a long success story.

For many years, A-Machining has been holding global market leadership in one of its segments and is recognized as innovation leader. However, through this rather stable position, negative consequences have evolved over time. Silos had formed around the individual departments within the company, which made it difficult to exchange information with other teams and areas of expertise. This negative effect is reinforced by the spatial distance between some departments. For example, while the mechanical engineers work in the main building, the departments of mechatronics and testing are situated in another building across the street. Three years ago, the culture of conversation was shaped by disputes about resources and self-serving behavior affecting innovative capacity and efficiency. In order to defend its global market position and to cope with the dynamic corporate environment in the future, it was time for the company to change the culture of collaboration. Therefore, A-Machining decided to restructure internal collaboration by tearing down the departmental silos and improving communication with the ultimate aim to increase development efficiency and shorten time-to-market. In November 2017 the machine constructor introduced AGILE as the new method for structuring its new product development projects. Good leadership with clear goals, autonomy and feedback, as well as good collaboration are the key characteristics of AGILE for the company. With the new method A-Machining aimed at overcoming department-centered thinking, motivating employees, coping with complexity and being able to plan the unknown. Accompanied by a group of consultants, two pilot projects were set up with the new structure and way of working. In advance, the employees were informed about the agile way of working in a day event. By the end of 2019 the company had ten agile projects with approximately 100 employees involved.

In the following, I want to give an overview of the key components and ceremonies of agile product development, as introduced by A-Machining, to offer a richer understanding of the comprehensive structure of such projects. The descriptions are derived from the book of Schröder (2017, pp. 44–77), which the company used as a guide for implementation. In addition, data from conversations and a PowerPoint document provided by the company, including the main building blocks of agile product development, were used to cover the adaptions of the methods in the company.

An agile team needs to cover the following three roles: (1) The Product Owner Team (POT), (2) the Work Team, and (3) the Agile Coach. The Product Owner Team comprises representatives from technology, the market, project management and production. The POT is in charge of the project plan. Their main duty lies in defining what has to be done and creating, maintaining and prioritizing the Product Backlog, i.e. the target specifications of the product. They are responsible for the economic success of the product and represent the interests of the customers in the team. The members of the Work Team originate from different departments of the company, such as mechanical engineering, test engineering, mechatronics, controlling and purchase. It is up to them to decide how to meet the requirements defined in the Product Backlog. They are responsible for the timely delivery and the quality of the product. The Agile Coach ensures compliance with the agile principles and procedures and moderates all agile ceremonies. Also, he or she shall protect the team from disruptions and remove any appearing obstacles and barriers.

In agile project management, the overall project plan gets divided into stages of three months – the closer the stage, the more detailed the plan. The stages in turn are divided into six sprints of 14 days each. This differentiation between rough and detailed planning allows the POT to deal flexibly with market changes and other uncertainties. Within one sprint the following agile ceremonies take place:

The Conclave – A new sprint starts with a planning meeting of the POT members. It is called Conclave, because the POT members need to agree unanimously on the plan for the next sprint. This meeting takes place in a 14-day-rhythm at the same time and at the same place. Its duration varies between 30 and 90 minutes depending on the phase of the project and the experience of the POT. A POT-Agile-Board, that typically consists of the four columns (1) Backlog (i.e. ToDo), (2) Work in Progress, (3) Done, and (4) Definition of Done, i.e. the specifications of a result and the prioritization of the desired outputs, helps the POT to structure and visualize the sprint plan. First, each member of the POT writes down the desired outcomes for the upcoming sprint on blue Post-Its and places them in the Backlog column. Importantly, the outcomes shall be measurable and precisely formulated and the descriptions shall concentrate on what to do and not how to do it. As the board shall be clearly arranged and easily comprehensible for the Work Team, the number of items should not exceed 15. Second, the representatives of technology, market, production and project management need to agree on a prioritization of the Backlog and arrange the items on the list according to their importance. This process is moderated by the Agile Coach, who closes the meeting by asking whether all participants believe that this Backlog will lead to a successful project. If the answer is yes, then the next ceremony can start.

The Sprint Planning – Directly after the Conclave, POT and Work Team meet to discuss the prepared Backlog list displayed on the POT-Agile-Board. Again, this meeting is moderated by the Agile Coach. In a first phase, the POT members present their expectations, and the Work Team members have the possibility to ask questions, resolve uncertainties or ambiguities, point to missing items and express doubts. The team members estimate time and effort needed to fulfil the requirements of every single item. If necessary, the blue Post-Its of the POT are corrected and changed or new items are added to the list. As a result, the original Product Backlog list is adapted to display and integrate the perspectives of both POT and Work Team. In a second phase, the Agile Coach asks the Work Team whether this workload can realistically be accomplished within two weeks. As the team members are usually involved in more than one project, they need to check their capacities for the upcoming weeks. In

the third phase, the Work Team communicates what is possible against the background of the updated Sprint Backlog and the team's capacity. By drawing a horizontal line with a yellow fine-liner on the POT-Agile-Board, the team indicates how many items can be achieved at the given capacity. The POT can react to this feedback by, for example, increasing capacities, moving items to the next sprint, or changing prioritization. The fourth phase is about confirming the final agreement on the Sprint Backlog. In a thumbs-up-ritual every member of POT and Work Team shows his or her commitment to the plan. Now, the POT leaves the meeting and in the fifth and final phase, the Work Team has time to work out how to best achieve the goals on the Sprint Backlog they committed themselves to. By using a Team-Agile-Board, they break down the requirements on the blue Post-Its into smaller activities and write them down on yellow Post-Its.

The Daily-Stand-up-Meeting - In this short daily team meeting in front of the Team-Agile-Board the members of the Work Team discuss barriers and challenges and try to solve any pending problems. Meeting duration is exactly 15 minutes to ensure efficiency. Devices, like an agile-clock displaying the laps of 15 minutes or a sandglass for managing the equally distributed time allotted for speaking, can be used to ensure compliance with the strict time constraint. If team members recognize a need to go deeper into a particular topic, they arrange informal, bilateral meetings outside the formal meeting structure wherein only those people, who are required for problem-solving, are involved. Especially in the introductory phase of agile project management, it can be useful to have the Agile Coach as a moderator in the Dailies. However, his or her presence is optional.

The Sprint Review (also DEMO, for demonstration) – After 14 days POT and Work Team meet for the Sprint Review. Depending on the project phase the review may last between one and two hours. In this meeting the team members present the results of the sprint to the POT for the first time. Thus, the moderation of the Agile Coach is an important factor. Visualization is key for an effective presentation. Therefore, a DEMO can also take place at a testing or production site including, for example, a prototype. However, usually the results are jointly prepared in a PowerPoint presentation with the use of pictures and descriptive language. A crucial part here is that the POT takes on its leadership function by showing real appreciation for the work done, thereby motivating the team. The results are either accepted by the POT and put in the "Done" column of the Sprint Board or remain on the board for revision in the next Conclave and Sprint Planning sessions.

The Retrospective (also RETRO) (30 min) – Directly after the Sprint Review, the team reflects on the last sprint in terms of teamwork, agile way of working and constraints in the project. This feedback session may be conceived as a regular lessons-learned workshop and represents an important means for resolving conflicts and continuous process improvements. Depending on the number of topics, this meeting may last about 30 minutes. However, it emerged from the interviews that the RETRO is hardly used. While the Work Team is busy with the Retrospective, the POT leaves for the Conclave, thereby starting the new sprint.

In the following paragraphs, the two pilot projects, which I re-named Project Colossus and Project Homestretch, will be shortly described. As of October 2019, Project Colossus had been running for five years. The project revolves around the fundamental revision of the machine manufacturer's flagship product, in whose segment the company is global market leader. It is a demanding and unusually large project for A-Machining, in which many different parts have to be precisely coordinated. Thus, in many ways, the company is breaking new ground with this project. The overarching aim of Project Colossus is to set an example with the new model. It should be innovative, differentiate itself from the competition with clear USPs, and strengthen the company's role as technology leader, which is associated with high pressure for success for the development team. Two years earlier, however, a direct benchmark test with the toughest rival in the product segment revealed that the product in development by far cannot compete with the competition. That was a devastating day for the development team. The market requirements for the product had changed and the product, as it was at that time, was not able to meet those expectations. As a consequence, top management decided to stop the project. To make a clear cut, all the prototypes were scrapped. In November 2017, Project Colossus was relaunched with the agile product development methodology. The team now comprises 23 members, including one Agile Coach and a Product Owner Team of four. The functional departments involved in the Work Team are mechanical engineering, mechatronics, testing, purchase, controlling, quality, and production. After an initial orientation phase, the team members soon accepted the new methods and exemplarily exercised self-organization. Thus, concerning compliance with the agile principles and methods, there was no need for the Agile Coach to intervene a lot. The development team started from scratch and had the opportunity to build the new product on the "green field". Notwithstanding the big change, motivation was high, and the team members had many innovative ideas, including real novelties in the market. The development time starting with the first line drawn in the concept phase and ending with the assembly of the first prototype was seven months – extraordinarily short for A-Machining. However, despite the good restart, the project team was faced with problems concerning the manufacturing costs of the product at the time the interviews were conducted. Consequently, many parts of the product would need to be reworked or sourced from other suppliers to be able to meet market prices. So far, three prototypes have been built in the course of the project and two more are planned.

Project Homestretch is about a program expansion in another product segment. The main requirement of this project is to develop a bigger version of a special machine including the corrections of reclamations in the existing product line and the integration of some additional features that mirror customer demands. In contrast to Project Colossus, the pressure to succeed for the project team and the desired

level of innovativeness is as high as in average projects at A-Machining. Project Homestretch comprises a rather big development team, with 17 team members in the beginning and 10 members in the final phase. Similar to Project Colossus, the functional departments involved in the Work Team are mechanical engineering, mechatronics, testing, purchase, controlling, quality, and production. The project had been running for one year when it was reorganized with the agile product development methodology. At that time, the concept phase was already completed, and the focus shifted to making the product ready for production, narrowing the corridor for subsequent changes. The introduction of the agile methodology was challenging, as the team members reacted rather cautiously to the considerable transformation of collaboration in the product development process. Especially in the initial phase, many interventions of the Agile Coach were necessary to ensure that the agile principles and methods were applied correctly and thoroughly by the team. However, within the course of the project no bigger problems appeared. At the time of the interviews, the project was nearly completed and therefore the meeting rhythm in the agile methodology was reduced.

3.3. Data collection

Especially through interviews researchers can access peoples' perceptions of certain events or experiences and can comprehend, how individuals make sense of their social world (Berg, 2001, p. 72). Semi-structured interviews form the core of the data collection strategy applied in studies using the Gioia Methodology (Gioia et al., 2013, p. 19). The degree of structure is often used as a means to classify research interviews in the literature (Qu & Dumay, 2011, p. 244). Most commonly, the following three types are distinguished: (1) structured, (2) unstructured and (3) semi-structured interviews. In a structured interview the questions and their sequence are predefined, ensuring consistency across all the interviews. Deviations from the interview script shall be avoided, which makes the process inflexible (Qu & Dumay, 2011, p. 244). Researchers, who utilize this technique, have already concrete ideas of what they want to find in the interviews (Berg, 2001, p. 69). Using this interviewing approach, researchers seek to minimize researcher bias and increase generalizability of their findings (Qu & Dumay, 2011, p. 244). A disadvantage is that structured interviews do not utilize the dialogical potentials for knowledge production that conversations hold and therefore only represent a passive picture of individuals' opinions and attitudes (Brinkmann, 2014, p. 286). Thus, structured interviews were considered inappropriate for this research project. The counterpart to the structured interview described above, are unstructured interviews. Researchers who apply this method, do not use a predefined set of questions but rather intuitively develop questions appropriate to the interviewee's reactions and statements in the interview situation (Berg, 2001, pp. 69–70), while at the same time keeping the central purpose of the research in mind (Qu & Dumay, 2011, p. 245). The role of the interviewer is to actively listen and not interrupt (Brinkmann, 2014, p. 286). One underlying assumption of this interview type is that the necessary questions to gather information on the research topic are unknown at the beginning, thus making a list of questions obsolete (Berg, 2001, p. 70). The aim of unstructured interviewing should be to gain insights into the individual perspectives of the interviewee (Qu & Dumay, 2011, p. 245). A potential disadvantage of this method is that the course of the interview is rather controlled by the interviewee than the interviewer and therefore issues deemed to be important by the researcher may be left out (Brinkmann, 2014, p. 286). As I had already an idea of what I wanted to know from the interviewees, the unstructured interview was not appropriate for data collection as well.

The semi-structured interview is the most common data collection method in qualitative research (Qu & Dumay, 2011, p. 246) and combines the predetermined questions and topics of the structured interview with the intuitive ad-hoc generation of questions of the unstructured interview (Berg, 2001, p. 70). Kvale and Brinkmann (2009, p. 3) define the semi-structured life world interview "... as an interview with the purpose of obtaining descriptions of the life world of the interviewee in order to interpret the meaning of the described phenomena". This definition highlights that interviews have the purpose of knowledge generation. An interview guideline, which is structured into themes, helps to conduct the interview in a systematic manner but also allows the interviewer to dig deeper and ask the respondent to elaborate in more detail on statements or topics of interest (Qu & Dumay, 2011, p. 246). Underlying assumptions of this approach are that questions should be formulated with the vocabulary of the interview subject and researchers should "approach the world from the subject's perspective" (Berg, 2001, p. 70). An important advantage is that hidden facets of individual or organizational behavior can be elucidated. Especially, when the aim of a study is to gain insights into individuals' perceptions of the social world, the semi-structured interview is an appropriate and effective method, as respondents have the possibility to answer in their own terms (Qu & Dumay, 2011, p. 246). As opposed to structured interviews, semi-structured interviews can make better use of dialogues to produce knowledge, as immediately following up on the interviewees' statements is allowed and encouraged, which provides the interviewer with the opportunity to participate in generating knowledge (Brinkmann, 2014, p. 286). An advantage over unstructured interviews is that the interviewer has more control over the thematic direction of the interview and can ensure that the focus remains on issues that are perceived to be important for the research project (Brinkmann, 2014, p. 286). However, semi-structured interviews need to be carefully planned and require the interviewer to be well trained in asking the right questions and correctly interpreting the answers of the respondents in the interview situation (Qu & Dumay, 2011, p. 247).

The flexibility of the semi-structured interview method as well as its suitability for examining attitudes, perceptions and

hidden facets of individual behavior convinced me to apply this interview type to examine the group processes in agile product development projects. Guided by the initial, rather general research question of the success factors of agile teamwork in product development projects in the manufacturing industry, I first familiarized myself with the basic pillars of agile product development, as practiced in the case company. As a next step, I reviewed the literature on new product development and cross-functional teams for relevant theories and empirical studies that aim at explaining team performance and new product success. The results of this review were sorted by topic and summarized in a Word document in order to formulate questions for the interview guideline, which were adapted to the context of agile product development.

The data were collected by conducting 14 semi-standardized interviews using an interview guideline with members of the two development teams of Project Colossus and Project Homestretch. The interview guidelines were adapted according to the roles of the interviewees. Therefore, there was a slight difference in the questions asked to members of the Product Owner Team (POT), the Agile Coach, and members of the Work Team. I started with interviewing two POT members and the Agile Coach of Project Colossus, next I spoke with three POT members and the Agile Coach of Project Homestretch. One and two weeks later, the Work Team members of the respective projects got their turn. I interviewed three members of Project Homestretch and four of Project Colossus. After the first round of interviews with each role, I further refined the interview guideline by adding questions about some significant events or topics mentioned by the first interviewees. Furthermore, after a first round of analysis I put increased emphasis on instances of decision-making and problem-solving as well as on disruptive factors in cross-functional coordination and collaboration. Nonetheless, every interview focused on the implementation of the agile methodology in the company and the course of the development project since then. The aspects covered in the interview guideline were the respondents expectations towards the new way of working, perceived changes in comparison to the prior project management method, challenges and successes in the project in general and with the agile methodology in particular, team cohesion and quality of collaboration, conflicts, (group) decision-making, responsibility, transparency, time and resource management, and autonomy. The interviews offered a good insight into the perceptions and attitudes of the interviewees towards agile product development and enabled me to grasp an understanding of what the pressing topics are in collaboration and coordination of cross-functional teams in new product development. As suggested by Berg (2001, p. 77), the language level of the respondents was taken into account in preparing the interviews and the interview guideline was tested and revised in advance. To ensure that the interviewees are able to easily comprehend my questions and to prevent misunderstandings and ambiguity, the interviews were carried out in German, the mother language of both interviewees and interviewer. All interviews took place within a time span of 15 days in October 2019 in a meeting room at a high table in the interviewees' company. On average the interviews took 45 minutes, the shortest lasting 20 minutes and the longest 60 minutes. I never conducted more than four interviews a day and made sure to make notes about particularly interesting or outstanding statements during the interviews. In addition, the interviews were recorded with a smartphone and fully transcribed usually on the same day. The interviewees were guaranteed anonymity in order to create a safe interview environment in which they can express themselves freely. For this reason, the transcripts have been anonymized.

For the purpose of data triangulation, other sources of relevant information, in the form of company documents (i.e. presentations on the introduction of agile product development, earlier interviews conducted by the company, and details of the projects) and notes of informal conversations at the company, were considered too. Furthermore, I used the notes I made during the interviews and memos on theoretical considerations and ideas I wrote during the initial phase of analysis to enrich my data.

3.4. Data analysis

Data analysis, i.e. coding the data, is considered the central process in grounded theory studies (Flick, 2009, p. 435). As already mentioned, data collection and analysis are carried out simultaneously within this type of research. The overarching objective of data analysis in the grounded theory methodology is theory development. Basically, coding in such studies has two sub-goals: (1) developing an understanding of the phenomenon under study, which requires an open-minded approach to analyzing the data, and (2) identifying an underlying structure, process or core category (Flick, 2009, p. 436). Strauss and Corbin (1998) differentiate between three procedures that may be used to work with texts in their approach to the grounded theory methodology: open, axial and selective coding. They understand coding as the central process behind theory building, in which data are ruptured, abstracted, and reassembled in novel ways (Flick, 2009, p. 307; Strauss & Corbin, 1998, p. 3). These procedures do not have a fixed sequence but usually start with open coding and get more abstract as data gathering and analysis proceed (Flick, 2009, p. 307). Open coding includes the identification and development of concepts and results in a list of codes and categories close to the data (Flick, 2009, p. 310; Strauss & Corbin, 1998, p. 74). With axial coding researchers aim at identifying relationships among the open codes and categories using the paradigm model, a general model that depicts the relations between a phenomenon, its causes and consequences, the context and the strategies of the people who are involved (Flick, 2009, p. 311; Strauss & Corbin, 1998, p. 114). Selective coding extends axial coding to a higher level of abstraction and involves the identification of a core concept or variable that relates to all categories, i.e. "the story of the case" (Flick, 2009, p. 312; Strauss & Corbin, 1998, p. 131).

Gioia et al. (2013, p. 20) distinguish between a first-order and a second-order analysis. Within the former, they check the data for relevant terms, codes, and categories. Thereby they try to stick closely to the informant terms, which leads to a huge number of categories already after the first ten or so interviews. This step can be compared to the process of open coding as described by Strauss and Corbin (1998). Accordingly, I began my analysis with sentence by sentence in-vivo coding of the first seven interview transcripts before conducting further interviews. In this first analysis I extracted interesting and relevant interview passages one-to-one and inserted them into an Excel sheet. In line with Gioia et al. (2013, p. 20), I went through these extracts, i.e. the in-vivo codes, again and tried to uncover similarities and differences among them. An analysis step similar to axial coding described by Strauss and Corbin (1998). To get an overview, I assigned a color to statements with a similar meaning and arranged them in my Excel sheet. Next, I tried to formulate first order concepts (Gioia et al., 2013, p. 20) to reduce the number of categories that emerged from the initial analysis. In their second-order analysis, Gioia et al. (2013, p. 20) move to the theoretical realm and seek to uncover themes and dimensions which they compound to reveal the larger story that explains what is going on in the data. In doing so, they put emphasis on concepts that seem to lack theoretical elaboration in the extant literature. So, as a next step I revisited the first-order concepts from a theoretical point of view and tried to derive a deeper structure that might offer an explanation for the way people work together in the agile teams under study. Simultaneously, I again consulted the literature on new product development and cross-functional teams to check whether there are extant theories that might explain the phenomena I observed during the interviews. Based on the interim results of my analysis, I familiarized myself more deeply with the coordination literature and found that there is a lack of research that integrates formal and informal or contextualized and emergent coordination mechanisms in cross-functional teams (Okhuysen & Bechky, 2009). As agile product development is tightly structured and thus naturally incorporates diverse formal coordination mechanisms, I wanted to find out more about informal mechanisms and the interplay of formal and informal mechanisms. As already mentioned, I thus decided to focus the subsequent interviews on instances of interaction, i.e. decision-making and problem-solving situations, on the one hand and to inquire about disruptive factors in interdisciplinary coordination on the other. After I transcribed the last seven interviews, I coded them line-by-line with the first-order concepts and second-order themes in mind. After coding the transcripts, I had support for the extant concepts and some new first-order concepts that added up to form new second-order themes. This was actually not surprising, because the last interviews added the perspective of the Work Team members to the data. Thus, I revisited the transcripts of the interviews with the POTs and the coaches again to make sure I did not overlook statements that turned out to be relevant. In fact, I found support for the new concepts and themes in these transcripts. Data collection in the Gioia Methodology usually ends, when the analysis has born a practicable collection of concepts and themes and theoretical saturation is reached (Gioia et al., 2013, p. 20; Glaser & Strauss, 1967). Because the number of my interview partners was limited from the start, the data collection ended with the last one. It would have been possible to speak to all interviewees again, but in my opinion that was not necessary. As a result of my analyses so far, I had a set of first-order concepts and second-order themes to further work with. In a next step, Gioia and colleagues try to further reduce the emergent second-order themes into "aggregate dimensions" (Gioia et al., 2013, p. 20). While trying to group my second-order themes and finding appropriate labels for the resulting aggregate dimensions, I once again consulted the literature and made an important discovery. I read some papers on cross-functional knowledge integration and realized that the success of agile teams may be determined by their ability to integrate diverse knowledge bases and that my data may provide important insights into how the process of knowledge integration unfolds in cross-functional teams. So, I went through my second-order themes again and partly reframed them to fit into the vocabulary used by scholars dealing with the phenomenon of knowledge integration. Afterwards, I distilled the second-order themes into five overarching theoretical dimensions. Having these, I was ready to build my data structure (see Figure 1 and Figure 2). The data structure visualizes how the researcher proceeded in the analysis, from raw data to concepts and themes. This graphical record of the analysis process is a vital element of validating thoroughness in qualitative research (Gioia et al., 2013, p. 20; Pratt, 2008; Tracy, 2010). In the following findings section, I will describe the aggregate dimensions and second order themes displayed in the data structure in more detail and provide example statements of the interviewees for each first order concept. The final process model that incorporates the relations between the dimensions will be presented and explicated in the last sub-section.

4. Knowledge integration in agile product development teams

The structure of the agile product development methodology with the 14-day sprints and the fixed meeting components would already provide a formal and timely framework to explain how knowledge is integrated in agile teams, however the analysis of the interviews revealed that the process of knowledge integration within this context is more nuanced, as team members permanently iterate between group-level and individual-level knowledge integration. At the group-level team members interact in formal and informal meetings to exchange coordination-related and problem-centered knowledge and concern themselves with a rather broad compilation of information and knowledge. Whereas, at the individual-level team members are primarily engaged with their individual task performance, processing the information they received from the interactions with their

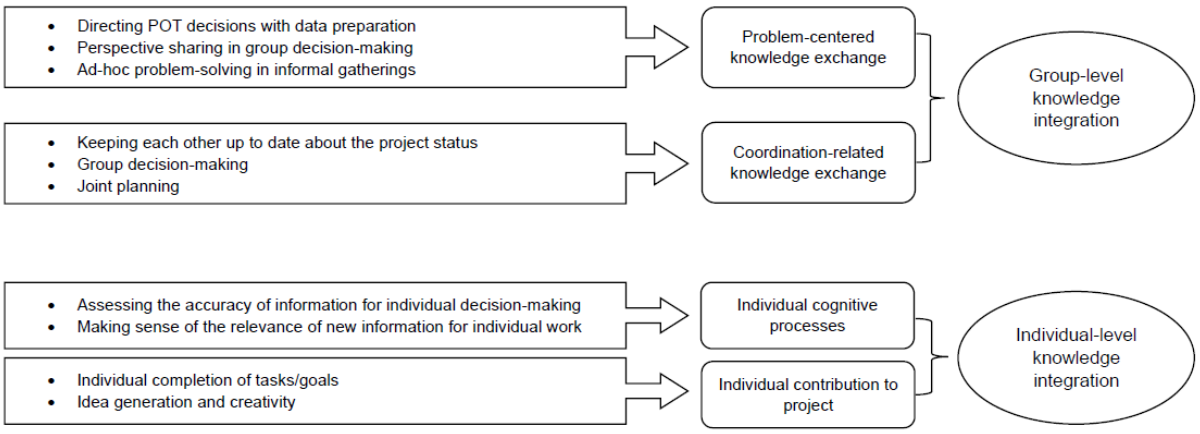

Figure 1: Data structure (part one)

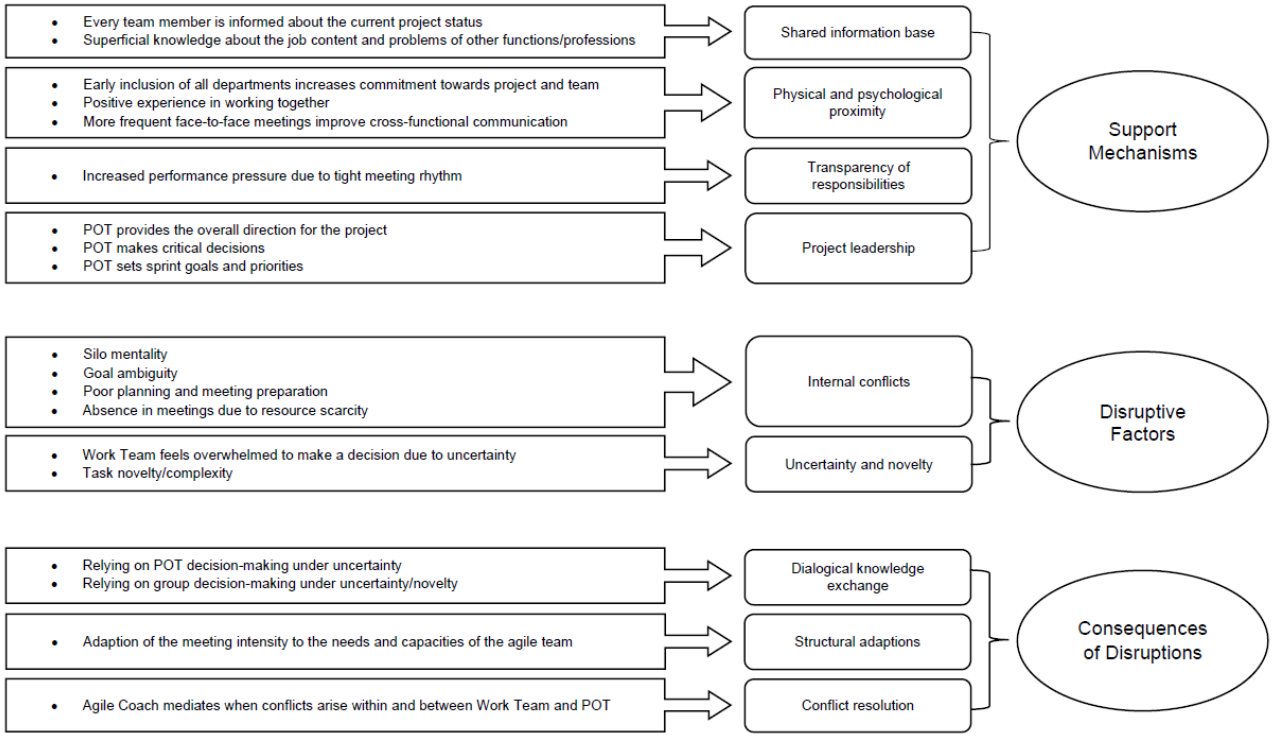

Figure 2: Data structure (part two)

colleagues and integrating them into their individual work. Thus, the data suggest that the process of knowledge integration is iterative in nature, meaning that the outcomes of the group-level provide the team members with input of different kinds for their individual work and that the team members in turn enrich the formal and informal interactions with their individual contributions. Figure 3 depicts this basic iterative model of knowledge integration in agile product development teams.

4.1. Group-level knowledge integration

The formal meetings at the transition from one sprint to the next serve as primary means for communication and information sharing across functions and roles throughout the project. They are the only occasion in which all members of the agile team come together: Agile Coach, Product Owner Team, and Work Team. Therein, team members have the possibility to share their ideas, present their results to their colleagues and give and receive feedback. Far more, the meetings are the main arena for large-scale problem-solving and group decision-making and allow for comprehensive discussions, thus representing the central platform for integrating different functional perspectives on bigger problems. Also, collective planning takes place within this formal context. Depending on the project phase and number of issues that may need to be discussed, these meetings may

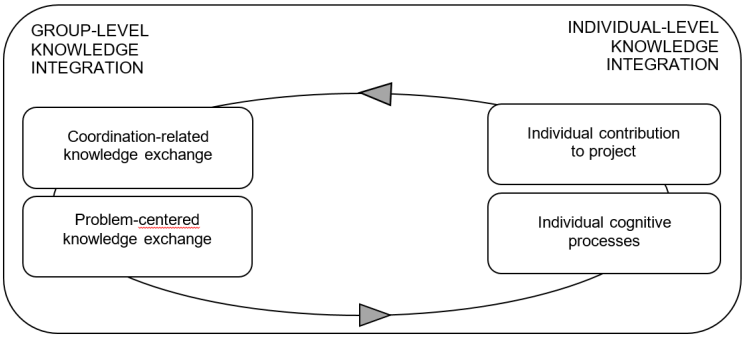

Figure 3: Basic iterations in the knowledge integration process

last up to four hours. However, also informal gatherings and the Daily-Stand-up-Meetings are used for knowledge sharing, whereby in these occasions, knowledge exchange is much more focused on problem-solving than in the biweekly sprint meetings. To sum up, within formal and informal interactions two different types of group-level knowledge integration take place: coordination-related knowledge exchange and problem-centered knowledge exchange. While the former mainly concerns cooperation, division of labor, planning and the overall direction of the project, the latter concentrates on exchange of expertise with the aim of cooperatively solving problems and creating innovations.

4.1.1. Coordination-related knowledge exchange

A fixed component of the biweekly sprint meetings is the presentation of results, in which team members share their achievements and encountered problems of the past two weeks. Therein the team members receive a broad range of information about the work content and obstacles of other departments and can develop a holistic picture of the overall project status. Also, the Daily-Stand-up-Meetings are used to share superficial information. Thus, keeping each other up-to-date about the project status represents a recurrent coordination and knowledge-integration practice in the agile teams under study. One of the engineers explained, that the regular meetings provide an effective platform for shortly discussing problems in a bigger round, because all the team members he would need for coordination are there, which makes the need to contact these people individually or to make an extra appointment obsolete, thus saving time.

> „Und so habe ich alle gleich einmal beieinander, die ich brauche, und dann kann ich das ganz schnell abwickeln zum Teil. [...] Ich sage mal, der ganze Besprechungsaufwand und Kommunikationsaufwand, den wir sonst gehabt haben, der ist jetzt minimiert worden."
>
> Mechanical Engineer, Project Colossus

However, at this point it is important to note that several interviewees highlighted the perceived inefficiency of information and knowledge sharing in the tight meeting structure. For example, the mechatronics engineer of project Homestretch claimed that most team members would only be affected by 20 percent of the content, and the other 80 percent are more of a "nice-to-know".

> „Richtig betreffen tut die meisten nur 20 Prozent. Aber die 20 Prozent sind halt wichtig. Da muss man die 80 Prozent Ineffizienz einfach als ‚aha, was die anderen Abteilungen leisten', annehmen."
>
> Mechatronics Engineer, Project Homestretch

With a general increase in group discussions, also the tendency to make decisions at the group-level rose. Thus, another common knowledge integration practice in agile teams is group decision-making, which primarily takes place in formal meetings. While decision-making is often related to problems or their solutions respectively, it has more of a coordination function, since decisions pave the further path of a project and provide the informative basis for the team members by which they arrange their individual efforts. The type of decisions that are made in the group are varied. It can be a decision between different problem-solving paths or design options, or smaller decisions where individual team members are unsure how to proceed and prefer to obtain the approval or support of the team. In the interviews, team members emphasized that they find their opinions to be a lot more substantial than they were before the introduction of agile product development. If all or many team members would agree then it would matter more than if only one person would stand in for an idea, they stated.

> „Es ist einfach auch ein Unterschied, ob jetzt ein ganzes Team eine Meinung hat und das vor dem POT präsentiert oder ob du allein mit deinem Chef das ausmachen musst. Und das hat einfach dann ein ganz anderes Gewicht, wenn das auf einmal ein ganzes Team trägt, die Entscheidung, oder die Idee [...]. Das macht auch einen wesentlichen Unterschied zu früher."
>
> Mechanical Engineer, Project Colossus

Obtaining different perspectives and making decisions together thus seems to increase commitment to the decisions

made and to reduce renegotiations of choices. As group decisions tend to create a feeling of security or support on the one hand, and group consensus on the other hand strengthens the perceived power of the group, decisions that could also be made by individual team members are pushed to the group-level. This stood out especially with Project Colossus. Possible reasons for this effect are elaborated in a later section on disruptive factors.

Another central coordination practice in agile product development is the detailed joint planning and the involvement of both POT and Work Team in doing so. A central responsibility of the POT is to integrate market demand and state of technology in their planning for the project and the Work Team. The consensus-oriented planning practice was new to the managers in the company and demanded them to invest considerably more time on planning, than they were used to. Because the POT should only specify what should be achieved and not how the team should perform the tasks, they had to learn to formulate the goals as clearly as possible and to provide the team members with autonomy and trust in the execution. While the POT formulates and proposes the goals for the upcoming sprint, the Work Team has the possibility to veto and stipulate adaptions. During Sprint Planning the team members commit themselves to the sprint goals set by the POT and the POT confirms that this is what they want. Thus, the outcome of such a planning meeting is a mutual agreement on the objectives for the next sprint. How important planning and goal setting is, was repeatedly emphasized in the interviews in both projects. One interviewee, for example, underlined that the quality of the product strongly depends on the quality of the goal specifications of the POT in the sprints.

> „[D]er Output vom Projekt hängt nicht nur vom Team ab, sondern sehr stark auch vom Product Owner Team. Also in der Qualität, in der sie die Aufgaben stellen, kommt auch die Qualität dann zum Schluss bei den Produkten raus."
>
> Simulation Engineer, Project Colossus

Once agreement on the sprint goals is achieved, the team autonomously splits the requirements of the POT into smaller activities and disperses them among the team members involved enabling the team members to work individually or in groups on their working packages. These activities are posted on the Team-Sprint-Board. As a result, the team members know at any time who is responsible for what and can also track the progress of work on the board. The predictable dependencies of the departments for the current sprint are already shown in the planning and activities. Thus, joint sprint planning fulfils important knowledge integration functions. First, the goals and activities formulated incorporate input from functional as well as management parties. Second, interdependencies of different areas of expertise are revealed. And third, the Sprint Boards visualize the outcomes of planning and serve as a point of reference and knowledge reservoir for both Work Team and POT.

4.1.2. Problem-centered knowledge exchange

As already implied, problem-centered knowledge exchange may take place in formal and informal interactions. A way of cross-functional knowledge integration that is practiced in the context of formal meetings is perspective sharing in group discussions. In group discussions the team members from different departments reveal their perspectives on problems or proposals for solutions and provide each other with feedback. This practice aims at enriching the individual understandings of the team members with other perspectives and valuing them for decision-making. In general, the interviewees agree that with the introduction of agile product development, problems and their solutions are discussed much more than before. Thus, topics are dealt with more intensively and the solutions are built on a broader knowledge base.

> „Und heute ist es so, man diskutiert das vielmehr im Team, wenn man ein Problem hat und [...] dann hat man auch gleich einmal viel mehr Meinungen. Das ist alles breiter aufgestellt dann, die Lösung oder die Vorschläge."
>
> Mechanical Engineer, Project Colossus

This tendency to discuss topics more in the group also means that the probability of overlooking important issues decreases and the potential of gaining new insights increases. Moreover, the interdisciplinary team setting enables some team members to contribute with completely different input, since they are involved in the project from the start.

> „Und der ist jetzt von Anfang an dabei und kann somit einen ganz anderen Input liefern, als er sonst könnte."
>
> Agile Coach, Project Homestretch

The formal meetings are of course not the only form of interaction and knowledge integration. Team members engage in ad-hoc problem-solving primarily within an informal context. While they are a fixed element in the agile product development methodology, the Daily-Stand-up-Meetings have a rather informal character. In these 15-minute meetings reporting of problems and quick coordination of expertise to resolve them are in the focus. Team members especially look for information that gets them ahead in their own work, but they also share their expertise on certain issues and support their colleagues in problem-solving, if desired.

> „[Ö]fter ist es so, dass man jemand anderen auch unterstützen kann, oder, dass du von jemand anderen unterstützt wirst."
>
> Testing Engineer, Project Colossus

Some team members also spontaneously seek out other opinions outside of the meetings while doing their tasks for the project. For example, the mechanical engineer of Project Homestretch explained, that he often consults the testing department when he designs a machine part, because they have the most experience with the machines.

> „Wenn ich zum Beispiel konstruiere, dann frage ich schon viel [bei Kollegen aus einer anderen Abteilung nach], weil die gerade die meiste Erfahrung haben mit den Maschinen."
>
> Mechanical Engineer, Project Homestretch

Sometimes, new dependencies or problems arise spontaneously, and team members inform their affected colleagues by making a phone call and engage in ad-hoc problem-solving.

> „Ein Beispiel ist, ein Konstrukteur ruft an: ‚Ich habe da jetzt etwas geändert, ich glaube das betrifft [ein spezifisches Bauteil]', [...] der hat das Gefühl, da tut sich was, dann fragt er, das ist ein Anruf, das kostet ihm 30 Sekunden und dann schau ich mir das an."
>
> Mechatronics Engineer, Project Homestretch

While group decision-making fulfills a coordinative function, POT decision-making is more problem centered. Not all decisions are meant to be made by the group or individual team members. The critical issues in the project, e.g. target specifications or make-or-buy decisions, are in the responsibility of the Product Owner Team. However, they need information from the Work Team as a basis for decision-making, as they cannot dispose of all the necessary details, due to their engagement in multiple projects and their departmental responsibilities. Thus, it is up to the Work Team to provide the POT with data for decision-making – a central knowledge integration practice, that guides the flow of information from team members to POT. The mechanical engineer of Project Colossus explained, that if any decisions are pending, they try to prepare the issue as best as possible for the sprint transition so that the POT can make a decision.

> „Also grundsätzlich ist es so, wir versuchen, [...] wenn irgendwelche Entscheidungen anstehen, die Thematik bestmöglich aufzubereiten für den Sprintübergang, damit das POT Team dann eine Entscheidung fällen kann. Die brauchen natürlich Infos, die können nicht alles wissen."
>
> Mechanical Engineer, Project Colossus

Thereby, the information asymmetry between POT and Work Team is an inherent point. Some of the interviewees stated, that the power of the Work Team is considerably higher because in total they know more about the product and the possible approaches to problem-solving. The difficulty for the POT then is to make speedy decisions based on the facts presented to them in the sprint meetings and to trust that the members of the Work Team have dealt with the topic sufficiently and rationally present the decision options. Occasionally, team members present three alternatives, but direct the decision to their favorite option by preparing this alternative more detailed than the others. Thus, team members carefully select the information they share with the POT in the sprint presentations.

4.2. Individual-level knowledge integration

As already mentioned, the process of knowledge integration appears to be iterative in nature. Thus, group-level and individual-level knowledge integration cannot be strictly separated from each other in terms of time, but rather happen simultaneously continuously providing each other with input from their respective integration outcomes. Especially during the sprint meetings high information-processing and sensemaking demands are posed to the team members. They are confronted with a lot of new information, problems, dependencies, as well as information about external market or technology changes brought in by the POT. The team members have to process these novelties accordingly and integrate their meaning in their own bounded worldview.

4.2.1. Individual cognitive processes

A central aspect of individual work that already starts in the meetings is information-processing. Thereby, making sense of the relevance of new information for one's own work represents an important knowledge integration practice at the individual level. The data have shown that the regular reporting of results and group discussions do not constitute an intensive engagement with each other's knowledge, rather team members principally take only the information and knowledge necessary for their own work with them from the sprint meetings. Especially in the beginning of the agile projects, information overload was an issue as the amount of information that team members had to process rose considerably compared to prior project work. As one team member stated, you get a lot of information, and at the beginning you do not know how to deal with it.

> „[M]an erfährt viel, bekommt sehr viel Information, und am Anfang weiß man auch nicht, wie man mit dem umgehen soll."
>
> Simulation Engineer, Project Colossus

Thus, the team members had to learn, how to process the significantly higher amount of information. A project's testing engineer explained how to deal with all the new information in one simple sentence. He stated that you would have to pull out what is relevant to you, implying that you should not really care about the rest.

> „[D]as was für dich relevant ist, musst du dir rausziehen."
>
> Testing Engineer, Project Colossus

The perceived value of information content also impacts the average attendance of the team members in the meetings. While attendance is mandatory, most of the team members decide on their own, if it is necessary for them to attend the meetings. For example, the controller of Project Colossus questions the usefulness of her presence in deep technical discussions because she sees no added value for her work.

> „Weil, wenn die die ganze Zeit wirklich nur über das Technische reden, was mache ich dann dort, wenn ich keinen Mehrwert habe?"

Controller, Project Colossus

Another individual cognitive process that plays an important role for knowledge integration, is the assessment of the accuracy of a piece of information for individual decision-making. Several interviewees, both Work Team and POT members, highlighted that a big challenge in making decisions is to have the right information at hand.

> „Ja eigentlich ist es die Herausforderung, dass man sicher die richtigen Informationen hat, wenn man eine Entscheidung treffen muss, auf welcher Basis, welchen Fakten, dass ich mir sicher bin, das ist so."
>
> Testing Engineer, Project Colossus

Some interviewees reported that they often rely on gut instinct when making decisions. It seems that the individual cognitive processing ability is often insufficient to include all information in the decision-making process, which is why team members rely on their intuition. Thereby they integrate their experience with the other information received.

4.2.2. Individual contribution to the project

Task completion, which represents the main part of individual work during a sprint, happens primarily outside of the meetings and other instances of interaction. The task assignments that are the outcome of Sprint Planning serve as important basis for the individual work phase. They represent an input that goes without further reflections on or discussions of meaning, because the task assignments are tailored to the skills of the team member who is responsible for completion and they were jointly developed in the sprint meeting. Thus, the specialists mainly work alone to fulfil the tasks agreed on for the current sprint. Each team member is responsible for how they come to performing their task.

> „[U]nd jetzt gibt es wirklich die verschiedenen Abteilungen [...], die nehmen ihre Aufgaben mit und es muss sich jeder im Team darum kümmern."
>
> Mechanical Engineer, Project Homestretch

In completing their task assignments, the team members integrate the obtained knowledge in the form of multiple perspectives, group or management decisions, solutions, and goals, from presentations, discussions, and planning in their individual contribution to the project.

Besides the fulfilment of tasks, integral parts of individual work are idea generation and creativity. In the beginning of projects, the team members usually have more freedom to elaborate on own ideas as the decision-making corridor is more open. However, as the project proceeds, choices get more and more constrained. Furthermore, the predefined goals of the POT determine the direction of individual work. As a member of the POT of Project Homestretch explained, some team members may just do what the tasks say, and in the past, they did what they thought, and maybe it was more.

> „Vielleicht ist es bei manchen Teammitgliedern auch so, sie machen das, was auf den Aufgaben draufsteht und früher haben sie das gemacht, was sie sich gedacht haben, und es war vielleicht sogar mehr."
>
> Member of the POT, Project Homestretch

Whereas a team member explained that he just has no time to be creative. Due to the high number of projects, he is engaged in, he can only manage to get done what he needs to, but not more.

> "Das geht sich aber auf der Projektdichte nicht aus. Da ist man froh, dass man das liefern kann, was man muss. Und die Pflicht, die geht sich gerade noch aus, für die Kür hätte ich gerne mehr Zeit."
>
> Mechatronics Engineer, Project Homestretch

However, the manner in which to achieve desired outcomes is rarely given. Only the expected result is provided, how the team members get there, is always up to them. So, from that point of view, they can live out their creativity in finding solutions to predefined problems in the later project phases.

4.3. Support mechanisms

The data analyzed suggests that knowledge integration in agile teams is supported by formal and informal coordination mechanisms. These support mechanisms tend to reduce the need for coordination-related knowledge exchange in agile teams by substituting for dialogical knowledge exchange. Figure 4 on the next page illustrates the mechanisms identified and their effect on group-level knowledge integration. In the following, I will explain in more detail how a shared information base, physical and psychological proximity, transparency of responsibilities, and project leadership may reduce the need for knowledge exchange in agile product development teams.

4.3.1. Shared information base

At the end of a sprint cycle, in the DEMO, the results of the work of the last 14 days are presented by the team members. This is usually aided by a PowerPoint document, which the members of the Work Team fill with pictures and descriptions to visualize their outcomes. In this phase of the meeting, comments or discussions are usually not desired. Rather, the mutual presentations should keep the team members on an equal level of information regarding project status and current topics. The results presented may include completed tasks, e.g. technical solutions, design solutions, or calculations, or also encountered problems that held the team members off from successfully delivering their contributions. During the sprint cycle achievements and pending problems are shared in the Daily-Stand-up-Meetings. It follows that every team member is informed about the current project status, providing that they attend the meetings. This shared

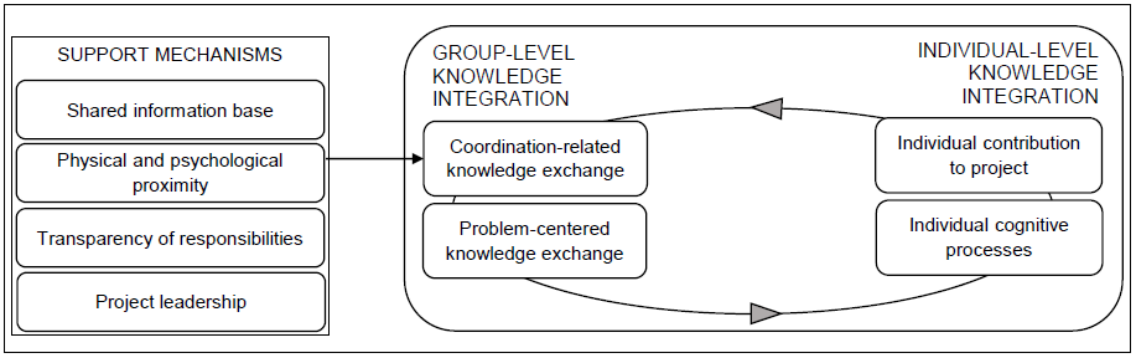

Figure 4: Support mechanisms in the knowledge integration process

level of information is particularly appreciated by many team members. While in the past they often had to get information themselves to get ahead, the flow of information has improved with the implementation of agile methods.

> „Naja, die Erwartung ist die, [...] dass man in einem ständigen Austausch mit der Gruppe ist [...], dass jeder Bescheid weiß, wie ist der Stand auch bezüglich Gesamtprojekt."
>
> Member of the POT, Project Colossus

The mutual presentations seem to reduce the need for subsequent bilateral exchange of knowledge by providing sufficient information to enable task completion in individual work. In the quote below, a team member highlights the advantages of a shared level of information about the project status in coordinating work. If certain tasks demand closer collaboration with colleagues from other departments, he does not have to explain again what the task is about or what his problems are, because the other team members were there at the sprint meetings and already know.

> "[D]ann muss ich dem das nicht lange erklären, um was es geht und einen Termin wieder ausmachen, sondern der war dabei direkt beim Sprint oder bei den Dailys, der weiß sofort um was es geht, der nimmt sich die Aufgabe mit."
>
> Mechanical Engineer, Project Colossus

Furthermore, the regular exchange of information in informal and formal interactions improves individuals' knowledge about the work and respective challenges of their team colleagues. This superficial knowledge about the job content of other functions appears to support coordination in agile teams. The interviewees emphasized that they are able to get to know and understand their colleagues, their work and the associated challenges much better through the regular gatherings and exchanges. In other words, mutual understanding is promoted. As one of the team members explained, in former times he often did not know what his colleagues were doing, and they did not know what he was doing. As a result, they talked at cross-purposes or their components did not fit together at all.

> „Die haben oft, also wir haben nicht gewusst was sie machen und sie haben nicht gewusst was wir machen und da hat man dann oft aneinander vorbeigeredet oder es hat dann einfach überhaupt nicht zusammengepasst."
>
> Mechanical Engineer, Project Colossus

However, the frequent communication across functions reduced problems of this kind and increased interaction further effectuated that team members not only know about each other's problems, but also about their skills, which makes it easier for them to locate experts to consult in the event of difficulties.

> „Und jetzt weiß ich was die Probleme vom Konstrukteur sind und der weiß was meine Probleme sind. Und ich weiß auch, was die Konstrukteure können."
>
> Simulation Engineer, Project Colossus

The regular meetings in agile project management resulted in the development of a shared information base about the overall project, wherein every team member is informed about the current project status and disposes of superficial knowledge about the job content and problems of other functions or professions. The information shared in meetings is rather broad than deep and supports the process of knowledge integration in reducing the need for subsequent bilateral exchanges on general topics, making collaboration more focused. Furthermore, the regular mutual communication of achievements and problems in the meetings facilitates the development of transactive memory, i.e. who knows what in the team, thereby reducing effort to identify experts for joint problem-solving.

4.3.2. Physical and psychological proximity

With the implementation of agile product development, the collaboration in the project teams inevitably became closer due to the regularity of the obligatory face-to-face meetings. While many project members stated that collaboration within the development teams has always been good, the implementation of the agile product development methodology brought about improvements in teamwork,

they agreed. Through the frequent face-to-face meetings, physical proximity of the team members was increased, and the employees of the different departments developed a stronger sense of belonging to the team and a feeling of togetherness. Thus, also psychological proximity to the team and the product grew. The majority of the interviewees sees the regular information and experience exchange between the departments and the short, rather informal coordination in the Daily-Stand-up-Meetings as a great benefit of the agile methodology. As one of the interviewees stated, the change in the intensity of working together is a big advantage, because you are just so much closer to each other and you are in constant exchange with colleagues from other departments.

> „Ich meine, die Zusammenarbeit ist natürlich schon ein großer Vorteil von der Veränderung her, weil man einfach viel mehr beieinander ist und sich ständig austauscht."
>
> Testing Engineer, Project Colossus

Some departments that are located in a separate building have previously been perceived as external service providers that only fulfill order after order without caring much about product or project success. Now, as one team member stated, they are much more involved in the project and can identify more with the product as a whole and are perceived as real team members.

> „Und jetzt sind die viel mehr integriert und haben einfach auch einen ganz anderen Bezug zu dem. Die werden ganz anders eingebunden. Und ja, ich glaube, dass man sich dann auch ganz anders identifiziert mit dem Projekt."
>
> Mechanical Engineer, Project Colossus

Moreover, the early involvement of these departments in the development team changed their sense of responsibility for the success of the project. As one technical services engineer stated, if you get involved, you can really make a difference.

> „Ja, weil man hat mehr Verantwortung für das Projekt, weil man einfach auch, wenn man sich einbringt, kann man echt was weiterbringen."
>
> Simulation Engineer, Project Colossus

Importantly, a member of a Product Owner Team observed that after you meet regularly, differing departmental worlds simply grow together through experience exchange and mutual support. He further suggested that this regularity is what drives the project forward.

> „Nachdem man sich aber regelmäßig trifft, wachsen die Welten einfach zusammen. Einfach dieses regelmäßige Treffen, Austauschen, ‚he was machst du gerade, was brauchst du gerade und was brauchen wir gerade'. Da bringt einfach diese Regelmäßigkeit das Projekt voran glaube ich."
>
> Member of the POT, Project Homestretch

Summing up, the frequent face-to-face meetings in agile product development increased physical proximity, thereby making collaboration in the teams closer and improving communication across functions. Moreover, the early inclusion of all departments increased team members' commitment towards the project and the team. However, positive experience in working together plays a crucial role in determining teamwork quality. Finally, the close collaboration in the agile teams and team member familiarity also led to the development of psychological proximity. This perceived proximity to other team members and the product to be developed serves as support mechanism for knowledge integration, as it increases team members' sense of responsibility and aids in aligning the different contributions of the team members by providing a common point of reference.

4.3.3. Transparency of responsibilities

Many team members reported that the transparency of the division of responsibilities, reinforced by the visualization on the Sprint Board, and the regular meetings make them feel more obliged to complete the tasks within the mandated 14-day time horizon or at least to think about an approach to solving a problem, because nobody wants to stand there empty-handed during the presentation at the end of the sprint.

> „Und so ist halt schon, das steht am Board und in zwei Wochen kommen wir wieder zusammen und dann muss halt was präsentiert werden und wenn der nichts hat, schaut es auch ein wenig blöd aus und das [...] will dann eigentlich auch jeder im Grunde vermeiden, dass er dann mit leeren Händen dasteht."
>
> Mechanical Engineer, Project Colossus

The Project Colossus controller stated that she feels more pressured to deliver results in the agile way of working or has a guilty conscience towards the team if she is not able to deliver.

> „Ich fühle mich viel mehr unter Druck gesetzt. Weil, man sieht sich viel öfter. Also unter Druck gesetzt, ich habe ein schlechteres Gewissen, weil ich weiß am Montag ist das wieder, ich muss das da wirklich bringen."
>
> Controller, Project Colossus

So, with the meeting intensity also the performance pressure rose for the team members. Due to the 14-day sprints the time span to complete tasks is rather short. On top of that, joint planning and distribution of tasks, including the display on the Team-Sprint-Board, increased the transparency of responsibilities, which made the team members feel more obliged to complete the tasks until the next sprint to avoid the embarrassment of not being able to deliver results for the team. The transparency of responsibilities supports knowledge integration in providing the team members with a clear

allocation of tasks, making subsequent discussions over duties largely obsolete. Furthermore, the Team-Sprint-Board incorporates the mutual expectations of the team concerning the performance in the upcoming two weeks and might be seen as the visualization of the team's aspiration level.

4.3.4. Project leadership

While the central idea of the agile methodology is to set up self-leading teams and provide them with more autonomy in carrying out the development project, the interview data have shown that the members of the Work Team largely attribute the role of the leaders to the Product Owner Team. Several interviewees stated, that one of the main and most important responsibilities of the Product Owner Team is to provide the overall direction for the project. In other words, they develop the vision for the product.

> „Das POT trifft die Entscheidungen und gibt einfach die Richtung vor und motiviert dann die Leute und wenn man wo schauen muss ist, dass das POT passt."
>
> Controller, Project Colossus

Another important task of the POT is to set sprint goals and priorities. This project management task is important, as it provides the Work Team members with a frame for their individual tasks. With the preestablished goals in mind, the team members then only have to decide how to meet the requirements of the POT. This intended split between what to do and how to do it, represents a central feature in agile product development. As a POT member of Project Homestretch mentioned, setting precise goals and being present in meetings are crucial factors that determine the quality of collaboration in agile teams.

> „[W]as man halt lernen muss ist das Thema konkrete Ziele auch setzen für die Sprintübergänge und dort einfach schauen, dass man laufend präsent ist."
>
> Member of the POT, Project Homestretch

When asked about the distribution of decision-making authority, most of the interviewees talked about some kind of rule of thumb that indicates, whether they can make the decision themselves, if it should be a group decision, or if the POT needs to concern itself with the topic. This rule of thumb developed over time. In the beginning the team members were sometimes uncertain about how far their autonomy in decision-making would go. In general, they agree that details or approaches to problem-solving are meant for single or group decisions and the POT is in charge of making the critical decisions in the projects, like target specifications or make-or-buy decisions.

> „Ja ich sage mal, wenn es um Details geht oder wie ich zu der Lösung komme, das ist eher das, wo wir als Team entscheiden, aber wenn es um Grundsatzentscheidungen geht, gerade auch was in einem Pflichtenheft steht, das muss halt einfach vom POT kommen."
>
> Mechanical Engineer, Project Colossus

As already mentioned, the members of the POT often have to rely on the information they get from the Work Team. So, trust in the integrity and thoroughness of the data is important when it comes to POT decision-making. A POT member mentioned that he especially has an eye on the way the results of a sprint are presented to him. Outgoing individuals, who can well present themselves, quickly sell their favored solution as the only truth on the planet. However, he added that in the end he bears responsibility for technical decisions and intensely relies on his gut feeling when making decisions of this kind.

> "Technisch gesehen liegt die Verantwortung ganz klar bei mir und ich verlass mich das extrem viel aufs Bauchgefühl und wie mir was präsentiert wird, wie sattelfest die Leute da sind, wie einig sie sich sind."
>
> Member of the POT, Project Colossus

The Product Owner Team clearly takes over the leadership role in the two agile teams under study. However, it is important to mention that there are mostly one or two persons in the POT, who drive the leadership role and motivate people. Nonetheless, the POT seeks to appear as one management (not leader) team on the outset. Project leadership, as described above, supports the knowledge integration process by reducing the need for coordination-related knowledge exchange. In providing an overall direction for the project and setting sprint goals and priorities, the POT fulfils central coordination mechanisms. Furthermore, the distribution of decision-making authority reduces the need for knowledge exchange between the agile roles.

4.4. Disruptive factors

Data analysis revealed that there are two types of disruptive factors affecting group-level knowledge integration in different ways. On the one hand, team internal conflicts of different causes temporarily diminish the effects of the support mechanisms and thereby increase the need for coordination-related knowledge exchange. On the other hand, environmental uncertainty and task novelty directly intensify problem-centered and coordination-related knowledge exchange in the agile teams. Figure 5 depicts these effects. The empirical background to the disruptive factors is described on the following pages.

4.4.1. Internal conflicts

The Agile Coach of Project Colossus stated that most of the conflicts or troubles in the team are triggered by poor planning and meeting preparation. The Agile Coach of Project Homestretch agrees in stating that as soon as the POT is weakly prepared and discusses important issues during Sprint Planning in front of the team, the mood in the Work Team deteriorates, immediately giving rise to conflicts.

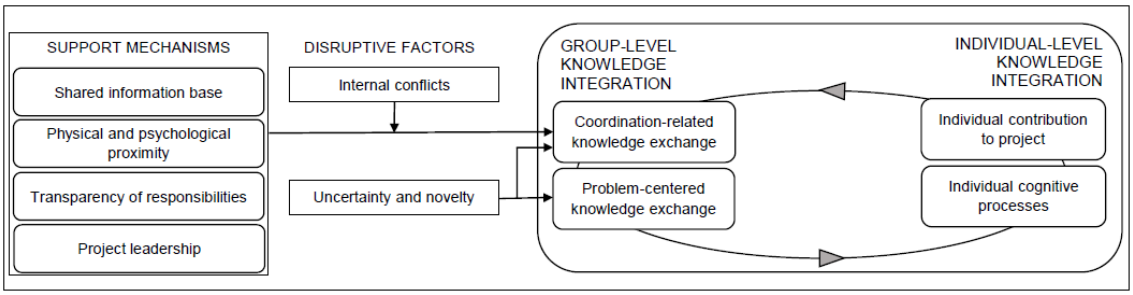

Figure 5: Disruptive factors in the knowledge integration process

> „Wir stellen einfach fest, ich sage jetzt mal, behaupte ich, in 85% der Fälle, wenn wir merken es sind Probleme im Team, Unruhen im Team, ja, also quasi auf gut Deutsch, irgendetwas passt nicht, wenn man es zurückverfolgt auf die Ursache des Problems, kommt man eigentlich darauf, dass man sagt ok, es liegt irgendwo in der Planung vom POT."
>
> Agile Coach, Project Colossus

Especially in the initial phase of agile product development, conflicts in the project teams arose, because the POT formulated the goals or tasks in too much detail on the Sprint Board. The mechanical engineer of project Homestretch mentioned, that this detailedness represented the biggest area of conflict in the beginning.

> „Die größeren Reibungspunkte waren vielleicht, dass die Aufgaben zu detailliert waren auf dem Sprintboard."
>
> Mechanical Engineer, Project Homestretch

As a result, the tasks or goals on the Sprint Board had to be negotiated and the supportive character of goal setting in project leadership got weakened, thereby increasing the need for coordination-related knowledge exchange in the teams. A related but independent topic is goal ambiguity. Especially in Project Colossus, goal ambiguity was a trigger for conflict between the Work Team and the POT. The cost of the machine was checked at a very late stage, with the result that the current design was far too expensive. The Agile Coach explained that there was a lack of understanding for this issue in team Colossus. The members of the Work Team were confused, because in their point of view, they have done exactly what the POT demanded them to do and have built the best machine possible.

> „Für ein Team ist das, die sehen das aus einer eigenen Brille und aus einer eigenen Sicht und im Team ist das teilweise auch, glaub ich, auf Unverständnis gestoßen, dass man gesagt hat ‚he was wollt ihr, wir haben genau das gemacht, eigentlich was ihr von uns wolltet, wir haben (aus deren Sicht) die beste Maschine gebaut, wieso jetzt auf einmal dieses Thema¿"
>
> Agile Coach, Project Colossus

The requirement of the project management was to build a highly innovative and novel machine that would outclass all previous models and the competition. Without paying much attention to the costs, the team developed the machine true to the goal that was set, to find out later that the machine is great, but far too expensive. As a result, there was a decrease in motivation, because the team members had to revise some of the key functions and design elements, of which they were proud. It took time for the team members to make sense of the cost issue. As a trigger of internal conflicts, goal ambiguity has detrimental effects on the knowledge integration process. While project leadership and accompanied goal setting are intended to support knowledge integration in agile product development teams by providing the team members with a point of reference by which they can carry out their project work, goal ambiguity diminishes this effect leading to undesired outcomes and project delay.

Due to resource scarcity many team members decide to absent themselves from the biweekly or Daily-Stand-up-Meetings, if they see no value for their own work or do not have any topics for the current sprint. Furthermore, the number of tasks for the overarching departments fluctuates considerably over the course of the project. Sometimes there are no topics for commercial departments or technical services in a sprint transition. As a result, the purchaser of Project Homestretch and others omit sprints, if they believe their attendance is not important and that their function or department has no issues.

> „Wenn ich sehe, dass es nicht so wichtig ist, oder glaube, dass der Einkauf keine Themen hat, dann bin ich in diesem Sprint nicht dabei."
>
> Purchaser, Project Homestretch

However, the team members should be present in the meetings in order to keep their general information about the project up-to-date or to be available for their colleagues if a topic that affects them arises spontaneously. Due to the strict implementation phase of the agile methodology, continuous presence is considered an important group norm. As described above, the Daily-Stand-up-Meetings and the other more formal meetings in the agile structure provide an arena

in which all the team members including the POT are available and topics with many interdependencies between departments and roles may be discussed straightforward. Thus, it is the (continuous) absence of project team members that leads to conflicts in the team.

> „Und das Nervige ist einfach, wenn die Leute, die man wirklich braucht, nicht da sind. Das ist das Nervige. Das ist dann, wo du dann auch den Unmut von den anderen spürst."
>
> Controller, Project Colossus

As the controller in Project Colossus put it, the absence is particularly negative and leads to resentment in the team if the person had been needed in the meeting and did not tell anybody that he or she would not appear. One of the interviewees expressed his frustration by implying that next time he would fail to appear, too.

> „Manchmal denke ich mir ok, jetzt ist der schon wieder nicht da, dann gehe ich das nächste Mal auch nicht, dann habe ich halt auch nicht Zeit, weil irgendein anderer Termin ist, oder was."
>
> Mechanical Engineer, Project Homestretch

The continuous absence of team members in the meetings impedes the development and maintenance of a shared information base, thus detracting the impact of the support mechanisms on group-level knowledge integration and increasing the need for dialogical knowledge exchange, because the absent team members need to be kept up-to-date outside the meeting context, if interdependencies arise. The aforementioned violation of group norms may further lead to conflicts that require the intervention of the Agile Coach and might necessitate the negotiation of interests among team members.

Finally, as already mentioned in the case description, silo mentality has been a major problem and conflict trigger at A-Machining for many years. While the introduction of agile product development and project management methods brought about some improvements, there are still occasions, in which department-centered thinking leads to conflicts in the project teams. As a POT member in project Homestretch explained, certain departments want to achieve their goals more vehemently than others without looking at the big picture.

> „Weil gewisse Abteilungen einfach ihre Ziele etwas vehementer erreichen wollen, ohne aufs große Ganze zu schauen."
>
> Member of the POT, Project Homestretch

This silo mentality is problematic, as it interrupts the knowledge integration process by increasing the potential for conflicts and subsequently the need for conflict resolution and coordination-related knowledge exchange. Moreover, conflicts possibly deviate team members' attention from the functional issues and problems in the project and might impair communication among the team members.

4.4.2. Uncertainty and novelty

While in agile product development the team members generally have more autonomy in decision-making, there are some decisions that the POT has to make because the team wants to protect themselves due to uncertainty or sees their competencies exceeded. Especially, the Work Team of Project Colossus is confronted with a high level of uncertainty and great pressure for success, putting extraordinary importance on decisions. The mechanical engineer of Project Colossus explained that the team would feel overwhelmed with certain decisions and could only prepare data for decision-making in the best possible way, but the decision would have to come from the POT. In such cases, the team would feel unable to make decisions.

> "Ja also man muss vielleicht schon sagen, dass sich hier und da vielleicht das Team bei gewissen Entscheidungen manchmal schon überfordert fühlt oder sagt einfach: ‚Das können wir nicht machen, wir können das nur bestmöglich aufbereiten, aber die Entscheidung muss dann vom POT kommen.' Also das kommt schon hier und da mal vor, dass wir uns dann nicht mehr in der Macht fühlen [...], dass wir gewisse Entscheidungen treffen."
>
> Mechanical Engineer, Project Colossus

Another externally triggered factor that might affect the knowledge integration process is task novelty. On the one hand, the new product development method triggered uncertainty. Agile methods were new to all team members and they did not really know what to expect. On the other hand, the Colossus project in particular was of an unusually large dimension for the company and many new ideas and concepts were created for it. The team's mechanical engineer described that because a lot was new, it was not clear at the beginning what challenges they would have and what else would come.

> „Und ja es war einfach dadurch, dass viel neu war, war es halt am Anfang noch nicht so klar, was wir für Herausforderungen haben und was dann noch alles dazu kommt."
>
> Mechanical Engineer, Project Colossus

When confronted with uncertainty and novelty, collaboration in the team tends to get closer and interactions more frequent. For example, the Daily-Stand-up-Meetings are more intensively used to quickly discuss pending problems and coordinate interdependencies and informal coordination and collaboration outside the fixed meeting components accelerates to uphold high team performance. Uncertainty and novelty disrupt the routine knowledge integration process because they require additional coordination efforts and cause team members to engage in verbal knowledge exchange more intensively. Furthermore, joint problem-solving is forced and the support mechanisms are largely insufficient to compensate for uncertainty and novelty.

4.5. Consequences of disruptions

Depending on the type of disruption, the strategies for coping with disruptive factors in the knowledge integration process vary. In general, the data revealed that in uncertain and novel contexts, there is a significant increase in dialogical knowledge exchange for both coordination and problem-solving. Moreover, the Agile Coaches might undertake structural adaptions, i.e. adapting the intensity of the meeting structure, to cope with uncertainty and prevent conflicts. Finally, to resolve internal conflicts, the Agile Coaches of the teams intervene with moderation techniques in discussions or individual talks to keep conflicts and their respective aftermaths in check.

4.5.1. Dialogical knowledge exchange

The phases of interaction, wherein team members meet face-to-face, become more important and more intensive as environmental uncertainty and the degree of task novelty rise. Team members increasingly rely on dialogic coordination to cope with the uncertainties that support mechanisms cannot resolve. Especially, group problem-solving and decision-making, formally as well as informally, are intensely used knowledge integration practices under these circumstances. In this regard, an engineer stated that if you are not sure what is right, then the whole thing becomes a team decision and you don't have to make it all by yourself.

> „Wenn man sich nicht sicher ist, was das Richtige ist, dann wird das Ganze eine Teamentscheidung und man muss das Ganze nicht allein fällen."
>
> Mechatronics Engineer, Project Homestretch

Overall, the level of uncertainty plays a major role in the individual perceptions about the meeting intensity and the necessity of escalating decision to the group-level. If uncertainty is high, for example because the project is at an early stage or unanticipated problems arise, the team members find the high frequency of the meetings more appropriate as if uncertainty is rather low, for example, because the project is close to the end and project work resembles more of a routine work without the need for extensive problem-solving. In uncertain or novel contexts, team members especially value the information content of the meetings and the possibility to discuss issues in a bigger round without having to call for additional meetings. However, as uncertainty turns rather low, for example in later project phases, in projects with a comparable low degree of novelty, or if team members do not have any tasks in the present sprint, the perceived appropriateness of meetings is low, as team members perceive the value of the information content of the meetings largely as "nice-to-know" but not necessary for their individual work. As already shown, attendance in the meetings is an important factor to enable knowledge sharing and integration in the project, because absence in the meetings negatively impacts collaboration by triggering detrimental conflicts in the team. Furthermore, the coordination practices that take place within the formal meetings may only be effective if all of the project team members attend the meeting.

In addition, the data revealed that team members increasingly rely on POT decision-making under uncertainty. Thus, under uncertain or novel circumstances, decision-making competencies are increasingly shifted to the POT and discussions within the biweekly meetings are more intense, seeking to cover all contingencies and to create a feeling of safety. A POT member in project Homestretch explained that at the beginning, where uncertainty about role competences and responsibilities was high, the Work Team tried to shift the decisions to the POT, which was a bit of a challenge, as agile product development foresees increased decision-making autonomy for the Work Team.

> "Das war am Anfang [...], ich will nicht sagen ein Hindernis, aber da hat das Team dann versucht eher den Weg zum Product Owner Team zu suchen. Bis sich das eingelebt hat, [...] dass sie viele Entscheidungen fällen dürfen, das hat etwas gedauert und war am Anfang [...] eine Herausforderung."
>
> Member of the POT, Project Homestretch

While project leadership has a fundamentally supportive effect in dealing with uncertainty, postponing the locus of decision-making from the Work Team to the POT also involves an increased effort for knowledge exchange, as the POT must first be brought up-to-date in order to enable a qualified decision.

4.5.2. Structural adaptions

In general, the degree of intensity to which the agile ceremonies are practiced, varies across the projects in the company and also within the course of single projects. The Agile Coach of project Homestretch stresses the importance of freedom of design, because not every phase of a project requires intensive communication on a daily basis and not every project demands the same intensity in collaboration. In the end, the POT and the Work Team should agree on a suitable adaption, that enables the exploitation of the advantages of the agile system, while at the same time fostering the acceptance of its formal elements.

> „Und das ist ganz wichtig bei jeder Methode. Man sollte sich Freiräume lassen, dass man sagt man macht es so oder so, wie es das Team oder das POT oder wie sie es gemeinsam sehen. Nur dann hat es einen Sinn, weil alles mit einem starren System ist dann wieder komplett falsch. [...] Vor allem für die Leute, die können sich nicht damit anfreunden."
>
> Agile Coach, Project Homestretch

Structural adaptations are an important mechanism to prevent conflicts that arise, for example, from being absent from meetings. With a lower number of meetings, the need for information of the team members tends to increase and

they perceive their presence in meetings to be more meaningful. Furthermore, by adjusting the intensity of the meetings and other ceremonies, the Agile Coaches can create an efficient framework for the required increase in dialogical knowledge exchange under uncertainty.

4.5.3. Increased effort for conflict resolution

For most of the Agile Coaches their role in agile projects represents only a small part of their total engagement in the company. They often have their main obligations in functional departments or in other agile projects as members of the Work Team. Therefore, a main challenge for them was to learn to keep a professional distance and avoid interfering in technical discussions with their own expertise. Nonetheless, the Agile Coach is deeply involved in the project as he or she is the first contact person for both Work Team and POT in case of collaboration problems or issues with the frame conditions and moderates all the agile ceremonies, even the Daily-Stand-up-Meetings if necessary. In applying moderation tricks, the Agile Coach can drive discussions, direct conversations, obtain contributions and resolve emerging conflicts at an early stage. The Agile Coach of Project Homestretch sees the value of his role especially in being a loyal or neutral person that mediates between the Work Team and the POT. Similarly, as quoted below, the Agile Coach of Project Colossus highlights the effectiveness of his interventions enabled by the professional distance to the development project and the general perception of his role as outsider.

> „Du bist eben der externe Faktor, der das steuern kann. Das lassen auch die Leute relativ schön zu."
>
> Agile Coach, Project Colossus

A central responsibility of the Agile Coach is to maintain a trouble-free workflow by removing barriers that distract the members of the project team from doing their work. The sources of such disturbances are manifold. As the Agile Coach of Project Colossus described, sometimes it is a matter of communication or collaboration between or within functions, or there are latent conflicts between Work Team and POT that need to be resolved. In the statement below, a member of the POT of Project Colossus highlights the general importance of the Agile Coach in the agile game and particularly points to the crucial ability of sensing conflicts or problems before they come to the surface.

> „[D]er agile Coach, der ist sozusagen eigentlich das Zünglein an der Waage. Für mich auch eine der absolut wichtigsten Personen in dem Spiel. Es hängt extrem von dem ab, was der spürt, ob er merkt, he da hat es was."
>
> Member of the POT, Project Colossus

As soon as the Agile Coach recognizes an emerging conflict, the first step is to directly confront the parties involved outside the meetings and to seek a clarifying conversation. If the coach senses a general tension in the team that cannot be attributed to individual team members, he has the opportunity to adapt the intensity of the agile ceremonies to improve the mood in the team again.

> „[M]an merkt, ok, es ist einfach permanent irgendwo eine gewisse Spannung, die man vielleicht gar nicht so offensichtlich sieht, aber man spürt sie, dann such ich einfach Gespräche, mit denjenigen, mit den Personen [...]"
>
> Agile Coach, Project Colossus

To summarize, at first glance, the role of the Agile Coach seemed to resemble those of simple moderators and mediators as frequently installed in group decision-making or conflict situations. However, on closer examination the role reveals to be more complex than expected, fulfilling important boundary spanning functions, such as maintaining the flow of communication and mediating between functions, and determining the design of the formal structure in the agile project. Through timely and appropriate interventions of the Agile Coach the need for intensive dialogue to overcome conflicts, clarify meanings and create shared understandings of tasks and responsibilities may be reduced.

4.6. An iterative process model of knowledge integration in agile product development teams

Based on the data presented in the previous sections, I will now put the different aggregate dimensions into relation and develop the iterative process model of knowledge integration in agile product development teams to summarize my findings. The analysis of the interviews showed that the process of knowledge integration iterates between phases of interaction, i.e. group-level knowledge integration, in which team members meet formally as well as informally, and phases of individual work, i.e. individual-level knowledge integration, wherein team members are engaged with their individual task performance. The data imply that the outcomes of the phases of interaction provide the team members with input of different kinds for their individual work and that the team members in turn enrich the formal and informal interactions with their individual contributions. This basic iterative process is visualized on the right side of Figure 6 on the next page, demonstrating the iterations between group-level and individual-level knowledge integration and the associated exchange of inputs and contributions.

Using the formal sprint structure as a frame, I start explaining the iterative process of knowledge integration with the formal interaction of the project team members in the biweekly sprint meetings, including upstream contribution and downstream input. As already outlined, the meeting opens with the presentation of the results of the previous sprint. Thereby, individual team members

contribute with presenting their completed tasks, alternative ideas for solving pending problems, newly encountered problems, or other issues, which demand further discussion and a group or POT decision. Thus, team members first of all

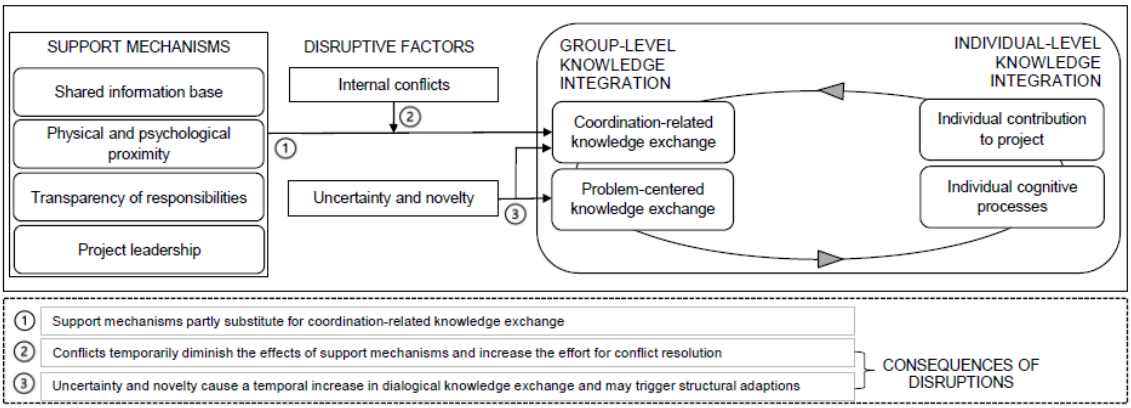

Figure 6: An iterative process model of knowledge integration in agile product development

get updated about the project status, including technical solutions, current problems of other team members, and pending decisions. After the initial presentation, there is time for group problem-solving and decision-making. Usually, this affects bigger problems or large-scale decisions, which require detailed examination through group discussion to be able to include as many different perspectives as possible in finding a solution or deciding which way to go. Individuals contribute to the discussion with their idiosyncratic viewpoint on the issue and their experience in working with the machine. As a result, team members may be provided with possible solutions or ideas for resolution, decisions, and also open questions, that may or may not affect their individual work. The outcomes of the previous meeting parts are then considered for The Conclave, the planning session of the POT, which happens detached from the Work Team. The individual POT members bring their demands and expectations into the discussion about the goal specifications for the next sprint, according to their roles as representatives of market, technology, production and project management. The outcome of this step are the goals for the next two weeks that are accepted by all POT members. Next, in Sprint Planning, the POT presents their agreed-on expectations for the upcoming sprint to the Work Team, whose members may question goals and priorities, or demand changes based on their experience and current workload. Once the goals are accepted by the two parties, the members of the Work Team engage in self-organized planning without the POT. The outcome of this planning stage are specific tasks that need to be completed individually or in groups until the next sprint transition in 14 days.

If a team member encounters problems or unexpected interdependencies within a sprint cycle, then additional, informal interaction might be necessary. The Daily-Stand-up-Meetings are in principle another planned part of the agile sprint cycle, but they are rather informal in nature. In these short meetings, team members primarily share their problems with others, point to interdependencies and contribute their experience and own perspectives to others' problems. In turn, the team members receive information on the sprint status, possible solutions to and multiple perspectives on smaller problems, and on whom to consult for given problems or direct offers of help. However, due to their history in working together, the team members have good knowledge about interdependencies and whom to consult for jointly solving smaller problems. Therefore, they directly contact the person they need to get ahead. Sometimes, it is just a phone call to point to a discovered interdependency between disciplines and sometimes, it is a coffee talk to exchange ideas and thoughts. However, informal interactions are always centered on a given problem and not on the project in general. Thus, knowledge exchange in these circumstances is highly focused and entails only the knowledge fragments necessary to solve a problem or coordinate an interdependency.

While individual work basically starts with the end of formal interaction, individual-level knowledge integration happens simultaneously with group-level knowledge integration. This is especially due to the individual information-processing and sensemaking demands posed to the team members during the sprint meeting. The team members are confronted with a lot of new information, problems, dependencies, as well as information about external market or technology changes brought in by the POT. The team members have to process these novelties accordingly and integrate their meaning in their own bounded worldview. These cognitive processes do not start after but during the meeting. Nonetheless, task completion, which represents the main part of individual work during a sprint including individual problem-solving and idea generation, happens primarily outside of interaction phases. The task assignments that are the outcome of Sprint Planning serve as important basis for the individual work phase. They represent an input that goes without further reflections on or discussions of meaning, because the task assignments are tailored to the skills of the team member who is responsible for completion and they were jointly developed in the meeting. As already shown in the description of the knowledge integration practices observed, team members principally take only the information and knowledge necessary for their own work with them from the sprint meetings. Therefore, they

integrate the obtained knowledge in the form of multiple perspectives, group or management decisions, solutions and goals, from presentations, discussions, and planning in fulfilling their task assignments. Also, insights gained from informal interaction are integrated by individuals. The informal interactions, dailies as well as ad-hoc coordination and collaboration, during a sprint and the associated feedback again include contributions and input in the form of smaller problems, own perspectives and discovered interdependencies, enriched with knowledge and experience from past informal interactions. The contributions of the individual team members, in particular their completed tasks, contain all inputs that were considered relevant by the responsible team member from formal and informal interactions over the last 14 days and are combined in the presentation of the results. However, the results presented are not necessarily solutions. Team members may also present bigger problems they encountered, prepared selection options for a decision too big to be made by a single team member, or also new ideas they have been working on to improve existing solutions. These individual contributions then are the basis for the other knowledge integration practices that take place within the biweekly formal meetings, e.g. group problem-solving and decision-making and planning.

While the basic iterative process model of knowledge integration in Figure 3 already shows well how phases of interaction and individual work iterate in the course of a sprint cycle and how inputs are transformed into contributions, there is still little information about the factors that influence knowledge integration practices. Thus, the support mechanisms that underly the knowledge integration process and the sensibility of knowledge integration practices to changes and its consequences shall now be integrated into the model. Support mechanisms like a shared information base, physical and psychological proximity, transparency of responsibilities, and project leadership reduce the need for dialogical knowledge exchange as they fulfil a coordination function. However, the reduction impact of the support mechanisms only applies to coordination-related knowledge exchange, which is implied by the long arrow in Figure 6. Data analysis further revealed that there are team internal as well as external factors that may trigger changes in coordination and collaboration demands that subsequently lead to an adaption of knowledge integration practices. External triggers for a change in coordination or collaboration demands can be subsumed under environmental uncertainty and task novelty. These factors raise the need for more intense interactions, formal as well as informal, and knowledge exchange. When confronted with novelty, collaboration in the team is closer and more frequent. If uncertainty is high, decision-making competencies are increasingly shifted to the POT and discussions within the biweekly meetings are more intense, seeking to cover all contingencies and to create a feeling of safety. In addition, consensus is desired to legitimize decisions and enable continuation of work with a better gut feeling. The Daily-Stand-up-Meetings are intensively used to quickly discuss pending problems and coordinate interdependencies and informal coordination and collaboration outside the fixed meeting components accelerates to uphold high team performance. In sum, the interaction phases, wherein team members meet face-to-face, become more important and more intensive as environmental uncertainty and the degree of task novelty rise. Team members increasingly rely on dialogic coordination to cope with the uncertainties that support mechanisms like shared information base or proximity cannot resolve. Especially, group problem-solving and decision-making, formally as well as informally, are intensely used knowledge integration practices under these circumstances. The Agile Coaches can accommodate the need for more dialogical knowledge exchange with undertaking structural adaptions, i.e. adapting the meeting intensity. Another important factor to cope with uncertainty is the leadership function of the Product Owner Team. In providing the Work Team with an overall direction and overtaking their responsibility to adequately integrate different viewpoints to solve problems and make decisions, they reduce the information-processing demands on the individual team members. As already implied, certain factors that arise as a result of individual work might trigger an increase in informal coordination and collaboration practices. For example, unexpectedly emerging interdependencies or newly encountered problems may necessitate the closer collaboration of different specialists. Essentially, these factors do not ascend due to changes in the team's external environment but represent the normal "surprises" when assumptions are tested and do not comply with initial expectations or requirements. Still, the typical trial-and-error processes in new product development that often come along with intensified collaboration and mutual adaptions increasingly occur in light of environmental uncertainty and task novelty. Thus, as shown in Figure 6, uncertainty and novelty lead to a temporal increase in dialogical knowledge exchange for both coordination and problem-solving. However, as soon as uncertainty and novelty reach a moderate level and things get back to "normal", the intensity of coordination and collaboration decreases and the phases of individual work in the knowledge integration process come to the fore again.

A team internal factor that changes coordination practices are conflicts, for example triggered by team member's absence in the formal meetings. While the physical presence of all team members in these meetings reduces the need for bilateral knowledge exchanges as problems occur or interdependencies are encountered, the absence of team members representing a concerned function does the opposite. This undesirable additional effort for coordination can, if absences accumulate, cause conflicts between team members and subsequently impair teamwork so that conflict resolution measures have to be taken by the Agile Coach. Usually, this goes hand in hand with a more intensive engagement with different viewpoints and negotiations of interests. Another major trigger for intensified discussions that could be avoided is unpreparedness of the POT. As the members of the Product Owner Teams are involved in multiple projects and concerned with management issues in their home departments,

their time for meeting preparation is limited. However, poor POT planning, which is often the result of ill-prepared members, raises additional coordination demands during the biweekly meetings. The formulation of ambiguous goals, the specification of detailed tasks instead of a rough direction, or the lack of certain points that are considered important in the current project phase by the team members lead to an increased need for discussion in Sprint Planning. As depicted in the model on Figure 6, internal conflicts seem to cancel the effect of the support mechanisms, which are supposed to reduce the exchange of coordination-related knowledge between different functions.

5. Discussion & Conclusion

5.1. Contributions

To address the scarcity of research on agile teams outside the software development context, the present study primarily aimed at exploring the success factors of agile teamwork. The organizational learning literature dealing with the change or increase in organizational knowledge provided the theoretical starting point for this examination. Two partly contradicting approaches to knowledge integration, the cross-learning perspective and the specialization perspective, were discussed. While the cross-learning view stresses that knowledge needs to be transferred between individuals and that knowledge boundaries need to be traversed with mechanisms and practices to enable knowledge integration (Majchrzak et al., 2012), proponents of the specialization approach argue that intensive knowledge exchange contradicts the very notion of specialization and poses excessive demands on individuals' cognitive capacities, thereby ignoring individuals' bounded rationality (Kieser & Koch, 2008). Empirical findings imply that extensive knowledge sharing is not key to successful knowledge integration, thereby contradicting the central conjecture of the cross-learning approach. Rather, extant empirical work suggests that certain practices and structural mechanisms reduce the need for knowledge sharing. However, recent studies on product development projects indicate that the need for and the engaging in knowledge sharing are dependent on the level of task novelty or innovativeness of the product or project (e.g. Schmickl & Kieser, 2008). While most of the studies on cross-functional teams in new product (e.g. Schmickl & Kieser, 2008) or new process development (e.g. Majchrzak et al., 2012) consider the overall level of uncertainty and its impact on knowledge integration, they do not show how knowledge integration mechanisms and practices change over time within the course of a project. Similarly, studies adopting a process perspective on knowledge integration in new product development do not account for changes in uncertainty or novelty over shorter periods of time (e.g. Enberg et al., 2006). To close this research gap, I examined the process of knowledge integration in two agile product development teams with special emphasis on changes in knowledge integration practices and mechanisms within the course of projects. Consequently, I developed an iterative process model of knowledge integration that shows (a) how agile teams integrate the diverse contributions of the individual team members into a new product, (b) how internal and external factors trigger alterations in knowledge integration practices, and (c) how agile teams adapt to the resulting changes in coordination and collaboration demands.

The present study contributes to the organizational learning literature on knowledge integration in three ways. First, the thesis adds important insights on the contingencies that determine the depth and content of knowledge exchange in project teams. Carlile and Rebentisch (2003, p. 1182) proposed that the complexity of knowledge integration in product development increases with the amount of dependencies between different specialized domains or departments in an organization. Accordingly, prior research suggests that the coordination of teams high in expertise diversity demands mechanisms beyond formal planning (Faraj & Sproull, 2000, p. 1555). Previous practice and process-based studies of knowledge integration, stress that excessive deep-level knowledge sharing is hardly found in multidisciplinary teams. These studies suggest that dialogical practices (Majchrzak et al., 2012) and structural support mechanisms like, modularization, prototyping and transactive memory (Kieser & Koch, 2008; Schmickl & Kieser, 2008), process specifications and presentation genres (Enberg, 2012), and a common digital space (Kellogg et al., 2006), facilitate knowledge integration by reducing the need for extensive knowledge exchange between specialists. Similarly, knowledge integration in the two agile development projects explored was aided by support mechanisms such as physical and psychological proximity, transparency of responsibilities, project leadership and a shared information base. These mechanisms, individually and in combination, reduced the need for extensive knowledge sharing among team members under "normal" circumstances, in which interdependencies were clear. Then, knowledge sharing in informal interactions was strongly problem-centered and team members only exchanged knowledge that was relevant for solving a particular problem. In formal interactions the focus was on creating a common knowledge base in the sense of an even distribution of general project information and overall project goals and not on creating common meaning. However, previous studies took the context in which the mechanisms are implemented and the practices are carried out as given, without considering the potentially detrimental effects of environmental or internal changes on the effectivity of knowledge integration over time. As Grant (1996, p. 115) stipulated, situational characteristics affect the appropriateness and relative expenditures of different knowledge integration mechanisms. He argues that, with increasing task uncertainty and task complexity, dependence on highly-interactive non-routine coordination mechanisms increases (Grant, 1996, p. 116). These effects could also be observed in the present study. Environmental uncertainty and task novelty were found to exceed the limit of effect of the support mechanisms and thereby increased the use of dialogue-based knowledge integration

Table 1: Integrating knowledge in dynamic environments

Current view	Neglected issues	Revised view
Depth of knowledge exchange		
Structural support mechanisms substitute for deep-level knowledge exchange between specialists	The level of uncertainty may change within the course of a project	Disruptions trigger changes in the mode and depth of knowledge integration within the course of a project, temporarily increasing the need for deep-level knowledge exchange among experts
	Coordination and collaboration demands of teams change over time	Effectiveness of structural support mechanisms varies over the course of a project, depending on the project phase, the project context, and the quality of teamwork
Locus of knowledge integration		
Most studies examine either individual, team, or organizational-level knowledge integration	The multilevel nature of knowledge integration	The dynamically changing degree of uncertainty determines the primary locus of knowledge integration → High uncertainty: group-level → Low uncertainty: individual-level
Breakdowns in the knowledge integration process		
Are triggered by epistemic uncertainty and lead to temporary intensifications of collaboration in integrating knowledge	Breakdowns may be caused by smaller events that disrupt the normal, taken-for-granted flow of practice	The knowledge integration process may be directly disrupted by uncertainty/novelty and indirectly by team internal conflicts for a limited time span
	Consequences beyond the changing nature of collaboration may be accompanied by breakdowns	Disruptions may lead to temporary changes in the nature of collaboration as well as in the characteristics of coordination

practices at the group-level that aimed at jointly developing meaning and maintaining innovative performance in ambiguous environments. Furthermore, in studying the two agile teams, I found factors beyond task characteristics that might disrupt the effectivity of routine knowledge integration mechanisms and practices. As shown, internal conflicts potentially interrupt the effects of the support mechanisms on the knowledge integration process. To restore their effectiveness, conflict resolution techniques, in which the interests of the conflicting parties were discussed, and an agreement was sought, were increasingly applied. These insights partly support the propositions of the cross-learning perspective, suggesting that individuals from different areas of expertise need to deeply engage with each other's knowledge, i.e. meanings, interests, and attitudes, to arrive at a common understanding that enables the integration of their knowledge (e.g. Carlile, 2004). Thus, the present study integrates insights of the specialization and the cross-learning perspective, in showing that certain internal and external disruptive factors trigger changes in the mode and depth of knowledge integration, temporarily increasing the need for deep-level knowledge exchange among experts. As outlined in the results section, the disruptive factors in my model increase the need for intense dialogical knowledge exchange for the time the factors persist, but once uncertainty, novelty or conflicts diminish, the practices rather quickly change back to the routine way of knowledge sharing that is aided by support mechanisms. It follows that the effectiveness of structural support mechanisms varies over the course of a project, depending on the project phase, the project context, and the quality of teamwork.

Second, this master's thesis adds empirical insights on the multi-level nature of the knowledge integration process and the dynamic changes in the locus of knowledge integration by showing how individual-level and group-level knowledge integration practices interact and mutually influence each other. Grant (1996, pp. 112–113) proposes that knowledge creation is an individual activity and that organizational knowledge is created through the interactions of individuals, thus implying the central role of the individual in integrating diverse strands of knowledge. In general, individual-oriented perspectives on knowledge creation posit that "individuals are the primary locus of knowledge" and therefore should be the basis for any attempt of understanding organizational knowledge creation and other knowledge processes (Felin & Hesterly, 2007, p. 197). While modelling the individual cognitive processes that underly individual knowledge integration was out of the scope of this thesis, I observed that a central component of individual work in the projects was the integration of new information with existing individual

knowledge stocks. Depending on the intensity of knowledge exchange in the phases of interaction the demands on individual knowledge integration, i.e. information-processing and sensemaking, fluctuated over the course of the projects. Furthermore, extant research adopting a process perspective on knowledge integration, proposes that the locus of integration varies over time. For example, Bruns (2013, p. 67) differentiates between expert practices, coordination practices, and collaborative practices. His process model of coordination in cross-domain collaboration depicts how team members swap between shared and domain-specific practices. In the context of new product development, Enberg et al. (2006, p. 158) also found that knowledge integration in teams happens through an iterative process of acting (alone) and interacting (with others). Similarly, Mengis et al. (2018, p. 601) found that the scientists in their study worked alone a considerable amount of time, with only short interactions if problems arose. However, "epistemic breakdowns", causing epistemic uncertainty, triggered adaptions of knowledge integration practices in the team of scientists. The primary mode of teamwork shifted from "working together-alone" with a focus on coordination to "drawing distinctions dialogically" as the breakdown occurred, to "working together-alone" again albeit with a focus on cooperation. This discovery is consistent with findings of other authors, implying that novelty and uncertainty effect the intensity of knowledge sharing in cross-functional teams (e.g. Majchrzak et al., 2012; Schmickl & Kieser, 2008). In line with previous studies and models, the process of knowledge integration across areas of expertise proposed here iterates between individual work and teamwork. In particular, the model shows that the process of knowledge integration in agile new product development teams iterates between interactions in which knowledge sharing and joint problem-solving take place and individual work, in which team members integrate the information and knowledge they gathered and processed in the interaction phases in their individual contributions. As opposed to the model of Mengis et al. (2018), the present process model takes into account the formal and repeating coordination elements inherent in new product development. In doing so, not only external factors, like uncertainty or novelty were identified to trigger changes in knowledge integration practices, but also internal conflicts that arise from absence in meetings, poor planning and preparation, silo mentality, and goal ambiguity, were found to change the intensity of knowledge exchange in the agile product development teams and thereby the primary locus of knowledge integration. Thus, triggers of change in coordination and collaboration practices do not necessarily stem from major epistemic breakdowns, but from everyday project ups and downs and violations of the expected way of working, which leads me to the final contribution of my study.

So, third and finally, this master's thesis enriches our understanding of the nature and consequences of breakdowns or disruptions in the knowledge integration process. Defined as "disruptions of the normal, taken-for-granted flow of practice when things don't go as expected" (Lok & Rond, 2013, p. 186), temporary breakdowns in sociomaterial practices are at the center of interest in studies that aim at theorizing through practical rationality (Sandberg & Tsoukas, 2011, p. 347). In investigating the dynamics of routines, Deken, Carlile, Berends, and Lauche (2016, p. 673) showed how routine work can break down in novel settings when significant differences in actors' understandings of ostensive patterns surface. Actors responded to these breakdowns through iterative episodes of routine work. In his papers on barriers to knowledge integration, Carlile (2002, 2004) already suggested that differences in knowledge, meaning and interests among actors represent boundaries that impede the integration of knowledge and become more severe as interdependencies become increasingly unknown. In introducing the concept of breakdowns to the study of interdisciplinary knowledge integration, Mengis et al. (2018) achieved similar results. The authors found that knowledge integration requires switching between different knowledge integration practices over time, which is particularly salient in light of epistemic breakdowns that are triggered by unsettling events that shake persisting understandings. The authors emphasize that in such cases, dealing with coordination issues is insufficient to maintain knowledge integration, rather collaborators need to engage in a dialogic process to handle the epistemic uncertainty they face (Mengis et al., 2018, p. 607). In a similar vein, uncertainty and novelty led to a disruption of the knowledge integration routines in the present study by overstraining the support mechanisms and increasing the need to jointly create meaning. Team members temporarily relied on group-level practices and POT decision-making to cope with these breakdowns. Consequently, the data presented here support the suggestion of Mengis et al. (2018, p. 608) that breakdowns play a critical role in temporarily changing the nature of collaboration. However, the breakdowns or disruptions in my study also effectuated a shift in decision-making and structural adaptions, thus additionally indicating temporary changes in the characteristics of coordination. Furthermore, as already discussed, not only uncertainty and novelty but also conflicts triggered changes in the knowledge integration process of the two agile teams under study, albeit in an indirect way. Silo mentality, goal ambiguity, poor meeting preparation and absence in meetings caused conflicts that were found to impair the substitutive effect of the support mechanisms, thereby unsettling the routine practices of knowledge integration. These conflicts surfaced differences (Deken et al., 2016) in interests and expectations of roles and responsibilities that ultimately led to breakdowns of the support mechanisms rather than the knowledge integration process itself. Thus, the model developed in this master's thesis enriches our knowledge about the factors that might disrupt knowledge integration with a team internal perspective that prior research in the field has neglected.

5.2. Boundary conditions and suggestions for future research

The findings of this study have to be seen in light of some limitations. The first boundary condition of the study arises

from its design. Due to the grounded theory approach with its explorative nature of the research questions and due to accessibility, a single case was chosen for the study. While single-case studies allow for detailed qualitative descriptions of natural situations under considerations of context, the ability to draw generalizations to other cases is limited, as data are subjective to a great extent and causal relationships cannot be readily established. However, the aim of this study was to generalize the findings to theory (Yin, 2003, p. 38). Thus, to be able to analyze the knowledge integration processes in the project teams in-depth, the limited generalizability of the findings to other cases was accepted. Consequently, future research is needed to test the proposed model. Moreover, due to time constraints of the author the interviews were conducted within a time span of three weeks, not allowing to record changes in expectations, attitudes, or mood of the interviewees over time. While the questions posed in the interviews were retrospective in nature and aimed at aiding the interviewees in reconstructing past events, the answers could be biased by the outcome, positive or negative, of these events. Future research should therefore conduct longitudinal studies to avoid contextual bias and be able to adequately account for causes for and results of changes in the process of knowledge integration.

The second boundary condition is due to selection and accessibility of interview partners. While an agile team consists of three different roles, the focus in this thesis was on the Work Team. Interviews with two POT members were cancelled due to time constraints and overall, the number of interviews with Work Team members exceeded those with POT members because of easier access. The findings may thus be biased by a deeper engagement with the issues mentioned by Work Team members and lacking a management perspective. For further research in this area, it is therefore advisable to expand the number of interviews and to include several perspectives from different roles. However, as prior empirical research suggests that leadership plays a significant role in knowledge processes (see Krogh, Nonaka, & Rechsteiner, 2012 for a review), I at least tried to superficially reflect on the potential effects of project management and leadership on the knowledge integration process. The data in the present study indicate that the locus of knowledge integration may shift from the individual work team members to the Product Owner Team under situations of uncertainty, as team members increasingly demand decisions by the POT. Thus, future research on the process of knowledge integration in agile new product development teams should be more responsive to the management and leadership role of the POT in knowledge processes. Furthermore, it would be interesting to elaborate more on how POT members handle the high information load in speedy decision-making. The data in this study imply that POT members are aware of their limited rationality and instead rely on gut feeling and the way things are presented to them in making decisions under time pressure.

A third boundary condition is the lack of attention on factors that might disrupt individual-level knowledge integration that arose as a result of the outshining results on the group-level. As already outlined, the team members are confronted with a lot of information in the formal sprint meetings, which can lead to information overload, especially when environmental uncertainty or task novelty are high. Excess information poses heavy demands on individuals' information-processing capacities and may overstrain their cognitive abilities (e.g. Kieser, 2001, p. 244), prompting them to focus increasingly on their own work and blanking out any other information that has no direct effect on the task to complete. Thus, information overload might impede cross-functional knowledge integration by forcing individual team members to isolate themselves and disregard less relevant information to maintain their cognitive capacity, even if the overlooked information could be useful and enrich individual problem-solving. Future research should pay more attention to the possible disruptions of individual-level knowledge integration and their implications for the process of knowledge integration. In general, future research should adopt a multi-level perspective in studying knowledge integration to account for the reciprocal effects between individual, group, and even organizational-level knowledge integration. For example, differences in the individual cognitive capacities of team members might influence the effective integration of new knowledge at the group-level (Felin & Hesterly, 2007, p. 212).

5.3. Practical implications

Effective knowledge integration across functions and areas of expertise is an important success factor of agile teamwork. In creating an innovation promotive climate and mindset, the agile way of working offers advantages over traditional forms of project management, if managed thoroughly. From the insights of this empirical study on the process of knowledge integration, some practical implications for companies planning to establish or already utilizing agile teams in their product development can be derived. The central suggestion that arises from the present thesis is that effectiveness and efficiency of the knowledge integration process can be influenced by actively managing the internal factors that tend to erode the effects of support mechanisms. This may be achieved by taking different measures. First, to reduce the internal conflicts that may lead to a disruption in the knowledge integration process, the tendency of team members to absent themselves from meetings due to resource scarcity should be reduced. While it is important to carry out the agile ceremonies strictly in the implementation phase in order to develop a kind of routine and promote the agile mindset, in later phases it is just as important to maintain flexibility by adapting the intensity of the agile methods to the present requirements of the team. The efficiency and effectiveness of formal interactions can vary over the course of a project, depending on the project phase, the project context, and the quality of teamwork. Therefore, the meeting intensity should always be adapted to the current needs of the team to better manage resources. Second, verbal knowledge exchange in meetings may be substituted with a better documentation

system to avoid conflicts that arise from a lack of knowledge. For example, the PowerPoint presentation of the DEMO could be used as record of the sprint transitions, so that team members that were unable to attend a meeting can be kept up-to-date, without additional dialogic effort. Precondition for the virtue of this approach are detailed reports on solutions or problems that are comprehensible without further explanations. Thus, team members would need to put more effort in setting up the presentation. A third measure could be to limit the involvement of team members in agile projects. The time spent in meetings represents a considerable resource expenditure for both experts and POT members that are involved in multiple projects. Absence and unpreparedness of team members might be reduced by providing them with more time through a limitation of project engagements. In addition, it is important to note that the agile methodology may not be appropriate for all kinds of projects. Development projects that are only concerned with small improvements and correction of defects that resemble more of routine work, often do not need the intense collaboration between different departments that comes along with agile project management. For these kinds of projects, it is advisable to only use individual methods from the agile toolbox, such as KANBAN to save time resources. However, in uncertain and novel contexts, in concept phases, and other phases were close coordination and collaboration are necessary, agile product development can achieve decisive advantages over conventional project management methods by promoting the exchange between departments and thus paving the way to higher innovation performance.

5.4. Conclusion

So, what is it that drives success in agile teams? From a knowledge-based view, it is the ability of team members to integrate their diverse strands of knowledge to create new knowledge, in the form of innovative new products or processes (Grant, 1996, p. 378; Iansiti & Clark, 1994, p. 557). My investigation of two agile product development teams in the manufacturing industry has shown, that the group is not necessarily at the center of the knowledge integration process, but that individual team members and their cognitive capacities play a central role in creating new knowledge. Furthermore, the study revealed that the effectiveness of both formal and informal support mechanisms is linked to contextual conditions. With the shift of the primary locus of knowledge integration from the individual to the group-level under uncertainty and novelty, mechanisms and practices beyond the routine way of knowledge integration become salient. In addition, it was shown that conflicts temporarily diminish the effect of support mechanisms, which subsequently disrupt the routine process of knowledge integration. However, actively managing potential disruptions of the knowledge integration process by flexibly adapting to externally as well as internally triggered changes in coordination and collaboration demands, can uphold innovative performance and give companies an important lead over their competition.

References

Allen, T. J. (1970). Communication networks in R & D Laboratories. *R&D Management*, *1*(1), 14–21.

Argyris, C., & Schön, D. A. (1978). *Organizational learning. Organization development series*. Addison-Wesley Pub. Co.

Bechky, B. A. (2003). Sharing Meaning Across Occupational Communities: The Transformation of Understanding on a Production Floor. *Organization Science*, *14*(3), 312–330.

Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Highsmith, J., … Thomas, D. (2001). *Principles behind the Agile Manifesto*.

Berg, B. L. (2001). *Qualitative research methods for the social sciences*. Allyn and Bacon.

Boland, R. J., & Tenkasi, R. V. (1995). Perspective Making and Perspective Taking in Communities of Knowing. *Organization Science*, *6*(4), 350–372.

Brinkmann, S. (2014). Unstructured and semi-structured interviewing. In *Oxford library of psychology. The Oxford handbook of qualitative research* (pp. 277–299). Oxford University Press.

Brown, J. S., & Duguid, P. (1991). Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation. *Organization Science*, *2*(1), 40–57.

Brown, J. S., & Duguid, P. (1998). Organizing Knowledge. *California Management Review*, *40*(3), 90–111.

Brown, S., & Eisenhardt, K. M. (1995). Product Development: Past research, present findings, and future directions. *Academy of Management Review*, *20*(2), 343–378.

Bruns, H. C. (2013). Working Alone Together: Coordination in Collaboration across Domains of Expertise. *Academy of Management Journal*, *56*(1), 62–83.

Bryant, A., & Charmaz, K. (2007). *The SAGE handbook of grounded theory*. Sage.

Bryman, A. (2000). *Quantity and quality in social research*. Routledge.

Carlile, P. R. (2002). A Pragmatic View of Knowledge and Boundaries: Boundary Objects in New Product Development. *Organization Science*, *13*(4), 442–455.

Carlile, P. R. (2004). Transferring, Translating, and Transforming: An Integrative Framework for Managing Knowledge Across Boundaries. *Organization Science*, *15*(5), 555–568.

Carlile, P. R., & Rebentisch, E. S. (2003). Into the Black Box: The Knowledge Transformation Cycle. *Management Science*, *49*(9), 1180–1195.

Clark, K. B., & Fujimoto, T. (1991). *Product development performance: Strategy, organization, and management in the world auto industry*. Harvard Business School Press.

Clark, S. M., Gioia, D. A., Ketchen, D. J., & Thomas, J. B. (2010). Transitional Identity as a Facilitator of Organizational Identity Change during a Merger. *Administrative Science Quarterly*, *55*(3), 397–438.

Corley, K. G., & Gioia, D. A. (2004). Identity Ambiguity and Change in the Wake of a Corporate Spin-off. *Administrative Science Quarterly*, *49*(2), 173–208.

Cronin, M. A., & Weingart, L. R. (2007). Representational gaps, information processing, and conflict in functionally diverse teams. *Academy of Management Review*, *32*(3), 761–773.

Cyert, R. M., & March, J. (1963). *A behavioral theory of the firm*. Prentice-Hall.

Cyert, R. M., & March, J. G. (1992). *A behavioral theory of the firm*. Blackwell.

Deken, F., Carlile, P. R., Berends, H., & Lauche, K. (2016). Generating Novelty Through Interdependent Routines: A Process Model of Routine Work. *Organization Science*, *27*(3), 659–677.

Demsetz, H. (1991). The theory of the firm revisited. In O. E. Williamson (Ed.), *The nature of the firm: Origins, evolution, and development* (pp. 155–178). Oxford Univ. Press.

Donnellon, A., Gray, B., & Bougon, M. G. (1986). Communication, Meaning, and Organized Action. *Administrative Science Quarterly*, *31*(1), 43.

Dougherty, D. (1992). Interpretive Barriers to Successful Product Innovation in Large Firms. *Organization Science*, *3*(2), 179–202.

Edmondson, A., & Nembhard, I. M. (2009). Product Development and Learning in Project Teams: The Challenges Are the Benefits. *Journal of Product Innovation Management*, *26*(2), 123–138.

Enberg, C. (2012). Enabling knowledge integration in coopetitive R&D projects — The management of conflicting logics. *International Journal of Project Management*, *30*(7), 771–780.

Enberg, C., Lindkvist, L., & Tell, F. (2006). Exploring the Dynamics of Knowledge Integration. *Management Learning*, *37*(2), 143–165.

Faraj, S., & Sproull, L. (2000). Coordinating Expertise in Software Development Teams. *Management Science*, *46*(12), 1554–1568.

Faraj, S., & Xiao, Y. (2006). Coordination in Fast-Response Organizations. *Management Science*, *52*(8), 1155–1169.

Felin, T., & Hesterly, W. S. (2007). The Knowledge-Based View, Nested Heterogeneity, and New Value Creation: Philosophical Considerations on the Locus of Knowledge. *Academy of Management Review*, *32*(1), 195–218.

Flick, U. (2009). *An introduction to qualitative research*. Sage.

Galbraith, J. R. (1973). *Designing Complex Organizations*. Addison-Wesley Pub. Co.

Galbraith, J. R. (1974). Organization Design: An Information Processing View. *Interfaces*, *4*(3), 28–36.

Galison, P. L. (2000). *Image and logic: A material culture of microphysics*. Univ. of Chicago Press.

Gioia, D. A., & Chittipeddi, K. (1991). Sensemaking and sensegiving in strategic change initiation. *Strategic Management Journal*, *12*(6), 433–448.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research. *Organizational Research Methods*, *16*(1), 15–31.

Gioia, D. A., Price, K. N., Hamilton, A. L., & Thomas, J. B. (2010). Forging an Identity: An Insider-outsider Study of Processes Involved in the Formation of Organizational Identity. *Administrative Science Quarterly*, *55*(1), 1–46.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*(2), 109–122.

Grunwald, R. (2003). *Inter-organisationales Lernen*. Deutscher Universitätsverlag.

Hargadon, A., & Sutton, R. I. (1997). Technology Brokering and Innovation in a Product Development Firm. *Administrative Science Quarterly*, *42*(4), 716.

Harrison, S. H., & Corley, K. G. (2011). Clean Climbing, Carabiners, and Cultural Cultivation: Developing an Open-Systems Perspective of Culture. *Organization Science*, *22*(2), 391–412.

Hasenzagl, R., & Link, K. (2017). Agil: Das neue Paradigma in der Unternehmensführung? *Austrian Management Review*(7), 47–55.

Hippel, E. v. (1988). *The Sources of Innovation*. MIT Press.

Hoegl, M., & Gemuenden, H. G. (2001). Teamwork Quality and the Success of Innovative Projects: A Theoretical Concept and Empirical Evidence. *Organization Science*, *12*(4), 435–449.

Iansiti, M., & Clark, K. (1994). Integration and Dynamic Capability: Evidence from Product Development in Automobiles and Mainframe Computers. *Industrial and Corporate Change*, *3*(3), 557–605.

Kellogg, K. C., Orlikowski, W. J., & Yates, J. (2006). Life in the Trading Zone: Structuring Coordination Across Boundaries in Postbureaucratic Organizations. *Organization Science*, *17*(1), 22–44.

Kieser, A. (2001). Trust as a Change Agent for Capitalism or as Ideology? A Commentary. *Organization Science*, *12*(2), 241–246.

Kieser, A., Beck, N., & Tainio, R. (2001). Rules and Organizational Learning: The Behavioral Theory Approach. In M. Dierkes, A. Antal, J. Child, & I. Nonaka (Eds.), *Handbook of organizational learning and knowledge* (pp. 598–626). Oxford Univ. Press.

Kieser, A., & Koch, U. (2008). Bounded Rationality and Organizational Learning Based on Rule Changes. *Management Learning*, *39*(9), 329–347.

Kim, D. J. (1993). The link between individual and organizational learning. *Sloan Management Review*, *35*(1), 37–50.

Kogut, B., & Zander, U. (1992). Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science*, *3*(3), 383–397.

Krogh, G., Nonaka, I., & Rechsteiner, L. (2012). Leadership in Organizational Knowledge Creation: A Review and Framework. *Journal of Management Studies*, *49*(1), 240–277.

Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. Sage.

Lave, J., & Wenger, E. (1991). *Situated Learning*. Cambridge University Press.

Lawrence, P., & Lorsch, J. (1967). *Organizations and Environments: Managing Differentiation and Integration*. Harvard Business School Press.

Lawrence, P., & Lorsch, J. (1969). *Organization and Environment: Managing Differentiation and Integration*. Richard D. Irwin.

Leonard-Barton, D. (1995). *Wellsprings of knowledge: Building and sustaining the sources of innovation*. Harvard Business School Press.

Lok, J., & Rond, M. d. (2013). On the Plasticity of Institutions: Containing and Restoring Practice Breakdowns at the Cambridge University Boat Club. *Academy of Management Journal*, *56*(1), 185–207.

Majchrzak, A., More, P. H. B., & Faraj, S. (2012). Transcending Knowledge Differences in Cross-Functional Teams. *Organization Science*, *23*(4), 951–970.

March, J. G., & Olsen, J. P. (1975). The uncertainty of the past: organizational learning under ambuguity. *European Journal of Political Research*, *3*(2), 147–171.

Mengis, J., Nicolini, D., & Swan, J. (2018). Integrating knowledge in the face of epistemic uncertainty: Dialogically drawing distinctions. *Management Learning*, *49*(5), 595–612.

Nag, R., Corley, K. G., & Gioia, D. A. (2007). The Intersection of Organizational Identity, Knowledge, and Practice: Attempting Strategic Change Via Knowledge Grafting. *Academy of Management Journal*, *50*(4), 821–847.

Nag, R., & Gioia, D. A. (2012). From Common to Uncommon Knowledge: Foundations of Firm-Specific Use of Knowledge as a Resource. *Academy of Management Journal*, *55*(2), 421–457.

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. *Organization Science*, *5*(1), 14–37.

Nonaka, I., & Takeuchi, H. (1995). *The knowledge creating company: How Japanese companies create the dynamics of innovation*. Oxford Univ. Press.

Okhuysen, G. A., & Bechky, B. A. (2009). 10 Coordination in Organizations: An Integrative Perspective. *Academy of Management Annals*, *3*(1), 463–502.

Pratt, M. G. (2008). Fitting Oval Pegs Into Round Holes. *Organizational Research Methods*, *11*(3), 481–509.

Qu, S. Q., & Dumay, J. (2011). The qualitative research interview. *Qualitative Research in Accounting & Management*, *8*(3), 238–264.

Rigby, D. K., Sutherland, J., & Takeuchi, H. (2016). Embracing Agile. *Harvard Business Review*, *94*(5), 40–50.

Sandberg, J., & Tsoukas, H. (2011). Grasping the Logic of Practice: Theorizing Through Practical Rationality. *Academy of Management Review*, *36*(2), 338–360.

Schmickl, C., & Kieser, A. (2008). How much do specialists have to learn from each other when they jointly develop radical product innovations? *Research Policy*, *37*(3), 473–491.

Schüppel, J., Müller-Stewens, G., & Gomez, P. (1998). The Knowledge Spiral. In D. Kleine, G. von Krogh, & J. Roos (Eds.), *Knowing in firms: Understanding, managing, and measuring knowledge* (pp. 223–239). Sage.

Schröder, A. (2017). *Agile Produktentwicklung: Schneller zur Innovation - erfolgreicher am Markt*. Hanser.

Schreyögg, G., & Eberl, P. (1998). Organisationales Lernen: Viele Fragen, noch zu wenige neue Antworten. *Die Betriebswirtschaft*, *58*(4), 516–536.

Shannon, C., & Weaver, W. (1949). *The Mathematical Theory of Communications*. University of Illinois Press.

Strauss, A. L., & Corbin, J. M. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publications.

Takeuchi, H., & Nonaka, I. (1986). The new new product development game. *Harvard Business Review*(1), 137–146.

Tracy, S. J. (2010). Qualitative Quality: Eight "Big-Tent" Criteria for Excellent Qualitative Research. *Qualitative Inquiry*, *16*(10), 837–851.

Yin, R. K. (2003). Case study research: Design and methods (3. ed.). *Applied social research methods series*, *5*.

Junior Management Science 7(2) (2022) 390-411

Junior Management Science

journal homepage: www.jums.academy

Personality as a Determinant of Customer Experience Quality and Value-in-Use in a Public Crisis Situation – The Case of E-Learning

Katharina Lohmann

Freie Universität Berlin

Abstract

This thesis was designed to investigate whether differences exist between students' personalities regarding their perceived Customer Experience Quality and their perceived Value-in-Use, for the case of e-learning. In particular, the personality dimension Introversion-Extraversion was investigated. Furthermore, it was examined whether students' Fear of the Coronavirus Disease 2019 moderates the relationship between Introversion-Extraversion and the perceived Customer Experience Quality as well as the relationship between Introversion-Extraversion and the perceived Value-in-Use. Using survey data on asynchronous e-learning as well as survey data on synchronous e-learning, multiple two-way ANOVAs were conducted. It was found that no significant differences for either asynchronous e-learning or synchronous e-learning regarding the Customer Experience Quality and the Value-in-Use exist between introverted and extraverted students. Furthermore, no significant interaction effect on the perceived Customer Experience Quality and Value-in-Use was found for introverted and extraverted students with either low or high Fear of the Coronavirus Disease 2019.

Keywords: COVID-19; e-learning; introversion-extraversion; customer experience quality; value-in-use; two-way ANOVA.

1. Introduction

University lectures are usually characterized by large numbers of participants and a predominantly frontal teaching approach. Students follow presentations in real time, take notes and use additional online materials only for preparation or follow-up of teaching content. The central tasks of the lecturer involve preparing and presenting his[1] lecture content, communicating with students and organizing his courses (Freie Universität Berlin, 2020a). In 2017, around 90 percent of university lecturers stated that they would supplement their teaching with digital elements, such as the use of learning management platforms. However, only 42 percent of lecturers actually used blended learning formats, i.e. communication consisting of classroom teaching combined with e-learning, frequently or occasionally. On the student side, 61 percent of the students surveyed were in favor of blended learning formats. Nevertheless, a preference for pure e-learning formats did not appear to exist since 68 percent of the students surveyed continued to find classic teaching tools, such as blackboards and books, motivating (Schmid, Goertz, Radomski, Thom, & Behrens, 2017).

Until the Coronavirus Disease 2019 (COVID-19), e-learning services at Freie Universität Berlin were also only used as an addition to classroom teaching. However, as a result of the risk posed by COVID-19, in 2020 the university was forced to conduct the entire summer semester digitally, with possibly more semesters with a large proportion of online courses to follow (Freie Universität Berlin, 2020e). For universities, such as Freie Universität Berlin, which offer teaching services, it is therefore important to know the factors that influence students' perception of the quality of their e-learning experiences and ultimately their perceived value through e-learning.

Since personality traits of individuals are relatively stable over time (Ashton, 2013; Johnson, 1997), individuals' personality could be a crucial factor in this respect. In particular, the fact whether a person is intro- or extraverted might influence the perception of their e-learning experience as Introversion-Extraversion influences how information is pro-

[1] Words used herein, regardless of the number and gender specifically used, shall be deemed and construed to include any other number, singular or plural, and any other gender, masculine, feminine or neuter, as the context requires.

DOI: https://doi.org/10.5282/jums/v7i2pp390-411

cessed by individuals and which communication and learning settings are preferred (Capretz & Ahmed, 2010; Felder & Silverman, 1988). Due to the fact that e-learning fundamentally changed used teaching and communication tools, this thesis will investigate the research question of whether significant differences between introverted and extraverted students regarding their perceived Customer Experience Quality and their perceived Value-in-Use exist, for the case of e-learning at Freie Universität Berlin. In doing so, synchronous and asynchronous e-learning services are compared with each other.

Given that the outbreak of COVID-19 was the reason for the digital summer semester 2020, this thesis will also investigate how students' perception of COVID-19 influences their Customer Experience Quality and Value-in-Use regarding e-learning. For instance, de Keyser, Verleye, Lemon, Keiningham, and Klaus (2020) already argued for the need for research on the impact of the COVID-19 pandemic on Customer Experiences. It was found that the fear of diseases triggers uncertainty and irrational behavior and can thus influence the way individuals think and behave (Pappas, Kiriaze, Giannakis, & Falagas, 2009). It was also found that COVID-19 can cause fear in people (Ahorsu et al., 2020). The mass media, social isolation, the fear of losing loved ones or the insecurity of losing one's job are only a few reasons for this (Pappas et al., 2009). In particular, introverted personalities are attributed a higher level of anxiety, compared to extraverted personalities (Dewaele & Furnham, 1999; Eysenck, 1965). Therefore, this thesis will investigate the research question of whether there is a significant interaction effect of Introversion-Extraversion and Fear of COVID-19 on students' Customer Experience Quality and Value-in-Use.

The second section of this thesis provides an overview of the theoretical background of the research topic under investigation. As part of this, an overview of the e-learning services of Freie Universität Berlin is given and the theoretical constructs of Value-in-Use, Customer Experience Quality, Introversion-Extraversion as well as of Fear of COVID-19 are introduced. In the third section, the hypotheses to be tested are derived on the basis of the reviewed literature and past study results which result in the research model of this thesis. After an overview of the quantitative method approach in the fourth section, in the fifth section the results are presented and interpreted with respect to the research questions. The thesis ends with the discussion part in the sixth section, including practical implications as well as an outlook for possible future research.

2. Theoretical Background

2.1. E-Learning

2.1.1. Distance Education and E-Learning

Although there is no general agreement about the definition of "Learning", many researchers define it as a permanent change in behavior resulting from practice or experience (Baron, 2001; Logan, 1970; Myers, 2014). Lachmann (1997) established a less absolute definition of learning and defines it as a process rather than a permanent change of behavior, whereby experiences can contribute to learning processes:

> *"Learning is the process by which a relatively stable modification in stimulus-response relations is developed as a consequence of functional environmental interaction via the senses" (Lachmann, 1997, p. 477).*

In recent decades, the use of information and communication technologies has provided new opportunities for learning services which also enabled the physical separation of teachers and students in higher education, for instance. These distance education systems deliver education at a distance by video, interactive audio or computers, while simultaneously meeting the needs of individuals (Zigerell, 1984). According to Guri-Rosenblit (2005), although "Distance Education" and "E-Learning" are not interchangeable terms, e-learning can be used for distance educational purposes. In literature, e-learning is often considered as a new generation of distance education and definitions from different perspectives exist. Emphasizing the technological aspects of e-learning, it refers to *"the use of electronic media for a variety of learning purposes that range from add-on functions in conventional classrooms to full substitution for the face-to-face meetings by online encounter."* (Guri-Rosenblit, 2005, p. 469). Thus, e-learning is a form of e-service whose delivery depends on information technologies (Rowley, 2006).

Ebner et al. (2013) take a closer look at the interaction aspects of e-learning. Although students have to sit alone in front of a technical device, they do not necessarily have to learn in isolation. Rather, there is an opportunity for exchange via digital communication channels. In this context, teachers act as e-moderators who control and structure the communication and exchange of learning groups in a targeted manner (Ebner et al., 2013). As parameters for forms of e-learning communication, Ebner et al. (2013) introduce three parameters: "Supervision", "Time Dimension" and "Relationship of the Participants", i.e. who communicates with whom. Teachers can provide supervision or offer courses without additional supervision. Regarding time, e-learning models can be synchronous, asynchronous, or a mix of the two. If communication with and between students takes place in real time, this is referred to as "synchronous" communication, which is the case with live lectures or live chats, for instance. On the other hand, communication can also take place "asynchronously", that is time-delayed, which is the case with e-mail exchanges or recorded video lectures, for instance. With regard to the relationship of participants, a distinction can be made between direct one-to-one exchange (1:1), the exchange of individual teachers with several students (1:n) and the exchange of a large number of participants on one platform (n:n) (Ebner et al., 2013).

2.1.2. E-Learning at Freie Universität Berlin

Freie Universität Berlin (FUB) defines e-moderation and the supervision of students, parallel to the provision of digital

teaching materials, as central success factors for digital teaching and mainly relies on Blackboard (BB) for asynchronous e-learning purposes and on real-time meetings via Cisco Webex (Webex) for synchronous e-learning purposes (Freie Universität Berlin, 2020c).

BB is the central learning management system at FUB and has been used to support teaching at the university since 2004 by organizing and structuring learning content. The platform is provided by an external firm and can be accessed from the computer as well as from mobile devices, for which a corresponding app can be downloaded. In the summer semester 2020, for instance, 1,738 lecturers and 22,255 students were enrolled in the 2,467 actively used BB courses during the semester (Freie Universität Berlin, 2020d). In BB, each student has a personal profile and the opportunity to access attended courses and the associated teaching material, such as recorded lectures, video tutorials, lecture slides or literature. The platform is frequently used for organizational purposes, such as the announcement of dates or other relevant announcements (Freie Universität Berlin, 2019). Furthermore, BB provides some communication networks for asynchronous communication forms. For instance, it is possible for teachers to set up working groups for students. In addition, wikis, discussion forums or office hours blogs can be created as a forum for collaboration and course participants have the possibility to send messages to teachers or to other students (Freie Universität Berlin, 2019). BB is therefore suitable for asynchronous 1:1 communication as well as for asynchronous 1:n or n:n communication (Ebner et al., 2013).

In addition to the resources provided and the time-shifted communication possibilities on BB, online lectures can be held in real time using software, such as Webex. Within the program, lecturers and students meet in a virtual room and are connected via video and audio transmission. The synchronous communication is particularly suitable for holding live lectures and live office hours, but also for student learning in the context of smaller group work. In addition to the possibilities of audio and video sharing, Webex offers the possibility of screen transmission and interaction within a live chat (Freie Universität Berlin, 2020a). Therefore, it is suitable for synchronous 1:1 communication as well as for synchronous 1:n and n:n communication (Ebner et al., 2013).

2.2. Value-in-Use

2.2.1. Service Dominant Logic

In order to capture the construct "Value" with regard to the research purpose of this thesis and make statements about how it is generated, different perspectives can be taken. The "Goods Dominant Logic", for instance, takes a supply-side perspective and focuses on tangible goods and their embedded value, which means that a supplier can determine the value of a good by embedding it in the goods it produces (Vargo & Lusch, 2004). The ultimate value a supplier receives in exchange for these tangible goods is called "Value-in-Exchange" (Sandström, Edvardsson, Kristensson, & Magnusson, 2008). However, especially for services that require the involvement of customers (Gabler, 2020), this perspective seems inadequate which is why Vargo and Lusch (2004) propose a shift in perspective towards a "Service Dominant Logic" (SDL). Its origins can be traced back to Carl Menger in the 19^{th} century, who already regarded value as subjective and argued that a good must be able to be used in order to satisfy needs. Otherwise, it is useless and no longer a good (Menger, 1968).

SDL adopts a customer-centric and relational perspective and shifts its focus from the product exchanged and physical resources to the process of exchange as well as to knowledge and skills as primary resources. From the perspective of SDL, the "Value-in-Use" (ViU) of a customer can be defined as the extent to which the customer's goals have been achieved through the use of a service, at which the role of the provider is to formulate value propositions (MacDonald, Wilson, Martinez, & Toossi, 2011; Payne, Storbacka, & Frow, 2008). Customer goals determine the relative salience of service features so that customers are aware of them and attach meaning to them (Puccinelli et al., 2009; Woodruff & Flint, 2006). As an active co-creator, the customer defines his individual ViU by integrating knowledge and skills within the framework of his usage processes (Vargo & Lusch, 2004). This co-creation interaction between supplier and customer implies that exchange is a relational construct from the SDL perspective (Vargo & Lusch, 2004). The case of e-learning illustrates this as students and instructors need to interact on electronic engagement platforms in order to exchange knowledge and skills (Kleinaltenkamp, Storck, Gumprecht, & Li, 2018).

2.2.2. Dimensions of Value-in-Use

In the literature, various approaches exist to conceptualize "Value". Zeithaml (1988), for instance, offers a one-dimensional model approach and conceptualizes value as the customer's overall assessment of the utility of a product based on what is given and what is received. In contrast, there are multidimensional model approaches that define value as a multidimensional construct and consider various value components (Sheth, Newman, & Gross, 1991; Sweeney & Soutar, 2001). Sweeney and Soutar (2001), for instance, developed the "Perceived Value (PERVAL) Model", distinguishing between a "Functional", "Emotional" and "Relational" value dimension which together make up the value of an offer. The functional value of an offer is made up of the two subcomponents "Price" and "Quality", whereby the utilitarian benefit results from the perceived performance of an offer and the perceived price-performance ratio. The emotional value refers to the hedonistic value of an offer, which results from the feelings or affective states that an offer generates (Sweeney & Soutar, 2001). The social or relational value involves the role of other actors, whose behavior, motivation and support can also be crucial in generating value for customers in terms of their personal status or recognition by others (Lemke, Clark, & Wilson, 2011; Sweeney & Soutar, 2001). Although Sweeney and Soutar's (2001) PER-

VAL model can be used as a starting point for the development of value measurement tools, it focusses on the assessment of value before usage. This is reflected, for instance, in the relational value dimension where the actual interaction between users is not taken into account, but only the recognition by others.

Following the SDL, other research focuses instead on customer value in terms of ViU and transfers the multidimensional understanding of value to customer usage processes (Bruns & Jacob, 2014, 2016; Kleinaltenkamp, Storck, et al., 2018). Referring to the concrete case of e-learning, Kleinaltenkamp, Storck, et al. (2018) conducted a qualitative study at a German university and identified the individual ViU dimensions of students and lecturers for the case of asynchronous e-learning[2]:

- *Task Simplification*
- *Pressure Reduction*
- *Flexibility*
- *Cost Decrease*
- *Hedonistic Benefit*
- *Motivation*
- *Personal Self-Fulfillment*
- *Proficiency*
- *Self-Portrayal*
- *Uncertainty Reduction*
- *Perceived Control*

As already stated, according to Sweeney and Soutar (2001), the value of an offer is made up of a functional, an emotional and a relational value dimension. The eleven ViU dimensions identified by Kleinaltenkamp, Storck, et al. (2018) contain more functional goals, such as "Task Simplification" through time savings and reduction of effort through the use of e-learning, "Proficiency", "Cost Decrease" as well as "Flexibility" regarding time and space. In addition, Kleinaltenkamp, Storck, et al. (2018) identified a "Hedonistic Benefit" dimension for the case of e-learning. Furthermore, they found that individuals felt affective benefits in the perceived control over processes and resources that affected their own work and also perceived an uncertainty reduction due to less misinformation and communication. Because of e-learning, students achieved a reduction in stress (Pressure Reduction) during their daily tasks as well as pleasure and interest in fulfilling their tasks (Motivation). On a personal level, e-learning contributed to the fulfilment of personal ambitions and interests (Personal Self-Fulfillment) and offered the opportunity to demonstrate abilities and achievements during one's own work (Self-Portrayal) (Kleinaltenkamp, Storck, et al., 2018).

Other actors' behavior, motivation and support might also lead to a perceived "Relational Value" regarding e-learning (Lemke et al., 2011; Sweeney & Soutar, 2001). Since both asynchronous and synchronous forms of e-learning are investigated in the context of this thesis, an additional relational ViU dimension is considered. Compared to asynchronous e-learning settings, particularly synchronous e-learning settings offer additional interaction possibilities so that students and lecturers are able to exchange knowledge in real time (Freie Universität Berlin, 2020a; Sandström et al., 2008).

2.3. Customer Experience Quality

Already in the late 1930s, Keynes (1936) argued that customers buy products to satisfy their desire for experiences. Researchers like Holbrook and Hirschmann (1982) were among the first to study experiential aspects of consumption and according to Holbrook (1999), the experiences that customers make create value. Furthermore, Pine and Gilmore (1998) coined the term "Experience Economy" and stated that experiences should be treated as an economic offering and that merely providing services is no longer sufficient. Becker and Jaakkola (2020) provide a definition for the construct of "Customer Experience" (CE) and define it as *"non-deliberate, spontaneous responses and reactions to particular stimuli"* (p. 637), deriving from various touchpoints with the provider. Those responses include the customer's subjective cognitive, affective, emotional, social and physical responses over the time of the customer journey (Verhoef et al., 2009).

CEs can vary in terms of the nature of the touchpoint, as it can be human, physical, digital or a combination thereof and depend on the specific stage of the customer journey in which they take place - pre-purchase, purchase or post-purchase. Furthermore, the individual context, that is the transitory personal state of the customer as well as the social context - emerging from social groups - and the environmental context - composed of natural, economic, public or political externalities - can be determining factors of CEs (de Keyser et al., 2020). With regard to the SDL, CEs are influenced both by elements that the provider can control and by elements over which the provider has no control which is why value is co-created and the provider's role is to deliver value propositions (Lemke et al., 2011; Vargo & Lusch, 2004; Verhoef et al., 2009). The "Customer Experience Quality" (CEQ) finally comprises the perceived superiority or excellence of the CE and is judged individually with respect to its contribution to ViU, i.e. abstract goals (Edvardsson, 2005; Lemke et al., 2011; Sandström et al., 2008). According to Lemke et al. (2011), with reference to Payne et al. (2008), the perceived CEQ is composed of different dimensions and *"goes beyond the notion of service quality"* (p. 859). These dimensions can be assigned to three encounters: the "Service Encounter", the "Communication Encounter" and the "Usage Encounter" (Lemke et al., 2011).

[2] See Table 2 for further details.

Service Encounter:
The Service Encounter includes categories regarding the quality of the service, for instance in terms of the quality of the offer itself or the accessibility and reliability of the service personnel. In order to measure the perceived Service Quality (SQ), Parasuraman, Zeithaml, and Berry (1988) developed the "SERVQUAL Scale" consisting of 22 items, which cover the following aspects (Parasuraman et al., 1988, p. 23):

1. *Tangibles,* which includes items regarding physical facilities, equipment and appearance of personnel.
2. *Reliability,* which contains items regarding the ability to perform the promised service dependably and accurately.
3. *Responsiveness,* which refers to the willingness to help customers and provide prompt service.
4. *Assurance,* which relates to the knowledge and courtesy of employees and their ability to inspire trust and confidence.
5. *Empathy,* which refers to caring, individualized attention the firm provides its customers.

With the increasing use of the Internet, the focus of service research has shifted from the investigation of offline CEs and offline CEQ to the investigation of e-CE and e-CEQ (Elsharnouby & Mahrous, 2015; Parasuraman, Zeithaml, & Malhorta, 2005; Zeithaml, Parasuraman, & Malhorta, 2002). Compared to offline communication, the specifics of online communication lie above all in the interactivity and multimedia, the lower degree of personal contact, the speed and range of information dissemination as well as the heterogeneity of users (Frohne, 2020; Rose, Hair, & Clark, 2011). Santos (2003) defines "E-SQ" as the *"customers' overall evaluation and judgment regarding the excellence and quality of e-service delivery in the virtual market place"* (p. 235). E-SQ can be assessed through measurement scales that include both the evaluation of system attributes, such as the ease of use or design, as well as the evaluation of service attributes, such as the availability of the service personnel (Elsharnouby & Mahrous, 2015; Gera, 2011; Parasuraman et al., 2005). For the special case of asynchronous e-learning, Udo, Bagchi, and Kirs (2011) modified the SERVQUAL scale in order to assess the SQ of e-learning. In particular, the authors renamed the "Tangibles" dimension to "Web Content" since online settings have fewer physical attributes and are more multimedia in nature.

Communication Encounter:
According to Lemke et al. (2011), the construct of CEQ not only includes the perceived SQ, but also the perceived quality of the relationship with the provider which can be assigned to the "Communication Encounter". In the context of this thesis, from the students' perspective, the provider is the course instructor. The perceived "Relationship Quality" (RQ) with a provider results from the overall assessment by the customer of the quality of the interaction and the strength of the relationship with the provider (de Wulf, Odekerken-Schröder, & Dawn, 2001; Gummesson, 1987). RQ is a multidimensional construct for which several conceptualization approaches exist (de Canniere, de Pelsmacker, & Geuens, 2009; de Wulf et al., 2001). Keating, Rugimbana, and Quazi (2003), for instance, adapted Page's (2000) conceptualization of RQ for the online retail context. The findings of their study show that trust in the provider, the perceived amount of effort a provider makes as well as how a provider values its customers, understands and communicates with them form the RQ construct. For the context of e-learning, the findings of Swan (2002) as well as of Kuo, Walker, Belland, and Schroeder (2013) illustrate the importance of interaction with course instructors as it contributed significantly to students' satisfaction with e-learning and learning from online courses in general.

Usage Encounter:
The e-learning services offered by FUB are delivered in settings where many students are present at the same time. Therefore, in the context of value co-creation, relational experiences with peer students also play a role for the overall CEQ (Schouten, McAlexander, & Koenig, 2007; Verleye, 2015). Lemke et al. (2011) summarize the RQ with other customers under the "Usage Encounter", with peer customers and their identities being involved in value creation. Several studies found that other customers' behavior influences CEs and can influence the satisfaction with a service as well as the emotional and behavioral responses toward a service significantly (Grove & Fisk, 1997; Hui & Bateson, 1991; Martin & Pranter, 1989; Thakor, Suri, & Saleh, 2008). With regard to e-learning, Swan (2002) discovered that student-student interaction influenced students' satisfaction as well as their perception of learning from an online course. Gomez-Rey, Barbera, and Fernandez-Navarro (2016) found that students considered interaction with other students among the most important variables when evaluating the quality of their e-learning experience.

2.4. Personality and Introversion-Extraversion

Customers' evaluation of their CE and their perceived ViU is subjective and is influenced by customer specifics, such as demographic characteristics or personality traits (Becker & Jaakkola, 2020; Holbrook & Hirschmann, 1982; Sandström et al., 2008; Verhoef et al., 2009), which can act as individual filters that include *"everything connected to the individual user"* (Sandström et al., 2008, p. 116). Focusing on the impact of "Personality" in this thesis, Allport (1937) was among the first researchers who shaped the field of personality psychology and defined an individual's personality as *"the dynamic organization within the individual of those psychophysical systems that determine his characteristic behavior"* (Allport, 1961, p. 28). He and other researchers stated that the construct of Personality is composed of several lesser elements, that is a unique pattern of personality traits, which cause individual differences between people (Baughman & Welsh, 1962; Guilford, 1959). According to Ashton (2013),

personality traits refer to *"differences among individuals in a typical tendency to behave, think, or feel in some conceptually related ways, across a variety of relevant situations and across some fairly long period of time."* (p. 27). Therefore, personality traits show up over a longer period of time and represent structures or habits within a person that imply similar reactions in similar situations (Allport, 1937, 1961; Johnson, 1997). Individuals even tend to look for situations and circumstances that enable trait expression, such as extraverted individuals like to spend time on events and parties that promote their gregariousness (Snyder, 1983).

One of the purposes of this thesis is to measure the "Introversion-Extraversion" (I-E) of students as part of their personality. One of the first researchers who introduced the concept of I-E in the early 1920's was Carl Gustav Jung, who considered I-E as differing orientations toward the world. According to him, extraverted people are more outward focused and oriented toward external, objective experiences. They are perceived as sociable and outgoing by others. Introverted people, in contrast, are more inward focused and oriented toward internal, subjective experiences. They tend to focus on their own thoughts and feelings which is why they are perceived by others as quiet and reserved (Jung, 1971).

Extraverted and introverted personalities also differ in the way they process information. Since extraverts are better in processing sensory information, compared to introverts, their brains search for external stimuli. Introverts, in contrast, show more brain activity in areas where problem-solving, introspection and complex thinking are carried out (Laney, 2005). Consequently, extraverted personalities learn best through active experimentation and communication, while introverted personalities learn primarily through introspective information processing and preferably alone (Felder & Silverman, 1988).

In order to measure individuals' personalities, according to McAdams (1997), it was Bernreuter (1931) who developed the first multi-trait personality inventory, which contained multiple scales to assess the six trait dimensions of "Neuroticism", "I-E", "Dominance-Submission", "Self-Sufficiency", "Confidence" and "Sociability". Over the following years, further personality inventories were developed differing in the number and nature of personality dimensions distinguished. However, almost every conceptualization contained I-E as a central personality dimension (Costa & Mccrae, 1989; Eysenck & Eysenck, 1975; Eysenck & Wilson, 1991; John, Donauhue, & Kentle, 1991).

Over time, the differentiation of five basic personality dimensions, known as the "Big Five", became established in literature (Tellegen, 1991). First mentioned by Goldberg (1981) and evidenced through research by Cattell (1946); Norman (1963); Tupes and Christal (1961) as well as by Wiggins (1968), the Big Five reduce the scope of trait dimensions to five basic bipolar dimensions: "Extraversion", "Agreeableness", "Conscientiousness", "Neuroticism" and "Openness" (McCrae, Costa, & Busch, 1986). Regarding the single dimensions, several conceptualizations have been developed over time, which partly differ regarding their associated traits (Eysenck & Eysenck, 1975; Hogan, 1983; Tellegen, 1985). In the context of this thesis, the focus is on I-E and it is referred to John and Srivastava (1999), Costa and Mccrae (1985, 1992) as well as to McCrae and Costa (1987), who consider gregariousness, assertiveness, activity, excitement-seeking, positive emotions and warmth as traits of Extraversion. Introverted personalities, in contrast, are characterized by more negative emotions as well as passive, quiet, reserved and aloof traits. Research has shown that these traits are common across different countries and cultures (McCrae & Terracciano, 2005; Schmitt, Allik, McCrae, & Benet-Martinez, 2007).

2.5. Public Crisis Situations

2.5.1. The COVID-19 Pandemic

How customers experience services and judge their experiences can be influenced not only by personality traits, but also by individually perceived contextual conditions (Chandler & Vargo, 2011). "Context" is defined as *"the relevant aspects of a situation, which are relevant for the resource-integrating activities"* (Löbler & Hahn, 2013, p. 259) and is usually only partly within the control of the provider (Chandler & Vargo, 2011; Löbler & Hahn, 2013; Lemon & Verhoef, 2016). External environments provide contextual conditions, which are dynamic and subject to continuous change (Chandler & Vargo, 2011; de Keyser et al., 2020; Lemon & Verhoef, 2016). Therefore, an *"extreme crisis can have a strong, negative, and enduring effect on the customer experience"* (Lemon & Verhoef, 2016, p. 79) and thus, on the CEQ. Extreme public crises are undesirable, unexpected phases of disorder that threaten societies and shatter their orders. According to Boin, Hart, Stern, and Sundelius (2017), the importance of the endangered values and structures determines how deeply a crisis situation is sensed by the public. The more important the endangered values and structures are, the stronger a crisis is perceived.

Pandemics are transboundary crises that threaten human health and are marked by uncertain developments regarding numerous areas of life (Boin et al., 2017). By the end of 2019, numerous cases of pneumonia with unknown cause were registered in the Chinese city of Wuhan. A short time later, a novel type of coronavirus was identified as the cause, called "COVID-19". After the virus had spread to most countries in the world within a few months and had been characterized by high infection rates as well as relatively high mortality rates, it was declared a pandemic by the World Health Organization (WHO) on March 11, 2020 (WHO, 2020a). On August 26, 2020, the WHO already reported about four million confirmed cases of COVID-19 for Europe alone (WHO, 2020b). Since the virus has developed into a pandemic that challenged fundamental social structures and values and also led to an international economic crisis, which is accompanied by a period of decline in business profitability and economic strength, it is considered an "Extreme Public Crisis" in the context of this thesis (Boin et al., 2017; Lemon & Verhoef,

2016; McKinsey & Company, 2020). Since both public and private organizations had to respond to the crisis in order to bring outbreaks of diseases under control (Boin et al., 2017), the context in which services are delivered and experienced rapidly changed with COVID-19. All countries affected by the pandemic 19 have taken social distancing measures to stop its spread, including shutting down schools and universities offering e-learning services (Berliner Vewaltung, 2020).

2.5.2. Psychological Consequences of COVID-19

Besides physical consequences for peoples' health, COVID-19 has also led to a personal crisis for many people, characterized by a period of emotional turmoil and illness (Boin et al., 2017). The psychological consequences are, for instance, reflected in emotions of fear of COVID-19 (Ahorsu et al., 2020). "Fear" is considered an adaptive, negative emotion which *"represents the reactive removal of oneself from a position of immediate risk"* (Harper, Satchell, Fido, & Latzmann, 2020, Implications of Results section). The fear of diseases results from both medical and social factors and is predominantly rooted in the past. People used to be afraid of infectious diseases as they are transmissible, imminent and invisible. The memories of former infectious diseases have led to an automatic response in human subconscious of a fear of infection (Pappas et al., 2009). Inconsistent health policies, isolation procedures, the fact that many people lose their jobs as well as the lack of opportunities for childcare and social contacts increase the fear of diseases (Manderson & Levine, 2020; Pappas et al., 2009). Nowadays, the easier import of exotic diseases into metropolitan regions, as a result of globalization, as well as the mass media are additional factors which shape the fear responses of people (ibd.).

On the one hand, the fear of an infectious disease prepares people both physically and mentally for an acute response to possible harm (Pappas et al., 2009). Furthermore, Harper et al. (2020) found that fear of COVID-19 is a predictor of positive behavior change towards more social distancing and has a protective effect during the pandemic. On the other hand, "Fear of COVID-19" was shown to increase anxiety, depression and stress as well as to decrease life satisfaction among people from different countries (Ahorsu et al., 2020; Reznik, Gritsenko, Konstantinov, Khamenka, & Isarlowitz, 2020; Satici, Gocet-Tekin, Deniz, & Satici, 2020; Soraci et al., 2020). People report fear of coming into contact with people who may be infected with COVID-19 and are afraid of infecting themselves (Lin, 2020). With high levels of fear, which is reflected, for instance, in sleep disorders, clammy hands or palpitations, Ahorsu et al. (2020) suspect individuals might not think clearly and rationally when they react to COVID-19. Thus, "Fear of COVID-19" can be a trigger for uncertainty, anxiety or irrational behavior (Pappas et al., 2009).

3. Derivation of Hypotheses and Research Model

In the following section, the presented theoretical constructs of ViU, CEQ, I-E and Fear of COVID-19 are linked to each other and hypotheses are formulated regarding their relationships in the context of e-learning at FUB. Based on the hypotheses, the research model underlying this thesis is developed.

Personality influences the perception of the external world, the decisions individuals make as well as their behavior (Allport, 1937; Ashton, 2013). Several studies also investigated learning style differences between different personality types, including the I-E dimension (Ahmed, Campbell, Jaffar, & Alkobaisi, 2010; Raju & Venugopal, 2014). It was found that introverted and extraverted individuals differ in how they process information which leads to different cognitive processes as well as to different preferred learning settings (Felder & Silverman, 1988; Raju & Venugopal, 2014). Since extraverted personalities are characterized as being outgoing, communicative and active, they are considered "Active Learners", who like to learn through experimentation and prefer working and discussing information in groups (Capretz & Ahmed, 2010; Felder & Silverman, 1988). Active learners tend to learn less in passive learning situations, such as lectures, and longer periods of solitude or a lack of interaction could lead to exhaustion for them (Felder & Silverman, 1988; Kroeger & Thuesen, 1988). Introverted personalities, in contrast, prefer reflection over action and receive energy from introspection which is why they are considered "Reflective Learners". Reflective learners need time to think about information and prefer working by themselves over working in groups. For them, constant interaction with others could lead to exhaustion (Capretz & Ahmed, 2010; Felder & Silverman, 1988; Kroeger & Thuesen, 1988).

I-E and CEQ:

With regard to how individuals experience e-learning, the results of a study with nursing students by Malloy (2007) confirmed that introverted students prefer an independent learning style, while extraverted students prefer an interactive one. According to Livingood (1995), extraverts are too impatient for e-learning and get frustrated not talking to others. Introverts, in contrast, get frustrated in traditional face-to-face learning settings as they prefer working alone and perceive that extraverted students dominate conversations. The results of a study by Ellsworth (1995) correspond to this and reveal that introverts even perceive the communication with course instructors as facilitated using computer-mediated communication.

The perceptual differences between the e-learning experience of introverts and extraverts found in these studies, lead to the first hypothesis (H_{1a}-H_{1c}) of this thesis, which is that introverted students evaluate their CEQ regarding e-learning differently compared to extraverted students, i.e. there are group differences between introverted and extraverted students.

H_{1a}: Introverted and extraverted students differ significantly with regard to their perceived SQ.

H_{1b}: Introverted and extraverted students differ

significantly with regard to their perceived RQ with peer students.

H_{1c}: Introverted and extraverted students differ significantly with regard to their perceived RQ with the course instructor.

I-E and ViU:

The fact that individuals tend to look for situations and circumstances that enable trait expression (Snyder, 1983), explains why introverted and extraverted personalities strive for different goals in life and thus also prefer different learning styles and settings. A long-term study with Canadian college students by Little, Lecci, and Watkinson (1992), for instance, showed that different manifestations of the Big Five personality dimensions are accompanied by different goals or "personal projects" that individuals pursue. They found that a significant correlation between "Extraversion" and "Project Visibility" exists. The more extraverted students were, the more they looked for contexts, situations or projects that contained a strong interpersonal component and whose results were visible to others. That individuals with a high degree of Extraversion are more likely to have relational intentions was also confirmed by Al-Hawari (2015), who investigated customers' perception of retail banking. Overall, these study results correspond with the proposition that extraverts are active learners, who strive for active experimentation and group discussions (Felder & Silverman, 1988).

For introverted personalities, Bishop-Clark, Dietz-Uhler, and FIsher (2007) discovered that introverted students reported more often than extraverted students about completing their tasks at their own pace when using e-learning services. Ellis (2003) came to a similar result and found that introverted students valued having enough time to think about ideas when using e-learning services - particularly in the case of asynchronous e-learning. These results also correspond with the results of Felder and Silverman (1988), who found that introverts are reflective learners, who strive for working by themselves and having enough time to reflect. Therefore, the second hypothesis (H_{2a}-H_{2q}) of this thesis is that:

H_{2a-q}: Introverted and extraverted students differ significantly in their perceived ViU (with regard to all ViU dimensions).

I-E and Fear of COVID-19:

COVID-19 is associated with several psychological stressors including the health threat to oneself and loved ones, the disruption of daily routines, the possible separation from family and friends, social isolation as well as financial losses. Although society as a whole is confronted with these difficulties, individuals differ according to their personality in how strongly they feel fear and react to a pandemic (Taylor, 2019). With regard to I-E, studies found that introverted and extraverted personalities possess different levels of arousal, which leads to different responses to certain situations (Dewaele & Furnham, 1999; Eysenck, 1965). Due to neurobiological differences, extraverted personalities tend to be generally under-aroused, while introverted personalities tend to be generally over-aroused. As a result, introverted personalities have a lower stress resistance compared to extraverted personalities and show stronger fear reactions across a range of situations (Dewaele & Furnham, 1999; Eysenck, 1965; Gange, Geen, & Harkins, 1979). In over-arousing situations, Eysenck (1981) states that *"introverts take longer to access information (. . .) from long-term memory or permanent storage"* (p. 203).

Mowen, Harris, and Bone (2004), for example, were able to prove the relationship between "Introversion" and "Fear" for the case of advertisements. Introverted personalities responded with higher fear to certain advertisements than extraverted personalities did. Shapiro and Alexanders (1969) investigated the particular relationship between "I-E", "Affiliation" (i.e. gregariousness) and "Anxiety" and found that in a high-anxiety condition, introverted students had less desire to affiliate than did either anxious extraverted students or non-anxious introverted students. In order to reduce anxiety, they found that anxious introverted students preferred solitude, while anxious extraverted students preferred the company of others. Several other studies also confirmed the relationship between Introversion and Fear and showed the higher likelihood of introverted personalities to notice threatening stimuli and their higher susceptibility to stress (Eysenck, 1981; Gray, 1970; Ragozzino & Kelly, 2011). Furthermore, it was found by various studies that introverted personalities in reaction try to avoid situations which they perceive as over-arousing and also prefer low-arousing conditions in learning settings (Dewaele & Furnham, 1999; Socan & Bucik, 1998).

E-learning gives students the possibility to study alone and at their own pace which should be especially suitable for introverted students with high fear of COVID-19 since they try to avoid over-arousing situations. The perceived CEQ and ViU of other groups might be lower in comparison. Therefore, it is hypothesized that:

H_{3a}: Student I-E and Fear of COVID-19 interact such that introverted students with high Fear of COVID-19 perceive a significantly higher SQ than introverted students with low Fear of COVID-19 as well as extraverted students with low and with high Fear of COVID-19.

H_{3b}: Student I-E and Fear of COVID-19 interact such that introverted students with high Fear of COVID-19 perceive a significantly higher RQ with peer students than introverted students with low Fear of COVID-19 as well as extra-verted students with low and with high Fear of COVID-19.

H_{3c}: Student I-E and Fear of COVID-19 interact such that introverted students with high Fear of COVID-19 perceive a significantly higher RQ with

the course instructor than introverted students with low Fear of COVID-19 as well as extraverted students with low and with high Fear of COVID-19.

H_{4a-q}: Student I-E and Fear of COVID-19 interact such that introverted students with high Fear of COVID-19 perceive a significantly higher ViU with regard to all ViU dimensions than introverted students with low Fear of COVID-19 as well as extraverted students with low and with high Fear of COVID-19.

Research Model:
Based on the reviewed literature and the formulated hypotheses, a conceptual research model was developed for the purpose of this thesis. The model proposed here represents the hypothesized relationships between I-E, Fear of COVID-19, CEQ and ViU (see Figure 1). In the further course of this thesis, the formulated hypotheses are tested. It is investigated whether there are group differences between introverted and extraverted students regarding their perceived CEQ and ViU. Furthermore, it is analyzed whether an interaction effect of I-E and Fear of COVID-19 on both CEQ and ViU exists.

4. Method

4.1. Choice of Method and Data Collection

Choice of Method:
In order to test the formulated hypotheses, a quantitative research approach was chosen employing Analysis of Variance (ANOVA). Quantitative data not only allowed for summary information on a variety of constructs and was suitable for testing hypotheses, but also allowed for a larger sample compared to most qualitative methods (Hair, Page, & Brunsveld, 2019). In concrete, a self-administered online survey was conducted with students of FUB. This method enabled a rapid data collection at relatively low cost and a large number of students could be surveyed simultaneously and from any location (ibd.). In addition, no personal contact with participants was necessary which was also favorable considering the pandemic situation prevailing at the time of data collection.

Choice of Case:
The case of FUB was chosen since the university concentrated solely on e-learning during the summer semester 2020, using both asynchronous and synchronous forms of e-learning. With about 34,000 students and a total of 16 different faculties and 150 different study programs, FUB is furthermore one of the largest German universities (Berlin.de, 2020; Freie Universität Berlin, 2020f). Since literature recommends conducting surveys only with people who have experience using the service of interest, FUB with its digital summer semester was also suitable as the case to be investigated (Marshall, 1996).

Pretest:
Before the actual data collection phase started, a pretest was conducted between August 10 and August 15, 2020 in order to ensure that all items in the questionnaire are as reliable and valid as can be determined (Anderson & Gerbing, 1991). Sheatsley (1983) states that 12 participants are enough to uncover major difficulties and weaknesses when pretesting a questionnaire. Therefore, seven business students, two law students and three communication science students - all from FUB - were asked for written feedback[3] regarding the questionnaire design, the instructions for completing the questions as well as difficulties in understanding and responding to the items. Since they were familiar with the e-learning context and all students of FUB, they reflected the target group of the actual questionnaire (Ferketich, Phillips, & Verran, 1993). Based on their feedback, some items were slightly reworded in order to improve both their reliability and validity. Overall, the students found the questionnaire to be well structured and had no difficulties in understanding it.

Data Collection:
The actual data collection phase started on September 10, 2020 and ended on October 18, 2020. For the data collection, a convenience sampling was used. This sampling method generally promises a high response rate, but at the price of a lower generalizability of the results compared to probability sampling methods (Hair et al., 2019). Nevertheless, due to the limited time and financial resources, it was considered an appropriate method for the purpose of this thesis.

Efforts were made to obtain the most representative sample of the student body possible. For this purpose, all 12 departments of FUB were contacted by e-mail and were informed about the research purpose of the thesis. Furthermore, they were asked for support by disseminating the questionnaire link via BB or mailing lists. With a few exceptions, every department offered their support and disseminated the link. In addition, the link was posted to various Facebook groups of faculties and of FUB as a whole. In order to maximize the response rate and to motivate students to participate, ten Amazon vouchers, worth € 20 each, were raffled and ten cents per completed questionnaire were donated to the "Deutsches Kinderhilfswerk", an NGO that is committed to equal educational opportunities for children as well as to combating child poverty in Germany (Deutsches Kinderhilfswerk, 2020).

Respondents were given equal opportunity to participate in the survey, as long as they fulfilled the criterion of the first screening question. Altogether, 470 usable questionnaires were returned - 229 questionnaires of the BB-questionnaire

[3] Due to COVID-19, the students were asked for written feedback instead of a face-to-face meeting.

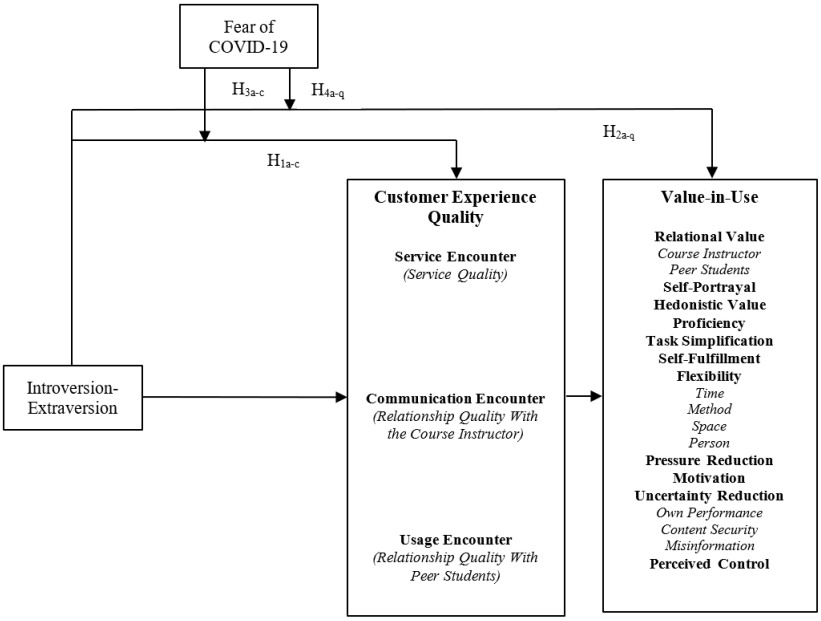

Figure 1: Research Model

version and 241 questionnaires of the Webex-questionnaire version.

4.2. Conception of the Questionnaire

The questionnaire was created and later completed by the students on "Unipark" - the academic program of "Questback" - which provides a standardized platform for creating online questionnaires. It started with a short introduction that contained a brief description of the questionnaire purpose and provided information about contact possibilities and data protection regulations. In addition, respondents were informed about the donation and the possibility of participating in the raffle.[4]

Due to the fact that the e-learning services of FUB include both asynchronous and synchronous offerings, two questionnaire versions were designed: one BB questionnaire version and one Webex questionnaire version. Although they referred to either BB or Webex in their instructions, both questionnaire versions contained the exact same questions and items so that comparisons could be made afterwards. Also, the instructions and layouts were identical in order to eliminate any kind of biases. The order of the questionnaires was structured by topic sections. Each section began with a couple of introductory instructions to prepare the respondent for the upcoming section. The questions asked related only to teaching and not to examination forms or situations since these are associated with particular emotions and coping strategies (Zeidner, 1995), which are not investigated in the context of this thesis. All theoretical constructs were measured by multi-item scales, with each construct being measured by at least three items in order to achieve an acceptable reliability (Hair et al., 2019). In addition, established scales were used to ensure the validity and reliability of each item (Schrauf & Navarro, 2005). Where it was necessary, items were slightly adjusted in their wording in order to adapt them to the research context of this thesis. Excluding the demographic questions and screening questions, the two questionnaire versions consisted of 125 items each, which were rated on either seven-point Likert scales (1 = "Strongly Disagree", 7 = "Strongly Agree") or five-point Likert scales (1 = "Strongly Disagree", 5 = "Strongly Agree"). A detailed overview of the measured theoretical constructs and the scales used can be found in Table 2 in the appendix.

Screening Questions:

As suggested by Sheatsley (1983), the survey started with two screening questions before the respondents were forwarded to the respective questionnaire version. The screening questions were easy to answer and did not yet contain any emotional components (ibd.). The first screening question asked whether the student participated in at least one online course at FUB during the summer semester 2020. This should ensure that only students of FUB, who already had experience with e-learning, were included in the final sample. Through the second screening question, an attempt was made to distribute the respondents as equally as possible to the two questionnaire versions. It was asked, whether the student's age is an even or an odd number. Students with an even age number received the BB questionnaire version, while those with an odd age number received the Webex questionnaire version.

[4]The questionnaire can be found in Appendix 2.

Customer Experience Quality:
The first topic section of both questionnaires contained items on the perceived CEQ regarding BB or Webex. As mentioned above, in order to measure the perceived SQ, the SERVQUAL scale is a suitable instrument since it has already been validated across various industries (Lemon & Verhoef, 2016). Udo et al. (2011) modified the SERVQUAL scale for the case of asynchronous e-learning which is why the items used here were taken from their study. Nevertheless, the wording of the 21 items was slightly adapted to be suitable for both synchronous and asynchronous forms of e-learning as investigated here. Since Santos (2003) identifies "Ease-of-Use" as a further dimension of e-SQ, four additional items for Ease-of-Use were adopted from Venkatesh (2000) and were added to the SQ dimension. In order to measure the RQ with peer students, seven items were taken from a study on students' experiences with e-learning by Paechter, Maier, and Macher (2010). The 11 items on the RQ with the course instructor were taken from a study by Keating et al. (2003) on RQ in the online retail context. Both the items on the RQ with the course instructor and on the RQ with peer students were slightly adapted in their wording.

Value-in-Use:
The second topic section of the questionnaires contained items on the ViU that resulted from the usage of e-learning services on BB or Webex. Since Kleinaltenkamp, Storck, et al. (2018) already conducted a qualitative study at a German university on ViU dimensions regarding e-learning, their identified dimensions were also considered to be appropriate for the context of this thesis. Due to the fact that the study was a qualitative study, suitable established scales were looked for in other literature. In addition, established scales for measuring a relational value dimension of e-learning were sought. The items for four constructs[5] were taken over from Bruns and Jacob (2016), who detected several ViU dimensions for the context of fitness apps. Since fitness apps are online applications as well, the items were considered appropriate for the research purpose of this thesis. The construct of "Self-Portrayal" was measured using three items from a study by Pura (2005) on the use of mobile services. The items on "Flexibility" were adopted from a study on the measurement of work autonomy by Breaugh (1985) as well as from a study on task interdependence and job design by Kiggundu (1983). The three items on "Flexibility Regarding Space" were developed in the style of Breaugh (1985) and were merely adapted in their wording to refer to the space instead of the method. The four items for the measurement of "Task Simplification" were adopted from a study on online shopping by Mathwick, Malhorta, and Rigdon (2001) as well as from a study by Bruns and Jacob (2016). In order to develop items for the construct of "Pressure Reduction", the "Perceived Stress Scale" by Cohen, Kamarck, and Mermelstein (1993) was used, which is the most widely psychological tool for measuring perceived stress. Four of the questions contained in the scale were reformulated into statements that correspond to the context of e-learning. From a study on factors influencing student's perception of e-learning by Tarhini, Al-Busaidi, Mohammed, and Magableh (2017), the three items on "Cost Decrease" were taken. In order to measure intrinsic motivation in field and laboratory settings, Guay, Vallerand, and Blanchard (2000) developed the "Situational Motivation Scale". Two of the three items on "Motivation" were taken from this scale. Since Kleinaltenkamp, Storck, et al. (2018) also consider the ease with which tasks are fulfilled as a value aspect of one's motivation, a third item regarding the perceived ease of studying was formulated and added to the "Motivation" scale. The construct of "Uncertainty Reduction" was measured by ten items, which were taken from different sources. The three items referring to the "Uncertainty Regarding One's Own Performance" were taken from the "Incompetence of Fear Scale" by Vollmeyer and Rheinberg (2000). The three items referring to the "Uncertainty Regarding Content Security" were adopted again from Tarhini et al. (2017). Since Kleinaltenkamp, Storck, et al. (2018) also consider the reduction of misinformation as a central value aspect of "Uncertainty Reduction", four items referring to the "Uncertainty Regarding Misinformation" were also included in the questionnaire, which were adapted from a study on behavioral intention to reuse e-learning systems in rural China by Li, Duan, and Alford (2012). The final three items of the ViU section were on "Perceived Control" and were taken from a study by Fu, Su, and Yu (2009), who intended to measure the perceived control of learners over e-learning games.

The wording of all items used in the ViU section was adapted to the context of this thesis, except the items for "Task Simplification" that were adopted exactly. In addition to the items measuring theoretical constructs, in the middle of the ViU section, the respondents' attention was tested by including an attention check statement to which respondents should respond with "Strongly Disagree".

For both the CEQ and ViU sections, respondents were asked to refer to the online course they found to be the best in the summer semester 2020 since emotional fluctuations in the answers should be minimized and the results should be comparable.

Introversion-Extraversion:
The third part of the questionnaire contained items on the construct of I-E. For the measurement of I-E several different approaches exist, even though no consensus has yet been reached in literature (McCrae & John, 1992). In the context of this thesis, the items on I-E of the "Big Five Inventory" (BFI) were used. The BFI is with only 44 items in total a brief multidimensional personality inventory whose short phrases enable above all time savings (John et al., 1991). Its eight items on I-E already proved to be valid in previous research (John, Naumann, & Soto, 2008) and were adopted without any adjustments in their wording.

[5] "Relational Value", "Hedonistic Benefit", "Proficiency", "Personal Self-Fulfillment".

Fear of COVID-19:
The penultimate section of the questionnaire was intended to measure "Fear of COVID-19" using "The Fear of COVID-19 Scale" (FCV-19S) by Ahorsu et al. (2020). The seven-item scale was developed through literature reviews, expert panels as well as interviews with pilot participants. Several studies, including Italian, Turkish and Eastern European samples, already proved that the scale is valid and reliable for assessing Fear of COVID-19 (Reznik et al., 2020; Satici et al., 2020; Soraci et al., 2020). The scale was adopted completely for this thesis, without any changes to the wording of the items. To not risk a breakoff or false response, both the I-E Scale and the FCV-19S were placed towards the end of the questionnaire since they both address more sensitive topics.

Demographic Questions:
As people's attitudes and behavior are affected by demographic variables and in order to be able to check the representativeness of the sample later, demographic questions formed the final section of the questionnaire (Sheatsley, 1983). Respondents were asked to provide information on their age, gender, program of study and income level. Placing the demographic questions at the end of the questionnaire was considered to be appropriate as those questions can be quite personal to some people who first need to build trust during the course of the questionnaire (Hair et al., 2019; Sheatsley, 1983).

4.3. Sample

Since the case under investigation of this thesis is FUB, the target population comprised the students of FUB. A non-randomized convenience sampling was carried out with an additional snowball procedure, where students of FUB forwarded the questionnaire link to other students of FUB (Ferber, 1977; Goodman, 1961). Of the BB sample, the majority of students was female (77 percent), enrolled in a Bachelor program (65 percent) and in the age groups between 18-20 (32 percent) and 21-30 (58 percent). The demographic characteristics of the Webex sample were similar: 71 percent of students were female and 59 percent were enrolled in a Bachelor program. Most students were in age groups 18-20 (12 percent) and 21-30 (76 percent). In both samples, the students were spread in roughly equal numbers across different academic fields. Table 1 provides an overview of most of the demographic variables surveyed. The high percentage of women in both samples does not compromise representativeness since the proportion of female students at FUB was 60.4 percent in the summer semester 2020 (Freie Universität Berlin, 2020b).

4.4. Data Analysis

Data Preparation:
Before both data sets were analyzed, they were prepared accordingly. 19 respondents in the BB data set and 15 respondents in the Webex data set failed the integrated attention check item. Since their concentration can be doubted when answering the items, they were excluded from the data sets. Additionally, in the BB data set a further response was excluded because all questions regarding one contruct were unanswered. In order to identify additional unengaged responses, the time each respondent took to complete the survey was examined more closely. In doing so, three respondents from the Webex data set and one respondent from the BB data set were excluded since their response time was more than three standard deviations above the mean response time ($\text{Mean}_{BB} = 805.65$, $\text{SD}_{BB} = 369.15$; $\text{Mean}_{Webex} = 868.48$, $\text{SD}_{Webex} = 414.19$)[6] (Anscombe, 1960). The final two samples consisted of 228 (BB data set) and 238 (Webex data set) units in total.

Reliability and Validity:
In order to evaluate the reliability and validity of the reflective measurement model, three criteria were considered: internal consistency reliability, convergent validity and discriminant validity (see Tables 3 and 4 in the appendix).[7]

For internal consistency reliability, Cronbach's alphas (CA) and composite reliabilities (CR) were examined. CA should be above the threshold of 0.7 (Nunnally, 1978). For all constructs, except the Cost Decrease construct in both data sets, CA values were satisfactory, ranging from 0.79 to 0.94 (BB data set) and from 0.78 to 0.95 (Webex data set). However, six items in each data set were dropped because their exclusion increased the CA of the respective construct (see Table 5). For those scales in the BB data set[8] and in the Webex data set[9] which consisted of only three items in total, no item was excluded to increase the CA of the respective construct. The CA values of these scales nevertheless exceeded 0.7.[10] To further meet the criteria of internal consistency reliability, the CR of each construct should exceed 0.6 (Bagozzi & Yi, 1988). All constructs in both data sets fulfilled this criterion.

Subsequently, convergent validity was examined. While some research considers a threshold of 0.5 for factor loadings (Bagozzi & Yi, 1988; Terho, Eggert, & Haas, 2015), other research considers a threshold of 0.6 (Kleinaltenkamp, Löbler, & Fennert, 2018). Therefore, in the context of this work, a threshold of 0.55 was used. All items which loaded less on the respective factor were excluded. In addition, it was checked whether the Average Variance Extracted (AVE) of each construct exceeded the threshold of 0.5 (Bagozzi & Yi,

[6] A value of 100 equals one minute.
[7] The statistical consulting service fu-stat of FUB was contacted in case of questions.
[8] Time Flexibility, Method Flexibility, Motivation, Uncertainty Reduction Regarding Own Performance, Uncertainty Reduction Regarding Content, Perceived Control.
[9] Time Flexibility, Method Flexibility, Space Flexibility, Motivation, Uncertainty Reduction Regarding Own Performance, Uncertainty Reduction Regarding Content.
[10] The Cost Decrease construct was excluded from both data sets due to low CA values as well as low factor loadings. One of three items in both data sets loaded with less than 0.3 on the factor Cost Decrease. As a consequence, the construct was removed from the research model.

Table 1: Demographic Data of Respondents

Sample Characteristics	BB Sample ($n = 228$)		Webex Sample ($n = 238$)	
	n	%	n	%
Female	173	76.55%	170	71.43%
Male	52	23%	65	27.31%
Diverse	1	0.44%	3	1.26%
18-20	73	32.30%	28	11.77%
21-30	130	57.52%	181	76.05%
Bachelor Program	147	65.04%	140	58.82%
Master Program	61	26.99%	75	31.51%
1.-2.	98	43.36%	90	37.82%
3.-4.	68	30.09%	77	32.35%
Medical Sciences	9	3.98%	13	5.46%
Natural Sciences/Mathematics	32	14.16%	42	17.65%
Humanities	36	15.93%	48	20.17%
Economic/Social Sciences	59	26.11%	60	25.21%
Education	46	20.35%	40	16.81%
Psychology	10	4.43%	5	2.10%
Other field of study	34	15.04%	30	12.61%
Part-Time (< 40 hours a week)	105	46.67%	129	54.43%
Unemployed (not looking for work)	29	12.89%	25	10.55%
Unemployed (looking for work)	56	24.89%	53	22.36%

1988). In both data sets, all constructs exhibited an AVE above 0.5, except the Fear of COVID-19 construct in both data sets ($AVE_{BB} = 0.49$, $AVE_{Webex} = 0.45$). Therefore, for the BB data set, the item with the lowest factor loading (Fear_Unconf: 0.63) was excluded which increased the AVE to 0.51. For the Webex data set, even if several items had been excluded, the AVE would not have increased to 0.5 which is why no item was excluded. Nevertheless, the construct was used for further analysis since all other quality criteria were met (Zerres, 2010), and the scale also achieved a sufficient AVE in the study by Ahorsu et al. (2020), who developed the scale.

In order to assess discriminant validity, in a next step the Fornell and Larcker (1981) criterion was used. Tables 6 and 7 show the square roots of the AVE for all constructs and the construct correlations. In the BB data set, the correlation (0.76) of the two constructs SQ and RQ (With The Course Instructor) exceeded the square root of the AVE. The reason for this might be that several items of the SQ scale also referred to the course instructor. As a consequence, three items (SQ Needs, SQ Concern, SQ Encourages) of the SQ scale, which cross-loaded with values between 0.66 and 0.72 on the RQ scale, were excluded. In the Webex data set, the correlation (0.73) of the two constructs SQ and RQ (With The Course Instructor) also exceeded the square root of the AVE. Therefore, two items (SQ Needs, SQ Helps) of the SQ scale, which both also cross-loaded with values of 0.74 on the RQ construct, were excluded. After the exclusions, the square roots of the AVE for each construct exceeded the construct correlations indicating a satisfactory degree of discriminant validity.[11]

For the further statistical analyses, the remaining items were used to calculate mean values for the respective constructs. In addition, a median split was carried out for the constructs of I-E and Fear of COVID-19 to achieve nominal measurement levels for both constructs that are suitable for conducting variance analyses.

5. Results

Using JMP Pro 15, multiple two-way ANOVAs were performed to explore the hypothesized group differences between introverted and extraverted students as well as the hypothesized interaction effects.

Prior to conducting the two-way ANOVAs, the assumption of homogeneity of variances was checked using Levene's Test for Homogeneity of Variance. For the Webex data set, the assumption of homogeneity of variances was satisfied across all dependent variables ($p > .05$). For the BB data set, the p-value of Levene's Test was significant ($p < .05$) for the dependent variables of Relational Value (Peer Students) and Space Flexibility indicating variance heterogeneity. Nevertheless, ANOVAs rather than non-parametric tests were conducted since group sizes were relatively equal and greater than 12, making the ANOVA robust to the assumption violation (Kohr & Games, 1974).

[11] Due to the item exclusions, the AVE of the SQ construct in the Webex data set dropped to 0.49. However, since the threshold value of 0.5 was only slightly undercut and all other quality criteria were fulfilled, no further items were excluded.

Furthermore, the assumption of normality was tested for both samples using the Anderson-Darling Test. For both data sets, the p-value of the Anderson-Darling Test was significant ($p < .05$) across all dependent variables indicating a violation of the normality assumption. Since the ANOVA is robust to violations of normality, the data analysis was continued as planned (Blanca, Alarcon, Arnau, Bono, & Bendayan, 2017).

For the BB data set, 20 two-way ANOVAs were conducted. The results showed that differences in mean values between introverted and extraverted students as well as between introverted students with high fear of COVID-19, introverted students with low fear of COVID-19, extraverted students with high fear of COVID-19 and extraverted students with low fear of COVID-19 exist. Nevertheless, these differences were minor and non-significant for all CEQ encounters and all ViU dimensions as indicated by the F-tests of the ANOVAs ($p > .05$). Tables 8 and 9 provide a detailed overview of the results. Regarding the CEQ, introverted students on average reported a more positive SQ ($M_{SQ} = 5.68$), RQ with peer students ($M_{RQPS} = 3.25$) and RQ with the course instructor ($M_{RQCI} = 4.74$) than extraverted students did ($M_{SQ} = 5.52$, $M_{RQPS} = 3.15$, $M_{RQCI} = 4.72$). Regarding 12 out of 17 ViU dimensions, introverted students on average also reported a higher ViU than extraverted students did. Looking at I-E and Fear of COVID-19 in interaction, introverted students with low fear of COVID-19 on average reported the most positive CEQ ($M_{SQ} = 5.81$, $M_{RQPS} = 3.41$, $M_{RQCI} = 4.8$). Regarding eight out of 17 ViU dimensions, introverted students with high fear of COVID-19 on average also reported the highest ViU. Nevertheless, all formulated hypotheses (H_{1a-c}, H_{3a-c}, H_{2a-q}, H_{4a-q}) had to be rejected as the two-way ANOVAs were non-significant ($p > .05$).

20 two-way ANOVAs were also conducted for the Webex data set. The differences in the mean values were also only minor and non-significant for all CEQ encounters and all ViU dimensions as indicated by the F-tests of the ANOVAs ($p > .05$). Tables 10 and 11 provide a detailed overview of the results. Regarding the CEQ, introverted students reported a more positive SQ (M_{SQ}=5.66), RQ with peer students (M_{RQPS}=3.13) and RQ with the course instructor (M_{RQCI}=4.89) than extraverted students did (M_{SQ}=5.63, M_{RQPS}=2.94, M_{RQCI}=4.74). Regarding nine out of 17 ViU dimensions, introverted students on average reported a higher ViU than extraverted students did. Looking at I-E and Fear of COVID-19 in interaction, introverted students with low fear of COVID-19 on average reported the most positive SQ (M_{SQ}=5.68). Introverted students with high fear on average reported the most positive RQ with peer students (M_{RQPS}=3.28) and RQ with the course instructor (M_{RQCI}=4.99). Regarding 11 out of 17 ViU dimensions, extraverted students with low fear of COVID-19 on average reported the highest ViU. Nevertheless, all formulated hypotheses (H_{1a-c}, H_{3a-c}, H_{2a-q}, H_{4a-q}) had to be rejected for the Webex data set as well since the two-way ANOVAs were non-significant ($p > .05$).

Since no significant group differences regarding the perceived CEQ and the perceived ViU were found for the independent variables of I-E and Fear of COVID-19, an additional look was taken at the demographic variables of "Gender" and "Academic Field". For both data sets, no significant gender differences between female and male students were found regarding their perceived CEQ and ViU (see Tables 17 to 19). For Academic Field, for the BB data set, the only significant pairwise difference was found between humanities students and economic and social sciences students regarding their perceived RQ with the course instructor (see Tables 12 and 13). Humanities students reported a slightly significantly (p=.04) higher RQ (M_{RQCI}=5.29) than economics and social sciences students did (M_{RQCI}=4.48). For the Webex data set, significant pairwise differences were found between humanities students and economic and social sciences students as well as between humanities students and medical students regarding their perceived RQ with the course instructor (see Tables 14 to 16). Humanities students reported a significantly (p=.01) higher RQ (M_{RQCI}=5.48) than economics and social sciences students did (M_{RQCI}=4.57). They also reported a significantly (p=.008) higher RQ than medical students did (M_{RQCI}=3.94). In addition, a significant pairwise difference was found between medical students and economic and social sciences students regarding the ViU dimension of Self-Fulfillment. Medical students reported a slightly significantly (p=.046) higher Self-Fulfillment (M_{SF}=4.52) than economic and social sciences students did (M_{SF}=3.08).

6. Discussion

6.1. General Discussion

This thesis was designed to investigate whether differences exist between students' personalities regarding their perceived CEQ and their perceived ViU for the case of e-learning. In particular, the personality dimension I-E was investigated. Furthermore, it was examined whether students' Fear of COVID-19 constitutes a moderating variable that affects the relationship between I-E and CEQ as well as the relationship between I-E and ViU.

6.1.1. Interpretation of Results

Although the differences were not significant, regarding asynchronous e-learning introverted students on average perceived both a more positive CEQ regarding all three encounters as well as a higher ViU regarding most dimensions. This result is consistent in its tendency with the results of past studies (Bishop-Clark et al., 2007; Ellis, 2003; Felder & Silverman, 1988), which found that introverts, as reflective learners, prefer sufficient time for reflection as well as self-study. In particular, with respect to the Proficiency dimension, which reflects the perceived learning success, but also with respect to the Flexibility dimensions and the Motivation dimension, introverted students on average reported a higher ViU than extraverted students did.

Although not significant either, the interaction of I-E and Fear of COVID-19 regarding most ViU dimensions pointed in the direction of the formulated interaction hypotheses and introverted students with high of Fear of COVID-19 on average

perceived a higher ViU regarding most ViU dimensions than the other investigated groups did. Presumably, introverted students with high Fear of COVID-19 had the strongest desire for withdrawal out of all four groups. Therefore, they were able to achieve their goals regarding e-learning the most out of all four groups through time-shifted communication and independent study of course materials posted online.

In contrast, regarding the previously evaluated CEQ, introverted students with low Fear of COVID-19 on average rated their CEQ the most positive out of all four groups. Thus, high Fear of COVID-19 particularly seems to play a role for introverted students in their goal achievement - for instance as an opportunity for uninterrupted learning - and less in their evaluation of their CE. Qualitative studies need to be conducted to find out the underlying reasons for this finding.

For synchronous e-learning, also no significant group differences were found between introverted and extraverted students regarding their perceived CEQ and ViU. For synchronous e-learning, introverted students on average reported a more positive CEQ with regard to all three encounters than extraverted students did. However, for only about half of the ViU dimensions they on average perceived a higher ViU than extraverted students did. This could be due to the fact that synchronous e-learning, compared to asynchronous e-learning, addresses more strongly the needs of extraverted students, who prefer active communication. It allows for participation in lectures, face-to-face communication via video and audio communication and therefore, shows parallels to face-to-face teaching.

Although not significant either, the interaction of I-E and Fear of COVID-19 regarding most ViU dimensions surprisingly pointed in the opposite direction of the formulated interaction hypotheses. Extraverted students with low Fear of COVID-19 on average perceived a higher ViU regarding most ViU dimensions than the other three groups did. The combination of Extraversion and low Fear of COVID-19 may have resulted in a higher openness to direct social contacts and a stronger desire for active interaction. Especially live lectures and live chats via Webex provide an opportunity to satisfy this need.

In contrast, regarding the CEQ, introverted students with low Fear of COVID-19 on average rated the highest SQ and introverted students with high Fear of COVID-19 on average rated the highest RQ with peer students and the highest RQ with the course instructor. It appears that extraverted students with low Fear of COVID-19 are most likely to achieve their goals regarding e-learning, but do not evaluate the quality of their CE more positively than the other groups. Again, qualitative studies are needed here that additionally could investigate the relationship of the constructs of CEQ and ViU.

A descriptive comparison of the results for asynchronous and synchronous e-learning shows that in both cases there are only minor and non-significant differences regarding the CEQ and the ViU across groups. With regard to the CEQ, the SQ on average was rated most positively across groups, for both asynchronous and synchronous e-learning. This indicates the perception of the e-learning services of FUB as reliable and the accessibility and helpfulness of the course instructors. The rating of the perceived RQ with the course instructor on average was also positive across groups in both data sets. Nevertheless, the relational value indicated by the students surveyed in both data sets as well as the social acceptance by others (Self-Portrayal) were among the ViU dimensions with the lowest rating. Across groups, the mean values for these constructs were in the lower half of the rating scale and ANOVA results showed no significant group differences. Consequently, although students felt that course instructors were available when needed, a strong relationship could not be established with them and acquired knowledge could not be adequately demonstrated. It appears to make no difference here whether students are introverted or extraverted, whether they have low or high Fear of COVID-19 and whether it is asynchronous or synchronous e-learning. A reason for this might be that the transition to e-learning happened quite suddenly and completely eliminated any kind of face-to-face teaching or real-life contact opportunities for students. It is possible that all students - including the introverted - felt overwhelmed by this complete transition to self-study and the sudden removal of any direct communication opportunities with course instructors and peer students. Perhaps it was not just about a lack of simple communication, but rather about comprehension or organizational problems that occurred across personalities.

In contrast, the groups in both data sets reported a comparatively high ViU with regard to Space Flexibility indicating a perceived benefit in being able to pursue their learning activities regardless of location. Furthermore, students of both data sets were able to achieve Proficiency and thus learning progress through e-learning and e-learning was able to reduce their uncertainty of obtaining misinformation (Uncertainty Reduction Regarding Misinformation). Consequently, it can be concluded that students perceived the quality of the taught course contents as relatively high in both asynchronous and synchronous e-learning.

6.1.2. Substantive Explanations

Overall, it was found that no significant group differences existed in terms of reported CEQ and perceived ViU between introverted and extraverted students with either high or low fear of COVID-19. Furthermore, the students surveyed evaluated their CEQ only in the upper-middle range and also perceived only a moderately high ViU with respect to most dimensions (see Tables 9 and 11).

One possible explanation for this result, as mentioned earlier, is that all students surveyed experienced comprehension problems or organizational problems due to the complete transition to e-learning. Across personalities, students may have felt overwhelmed with the self-study associated with e-learning and experienced information overload. The mid-range ratings of both the CEQ and the perceived ViU suggest a need for improvement of the e-learning services of FUB. Reconstructive qualitative research methods, e.g. interviews, could help to specify this need and identify the exact

reasons for the mid-range ratings of the CEQ and the perceived ViU.

Furthermore, differences in the evaluation of the CEQ and the perceived ViU could depend on factors other than I-E and Fear of COVID-19, such as the topic of the evaluated course, examination requirements, the personality of the course instructor or other personality dimensions of the Big Five Model.

Examination requirements, for instance, could have an impact on achieving goals, such as the flexibility of time and methods, students' motivation as well as the perceived pressure reduction. Students with assignments due weekly may achieve these goals to a lesser degree than students with only one assignment due at the end of a course.

The personality of the course instructor could also lead to differences regarding the evaluation of e-learning services. The results of the one-way ANOVAs for the independent variable of Academic Field showed, with regard to both asynchronous and synchronous e-learning, that Humanities students perceived a significantly higher RQ with their course instructors, compared to students of some other academic fields. This could be an indication that the personality of the course instructor and his course design have an impact on the evaluation of the RQ by students. It may be that course instructors from the department of Humanities resort to particular methods to strengthen their RQ with students. Future studies could try to find this out with the help of qualitative interviews.

This thesis has limited itself to the assessment of one particular personality dimension. However, personality traits other than students' I-E might influence the perceived CEQ and ViU regarding e-learning. For instance, it could be possible that students who score high on the Big Five personality dimension of "Openness to Experience" are more open to new teaching methods and value trying new things more than students who score low on this personality dimension. Since high openness is characterized by intellectual curiosity and a preference for diversity, it is possible that the sudden transition to digital teaching formats and the required self-study are perceived more positively by open students than by less open students, who try to avoid change (McCrae et al., 1986).

Another reason for differences regarding the perceived CEQ and ViU might be the housing situation of students. Many students, especially in Berlin, live in shared apartments which is why introverted students with high Fear of COVID-19, for instance, might not have perceived e-learning as quiet self-study with a lot of time to reflect at all. Living with others might have interfered with their learning and also might not have contributed to their avoidance of social contacts. Extraverted students, on the other hand, possibly perceived a shared apartment and the social contacts associated with it as compensation for not having social contacts at university. Therefore, their CEQ was not significantly more negative and their perceived ViU was not significantly lower than that of introverted students.

The digital summer semester of 2020 and the transition to e-learning served as a protective measure to contain the pandemic and to prevent new infections. Although Harper et al. (2020) state that negative emotions protect against further spread of COVID-19 and lead to more social distancing, which in turn should contribute to an appreciation of e-learning, it may be possible - in light of the results of this thesis - that high Fear of COVID-19 had little effect on the need for protection among students. The age group in which most students in both samples are located does not belong to the COVID-19 risk group, which is generally above 50 to 60 years of age (Rober Koch Institut, 2020). Therefore, even introverted students with high Fear of COVID-19 may not have felt themselves to be at high risk and therefore may not have developed a strong need for protection. Whether the perceived need for protection acts as a mediator between Fear of COVID-19 and the CEQ or the ViU regarding e-learning, could be investigated using a more age-diverse sample that also includes COVID-19 risk groups. E-learning courses offered by larger companies, for instance, would be a suitable case for such a research project.

6.1.3. Statistical Explanations

One statistical explanation for the results of this thesis could be the loss of information that may have occurred due to the two median splits that were performed. The median splits performed on the constructs of I-E and Fear of COVID-19 did not account for mid-range scores and students who located themselves in the middle of the five-point scale of I-E, for instance, were assigned to one of the two extremes. Future studies could alternatively divide their samples into three groups so that moderately introverted or extraverted personalities are included in the analysis.

In addition, although all students surveyed were instructed to refer to what they considered to be the best course in the digital summer semester 2020, all students surveyed here nonetheless referred to different courses that varied in terms of size, course contents, exam requirements and course instructors. Although both samples exhibited a high degree of representativeness, the comparability of the information provided by individual students nevertheless suffered as a result. For instance, some students may have participated in the survey prior to receiving exam results, while others may have already received negative feedback and participated in the survey with that in mind. As a result, the survey could be repeated with students from the same course, controlling for the influence of several other possible influencing variables.

Following the last point mentioned, the internal validity of the results of this thesis might have suffered from self-selection bias (Woolridge, 2002). According to Sen and Lerman (2007), the goal of utilitarian consumption is to maximize utility which is why negative experiences, especially in utilitarian consumption, are weighted more heavily than positive ones. Due to this, perhaps only those students who evaluated their CEQ regarding e-learning as negative or moderate and whose level of goal attainment was only moderate as well participated in the survey.

Besides the results of this thesis, the results and external validity of the results of past studies of the subject area are also not free from criticism. A comparison with past studies, which were used to develop the hypotheses of this thesis, shows that some of them worked with much smaller sample sizes. In two studies, the sample size was even n<30 (Bishop-Clark et al., 2007; Ellis, 2003). Furthermore, in two studies either no demographic variables were surveyed (Raju & Venugopal, 2014), or no demographic information about the sample was provided in the report of study results (Bishop-Clark et al., 2007).

6.2. Theoretical Contributions

Although all of the hypotheses formulated in this thesis had to be rejected, the thesis nevertheless makes an important theoretical contribution to service research. Not only in the context of the COVID-19 crisis, but also in the context of increasing globalization and technologization, e-learning services play an increasingly important role for companies as well as for universities and change the co-creation interaction between universities and students. Therefore, there is not only interest in finding out how customers or students perceive and evaluate e-learning services, but also in finding out which factors influence these evaluations. This thesis investigated both of these questions and examined the influence of the personality dimension I-E on both the perceived CEQ and the perceived ViU regarding e-learning services.

In addition, perceptions of e-learning services were examined in light of the COVID-19 pandemic, which constitutes a public crisis situation. The pandemic represents an external condition over whose development service providers have little, if any, influence and whose exact impact on various service sectors remains unexplored. The results of this thesis show that introverted and extraverted personalities with either high or low fear of COVID-19 do not differ significantly in their perceptions regarding e-learning services. Even introverted personalities with high fear of COVID-19 did not perceive e-learning services significantly more positively than the other investigated groups. Unlike past studies that do not consider fear and the need for protection from COVID-19 as separate constructs (Harper et al., 2020), this thesis provides indication that the need for protection mediates between different personalities' perceptions of fear and their perceptions of services.

Furthermore, this thesis serves as a basis for investigations of the influence of other personality dimensions of the Big Five Model on the perception of e-learning services and encourages the multidimensional assessment of personalities.

Finally, this thesis worked with relatively large sample sizes of n_{BB}=228 and n_{Webex}=238 which contributes to the reliability of the results. Also, unlike previous studies that focused primarily on asynchronous e-learning (Bishop-Clark et al., 2007; Ellis, 2003), this thesis made descriptive comparisons between asynchronous and synchronous e-learning and tested the formulated hypotheses for both forms of e-learning. Although none of the hypotheses could be confirmed, a comparison of the group mean values nevertheless showed different tendencies between the two forms of e-learning.

6.3. Practical Implications

Against the background of the rapid required transition from face-to-face teaching to full e-learning starting with the summer semester 2020 and the fact that this survey was already conducted after the completion of the first digital semester, the CEQ rated in the upper middle range and the mid-range perceived ViU of the students should not be equated with a negative judgment about e-learning in general, but may rather be due to the sudden transition for both students and course instructors. Therefore, the information provided by university students in the survey should be considered as both an incentive and an opportunity for improvement of e-learning services.

First, it can be stated that, based on the results, it is not advisable to make an abrupt and complete transition of learning settings and that, in general, a blended learning approach could combine the advantages of face-to-face teaching, such as building relationships with peer students and course instructors, and the advantages of e-learning, such as not being tied to a specific location when studying or reducing the uncertainty of getting wrong information. However, due to the fact that universities have virtually no control over the development of the COVID-19 pandemic, these recommendations are currently more forward-looking. Immediate resumption of face-to-face teaching or blended learning formats are still not possible at the time of writing this thesis.

Second, under the given conditions, course instructors should acquire sufficient knowledge of the e-learning services offered and actively encourage a lot of exchange and discussion via Webex consultations, written communication, or group work in order to increase the perceived relational value of students and to offer support in organizing self-study. As part of this, efforts could also be made to allow individual students to choose to speak in Webex lectures or to allow them to create wikis in order to provide opportunities for self-presentation of acquired knowledge (Self-Portrayal).

Third, course instructors should introduce students, regardless of their personalities and preferred learning settings, to the proper use of the information provided online in order to prevent information overload or even overwhelm with e-learning. Especially in light of the fact that the end of the pandemic is not yet foreseeable at the time of this thesis, an effective and efficient design of self-study is essential to prevent a deterioration of academic performance in possible further digital semesters.

Finally, course instructors should not only conduct evaluations after online courses have been completed, but should also ask about expectations and desires before courses begin in order to identify determinants for the perception and evaluation of e-learning services and to address students' desires accordingly. In this regard, this thesis suggest that the focus should not be on the I-E personality dimension. Also, Fear of

COVID-19 does not lead to significant differences regarding students' perception of their CEQ and their ViU.

6.4. Limitations and Further Research

In addition to the suggestions for future research that can be derived from the interpretation of results, other future research opportunities arise from the findings of this thesis.

First, only the case of FUB was investigated in this thesis, so the results may not be generalizable. Future studies, in addition to e-learning in corporate contexts, could examine other national contexts with different education systems and different policies to address the COVID-19 crisis. For instance, the CEQ and the perceived ViU regarding e-learning could differ between countries such as the U.S., with relatively high tuition fees, and countries such as Germany, with comparatively low tuition fees. Furthermore, it could be investigated how students living in countries with few restrictions due to COVID-19 experience e-learning compared to students living in countries with many restrictions due to COVID-19.

Second, due to the fact that only quantitative methods were used in this thesis, deeper insight into personal reasons for certain students' perceptions of e-learning cannot be provided. Future research could overcome this by conducting in-depth interviews with students regarding their experiences with e-learning services.

Third, students here were only surveyed at one point in time. Since no one yet knows how long the COVID-19 crisis will last, long-term studies should be initiated to examine how students' perceptions of their e-learning experience change over the period of the crisis. Future studies could also comparatively examine, based on previous teaching evaluations, whether or not the perceived CEQ and ViU differ between face-to-face teaching, before the onset of the COVID-19 crisis, and e-learning, after the onset of the COVID-19 crisis. For instance, has the CEQ for extraverted students worsened since the full transition to e-learning? Could it be that introverted students, even if they only reported scores in the upper-middle range here, perceived even less value regarding face-to-face teaching? A combination of qualitative and quantitative methods seems promising here to identify factors that influence students' perceptions of e-learning and to help universities develop appropriate strategies to improve their e-learning services.

6.5. Conclusion

The impact of the perception of the COVID-19 crisis on different service sectors of the economy is an unexplored topic to date. The pandemic forced universities to suddenly make a complete transition to offering e-learning services which has not been necessary yet in this form in the past and changed the co-creation interaction between universities and students.

The purpose of this thesis was to investigate, for the case of e-learning services at FUB, whether students evaluate their CEQ differently depending on their personality and whether they perceive a different ViU. Specifically, the personality dimension I-E was considered and ANOVAs found that introverted and extraverted students do neither significantly differ with respect to their perceived CEQ, nor significantly differ with respect to their perceived ViU. In addition, the thesis investigated whether Fear of COVID-19 forms a moderator between students' I-E and their CEQ as well as their ViU. No significant interaction effects were found between I-E and Fear of COVID-19 on the three CEQ encounters, nor on the individual ViU dimensions. Due to this, all research questions posed in this thesis had to be negated and all formulated hypotheses had to be rejected.

One reason for the results of this thesis could be that other factors than I-E and Fear of COVID-19 are responsible for differences in the evaluation of both the CEQ and the ViU regarding e-learning, such as the topic of the respective course, examination requirements, the personality of the course instructor, the housing situation of students or other personality dimensions of the Big Five Model. Also, additional mediator variables could exist between Fear of COVID-19 and the perceived CEQ or between Fear of COVID-19 and the perceived ViU.

The thesis contributes to the service domain of e-learning and sheds light on the impact of the COVID-19 crisis on e-learning services, which is a novel area of research. It proposes that whether a person is introverted or extraverted and whether a person has high or low Fear of COVID-19 makes no significant difference to the evaluation of e-learning services.

Future research could use qualitative methods to help identify other influencing factors and explore other national and corporate contexts for the purpose of generalizability. In addition, long-term studies would offer the opportunity to compare students' perceptions of e-learning during and after the pandemic.

Overall, the students surveyed rated their CEQ in the upper-middle range and perceived a moderately high ViU with respect to most ViU dimensions. Course instructors should therefore, in this novel situation, acquire sufficient knowledge about e-learning themselves and encourage mutual support as well as teach students to independently use the provided course materials so that there is no overload with e-learning. In addition, students' expectations and wishes should already be asked before online courses start in order to identify possible influencing factors for final evaluations. Not only for the further course of the pandemic, but also afterwards, regarding e-learning an approach should be pursued that takes into account the interests of the students without neglecting the available resources of the universities as service providers as well as external framework conditions, such as political regulations. Thus, there is a chance that universities and students can continue to benefit from the advantages of e-learning even after overcoming the COVID-19 crisis.

References

Ahmed, F., Campbell, P., Jaffar, A., & Alkobaisi, S. (2010). Learning & personality types: A case study of a software design course. *Journal of Information Technology Education*(9), 237–252.

Ahorsu, D. K., Lin, C. Y., Imani, V., Saffari, M., Griffiths, M. D., & Pakpour, A. H. (2020). The fear of COVID-19 scale: development and initial validation. *International Journal of Mental Health and Addiction*, 1–9.

Al-Hawari, M. A. (2015). How the personality of retail bank customers interferes with the relationship between service quality and loyalty. *International Journal of Bank Marketing*(33), 41–57.

Allport, G. W. (1937). *Personality: a psychological interpretation*.

Allport, G. W. (1961). *Pattern and growth in personality*.

Anderson, J. C., & Gerbing, D. W. (1991). Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. *Journal of Applied Psychology*(76), 732–740.

Anscombe, F. J. (1960). Rejection of Outliers. *Technometrics*(2), 123–146.

Ashton, M. C. (2013). *Individual Differences and Personality*.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*(16), 74–94.

Baron, R. A. (2001). *Essentials of Psychology*.

Baughman, E. E., & Welsh, G. S. (1962). *Personality: A behavioral science*.

Becker, L., & Jaakkola, E. (2020). Customer experience: fundamental premises and implications for research. *Journal of the Academy of Marketing Science*(48), 630–648.

Berlin.de. (2020). *Universitäten.* Retrieved from https://www.berlin.de/berlin-im-ueberblick/wissenschaft/lehr-und-forschungsstandort/universitaeten-und-hochschulen/

Berliner Vewaltung. (2020). *Aktuelle Informationen zu Corona-Maßnahmen an Hochschulen und Forschungseinrichtungen.* Retrieved from https://www.berlin.de/sen/wissenschaft/aktuelles/news/2020/artikel.908920.php

Bernreuter, R. G. (1931). *The personality inventory*.

Bishop-Clark, C., Dietz-Uhler, B., & FIsher, A. (2007). The Effects of Personality Type on Web-Based Distance Learning. *Journal of Educational Technology Systems*(35), 491–506.

Blanca, M. J., Alarcon, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*(29), 552–557.

Boin, A., Hart, P., Stern, E., & Sundelius, B. (2017). *The politics of crisis management: Public leadership under pressure*. Cambridge University Press.

Breaugh, J. A. (1985). The Measurement of Work Autonomy. *Human Relations*(38), 551–570.

Bruns, K., & Jacob, F. (2014). Value-in-use and mobile technologies. *Business & Information Systems Engineering*(6), 349–359.

Bruns, K., & Jacob, F. (2016). Value-in-Use: Antecedents, Dimensions, and Consequences. *Journal of Research and Management*(38), 135–149.

Capretz, L. F., & Ahmed, F. (2010). Making sense of software development and personality types. *IEEE IT Professional*(12), 6–13.

Cattell, R. B. (1946). *Description and measurement of personality*. World Book.

Chandler, J. D., & Vargo, S. L. (2011). Contextualization and Value-in-Context: How Con-text Frames Exchange. *Marketing Theory*(11), 35–49.

Cohen, S., Kamarck, T., & Mermelstein, R. (1993). A global measure of perceived stress. *Journal of Health and Social Behavior*(24), 386–396.

Costa, P. T., & Mccrae, R. R. (1985). *The NEO Personality Inventory manual*. Psychological Assessment Resources.

Costa, P. T., & Mccrae, R. R. (1989). *The NEO-PI / NEO-FFl manual supplement*. Psychological Assessment Resources.

Costa, P. T., & Mccrae, R. R. (1992). *NEO PI-R professional manual*. Psychological Assessment Resources.

de Canniere, M. H., de Pelsmacker, P., & Geuens, M. (2009). Relationship quality and the theory of planned behavior models of behavioral intentions and purchase behavior. *Journal of Business Research*(62), 82–92.

de Keyser, A., Verleye, K., Lemon, K. N., Keiningham, T. L., & Klaus, P. (2020). Moving the Customer Experience Field Forward: Introducing the Touchpoints, Context, Qualities (TCQ) Nomenclature. *Journal of Service Research*, 1–23.

de Wulf, K., Odekerken-Schröder, G., & Dawn, I. (2001). Investments in consumer relationships: A cross-country and cross-industry exploration. *Journal of Marketing*(65), 33–50.

Deutsches Kinderhilfswerk. (2020). *Spenden Sie für Ihr wichtigstes Thema!* Retrieved from https://www.dkhw.de/spenden/spendenprojekte/alle-spendenprojekte/

Dewaele, J. M., & Furnham, A. (1999). Extraversion: The Unloved Variable in Applied Linguistic Research. *Language Learning*(49), 509–544.

Ebner, M., Schön, S., Bäuml-Westebbe, G., Buchem, I., Lehr, C., & Egloffstein, M. (2013). Kommunikation und Moderation. In M. Ebner & S. Schön (Eds.), *Lehrbuch für Lernen und Lehren mit Technologien (L3T)*. epubli GmbH.

Edvardsson, B. (2005). Service quality: beyond cognitive assessment. *Managing Service Quality*(15), 127–131.

Ellis, A. E. (2003). Personality Type and Participation in Networked Learning Environments. *Education Media International*(40), 101–114.

Ellsworth, J. H. (1995). Using computer-mediated communication in teaching university courses. In Z. L. Berge & M. P. Collins (Eds.), *Computer-mediated Communication and the Online Classroom, Volume I: Overview and Perspectives.* Hampton Press.

Elsharnouby, T. H., & Mahrous, A. A. (2015). Customer participation in online co-creation experience: the role of e-service quality. *Journal of Research in Interactive Marketing*(9), 313–336.

Eysenck, H. J. (1965). Extraversion and the acquisition of eyeblink and GSR conditioned responses. *Psychological Bulletin*, *63*, 258–270.

Eysenck, H. J. (1981). *Learning, memory and personality*. A model for personality Berlin: Springer.

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire (adult and junior)*. London, Hodder & Stoughton.

Eysenck, H. J., & Wilson, G. (1991). *The Eysenck Personality Profiler*. Guildford, Psi-Press.

Felder, R. M., & Silverman, L. K. (1988). Learning and Teaching Styles in Engineering Education. *Engineering Education*, *78*, 674–681.

Ferber, R. (1977). Research by Convenience. *Journal of Consumer Research*, *4*, 57–58.

Ferketich, S., Phillips, L., & Verran, J. (1993). Development and administration of a survey instrument for cross-cultural research. *Research in Nursing & Health*, *16*, 227–230.

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Errors. *Journal of Marketing Research*, *18*, 39–50.

Freie Universität Berlin. (2019). *Learning Management System: Blackboard.* Retrieved from https://wikis.fu-berlin.de/display/eteaching/LearningManagementSystem:Blackboard#deck1-35202

Freie Universität Berlin. (2020a). *Fachbereiche und wissenschaftliche Einrichtungen der Freien Universität Berlin.* Retrieved from https://www.fu-berlin.de/einrichtungen/fachbereiche/index.html

Freie Universität Berlin. (2020b). *Gemeinsam verantwortlich handeln: Informationen für Universitätsangehörige zum Coronavirus (2019-nCoV).* Retrieved from https://www.fu-berlin.de/sites/coronavirus/index.html

Freie Universität Berlin. (2020c). *Kennzahlen.* Retrieved from https://www.cedis.fu-berlin.de/online-lehren-lernen/statistiken/index.html

Freie Universität Berlin. (2020d). *Online-Betreuung der Studierenden.* Retrieved from https://www.cedis.fu-berlin.de/online-lehren-lernen/online-betreuung/index.html

Freie Universität Berlin. (2020e). *Studierenden-und Prüfungsstatistik Sommersemester 2020.* Retrieved from https://www.fu-berlin.de/studium/studienorganisation/immatrikulation/weitere-angebote/statistik/daten/SoSe20_Studierendenstatistik_Gesamtstatistik_Stand-Juni-2020.pdf

Freie Universität Berlin. (2020f). *Vorlesungen digitalisieren.* Retrieved from https://www.cedis.fu-berlin.de/online-lehren-lernen/vorlesung/index.html

Frohne, G. (2020). *Erfolgsfaktoren von Online Customer Experience: Die Gestaltung von internetbasierten Kundenerlebnissen in der Kon-*

sumgüterbranche. Wiesbaden, Springer Gabler.

Fu, F. L., Su, R. C., & Yu, S. C. (2009). EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education*, *52*, 101–112.

Gabler, W. (2020). *Diensteleistungsmarketing.* Retrieved from https://wirtschaftslexikon.gabler.de/definition/dienstleistungsmarketing-27309

Gange, J. J., Geen, R. G., & Harkins, S. G. (1979). Autonomic differences between extraverts and introverts during vigilance. *Psychophysiology*, *16*, 392–397.

Gera, R. (2011). Modelling e-service quality and its consequences in India: an SEM approach. *Journal of Research in Interactive Marketing*, *5*, 203–225.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of Personality and Social Psychology*. Beverly Hills, CA: Sage.

Gomez-Rey, P., Barbera, P., & Fernandez-Navarro, F. (2016). Measuring teachers and learners' perceptions of the quality of their online learning experience. *Distance Education*, *37*, 146–163.

Goodman, L. A. (1961). Snowball Sampling. *The Annals of Mathematical Statistics*, *32*, 148–170.

Gray, J. A. (1970). The psychophysiological basis of introversion-extraversion. *Behaviour Research and Therapy*, *8*, 249–266.

Grove, S. J., & Fisk, R. P. (1997). The impact of other customers on service experiences: a critical incident examination of "getting along". *Journal of Retailing*, *73*, 63–85.

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion*, *24*, 175–213.

Guilford, J. P. (1959). *Personality*. New York, NY, McGraw-Hill.

Gummesson, E. (1987). The new marketing - Developing long-term interactive relationships. *Long Range Planning*, *20*, 10–20.

Guri-Rosenblit, S. (2005). 'Distance education' and 'e-learning': Not the same thing. *Higher Education*, *49*, 467–493.

Hair, J. F., Page, M., & Brunsveld, N. (2019). *Essentials of Business Research Methods*. New York, NY, Routledge.

Harper, C. A., Satchell, L. P., Fido, D., & Latzmann, R. D. (2020). Functional Fear Predicts Public Health Compliance in the COVID-19 Pandemic. *International Journal of Mental Health and Addiction*.

Hogan, R. T. (1983). A socioanalytic theory of personality. In M. Page (Ed.), *1982 Nebraska Symposium on Motivation.* Lincoln: University of Nebraska Press.

Holbrook, M. B. (1999). *Customer value—A framework for analysis and research*. London and New York, Routledge.

Holbrook, M. B., & Hirschmann, E. C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *Journal of Consumer Research*, *9*, 132–140.

Hui, M. K., & Bateson, J. E. G. (1991). Perceived Control and the Effects of Crowding and Consumer Choice on the Service Experience. *Journal of Consumer Research*, *18*, 174–184.

John, O. P., Donauhue, E. M., & Kentle, R. L. (1991). *The „Big Five" Inventory – Versions 4a and 54. Journal of Personality and Social Psychology*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm Shift to the Integrative Big Five Trait Taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of Personality.* New York, NY: Guilford Press.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research. 2 ed.* New York, NY: Guilford Press.

Johnson, J. A. (1997). Chapter 3 - Units of Analysis for the Description and Explanation of Personality. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of Personality Psychology.* San Diego, CA: Academic Press.

Jung, C. G. (1971). *Collected Works of C.G. Jung, Volume 6: Psychological Types*. Princeton, NJ, Princeton University Press.

Keating, B., Rugimbana, R., & Quazi, A. (2003). Differentiating between service quality and relationship quality in cyberspace. *Managing Service Quality*, *13*, 217–232.

Keynes, J. (1936). *The general theory of employment interest and money*. London, Macmillan.

Kiggundu, M. N. (1983). Task interdependence and job design: Test of a theory. *Organizational behavior and human performance*, *31*, 145–172.

Kleinaltenkamp, M., Löbler, B., H. und Esse, & Fennert, N. (2018). Social Attribution of Value–the Case of Electronic Engagement Platforms. In C. Arnold & H. Knödler (Eds.), *Die informatisierte Service-Ökonomie.* Wiesbaden: Springer Gabler.

Kleinaltenkamp, M., Storck, F., Gumprecht, P., & Li, J. (2018). The Impact of Psychological Ownership on Value in Use and Relational Outcomes. *Journal of Service Management Research*, *2*, 50–66.

Kohr, R. L., & Games, P. A. (1974). Robustness of the Analysis of Variance, the Welch Procedure and a Box Procedure to Heterogeneous Variances. *The Journal of Experimental Education*, *43*, 61–69.

Kroeger, O., & Thuesen, J. M. (1988). *Type Talk*. New York, NY, Delacorte.

Kuo, Y. C., Walker, A. E., Belland, B. R., & Schroeder, K. E. (2013). A Predictive Study of Student Satisfaction in Online Education Programs. *International Review of Research in Open and Distributed Learning*, *14*, 16–39.

Lachmann, S. J. (1997). Learning Is a Process: Toward an Improved Definition of Learning. *The Journal of Psychology*, *131*, 477–480.

Laney, M. O. (2005). *The Hidden Gifts of the Introverted Child: Helping Your Child Thrive in an Extroverted World*. New York, NY, Workman Publishing.

Löbler, H., & Hahn, M. (2013). Measuring value-in-context from a service-dominant logic's perspective. In N. K. Malhorta (Ed.), *Review of Marketing Research.* Emerald Group Publishing Limited.

Lemke, F., Clark, M., & Wilson, H. (2011). Customer experience quality: an exploration in business and consumer contexts using repertory grid technique. *Journal of the Academy of Marketing Science*, *39*, 846–869.

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, *80*, 69–96.

Li, Y., Duan, Y., & Alford, P. (2012). An empirical study on behavioural intention to reuse e-learning systems in rural China. *British Journal of Educational Technology*, *43*, 933–948.

Lin, C. Y. (2020). Social reaction toward the 2019 novel coronavirus (COVID-19). *Social Health and Behavior*, *3*, 1–2.

Little, B. R., Lecci, L., & Watkinson, B. (1992). Personality and Personal Projects: Linking Big Five and PAC Units of Analysis. *Journal of Personality*, *60*, 501–525.

Livingood, J. (1995). Revenge of the Introverts. *Computer-Mediated Communication Magazine*, *2*, 8.

Logan, F. A. (1970). *Fundamentals of Learning and Motivation*. WMC Brown Company Publishers.

MacDonald, E. K., Wilson, H., Martinez, V., & Toossi, A. (2011). Assessing value-in-use: A conceptual framework and exploratory study. *Industrial Marketing Management*, *40*, 671–682.

Malloy, H. D. (2007). *The relationship of learning style, academic achievement, and psychological type of baccalaureate nursing students* (Unpublished doctoral dissertation). Saint Louis University.

Manderson, L., & Levine, S. (2020). COVID-19, risk, fear, and fall-out. *Medical Anthropology*, *39*, 367–370.

Marshall, M. N. (1996). Sampling for qualitative research. *Family Practice*, *13*, 522–526.

Martin, C. L., & Pranter, C. A. (1989). Compatibility management: customer-to-customer relationships in service environments. *Journal of Services Marketing*, *3*, 5–15.

Mathwick, C., Malhorta, N. K., & Rigdon, E. (2001). Experiential value: conceptualization, measurement and application in the catalog and Internet shopping environment. *Journal of Retailing*, *77*, 39–56.

McAdams, D. P. (1997). Chapter 1 - A conceptual history of personality psychology. In R. Hogan, J. Johnson, & S. Brigss (Eds.), *Handbook of Personality Psychology.* San Diego, CA: Academic Press.

McCrae, R. R., & Costa, P. T. (1987). Validation of the Five-Factor Model of Personality Across Instruments and Observers. *Journal of Personality and Social Psychology*, *52*, 81–90.

McCrae, R. R., Costa, P. T., & Busch, C. M. (1986). Evaluating comprehensiveness in personality systems: The California Q-Set and the five-factor model. *Journal of Personality*, *54*, 430–446.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*, 175–215.

McCrae, R. R., & Terracciano, A. (2005). Universal Features of Personality Traits From the Observer's Perspective: Data From 50 Cultures. *Journal of Personality and Social Psychology*, *88*, 547–561.

McKinsey & Company. (2020). *COVID-19: Briefing materials - Global health and crisis response.* Retrieved from https://www.mckinsey.com/~/media/McKinsey/BusinessFunctions/Risk/OurInsights/COVID19Implicationsforbusiness/COVID19July23/COVID-19-Facts-and-Insights-July-23.pdf

Menger, C. (1968). *Grundsätze der Volkswirtschaftslehre*. Aalen, Scientia.

Mowen, J. C., Harris, E. G., & Bone, S. A. (2004). Personality Traits and Fear Response to Print Advertisements: Theory and an Empirical Study. *Psychology & Marketing*, *21*, 927–943.

Myers, D. G. (2014). *Psychologie*. Berlin, Heidelberg, Springer.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology*, *66*, 574–583.

Nunnally, J. C. (1978). *Psychometric Theory*. New York, NY, McGraw-Hill.

Paechter, M., Maier, B., & Macher, D. (2010). Students' expectations of, and experiences in e-learning: Their relation to learning achievements and course satisfaction. *Computers & Education*, *54*, 222–229.

Page, N. (2000). *Developing subjective and objective measures of relationship quality and investigation the discriminant validity between relationship quality and service quality* (Unpublished doctoral dissertation). University of South Australia.

Pappas, G., Kiriaze, I. J., Giannakis, P., & Falagas, M. E. (2009). Psychosocial consequences of infectious diseases. *Clinical Microbiology and Infection*, *15*, 743–747.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing*, *64*, 12–40.

Parasuraman, A., Zeithaml, V. A., & Malhorta, N. K. (2005). E-S-QUAL: A Multiple-Item Scale for Assessing Electronic Service Quality. *Journal of Service Research*, *7*, 213–234.

Payne, A. F., Storbacka, K., & Frow, P. (2008). Managing the co-creation of value. *Journal of the Academy of Marketing Science*, *36*, 83–96.

Pine, B. J., & Gilmore, J. H. (1998). Welcome to the experience economy. *Harvard Business Review*, *76*(97-105).

Puccinelli, N. M., Goodstein, R. C., Grewal, D., Price, R., Raghubir, P., & Stewart, D. (2009). Customer experience management in retailing: understanding the buying process. *Journal of Retailing*, *85*, 15–30.

Pura, M. (2005). Linking perceived value and loyalty in location-based mobile services. *Managing Service Quality*, *15*, 509–538.

Ragozzino, R. L., & Kelly, W. E. (2011). Typing the worrier: relationship between worry and Jung's personality types. *Education*, *131*, 791–797.

Raju, P. G., & Venugopal, M. (2014). Personality & Learning Styles-Lessons for Indian Corporate Trainers. *The Indian Journal of Industrial Relations*, *49*, 734–750.

Reznik, A., Gritsenko, V., Konstantinov, V., Khamenka, N., & Isarlowitz, R. (2020). Fear in Eastern Europe: Validation of the Fear of COVID-19 Scale. *International Journal of Mental Health and Addiction*.

Rober Koch Institut. (2020). *Informationen und Hilfestellungen für Personen mit einem höheren Risiko für einen schweren COVID-19-Krankheitsverlauf.* Retrieved from https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Risikogruppen.html

Rose, S., Hair, N., & Clark, M. (2011). Online Customer Experience: A Review of the Business-to-Consumer Online Purchase Context. *International Journal of Management Reviews*, *13*, 24–39.

Rowley, J. (2006). An analysis of the e-service literature: towards a research agenda. *Internet Research*, *16*, 339–359.

Sandström, S., Edvardsson, B., Kristensson, P., & Magnusson, P. (2008). Value in use through service experience. *Managing Service Quality*, *18*, 112–126.

Santos, J. (2003). E-service quality: a model of virtual service quality dimensions. *Managing Service Quality*, *13*, 233–246.

Satici, B., Gocet-Tekin, E., Deniz, M. E., & Satici, S. A. (2020). Adaptation of the Fear of COVID-19 Scale: Its association with psychological distress and life satisfaction in Turkey. *International Journal of Mental Health and Addiction*.

Schmid, U., Goertz, L., Radomski, S., Thom, S., & Behrens, J. (2017). *Monitor Digitale Bildung: Die Hochschulen im digitalen Zeitalter*. Bertelsmann Stiftung.

Schmitt, D. P., Allik, J., McCrae, R. R., & Benet-Martinez, V. (2007). The Geographic Distribution of Big Five Personality Traits: Patterns and Profiles of Human Self-Description Across 56 Nations. *Journal of cross-cultural psychology*, *38*, 173–212.

Schouten, J. A., McAlexander, J. H., & Koenig, H. F. (2007). Transcendent customer experience and brand community. *Journal of the Academy of Marketing Science*, *35*, 357–368.

Schrauf, R. W., & Navarro, E. (2005). Using Existing Tests and Scales in the Field. *Sage Publications*, *17*, 373–393.

Sen, S., & Lerman, D. (2007). Why are you telling me this? An examination into negative consumer reviews on the Web. *Journal of Interactive Marketing*, *21*, 76–94.

Shapiro, K. J., & Alexanders, I. E. (1969). Extraversion-introversion, affiliation, and anxiety. *Journal of Personality*, *37*, 387–406.

Sheatsley, P. B. (1983). Questionnaire Construction and Item Writing. In P. H. Rossi, J. D. Wright, & A. B. Anderson (Eds.), *Handbook of Survey Research.* San Diego, CA: Academic Press, Inc.

Sheth, J., Newman, B. I., & Gross, B. L. (1991). Why we buy what we buy: A theory of consumption values. *Journal of Business Research*, *22*, 159–170.

Snyder, M. (1983). The influence of individuals on situations: Implications for understanding the links between personality and social behavior. *Journal of Personality*, *51*, 497–516.

Socan, G., & Bucik, V. (1998). Relationship between speed of information-processing and two major personality dimensions - extraversion and neuroticism. *Personality and Individual Differences*, *25*, 35–48.

Soraci, P., Ferrari, A., Abbiati, F. A., del Fante, E., de Pace, R., Urso, A., & Griffiths, M. D. (2020). Validation and Psychometric Evaluation of the Italian Version of the Fear of COVID-19 Scale. *International Journal of Mental Health and Addiction*.

Swan, K. (2002). Building learning communities in online courses: The importance of interaction. *Education, Communication & Information*, *2*, 23–49.

Sweeney, J. C., & Soutar, G. N. (2001). Consumer perceived value: The development of a multiple item scale. *Journal of Retailing*, *77*, 203–220.

Tarhini, A., Al-Busaidi, K. A., Mohammed, A. B., & Magableh, M. (2017). Factors influencing students' adoption of e-learning: a structural equation modeling approach. *Journal of International Education in Business*, *10*, 164–182.

Taylor, S. (2019). *The Psychology of Pandemics: Preparing for the Next Global Outbreak of Infectious Disease*. Cambridge Scholars Publishing.

Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorder.* Erlbaum.

Tellegen, A. (1991). Personality Traits: Issues of Definition, Evidence, and Assessment. In D. Cicchetti & W. M. Grove (Eds.), *Thinking clearly about psychology: Essays in honor of Paul E. Meehl.* University of Minnesota Press.

Terho, H., Eggert, A., & Haas, W., A. andd Ulaga. (2015). How sales strategy translates into performance: The role of salesperson customer orientation and value-based selling. *Industrial Marketing Management*, *45*, 12–21.

Thakor, M. V., Suri, R., & Saleh, K. (2008). Effects of service setting and other consumers' age on the service perceptions of young consumers. *Journal of Retailing*, *84*, 137–149.

Tupes, E. C., & Christal, R. E. (1961). . *Recurrent personality factors based on trait ratings*. Lackland Air Force Base, TX: Personnel Laboratory, Aeronautical Systems Division, Air Force Systems Command, United States Air Force.

Udo, G. J., Bagchi, K. K., & Kirs, P. J. (2011). Using SERVQUAL to assess the quality of e-learning experience. *Computers in Human Behavior*, *27*, 1272–1283.

Vargo, S. L., & Lusch, R. F. (2004). Evolving to a new dominant logic for marketing. *Journal of Marketing*, *68*, 1–17.

Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology accep-

tance model. *Information Systems Research*, *11*, 342–365.

Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A., Tsiros, M., & Schlesinger, L. A. (2009). Customer Experience Creation: Determinants, Dynamics and Management Strategies. *Journal of Retailing*, *85*, 31–41.

Verleye, K. (2015). The co-creation experience from the customer perspective: its measurement and determinants. *Journal of Service Management*, *26*, 321–342.

Vollmeyer, R., & Rheinberg, F. (2000). Does motivation affect performance via persistence? *Learning and Instruction*, *10*, 293–309.

WHO. (2020a). *COVID-19 situation in the WHO European Region.* Retrieved from https://who.maps.arcgis.com/apps/opsdashboard/index.html#/ead3c6475654481ca51c248d52ab9c61

WHO. (2020b). *Pandemie der Coronavirus-Krankheit (COVID-19)*. Retrieved from https://www.euro.who.int/de/health-topics/health-emergencies/coronavirus-covid-19/novel-coronavirus-2019-ncov

Wiggins, J. S. (1968). Personality structure. *Annual Review of Psychology*(19), 293–350.

Woodruff, R. B., & Flint, D. (2006). Marketing's service-dominant logic and customer value. In R. F. Lusch & S. L. Vargo (Eds.), *The service-dominant logic of marketing: Dialog, debate, and directions.* Sharpe.

Woolridge, J. M. (2002). *Econometric analysis of cross section and panel data*. MIT Press.

Zeidner, M. (1995). Adaptive coping with test situations: A review of the literature. *Educational Psychologis*(30), 123–133.

Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. *Journal of Marketing*(52), 2–22.

Zeithaml, V. A., Parasuraman, A., & Malhorta, N. K. (2002). Service Quality Delivery through Web Sites: A Critical Review of Extant Knowledge. *Journal of the Academy of Marketing Science*(30), 362–275.

Zerres, C. (2010). Empirische Untersuchung. In C. Zerres (Ed.), *Preis-Promotions im Automobilverkauf.* Gabler.

Zigerell, J. (1984). *Distance Education: An Information Age Approach to Adult Education*. ERIC Clearingshouse.

Junior Management Science 7(2) (2022) 412-428

Junior Management Science

journal homepage: www.jums.academy

From Wuhan into the Balance Sheet: Accounting for Provisions due to COVID-19 According to the German Commercial Code (HGB) and IFRS

Von Wuhan in die Bilanz: Die Bildung von Rückstellungen aufgrund von Covid-19 nach HGB und IFRS

Lars Busch

Heinrich-Heine-Universität Düsseldorf

Abstract

The Corona pandemic has had a major impact on the social lives of people around the world since the beginning of 2020, leading to increased economic uncertainty. This uncertainty in turn has an impact on the accounting of companies.

The following paper provides an overview of the multifaceted effects of Covid-19 on a company's provisions, which represent a significant balance sheet item. Provisions that are significantly affected are provisions for impending losses, restructuring provisions, guarantee and goodwill provisions, and pension provisions. These are explained in more detail with regard to recognition and measurement according to HGB (German GAAP) and IFRS. The Covid 19-related effects, which can be both positive and negative, are clarified through corresponding practical examples. In addition, there is a brief presentation of other types of provisions that can be influenced by Covid-19.

This work makes it clear that the pandemic as a crisis will continue to influence the balance sheets of numerous companies in the coming years.

Zusammenfassung

Die Corona-Pandemie hat seit Anfang des Jahres 2020 große Auswirkungen auf das gesellschaftliche Leben der Menschen auf der ganzen Welt und führt damit auch zu einer erhöhten wirtschaftlichen Unsicherheit. Diese Unsicherheit wirkt sich wiederum auf die Rechnungslegung der Unternehmen aus.

Die folgende Arbeit liefert einen Überblick über die vielseitigen Auswirkungen von Covid-19 auf die Rückstellungen eines Unternehmens, welche einen bedeutenden Bilanzposten darstellen. Als wesentlich betroffene Rückstellungen werden Drohverlustrückstellungen, Restrukturierungsrückstellungen, Garantie- und Kulanzrückstellungen sowie Pensionsrückstellungen identifiziert. Diese werden hinsichtlich Ansatz und Bewertung nach HGB und IFRS näher erläutert. Durch entsprechende Praxisbeispiele werden die Covid-19-bedingten Auswirkungen, welche sowohl positiv als auch negativ sein können, verdeutlicht. Außerdem erfolgt eine kurze Vorstellung weiterer Rückstellungsarten, welche durch Covid-19 beeinflusst werden können.

Diese Arbeit macht deutlich, dass die Pandemie als Krise die Bilanzen zahlreicher Unternehmen auch in den kommenden Jahren beeinflussen wird.

Keywords: Corona; Covid-19; HGB; IFRS; Rückstellungen.

1. Einleitung

„The covid-19 pandemic pushed economies into a great lockdown that saved lives but also triggered the worst recession since the Great Depression"(Gopinath, 2020, Chefökonomin des IMF).

Die Corona-Pandemie hat seit Anfang des Jahres 2020 das gesellschaftliche Leben der Menschen auf der ganzen Welt fest im Griff. Zum Schutz der Gesundheit und einer Eindämmung der Infektionszahlen wurden in Deutschland unterschiedliche Maßnahmen veranlasst, wie z. B. Schlie-

DOI: https://doi.org/10.5282/jums/v7i2pp412-428

ßungen von Geschäften und Produktionsstätten, Veranstaltungsverbote, Reisewarnungen und Kontaktverbote. Diese entfalteten erhebliche negative wirtschaftliche Auswirkungen und stellen die Unternehmen vor eine Vielzahl an neuen Herausforderungen, wie z. B. Schwierigkeiten bei der Erfüllung gesetzlicher oder vertraglicher Verpflichtungen, einem Rückgang der Nachfrage und daraus resultierend eine deutlich erhöhte wirtschaftliche Unsicherheit. Diese Unsicherheit wirkt sich auch auf die Rechnungslegung der Unternehmen aus. Um dieser Unsicherheit entgegenzuwirken, veröffentlichte das IDW unterschiedliche Hinweise zu den wirtschaftlichen Folgen des Coronavirus.

In diesem Zusammenhang empfiehlt sich eine intensive Betrachtung der Auswirkungen von Covid-19 auf die Rückstellungen eines Unternehmens, welche 2017 durchschnittlich 16,20 % der Bilanzsumme ausmachten (darunter: 5,7 % Pensionsrückstellungen) und dementsprechend von großer Bedeutung sind (vgl. Deutsche Bundesbank, 2020, S. 18). In der vorliegenden Arbeit wird auf diesen Zusammenhang näher eingegangen.

Zu Beginn von Kapitel 2 wird der Begriff der Rückstellungen sowie deren Ansatz und Bewertung nach HGB und IFRS erläutert. Im Anschluss wird der Frage nachgegangen, welchen Einfluss Covid-19 auf die Unternehmen entfaltet. Ziel ist die Identifizierung der Rückstellungen, die wesentlich durch die Pandemie betroffen sein werden. Im Anschluss wird die Frage diskutiert, ob die Pandemie als wertbegründendes oder wertaufhellendes Ereignis zu beurteilen ist. Hiervon ist abhängig, ab welchen Bilanzstichtagen die Pandemieauswirkungen im Jahresabschluss berücksichtigt werden müssen.

In Kapitel 3 werden in einem ersten Schritt die in Kapitel 2 identifizierten Corona-relevanten bedeutenden Rückstellungen hinsichtlich Ansatz und Bewertung getrennt nach HGB und IFRS dargestellt. Im Anschluss daran wird der Frage nachgegangen, welche konkreten Auswirkungen Covid-19 auf die betreffenden Rückstellungen haben kann. Es folgen Praxisbeispiele, die diese Auswirkungen verdeutlichen sollen. Zum Ende des dritten Kapitels stellt der Verfasser weitere mögliche Rückstellungen vor, die im Einzelfall durch die Pandemie beeinflusst werden können. In dieser Arbeit wird das generische Maskulin verwendet. Weibliche und anderweitige Geschlechteridentitäten werden dabei ausdrücklich mitgemeint.

2. Relevanz von Covid-19 als zu bilanzierendes Risiko im Rahmen der Rückstellungen

2.1. Allgemeine Bilanzierungsgrundsätze

2.1.1. Rückstellungen nach HGB

Rückstellungen sind Passivposten, mit denen bestimmte Verpflichtungen, welche zu künftigen Ausgaben führen, gewinnmindernd erfasst werden (vgl. Schubert, 2020, Rn. 1). Gemeinsam mit den Verbindlichkeiten ergeben Rückstellungen die bilanzrechtlichen Schulden eines Unternehmens. Im Gegensatz zu den Verbindlichkeiten beschreiben Rückstellungen eine Verpflichtung, die dem Grunde und/oder der Höhe nach ungewiss ist (vgl. Baetge, Kirsch & Thiele, 2017, S. 417).

Eine allgemeine Definition der Rückstellung ist im HGB nicht enthalten, der Umfang des Begriffs „Rückstellung“ hängt vielmehr von dem zugrunde gelegten Zweck der Bilanz ab. Grundsätzlich lassen sich eine statische Interpretation und eine dynamische Interpretation differenzieren.

Die dynamische Bilanzinterpretation dient einer periodengerechten Erfolgszurechnung und impliziert die Bildung von sogenannten Aufwandsrückstellungen. Im Zusammenhang mit den Auswirkungen von Covid-19 ist die dynamische Sicht jedoch von geringer Bedeutung, sodass im Folgenden auf weitere Erläuterungen verzichtet wird.

Im Rahmen einer statischen Bilanzinterpretation wird der Gläubigerschutz als übergeordneter Zweck betrachtet. Dies gelingt durch eine Darstellung des Vermögens, welches den Gläubigern zur Befriedigung ihrer Ansprüche zur Verfügung steht (vgl. Coenenberg, Haller & Schultze, 2018, S. 433). Folglich beinhaltet die Bilanz eines Unternehmens sämtliche Verpflichtungen gegenüber Dritten (Außenverpflichtungen), die das Haftungsvermögen schmälern. Eine Außenverpflichtung beschreibt einen rechtlichen oder faktischen/wirtschaftlichen Leistungszwang gegenüber Dritten (vgl. Schubert, 2020, Rn. 1).

Eine rechtliche Verpflichtung kann sich aus Zivilrecht oder aus öffentlichem Recht ergeben. Wenn sich ein Unternehmen aus tatsächlichen oder wirtschaftlichen Gründen einer Verpflichtung nicht entziehen kann oder möchte, wird eine faktische Verpflichtung angenommen. Dies setzt jedoch nachvollziehbare und objektive Anhaltspunkte voraus, aus denen deutlich wird, dass das Unternehmen dieser Verpflichtung tatsächlich nachkommen wird (vgl. Ballwieser, 2020, Rn. 11; Nagengast & Boecker, 2016).

Es handelt sich um Verpflichtungen, die bezüglich ihrer Höhe, ihres zeitlichen Eintretens und/oder ihres Bestandes nach ungewiss sind, deren Aufkommen aber hinreichend wahrscheinlich ist. Diese letztlich noch nicht konkretisierten Schulden mit Verpflichtungscharakter werden als Rückstellungen für ungewisse Verbindlichkeiten bezeichnet. Die Bestimmung der Wahrscheinlichkeit für den Einzelfall erfordert eine Berücksichtigung der Erfahrungen aus der Vergangenheit und impliziert, dass ernsthaft mit einer Inanspruchnahme gerechnet werden muss (vgl. Bertram, 2019, Rz. 43).

Laut Coenenberg et al. (2018, S. 434) ergibt sich die Notwendigkeit zur Bildung von Rückstellungen für ungewisse Verbindlichkeiten außerdem aus dem Grundsatz der Vorsicht und dem daraus abgeleiteten Imparitätsprinzip i. V. m. dem Vollständigkeitsgebot nach § 246 Abs. 1 HGB. Demnach müssen am Bilanzstichtag abzusehende Risiken und Schulden eines Unternehmens erfasst werden, auch wenn sie noch nicht juristisch festgestellt, sondern lediglich wirtschaftlich verursacht sind (vgl. Thommen et al., 2020a, S. 238; Schwirkslies-Filler, 2019, S. 12).

Wie bereits erwähnt gibt es keine allgemeine handelsrechtliche Definition für den Passivposten „Rückstellungen“, dafür zählt § 249 HGB abschließend die Zwecke auf, für die Rückstellungen gebildet werden dürfen (vgl. Schubert, 2020,

Rn. 6). Gemäß § 249 Abs. 1 S. 1 HGB hat jeder Kaufmann für ungewisse Verbindlichkeiten und für drohende Verluste aus schwebenden Geschäften Rückstellungen zu bilden. Zusätzlich zu diesen, auf rechtlichen Verpflichtungen beruhenden, Rückstellungen enthält § 249 Abs. 1 S. 2 Nr. 2 HGB Gewährleistungen ohne rechtliche Verpflichtung, zu denen der Kaufmann i. d. R. aus wirtschaftlichen Gründen verpflichtet ist. Inhaltlich handelt es sich somit ebenfalls um Rückstellungen für Außenverpflichtungen.

Außerdem beinhaltet § 249 Abs. 1 S. 2 Nr. 1 HGB Rückstellungen für Innenverpflichtungen, sogenannte Aufwandsrückstellungen (vgl. Ballwieser, 2020, Rn. 7). Diese werden aus Mangel an Bedeutung im Zusammenhang mit Covid-19 hier nicht weiter betrachtet. Nach § 249 Abs. 2 S. 1 HGB dürfen für andere Zwecke keine Rückstellungen gebildet werden.

Die Bewertung von Rückstellungen erfolgt gem. § 253 Abs. 1 S. 2 HGB „in Höhe des nach vernünftiger kaufmännischer Beurteilung notwendigen Erfüllungsbetrages“. Die Verwendung des Begriffes „Erfüllungsbetrag“ verdeutlicht, dass – unter Einschränkung des Stichtagsprinzips – künftige Preis- und Kostenänderungen bei der Rückstellungsbewertung ausdrücklich einzubeziehen sind. Künftige Änderungen sind zu schätzen und dürfen erst berücksichtigt werden, wenn ausreichend objektive Hinweise vorliegen (vgl. Schwirkslies-Filler, 2019, S. 22). Aufgrund dieses nicht unerheblichen Ermessensspielraumes, handelt es sich bei Rückstellungen um ein interessantes Instrument der Bilanzpolitik (vgl. Baetge et al., 2017, S. 427; Weber & Weißenberger, 2015, S. 146).

Wenn der Betrag einer Rückstellung unsicher ist und somit geschätzt werden muss, erfolgt dies unter Berücksichtigung des in § 252 Abs. 1 Nr. 4 HGB genannten Grundsatzes der Vorsicht.

Zusätzlich impliziert der Wortlaut „nach vernünftiger kaufmännischer Beurteilung“ laut Baetge, Kirsch und Thiele eine vollständige Auswertung sämtlicher zugänglicher Informationen, um alle vorhersehbaren Verpflichtungen bzw. Risiken zu identifizieren und anschließend in der Rückstellungsbewertung zu berücksichtigen. Diese Informationsauswertung soll für sachverständige Dritte schlüssig und nachvollziehbar sein (vgl. Baetge et al., 2017, S. 431).

Laut § 253 Abs. 2 S. 1 HGB sind Rückstellungen mit einer Restlaufzeit von mehr als einem Jahr mit dem durchschnittlichen Marktzinssatz der vergangen sieben Geschäftsjahre und bei Pensionsrückstellungen der vergangenen zehn Geschäftsjahre abzuzinsen. Hintergrund dieses Abzinsungsgebotes ist die Berücksichtigung des Finanzierungseffektes von Rückstellungen (vgl. Tanski, 2019, S. 231).

2.1.2. Rückstellungen nach IFRS

Bei Rückstellungen im Rahmen der IFRS handelt es sich nach IAS 37.10 um „Schulden, die bezüglich ihrer Fälligkeit oder Höhe ungewiss sind“. Charakteristisch für sämtliche Rückstellungen ist somit ein gewisser Unsicherheitsfaktor (vgl. Schrimpf-Dörges, 2020, Rn. 7). Gemeinsam mit den Verbindlichkeiten, welche hinsichtlich ihres Eintritts und ihrer Höhe sicher sind, bilden sie die Schulden eines Unternehmens (vgl. Thommen et al., 2020b, S. 259). Die Bilanzierung von Rückstellungen wird insbesondere in den Standards IAS 19 und IAS 37 behandelt, wobei sich die wesentlichen Vorschriften in IAS 37 befinden und deswegen im Folgenden näher erläutert werden.

IAS 37.14 beinhaltet drei Grundvoraussetzungen für den Ansatz von Rückstellungen, welche kumulativ vorliegen müssen:

(1) Einem Unternehmen ist aus einem Ereignis der Vergangenheit eine gegenwärtige rechtliche oder faktische Verpflichtung entstanden.
(2) Ein Abfluss von Ressourcen mit wirtschaftlichem Nutzen zur Erfüllung dieser Verpflichtung muss wahrscheinlich sein.
(3) Die Höhe dieser Verpflichtung muss verlässlich geschätzt werden können.

Laut IAS 37.15 muss unter Berücksichtigung aller substanziellen Hinweise mehr für als gegen das Bestehen einer gegenwärtigen Verpflichtung sprechen. In der Praxis wird hier eine Wahrscheinlichkeit von über 50 % erwartet, dabei kommt es insbesondere auf den Inhalt und nicht die Anzahl der Gründe an („more likely than not“) (vgl. Schrimpf-Dörges, 2020, Rn. 28; Pellens, Fülbier, Gassen & Selhorn, 2017, S. 511). Das Vorliegen einer gegenwärtigen rechtlichen oder faktischen Verpflichtung gegenüber Dritten impliziert, dass ausschließlich für Außenverpflichtungen Rückstellungen gebildet werden dürfen (vgl. IAS 37.20). Eine Passivierung von Innenverpflichtungen wird damit ausgeschlossen. Gleichzeitig muss die Verpflichtung aus vergangenen Ereignissen entstehen und darf damit im Umkehrschluss nicht von künftigen Entscheidungen abhängen (vgl. IAS 37.19; Schrimpf-Dörges, 2020, Rn. 45; Thommen et al., 2020b, S. 274).

Eine rechtliche Verpflichtung ergibt sich aus einem privatrechtlichen Vertrag oder gesetzlichen Vorschriften und ist ggf. auch zwangsweise durchsetzbar, das Unternehmen kann sich der Verpflichtung somit nicht entziehen. Bei einer faktischen Verpflichtung hat sich das Unternehmen die Pflicht zur Erfüllung selbst auferlegt. Diese liegt vor, wenn bei einer anderen Partei eine gerechtfertigte Erwartung geweckt wurde, dass das Unternehmen seiner Verpflichtung nachkommt (vgl. Roos, 2016).

Zusätzlich muss mit dem Abfluss von Ressourcen mit wirtschaftlichem Nutzen im Zusammenhang mit der Erfüllung der Verpflichtung gerechnet werden. Gemäß IAS 37.23 ist dies der Fall, wenn mehr dafür als dagegen spricht, was ebenfalls eine Wahrscheinlichkeit von über 50 % impliziert („more likely than not“). Die subjektive Einschätzung des Bilanzierenden hat hier großen Einfluss und kann damit über den Ansatz oder Nichtansatz einer Rückstellung entscheiden (vgl. Thommen et al., 2020b, S. 274).

IAS 37.10 verlangt für den Ansatz einer Rückstellung außerdem eine verlässliche Schätzung der Höhe der Verpflichtung. Diese Voraussetzung sollte allerdings nach IAS 37.25

in den wenigsten Fällen den Ansatz einer Rückstellung verhindern, denn die Identifizierung einer Bandbreite möglicher Werte ist hier bereits ausreichend (vgl. Roos, 2016).

Wenn eine oder mehrere dieser Grundvoraussetzungen nicht erfüllt sind, liegt möglicherweise eine Eventualschuld vor, welche nach IAS 37.27 nicht passiviert werden darf. Stattdessen erfolgt eine kurze Beschreibung im Anhang (vgl. IAS 37.68).

Für die Bewertung einer Rückstellung fordert IAS 37.36 die bestmögliche Schätzung der künftigen Ausgaben, die zur Erfüllung einer gegenwärtigen Verpflichtung benötigt werden („best estimate"). „Bestmöglich" impliziert in diesem Zusammenhang einen Betrag, welchen das Unternehmen bei vernünftiger kaufmännischer Betrachtung zahlen würde, um die Verpflichtung am Bilanzstichtag zu erfüllen oder einem Dritten zu übertragen (vgl. IAS 37.37). Für den Bilanzierenden entsteht hier ein subjektiver Ermessensspielraum, welcher durch eine notwendige Berücksichtigung der zugrundeliegenden Risiken und Unsicherheiten eingeschränkt wird (vgl. IAS 37.42). Unsicherheiten rechtfertigen gem. IAS 37.43 jedoch keine Überbewertung der Rückstellungen bzw. die Bildung von stillen Reserven. Zusätzlich sollen regelmäßig objektive Informationsquellen, wie z.B. Erfahrungen aus ähnlichen Sachverhalten sowie Einschätzungen unabhängiger Experten in die Entscheidungsfindung integriert werden (vgl. Schrimpf-Dörges, 2020, Rn. 53; Pellens et al., 2017, S. 514).

Existieren bei einzelnen Sachverhalten jeweils unterschiedliche Bewertungen eines möglichen Ressourcenabflusses, stellt laut IAS 37.40 der wahrscheinlichste Betrag die bestmögliche Schätzung dar (vgl. Ruhnke & Simons, 2018, S. 546).

Bei Rückstellungen mit einer Restlaufzeit von mehr als einem Jahr wird eine Überwertung vermieden, indem der Zins- und Zinseszinseffekt nach IAS 37.45 i. V. m. IAS 37.47 auf erwartete Ressourcenabflüsse angewendet wird. Die bestmögliche Schätzung entspricht folglich dem Barwert der Auszahlung, welcher für die Begleichung der Verpflichtung erforderlich ist (vgl. Pellens et al., 2017, S. 516).

Um ungewissen Entwicklungen gerecht zu werden, sollen künftige Änderungen von Preisen bzw. Kosten, welche objektiv nachweisbar sind, gem. IAS 37.48 in der Bewertung berücksichtigt werden. Bloße Ermessensentscheidungen des Managements reichen in diesem Zusammenhang nicht aus (vgl. Schrimpf-Dörges, 2020, Rn. 66; Baetge et al., 2017, S. 466).

2.2. Einfluss von Covid-19 als zu bilanzierendes Risiko im Rahmen der Rückstellungen

2.2.1. Auswirkungen von Covid-19 auf die Unternehmen im Hinblick auf Rückstellungen

Im Dezember 2019 kam es in der chinesischen Stadt Wuhan zu ersten Covid-19-Infektionen beim Menschen. In den folgenden Wochen stieg nicht nur die Anzahl der Fälle in China, sondern auch außerhalb Chinas kam es in 18 weiteren Ländern zu bestätigten Fällen, sodass die WHO am 30. Januar 2020 den Ausbruch des neuartigen Coronavirus zu einer gesundheitlichen Notlage von internationaler Tragweite erklärte. Diese Erklärung sollte sich bewahrheiten, denn bereits am 11. März 2020 erklärte der Generaldirektor der WHO den Corona-Ausbruch offiziell zu einer Pandemie. Zu diesem Zeitpunkt gab es bereits mehr als 118.000 bestätigte Fälle und insgesamt über 4.000 gemeldete Todesfälle (vgl. Mitteilung WHO).

Auch Deutschland steht nun seit dem Frühjahr 2020 vor einer besonderen Herausforderung mit weitreichenden Folgen für das gesellschaftliche und soziale Leben der Menschen. Mit dem Ziel der Eindämmung der Infektionszahlen und damit einer Entlastung des Gesundheitssystems wurden von der Bundesregierung eine Vielzahl an Maßnahmen erlassen, die das öffentliche Leben stark einschränken. Diese Einschränkungen wirken sich wiederum, in Abhängigkeit ihrer Betroffenheit in einem jeweils unterschiedlichen Ausmaß, auf den Jahresabschluss von Unternehmen aus und verdeutlichen eine gesamtwirtschaftliche Abhängigkeit vom öffentlichen Leben (vgl. Raffelhüschen, Bahnsen, Kohlstruck, Seuffert & Wimmesberger, 2021).

Bereits im April 2020 ging die IMF (vgl. International Monetary Fund, 2020) davon aus, dass die Weltwirtschaft durch die Auswirkungen der Corona-Pandemie ihren stärksten Rückgang seit dem zweiten Weltkrieg erleben wird. Anders als bei der Finanzmarktkrise 2009 erstrecken sich die Auswirkungen der Corona-Pandemie gleichermaßen auf Industrie und Dienstleister (vgl. Bardt & Hürther, 2020; Bardt & Hürther, 2020). Laut einer Studie von KANTAR im Auftrag des Bundesministeriums für Wirtschaft und Energie verzeichneten in Deutschland in April 2020 drei von vier Unternehmen negative wirtschaftliche Auswirkungen, nur bei sehr wenigen Unternehmen hat die Pandemie positive wirtschaftliche Auswirkungen. Aus makroökonomischer Sicht kommt es infolge des Einbruchs der globalen wirtschaftlichen Aktivitäten zu einem negativen Angebotsschock und gleichzeitig zu einem negativen Nachfrageschock (vgl. Grömling, Hüther, Beznoska & Demary, 2020).

Abbildung 1 zeigt das Ergebnis einer Unternehmensumfrage auf die Frage nach den Auswirkungen der Pandemie auf die Geschäfte der Unternehmen:

Auf der Angebotsseite betrifft dies insbesondere Unternehmen mit internationalen Produktions- und Lieferketten (vgl. Ragnitz, 2020). Es kann zu Störungen der Wertschöpfungsketten kommen, beispielsweise durch Schwierigkeiten beim Bezug von nationalen und internationalen Vorleistungen bzw. Zwischenprodukten sowie Betriebsschließungen und Produktionseinbrüchen (vgl. Grömling et al., 2020). Dadurch können bestehende Beschaffungs- oder Absatzgeschäfte verlustbringend werden oder es kommt zu Schwierigkeiten bei der Erfüllung gesetzlicher oder vertraglicher Verpflichtungen, was wiederum zu Strafen führen kann. Auch bei Dauerschuldverhältnissen, beispielsweise bei Mietverträgen, kann es zu einem Missverhältnis der Leistungen kommen (vgl. Zwirner, Busch & Krauß, 2020). Die aufgeführten Sachverhalte erfordern möglicherweise die Bildung einer sogenannten Drohverlustrückstellung.

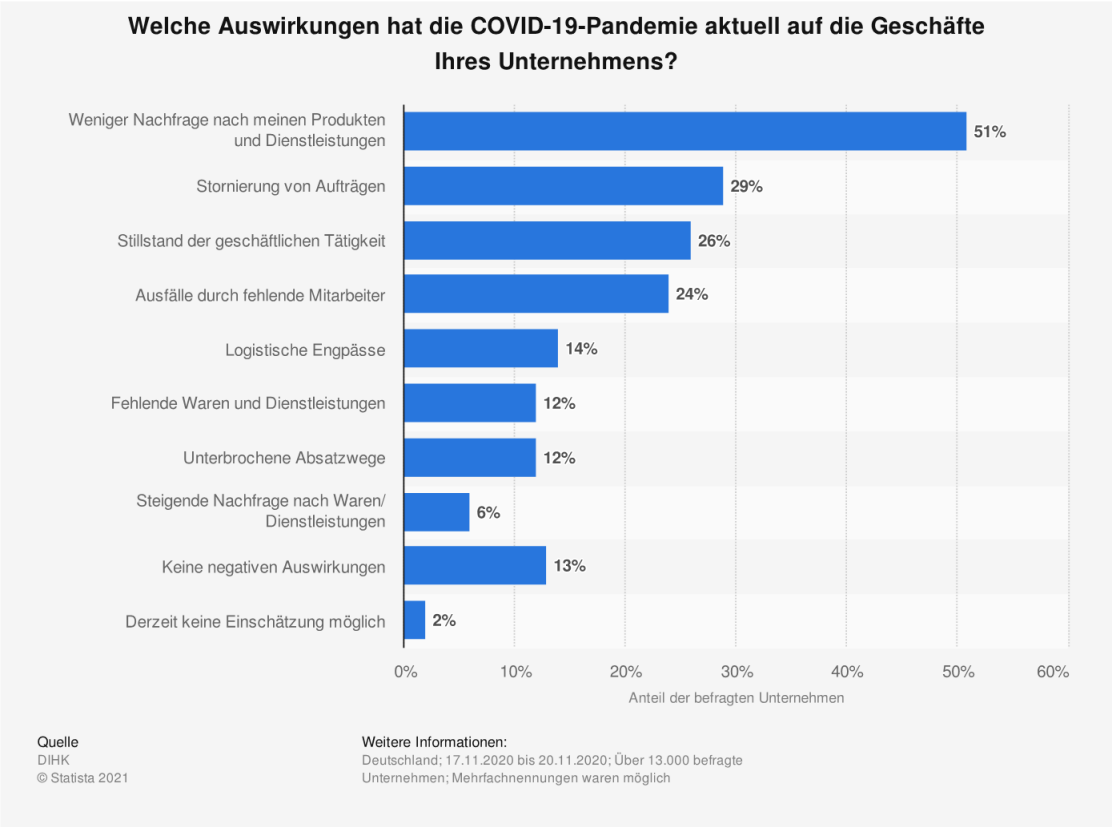

Abbildung 1: Umfrageergebnisse des DIHK zu den Auswirkungen des Coronavirus auf die Geschäfte deutscher Unternehmen im November 2020 (übernommen von: Statista 2021)

Auf der Nachfrageseite kommt es zu einem Rückgang der nationalen- und internationalen Nachfrage sowie logistischen Schwierigkeiten beim Absatz, bedingt durch eine erhöhte wirtschaftliche Unsicherheit, Einkommensverluste und blockierte Absatzwege (vgl. Bardt & Hürther, 2020).

Im April 2020 waren laut KANTAR in 8 % der Unternehmen Entlassungen geplant. Stand Juni 2020 planten laut einer Umfrage des DIHK bereits 20 % der befragten Unternehmen den Abbau von Mitarbeitern (vgl. Kantar, 2020; DIHK, 2020). Für die betreffenden Jahresabschlüsse wird die Bilanzierung von Restrukturierungs- und Abfindungsrückstellungen zu prüfen sein.

Im Falle einer Aufrechterhaltung der Einschränkungen aufgrund der Corona-Pandemie, gaben im April 2020 mehr als 50 % aller in einer Umfrage des ifo-Instituts befragten Unternehmen aus Deutschland an, dass sie nicht länger als 6 Monate überleben werden. Für den deutschen Einzelhandel stellen die derzeitigen Einschränkungen eine große Bedrohung dar, von den Unternehmen dieser Branche gaben fast 45 % an, nicht länger als 3 Monate überleben zu können (vgl. ifo-Institut, 2020).

Der stationäre Einzelhandel reagiert insbesondere mit Filialschließungen und damit einhergehenden Mitarbeiterentlassungen. Ähnliches ist im Bankensektor zu beobachten: Die Beratungsgesellschaft Investors Marketing geht davon aus, dass allein durch die Corona-Krise bis 2025 zusätzlich 3.500 Niederlassungen wegfallen werden, insgesamt werde ihre Zahl um 10.700 auf dann noch rund 16.000 zurückgehen (vgl. Kunz, 2020).

In den Jahresabschlüssen der betroffenen Unternehmen wird die Notwendigkeit von Drohverlustrückstellungen im Zusammenhang mit schließungsbedingten (verlustbehafteten) Filialmietverträgen sowie Restrukturierungs- und Abfindungsrückstellungen zu prüfen sein.

Diese beiden Branchen stehen beispielhaft für obige Auswirkungen. Andere bedeutende Branchen wie die Automobilindustrie und deren Zulieferer, die Chemieindustrie sowie die Luftfahrt sind nur einige, die ebenfalls erheblich von der Pandemie betroffen sind.

In der Pandemie zeigt sich, dass die Menschen, bedingt durch eine eingeschränkte Bewegungsfreiheit und einem zunehmenden Angebot von Home-Office, ihr Auto weniger nutzen, sodass die Nachfrage nach Benzin und die Abnutzung der Autos sinkt. Dies hat möglicherweise Auswirkungen auf die Garantie- und Kulanzrückstellungen eines Automobil-Unternehmens.

Somit kann grundsätzlich davon ausgegangen werden, dass die Folgen der Pandemie die wirtschaftliche Unsicherheit erhöhen und damit zu einer längerfristig angespannten wirtschaftlichen Lage beitragen werden, was sich wiederum auf die Marktzinssätze, Lohn-und Gehaltstrends sowie Karriere- und Rententrends auswirken wird. Infolgedes-

sen müssen Unternehmen möglicherweise ihre Pensionsrückstellungen anpassen (vgl. Müller, 2021).

2.2.2. Relevanz für ausgewählte Bilanzstichtage

Die Folgen der Corona-Pandemie sowie die Maßnahmen zur Eindämmung der Infektionszahlen können Einfluss auf die Rechnungslegung (HGB/IFRS) und damit auf den Ansatz und die Bewertung von Rückstellungen betroffener Unternehmen haben. Fraglich ist nun, ab welchem Zeitpunkt diese Auswirkungen bei der Aufstellung des Jahres- oder Konzernabschlusses zu berücksichtigen sind.

Im Folgenden wird zur Vereinfachung die Relevanz von Covid-19 für Bilanzstichtage zum 31.12.2019 und ab dem 31.03.2020 betrachtet. Zusätzlich erfolgt die Annahme einer Unternehmensfortführung nach § 252 Abs. 1 Nr. 2 HGB bzw. eine „Going concern-Annahme“ nach IAS 1.25.

Im Rahmen der HGB-Rechnungslegung wird zur Beantwortung dieser Frage in § 252 Abs. 1 Nr. 3 HGB, als Folge des Stichtagsprinzips, zwischen wertaufhellenden und wertbegründenden Ereignissen differenziert. Grundsätzlich müssen laut dem Stichtagsprinzip alle Ereignisse, die vor dem Abschlussstichtag entstanden sind und zu diesem Zeitpunkt bekannt waren, bei der Erstellung des Jahresabschlusses berücksichtigt werden. Wenn ein relevantes Ereignis mit bilanziellen Konsequenzen vor dem Bilanzstichtag zum 31.12.2019 verursacht wurde und bis zum Tag der Aufstellung des Jahresabschlusses bekannt geworden ist, müsste es als wertaufhellendes Ereignis berücksichtigt werden (vgl. Jonas, 2021).

Wenn ein relevantes Ereignis mit bilanziellen Konsequenzen im Zeitraum zwischen Abschluss- und Aufstellungsstichtag bekannt wurde, dabei aber erst nach dem Abschlussstichtag verursacht wurde, müsste es als wertbegründendes Ereignis identifiziert werden und darf als solches zum 31.12.2019 grundsätzlich nicht berücksichtigt werden (vgl. Störk & Büssow, 2020, Rz.38). Als Folge des Stichtagsprinzips findet es somit erst in der Folgeperiode Berücksichtigung (vgl. Zwirner et al., 2020). Für Kapitalgesellschaften gibt es unter Umständen nach §285 Nr. 33 HGB zusätzlich die Vorschrift, wertbegründende Informationen im Anhang zu erläutern (vgl. Rimmelspacher & Kliem, 2020; Baetge et al., 2017, S. 122).

Im Dezember 2019 gab es erste Covid-19-Infektionen bei Menschen, welche damals regional auf die Stadt Wuhan in China begrenzt waren. In den folgenden Monaten stieg die Zahl der globalen Infektionen rasant an, sodass durch die WHO am 11.03.2020 die offizielle Erklärung des Covid-19 Ausbruch zur Pandemie folgte (vgl. WHO-Mitteilung). Daraus ergeben sich laut den Fachlichen Hinweisen des IDW (2020 (a), S.1) für Bilanzstichtage zum 31.12.2019 grundsätzlich noch keine objektiven Anhaltspunkte für die Folgen der Corona-Pandemie auf den Jahresabschluss. Damit ist die Corona-Pandemie grundsätzlich als wertbegründendes Ereignis einzustufen und darf als solches nicht in den Abschlüssen berücksichtigt werden (vgl. Müller & Reinke, 2020).

Eine mögliche Ausnahme stellen jedoch Unternehmen mit engen Geschäftsbeziehungen nach Wuhan dar, bei denen Lieferketten unterbrochen und Vorprodukte nicht bezogen werden konnten. Diese unterschiedliche Einordnung verdeutlicht die Bedeutung einer notwendigen einzelfallspezifischen und unternehmensindividuellen Betrachtung. Darüber hinaus ist „Covid-19“ als laufender Prozess zu betrachten, welcher stets neue Ereignisse auslösen kann, was eine Qualifizierung zusätzlich erschwert (vgl. Rimmelspacher & Kliem, 2020).

Mögliche Implikationen für den Anhang und Lagebericht werden im Rahmen dieser Arbeit zur Vereinfachung vernachlässigt.

Mit Hinblick auf Bilanzstichtage ab dem 31.03.2020 ist das Ereignis „Covid-19“ grundsätzlich als wertaufhellend zu betrachten und führt als Folge des Stichtagsprinzips dazu, dass die Auswirkungen von Covid-19 sorgfältig und gewissenhaft auf Grundlage objektiver Tatsachen bei Ansatz und Bewertung von Rückstellungen zu berücksichtigen sind. Ungeachtet dessen ist eine einzelfallspezifische Betrachtung notwendig, da prinzipiell auch wertbegründende Ereignisse eintreten können (vgl. IDW, 2020a, S. 6; Rimmelspacher & Kliem, 2020).

Im Rahmen der IFRS-Rechnungslegung unterscheidet man, vergleichbar mit der HGB-Rechnungslegung, berücksichtigungspflichtige („adjusting events“) und nicht zu berücksichtigende Ereignisse („non-adjusting events“) nach dem Abschlussstichtag.

Laut IAS 10.3 (a) werden Ereignisse, die weitere substanzielle Hinweise zu Gegebenheiten liefern, die bereits am Abschlussstichtag vorgelegen haben, als berücksichtigungspflichtig eingestuft und führen nach IAS 10.8 zu einer Anpassung der erfassten Beträge. IAS 10.3 (b) beschreibt nicht zu berücksichtigende Ereignisse als Ereignisse, die Gegebenheiten anzeigen, die nach dem Abschlussstichtag eingetreten sind. Im Abschluss erfasste Beträge dürfen laut IAS 10.10 somit nicht angepasst werden, um nicht zu berücksichtigende Ereignisse nach dem Abschlussstichtag abzubilden.

Für eine Qualifizierung von Covid-19 und deren bilanzielle Konsequenzen als berücksichtigungspflichtiges bzw. nicht zu berücksichtigendes Ereignis ergeben sich laut den Fachlichen Hinweisen des IDW (2020b, S. 2) die gleichen Implikationen für Einzel- oder Konzernabschlüsse nach IFRS, wie für bereits beschriebene Abschlüsse nach HGB (vgl. Jonas, 2021; Rahe, 2020).

3. Relevante Rückstellungen nach HGB und IFRS im Zusammenhang mit Covid-19

3.1. Drohverlustrückstellungen

3.1.1. Allgemeine Grundsätze nach HGB

Eine Notwendigkeit zur Bildung von Rückstellungen für drohende Verluste aus schwebenden Geschäften (im Folgenden: „Drohverlustrückstellung“) ergibt sich aus § 249 Abs. 1 S. 1 HGB. Als Folge des Vorsichtsprinzips aus § 252 Abs. 1 Nr. 4 HGB dient die Drohverlustrückstellung der Verlustantizipation und einer vollständigen Schuldenerfassung, mit dem übergeordneten Ziel des Gläubigerschutzes (vgl. Schubert,

2020, Rn. 51; Ballwieser, 2020, Rn. 57). Daraus ergibt sich eine zukunftsorientierte Drohverlustrückstellung.

Bei schwebenden Geschäften i. S. d. § 249 Abs. 1 S. 1 HGB handelt es sich um zweiseitig verpflichtende Verträge, welche auf einen Leistungsaustausch (Sach- oder Dienstleistung) gerichtet sind, wobei die Verpflichtungen bislang noch von keiner der beiden Seiten erfüllt worden sind (vgl. IDW RS HFA 4, Tz. 2). Das Vorliegen einer Außenverpflichtung ist somit zwingend notwendig (vgl. Schubert, 2020, Rn. 52). Schwebende Geschäfte können sich auf Absatz- oder Beschaffungsgeschäfte sowie Dauerschuldverhältnisse beziehen. Mit dem rechtswirksamen Vertragsschluss beginnt der Schwebezustand eines Geschäfts, wobei nach Merkt auch ein bindendes Angebot des Bilanzierenden ausreichen kann (vgl. Merkt, 2021, Rn. 19).

Grundsätzlich wird ein schwebendes Geschäft als Folge des Realisationsprinzips nicht bilanziert, da zunächst eine sogenannte Ausgeglichenheitsvermutung angenommen wird, nach der sich die gegenseitig geschuldeten Verpflichtungen wertmäßig ausgleichen (vgl. Schwirkslies-Filler, 2019, S. 102). Wenn diese Annahme nicht mehr haltbar ist und aufgrund konkreter Anhaltspunkte ein Verlust aus einem schwebenden Geschäft droht, muss entsprechend dem Imparitätsprinzip, eine Drohverlustrückstellung gebildet werden. Die bloße Möglichkeit eines Verlusteintritts genügt somit nicht (vgl. IDW RS HFA 4, Tz. 15; Schubert, 2020, Rn. 60).

Ein drohender Verlust ist ein Verpflichtungsüberschuss, welcher die Differenz aus einer Saldierung der künftigen Ausgaben und der künftigen Einnahmen beschreibt (vgl. Ballwieser, 2020, Rn. 57).

Die Bewertung von Drohverlustrückstellungen erfolgt gem. § 253 Abs. 1 S. 2 HGB in Höhe des Erfüllungsbetrags, der zur Antizipation der drohenden Verluste notwendig wird und einer vernünftigen kaufmännischen Beurteilung entspricht (vgl. IDW RS HFA 4, Tz. 28; Schwirkslies-Filler, 2019, S. 108). Dies beinhaltet auch eine Anpassung der drohenden Verluste anhand zugrunde gelegter Wertverhältnisse am Abschlussstichtag sowie vorhersehbarer zukünftiger Preis- und Kostenänderungen (vgl. IDW RS HFA 4, Tz. 38f.).

3.1.2. Allgemeine Grundsätze nach IFRS

Ähnlich der Drohverlustrückstellungen nach § 249 Abs. 1 S. 1 HGB erfolgt im Rahmen der IFRS nach IAS 37.66 die Bildung von Rückstellungen für belastenden Verträge (im Folgenden: „Drohverlustrückstellung“). Diese beziehen sich grundsätzlich auf schwebende Absatz- und Beschaffungsgeschäfte (vgl. IAS 2.31) sowie Dauerschuldverhältnisse.

Bei einem belastenden Vertrag handelt es sich gem. IAS 37.10 i. V. m. IAS 37.68 um eine rechtliche Verpflichtung, bei deren Erfüllung dem Unternehmen unvermeidbare Kosten entstehen, welche höher sind als der erwartete wirtschaftliche Nutzen.

Dieses Ungleichgewicht wird auch als Verpflichtungsüberschuss bezeichnet und ergibt sich aus einer Saldierung von Aufwendungen und Erträgen (vgl. Schrimpf-Dörges, 2020, Rn. 119).

Unvermeidbare Kosten beschreiben in diesem Zusammenhang den niedrigeren Wert aus den Nettokosten für die Erfüllung sowie den möglichen Straf- und Entschädigungszahlungen als Folge für die Nichterfüllung (vgl. IAS 37.68). Der Wortlaut „unvermeidbar“ impliziert außerdem, dass bei einer möglichen kostenfreien Auflösung des Vertrags keine Rückstellung gebildet werden darf.

Im Rahmen der Ermittlung von drohenden Verlusten bei schwebenden Geschäften erfolgt grundsätzlich eine absatzmarktorientierte Bewertung. Infolgedessen dürfen Vermögenswerte aus Beschaffungsgeschäften oder die daraus resultierenden fertigen Erzeugnisse und Leistungen nicht als Drohverlustrückstellung passiviert werden, wenn der mögliche Absatzpreis den Anschaffungs- oder Herstellungskosten entspricht (vgl. IAS 2.31; Schrimpf-Dörges, 2020, Rn. 120).

Laut IAS 37.69 erfolgt für Vermögenswerte in direktem Zusammenhang mit dem schwebenden Geschäft zunächst eine außerplanmäßige Wertminderung. Dadurch wird eine doppelte Erfassung von vertraglich bedingten Wertminderungen vermieden. Die Passivierung einer Drohverlustrückstellung erfolgt somit in Höhe der unvermeidbaren Nettokosten, welche den bisherigen Buchwert der Vermögenswerte übersteigen (vgl. Schrimpf-Dörges, 2020, Rn. 121).

Künftige operative Verluste oder Kosten für die Wiederherstellung des (früheren) Geschäftsbetriebes sind für die Bildung einer Rückstellung grundsätzlich ausgeschlossen, da diese weder auf eine gegenwärtige Verpflichtung aus einem vergangenen Ereignis noch auf bilanzierte Vermögenswerte zurückzuführen sind (vgl. IAS 37.10; Schrimpf-Dörges, 2020, Rn. 123; Pellens et al., 2017, S. 520).

3.1.3. Covid-19-bedingte Sachverhalte und Auswirkungen

Als Folge der Corona-Pandemie kommt es in vielen Branchen zu Störungen vertraglicher Rechtsbeziehungen bzw. schwebender Geschäfte, welche sich sowohl auf **Absatz- als auch Beschaffungsgeschäfte** auswirken können (vgl. IDW, 2020a, S. 10). Diese Störungen können möglicherweise zu einer wertmäßig erhöhten Leistungsverpflichtung oder zu einem wertmäßig gesunkenen Gegenleistungsanspruch führen. Darüber hinaus kann die verspätete Erfüllung bzw. Nichterfüllung von vertraglichen Verpflichtungen eines Absatzgeschäftes zu Schadensersatzverpflichtungen oder Konventionalstrafen führen (vgl. Henkel, Isensee & Esquivel, 2021; Rinker, 2020).

Aber auch Schadensersatzverpflichtungen aus Beschaffungsverträgen bei einer fehlenden hinreichenden Abnahme (z.B. bei take or pay-Verträgen) sind denkbar (vgl. Schütte & Götz, 2021).

Wenn einem Unternehmen beispielsweise aufgrund von Störungen in der Lieferkette oder stillgelegten Produktionsstätten höhere Kosten entstehen, können schwebende Absatz- und Beschaffungsgeschäfte verlustbringend werden (vgl. Fischer, 2020; Kirsch, 2021). Auch gesunkene Verkaufserlöse aufgrund von geplanten Notverkäufen, Abnahmeverpflichtungen nicht mehr benötigter Produkte bzw. Dienstleistungen oder Preissenkungen als Folge eines Nachfragerückgangs sind denkbar (vgl. Rimmelspacher & Kliem,

2020; Zwirner et al., 2020).

Folglich handelt es sich hier um Sachverhalte, bei denen die Ausgeglichenheitsvermutung nicht aufrechterhalten werden kann, sodass die die Bildung einer Drohverlustrückstellung nach § 249 Abs. 1 S. 1 HGB bzw. IAS 37.66-69 in Betracht kommt (vgl. Müller & Reinke, 2020).

Wenn ein Unternehmen die Ausgeglichenheitsvermutung aus einem schwebenden bzw. belastenden Geschäft durch die Folgen der Corona-Pandemie nicht aufrechterhalten kann, empfiehlt sich eine Überprüfung der zugrunde liegenden Verträge hinsichtlich höherer Gewalt, sog. **Force Majeure-Klauseln**, wodurch die Pflicht zur Leistungserfüllung ggf. entfällt (vgl. IDW, 2020a, S. 11; Zwirner et al., 2020).

Eine Force Majeure-Klausel wird regelmäßig in Lieferverträgen und allgemeinen Geschäftsbedingungen von Unternehmen integriert und führt zu einer Leistungsbefreiung der Parteien. Wenn die in der Klausel definierten Formen von höherer Gewalt eintreten, entfallen die vertraglichen Verpflichtungen der Parteien, wie Abnahme- bzw. Lieferverpflichtungen oder mögliche Schadensersatzverpflichtungen für die Nichterfüllung. Folglich könnte der Bilanzierende den Vertrag kostenfrei auflösen, wodurch die Pflicht zur Passivierung einer Drohverlustrückstellung entfiele (vgl. Burkhardt-Böck, 2020).

In diesem Zusammenhang stellt sich zunächst die Frage, inwiefern der Ausbruch der Corona-Pandemie mit der zugrunde gelegten Force Majeure-Klausel in Einklang zu bringen ist. Einige Klauseln zählen konkrete Ereignisse abschließend auf, bei denen höhere Gewalt anzunehmen ist, wie z.B. Naturkatastrophen, Krieg, Epidemien und Pandemien. Sofern in der Klausel nicht ausdrücklich definierte Ereignisse vorliegen, muss näher untersucht werden, ob höhere Gewalt entsprechend angenommen werden kann.

Grundsätzlich ist von höherer Gewalt auszugehen, wenn die Unmöglichkeit der Leistungserfüllung von keiner Partei verschuldet wurde und auf ein von außen einwirkendes Ereignis zurückzuführen ist, welches zum Zeitpunkt des Vertragsschlusses unvorhersehbar war und selbst durch äußerste Sorgfalt nicht abgewendet werden konnte (vgl. Weaver, 2020; Schellenberger & Baggen, 2020).

Covid-19 und die von der Politik erlassenen Maßnahmen zur Eindämmung der Infektionszahlen hatten weitreichende wirtschaftliche Folgen und waren zudem für die Parteien unvorhersehbar und nicht zu vertreten. Zusätzlich wurde der Ausbruch von Covid-19 am 11. März 2020 seitens der WHO offiziell zur Pandemie eingestuft (vgl. Mitteilung WHO). Folglich besteht unter Umständen die Möglichkeit, Vertragsausfälle aufgrund der Corona-Pandemie als höhere Gewalt einzustufen (vgl. Weaver, 2020).

Eine allgemeine Identifizierung als höhere Gewalt ist jedoch nicht möglich, sodass stets eine einzelfallbezogene Überprüfung notwendig wird.

Bei **Mietverhältnissen** handelt es sich um Dauerschuldverhältnisse innerhalb der Beschaffungsgeschäfte über nicht bilanzierungsfähige Leistungen. Wenn der Beitrag der Gegenleistung zum Unternehmenserfolg geringer ist als der Wert der vom Bilanzierenden zu erbringenden Leistung, kann die Ausgeglichenheitsvermutung nicht aufrechterhalten werden und es ist handelsbilanziell eine Drohverlustrückstellung zu bilden (vgl. Schubert, 2020, Rn. 76).

Der Beitrag der Gegenleistung zum Unternehmenserfolg kann jedoch regelmäßig nicht hinreichend objektiv ermittelt werden (z.B. wegen fehlender Ertragszurechenbarkeit), sodass die Passivierung einer Drohverlustrückstellung nach §249 Abs. 1 S. 1 HGB nur bei vollständig fehlender oder nicht nennenswerter Nutzungs- oder Verwertungsmöglichkeit der beschafften Leistung in Betracht kommt (vgl. IDW RS HFA 4, Tz. 32). Wird eine Filiale dagegen nicht geschlossen, scheidet – bei der gebotenen absatzmarktorientierten Bewertung – die Passivierung einer Drohverlustrückstellung grundsätzlich aus, da ein Verlust nicht zwingend durch den Mietvertrag verursacht sein muss. Etwas anderes kann sich ergeben, wenn in einem Mietvertrag eine Betriebspflicht enthalten ist (z. B. bei sog. Shop in Shop-Geschäften). Diese führt zu einer einheitlichen Betrachtung von Beschaffungs- und Absatzmarkt, sodass sämtliche aus dem Filialbetrieb resultierende Erträge und Aufwendungen einander gegenüberzustellen sind (vgl. Oser & Wirtz, 2021).

Die IFRS-Bilanzierung verfolgt indes einen anderen Weg. Demnach erfolgt nach IFRS 16 mit Vertragsbeginn i. d. R. die Bilanzierung eines Nutzungsrechts sowie einer Leasingverbindlichkeit seitens des Mieters. Wenn die Zahlungsverpflichtung aus dem Mietverhältnis auch ohne Nutzungsmöglichkeit der Räumlichkeiten bestehen bleibt, erfolgt eine außerplanmäßige Abschreibung des Nutzungsrechts nach IFRS 16.33 i. V. m. IAS 36 auf seinen erzielbaren Betrag (vgl. Burkhardt-Böck, 2020).

Vorübergehende Filialschließung als Folge der Lockdown-Maßnahmen

Als Folge der Corona-Pandemie kam es bundesweit zu Lockdowns und damit insbesondere zur Schließung des stationären Einzelhandels. Den fortlaufend zu leistenden Mietzahlungen steht nunmehr kein (gegenwärtiger) wirtschaftlicher Nutzen aus der Nutzungsüberlassung gegenüber. Für den Zeitraum, in dem dieses Ungleichgewicht vorliegt, stellt sich die Frage nach einer Drohverlustrückstellung. Wenn dieser Zeitraum der Nicht-Nutzung in Anbetracht der zukünftigen Nutzung und der zukünftigen Vorteile als nicht wesentlich erachtet wird, also nur eine temporäre Unausgewogenheit vorliegt, darf keine Drohverlustrückstellung gebildet werden. Für die Praxis bedeutet dies, dass es in handelsrechtlichen Jahresabschlüssen nur im Falle von kurzen Restlaufzeiten zu passivierungspflichtigen Drohverlustrückstellungen kommen kann. Bei der Ermittlung der Rückstellung ist zu beachten, dass durch den Lockdown verursachte und vom Vermieter anerkannte Mietkürzungen rückstellungsmindernd zu berücksichtigen sind (vgl. Atilgan, 2021). Eine mögliche Stundung der Mietzahlungen hat indes keinen Einfluss.

Im Umkehrschluss ist die handelsrechtliche Bildung einer Drohverlustrückstellung untersagt, wenn ein möglicher Verlust nicht auf die Schließung im Rahmen des Lockdowns zurückzuführen ist (vgl. Oser & Wirtz, 2021).

Dauerhafte Schließungen von Filialen als Folge von Covid-19
In Kapitel 2 wurde aufgezeigt, dass es insbesondere im stationären Einzelhandel und bei den Banken coronabedingt zu umfangreichen Filialschließungen kommen wird. Hier finden sich häufig Mietverträge mit langen Restlaufzeiten, die keine vorzeitige Kündigung vorsehen. Es kommt in diesen Fällen regelmäßig zu einem verlustbringenden Vertrag, der handelsrechtlich zu einer Drohverlustrückstellung führen wird. Den Mietzahlungen stehen keine entsprechenden Erträge gegenüber. Untervermietungen sind häufig nicht gestattet bzw. können die Mietzahlungen nicht kompensieren (vgl. Atilgan, 2021).

In einer Schließungsfiliale sind demzufolge in der Regel alle Mietzahlungen ab dem Schließungszeitpunkt mit ihrem Barwert zurückzustellen. Gegebenenfalls anfallende Untermieterträge sind zu saldieren. Dem Filialmietvertrag zuzuordnende Vermögenswerte, die einer außerplanmäßigen Abschreibung aus dem drohenden Verlust unterzogen werden müssten, liegen in der Regel nicht vor, da Anlage- und Vorratsvermögen grundsätzlich weiterverwendet, veräußert oder verschrottet und somit gesondert bewertet werden (vgl. Atilgan, 2021).

Für IFRS gelten im Fall von operate lease-Verträgen obige Ausführungen analog. In der Regel wird jedoch ein Nutzungsrecht im Sinne des IFRS 16 aktiviert sein. Mögliche Auswirkungen auf die Werthaltigkeit wären zu prüfen (vgl. IFRS 16.33 i. V. m. IAS 36).

Arbeitsverhältnisse sind ebenfalls schwebende Dauerschuldverhältnisse im Rahmen der Beschaffungsgeschäfte über nicht bilanzierungsfähige Leistungen. Grundsätzlich liegt ihnen eine Ausgeglichenheitsvermutung von Leistung und Gegenleistung zugrunde (vgl. Lüdenbach, 2020). Die Maßnahmen zur Eindämmung der Corona-Pandemie in Deutschland wirken sich negativ auf die Produktion von Waren und das Angebot von Dienstleistungen aus. Viele Arbeitgeber müssen daher auf das Instrument der Kurzarbeit zurückgreifen. Betroffene Arbeitnehmer erhalten dadurch reduzierte Löhne und Gehälter sowie zum Ausgleich des Netto-Verdienstausfalls teilweise **Aufstockungsbeträge zum Kurzarbeitergeld**. Aufstockungsbeträge zum Kurzarbeitergeld sind Bestandteil der Leistungs- und Entgeltpflichten aus dem Arbeitsverhältnis und werden folglich als laufender Personalaufwand angesehen (vgl. IDW, 2021, S. 10; Oser & Wirtz, 2021).

Übergeordnetes Ziel dieser Zahlungen ist die Fortführung des Beschäftigungsverhältnisses, um einen aufwendigen Neuaufbau der notwendigen Arbeitskraft nach Beendigung der Krise zu vermeiden und die vorhandene Kapazität aufrechtzuerhalten. Gleichzeitig entsteht dem Unternehmen ein künftiger Nutzen durch ein erleichtertes Hochfahren der Produktions- und Leistungsfähigkeit (vgl. Lüdenbach, 2020).

Folglich sind im Rahmen der Ausgeglichenheitsvermutung sämtliche Leistungen und Gegenleistungen über die erwartete (Rest-)Laufzeit der Arbeitsverträge zu berücksichtigen.

Es handelt sich somit lediglich um eine temporäre Unausgeglichenheit zulasten des Arbeitgebers, welche die Ausgeglichenheitsvermutung innerhalb des Arbeitsverhältnisses noch nicht widerlegt. Demzufolge liegt kein Verpflichtungsüberschuss vor, sodass die Bildung einer Drohverlustrückstellung nach § 249 Abs. 1 S. 1 HGB für gezahlte Aufstockungsbeträge ausscheidet. Für die Bildung einer Rückstellung nach IFRS gilt entsprechendes (vgl. IDW, 2021, S. 11; Oser & Wirtz, 2021; Lüdenbach, 2020).

3.2. Restrukturierungsrückstellungen

3.2.1. Allgemeine Grundsätze nach HGB

Notwendig für die Passivierung von Rückstellungen für Restrukturierungsmaßnahmen (im Folgenden: „Restrukturierungsrückstellungen“) ist das kumulative vorliegen sämtlicher Ansatzvoraussetzungen für Rückstellungen für ungewisse Verbindlichkeiten. Demnach muss eine Außenverpflichtung vorliegen, die am Bilanzstichtag rechtlich entstanden bzw. wirtschaftlich verursacht ist und die notwendigen Aufwendungen zur Verpflichtungserfüllung mit überwiegender Wahrscheinlichkeit zu entsprechenden Auszahlungen führen. Im Umkehrschluss darf es sich bei der aus der Restrukturierung resultierenden Verpflichtung um keine selbst auferlegte Innenverpflichtung handeln (vgl. Bonnecke, 2021; Oser & Wirtz, 2021).

Maßnahmen, welche mit einem Personalabbau verbunden sind, beinhalten unstreitig eine rückstellungspflichtige Außenverpflichtung und führen regelmäßig zu Abfindungszahlungen an bisherige Beschäftigte bzw. verpflichten das Unternehmen, dazu die aus der Restrukturierung ergehenden wirtschaftlichen Nachteile für die Arbeitnehmer auszugleichen.

Aufwendungen aus der vorzeitigen Auflösung von Miet-, oder Leasingverträgen sowie sonstigen einzelvertraglichen (Liefer- oder Abnahme-)Verpflichtungen begründen ebenfalls eine Außenverpflichtung. Im Rahmen einer Restrukturierung fallen einem Unternehmen auch regelmäßig Rechts- und Beratungskosten an, welche eine Verpflichtung gegenüber Dritten darstellen und somit restrukturierungsbedingten Aufwendungen zuzuordnen sind (vgl. Bonnecke, 2021).

Von den rückstellungsfähigen Aufwendungen abzugrenzen sind beispielsweise „stay boni“, dabei handelt es sich um Prämien für Mitarbeiter, um deren Arbeitsleistung bis zur geplanten Einstellung der Produktion und ggf. darüber hinaus zu garantieren. Diese Aufwendungen sind zukunftsbezogen und somit zum Bilanzstichtag wirtschaftlich noch nicht verursacht, womit eine Rückstellungsbilanzierung ausscheidet (vgl. Böckem & von Heynitz, 2010).

Aufwendungen im Zusammenhang mit der künftigen Geschäftstätigkeit, wie z.B. Kosten für die Verlegung von Betriebsteilen durch die Verlagerung von Produktionsmaschinen oder Umschulungskosten für Mitarbeiter, die künftig in anderen Geschäftsbereichen eingesetzt werden sollen, stellen interne Verpflichtungen dar und sind dementsprechend nicht rückstellungsfähig (vgl. Bonnecke, 2021).

Die rechtliche Entstehung bzw. wirtschaftliche Verursachung bis zum Abschlussstichtag wird im Rahmen der Restrukturierungsrückstellung als erfüllt angesehen, wenn

die entsprechende Beschlussfassung der zuständigen Organe hinreichend konkret ist. „Konkret“ bedeutet in diesem Zusammenhang, dass die tatsächliche Durchführung der Restrukturierung ernsthaft erscheint oder die Unterrichtung des Betriebsrats kurz bevorsteht. Die Kommunikation mit dem Betriebsrat/den Arbeitnehmern bis zum Bilanzstichtag ist somit keine notwendige Bedingung, denn ein Nachweis bis zum Ende des Wertaufhellungszeitraums reicht handelsrechtlich aus (vgl. Bonnecke, 2021; Nagengast & Boecker, 2016).

Die Bewertung von Rückstellungen für Restrukturierungsmaßnahmen erfolgt gem. § 253 Abs. 1 S. 2 HGB mit „dem nach vernünftiger kaufmännischer Beurteilung notwendigen Erfüllungsbetrag“.

Ein Restrukturierungsplan besteht regelmäßig zu einem wesentlichen Anteil aus Maßnahmen im Personalbereich, die dazugehörigen Aufwendungen werden in einem Sozialplan zusammengetragen. In die Bewertung fließen daher jene Kosten ein, deren Ursache in den Vereinbarungen des Sozialplans liegen, wie z.B. Abfindungen an Arbeitnehmer bei Arbeitsplatzverlust. Weitere typische Beispiele sind u.a. Ausgleichszahlungen bei Verdienstminderungen, Rentenausgleichszahlungen sowie Umzugs- oder Fahrtkosten (vgl. Bonnecke, 2021).

Grundsätzlich müssen im Rahmen der Bewertung von längerfristigen Rückstellungen auch künftige Preissteigerungen und eine mögliche Abzinsung berücksichtigt werden. Restrukturierungsmaßnahmen erfolgen allerdings i. d. R. kurzfristig, sodass hier auf eine weitere Erläuterung verzichtet wird (vgl. Bonnecke, 2021; Nagengast & Boecker, 2016).

3.2.2. Allgemeine Grundsätze nach IFRS

Im Rahmen der IFRS werden Restrukturierungsmaßnahmen als Programm beschrieben, das vom Management geplant und kontrolliert wird, und wesentliche Veränderungen in einem vom Unternehmen abgedeckten Geschäftsfeld oder der Art und Weise, in welcher dieses Geschäft durchgeführt wird, herbeiführt (vgl. IAS 37.10). Typische Restrukturierungsmaßnahmen werden, anders als im HGB, konkretisiert und in IAS 37.70 beispielhaft genannt: der Verkauf oder die Aufgabe eines Geschäftszweigs, die Schließung eines Standorts oder die Verlagerung der Geschäftsaktivitäten in andere Regionen bzw. Länder, eine Veränderung der Managementstrukturen sowie eine fundamentale Neuorganisation mit wesentlichen Auswirkungen auf den Schwerpunkt der Geschäftstätigkeit des Unternehmens.

Die Bildung einer Rückstellung für derartige Restrukturierungsmaßnahmen setzt eine kumulative Erfüllung der allgemeinen Ansatzkriterien für Rückstellungen voraus (vgl. IAS 37.71; Zwirner & Boecker, 2021). Demnach muss gem. IAS 37.14 eine rechtliche oder faktische Verpflichtung vorliegen, die aus einem Ereignis der Vergangenheit entstanden ist und zur Erfüllung dieser Verpflichtung wahrscheinlich zu einem Abfluss von Ressourcen mit wirtschaftlichem Nutzen führt (vgl. IDW, 2020a, S. 21). Zusätzlich muss die Höhe der Verpflichtung verlässlich geschätzt werden können. Folglich ist das Vorliegen einer Außenverpflichtung eine gemeinsame Voraussetzung des HGB und der IFRS.

Eine rechtliche Außenverpflichtung entsteht durch Vertragsschluss (in Verkaufsfällen) (vgl. IAS 37.78). Im Rahmen von Restrukturierungen ist die faktische Verpflichtung von besonderer Bedeutung, da zum Bilanzstichtag regelmäßig noch keine rechtliche Verpflichtung vorliegt (vgl. Nagengast & Boecker, 2016). Folglich enthält IAS 37.72 konkrete Vorgaben für die Entstehung einer rückstellungsfähigen faktischen Verpflichtung:

Diese entsteht nur, wenn zum Bilanzstichtag ein detaillierter, formaler Restrukturierungsplan vorliegt, welcher Informationen zu dem betroffenen Geschäftsbereich, den wichtigsten Standorten und der Anzahl der abzufindenden Arbeitnehmer enthält. Außerdem sind Angaben zu den entstehenden Ausgaben und dem geplanten Umsetzungszeitpunkt zu machen. Die Gesamtheit dieser Informationen muss kommuniziert oder umgesetzt werden, um bei den Betroffenen eine gerechtfertigte Erwartung zu wecken, dass der Restrukturierungsplan auch tatsächlich durchgeführt wird (vgl. IDW, 2020a, S. 21; Zwirner & Boecker, 2021). Die Verpflichtungen für abzufindende Arbeitnehmer sind in Übereinstimmung mit den Anforderungen des IAS 19 anzusetzen (vgl. Kirsch, 2021; Böckem, 2020).

Im Rahmen der IFRS gibt es somit hohe formale Anforderungen an eine faktische Verpflichtung, sodass handelsrechtlich ein größerer Spielraum für den Ansatz von Restrukturierungsrückstellungen vorliegt. Die IFRS-Rechnungslegung verlangt zwingend einen Informationsaustausch mit den Betroffenen bis zum Bilanzstichtag, was handelsrechtlich nicht gefordert wird.

Wenn die in IAS 37.14 genannten Ansatzvoraussetzungen nicht kumulativ vorliegen, ist lediglich eine Eventualschuld im Anhang zu erläutern. Interne Beschlüsse reichen somit für die Bildung einer Restrukturierungsrückstellung nicht aus.

Nach IAS 37.36 erfolgt die Bewertung von Restrukturierungsrückstellungen in Höhe der bestmöglichen Schätzung der Ausgaben, die zur Begleichung der gegenwärtigen Verpflichtung zum Abschlussstichtag notwendig sind. Diese enthält laut IAS 37.80 ausschließlich Kosten, die direkt im Zusammenhang mit der Restrukturierungsmaßnahme stehen. Dazu gehören z. B. direkt zurechenbare Beratungskosten sowie die antizipierten Kosten vom Zeitpunkt der Einstellung des Geschäftsbetriebes bis zum endgültigen Abgang (vgl. Coenenberg et al., 2018, S. 473).

Im Umkehrschluss dürfen diese Ausgaben nicht im Zusammenhang mit den laufenden Aktivitäten des Unternehmens stehen (vgl. Pütz, 2020). Gleiches gilt für Kosten, die der künftigen Geschäftstätigkeit zuzuordnen sind, wie z.B. Umschulung oder Marketinginvestitionen in Bezug auf neue Geschäftsfelder (vgl. IAS 37.81; Pütz, 2020; Lüdenbach, 2019, S. 207).

Der bestmögliche Schätzwert entspricht bei unterschiedlichen Eintrittswahrscheinlichkeiten verschiedener Werte dem Erwartungswert und liegt damit i. d. R. unter dem zugrunde gelegten Wert in der Handelsbilanz, in welcher das Vorsichtsprinzip berücksichtigt wird (vgl. Nagengast &

Boecker, 2016).

3.2.3. Auswirkungen von Covid-19 auf die Restrukturierungsrückstellungen

Die Wirtschaft ist durch eine steigende Dynamik, zunehmende Digitalisierung sowie immer kürzere Konjunkturzyklen geprägt. Infolge dessen entwickelt sich das Marktumfeld der Unternehmen rasch weiter und fördert dadurch die Wettbewerbsintensität, sodass die Unternehmen dazu gezwungen sind, ihre Organisations- und Ablaufstrukturen regelmäßig zu überprüfen, um nachhaltig gewinnbringend agieren zu können und konkurrenzfähig zu bleiben (vgl. Bolik & Schuhmann, 2016). Die Corona-Pandemie verstärkt solche Entwicklungen signifikant und drängt viele Geschäftszweige oder sogar gesamte Betriebe in dauerhafte Verlustsituationen (vgl. Zugehör, 2021; Zwirner & Boecker, 2021; Rinker, 2020). Es droht insbesondere die Nichteinhaltung von Financial Covenants. Die betroffenen Unternehmen werden – auch auf Druck der finanzierenden Banken – versuchen, dieser Entwicklung durch geeignete Maßnahmen, insbesondere Restrukturierungen, entgegenzutreten.

Schaut man in die Geschäftsberichte der DAX-30 Unternehmen, wird deutlich, dass insbesondere die Banken und die Luftfahrtindustrie coronabedingt umfangreiche Restrukturierungsprogramme aufgesetzt haben. So hat die Commerzbank bereits im August 2020 angekündigt, von ihrem insgesamt 1000 Filialen jene 200 nicht mehr zu öffnen, die im Frühjahr wegen der Pandemie geschlossen worden waren; sie hätten eigentlich bis 2023 weiter bestehen sollen. In Deutschland streicht die Commerzbank somit 10.000 Stellen, dies entspricht jedem dritten Arbeitsplatz. Die Deutsche Bank plant, bis 2021 die Zahl ihrer Dependancen von 500 auf 400 zu senken und bei ihrer Tochter Postbank 50 Filialen zu schließen (vgl. Freiberger, 2020).

Der Lufthansa-Konzern geht aktuell davon aus, dass die Nachfrage nach Flugreisen frühestens im Jahr 2024 wieder das Niveau von vor der Krise erreicht. Die Lufthansa Group hat daher ein umfassendes Restrukturierungsprogramm beschlossen, welches auch die bereits laufenden Restrukturierungsprogramme der Airlines und Servicegesellschaften umfasst. Das Programm sieht den Abbau von 22.000 Vollzeitstellen in der Lufthansa Group vor (vgl. Cockpit, 2020).

Die angeführten Praxisbeispiele machen deutlich, dass es als Folge der coronabedingten Restrukturierungs- und Sanierungsmaßnahmen zu einem erheblichen Anstieg der Restrukturierungsrückstellungen kommen wird.

3.3. Garantie- und Kulanzrückstellungen

3.3.1. Allgemeine Grundsätze nach HGB

Bei Garantierückstellungen handelt es sich um Rückstellungen für ungewisse Verbindlichkeiten gem. § 249 Abs. 1 S. 1 HGB. Somit ergibt sich eine Passivierungspflicht, wenn eine Außenverpflichtung besteht oder künftig entsteht, die rechtlich oder wirtschaftlich im vorangegangenen Geschäftsjahr verursacht wurde und mit einer tatsächlichen Inanspruchnahme des Unternehmens ernsthaft gerechnet werden muss.

Konkrete Gewährleistungsansprüche, welche vom Kunden geltend gemacht wurden, sind in diesem Zusammenhang nicht erforderlich (vgl. Schubert, 2020, Rn. 24).

Die Verpflichtung im Rahmen der Garantierückstellungen kann sich aus gesetzlichen Gewährleistungsverpflichtungen (z.B. Kaufvertrag) oder aus Eigenschaftszusicherungen und eigenständigen Gewährleistungszusagen (z.B. Herstellergarantie) ergeben (vgl. Winnefeld, 2015, Rn. 1205).

Davon zu unterscheiden sind Rückstellungen für übernommene Gewährleistungen ohne rechtliche Verpflichtung (Kulanzrückstellungen), welche gem. § 249 Abs. 1 S. 2 Nr. 2 HGB ebenfalls zu passivieren sind. Notwendige Voraussetzung ist jedoch, dass sich der Unternehmer den Verpflichtungen aus geschäftlichen/wirtschaftlichen Gründen nicht entziehen kann, beispielsweise aus Imagegründen oder zur Pflege des Kundenstamms. Zusätzlich muss der Gewährleistungsaufwand in direktem Zusammenhang mit einem früheren Rechtsgeschäft (z.B. Lieferung oder Leistung) stehen.

Werden bis zum Bilanzstichtag bzw. zum Zeitpunkt der Bilanzerstellung konkrete Garantie- und Kulanzansprüche geltend gemacht, erfolgt ihre Berücksichtigung in Form von Einzelrückstellungen.

Zusätzlich erfolgt die Berücksichtigung von künftigen, noch nicht bis zu diesem Zeitpunkt bekannt gewordenen, Ansprüchen als Pauschalrückstellung. Erforderlich für die Bildung einer Pauschalrückstellung ist, dass auf Grundlage von Vergangenheitserfahrungen des einzelnen Betriebs sowie Branchenerfahrungen, ernsthaft mit der Inanspruchnahme von Garantie- und Kulanzleistungen gerechnet werden muss (vgl. Winnefeld, 2015, Rn. 1207; Krudewig, 2018).

In der Praxis erfolgt laut Winnefeld (2015, Rn. 1207) regelmäßig die Anwendung eines „gemischten Verfahrens“, welches im Rahmen der Rückstellungsbildung eine Einzel- und Pauschalbewertung erlaubt.

Die Höhe der Rückstellungen richtet sich nach den voraussichtlich anfallenden Kosten des Unternehmens, um die Garantie- bzw. Kulanzleistungen zu erfüllen. Bei einer Laufzeit der Rückstellung von mehr als einem Jahr erfolgt eine Abzinsung mit dem durchschnittlichen Marktzinssatz der vergangenen 7 Geschäftsjahre gem. § 253 Abs. 2 S. 1 HGB. Mögliche Gewinnzuschläge dürfen nicht berücksichtigt werden (vgl. Schwirkslies-Filler, 2019, S. 115).

3.3.2. Allgemeine Grundsätze nach IFRS

Die Erfassung von Garantieverpflichtungen erfolgt im Rahmen der IFRS als Rückstellung, wenn die Voraussetzungen nach IAS 37.14 kumulativ vorliegen. Demnach muss eine gegenwärtige rechtliche Verpflichtung bestehen, die aus einem vergangenen Ereignis (z.B. Kaufvertrag) entstanden ist und wahrscheinlich zu einem Abfluss von Ressourcen mit wirtschaftlichem Nutzen führt. Zusätzlich muss die Höhe dieser Verpflichtung verlässlich geschätzt werden können.

Im Rahmen der Garantieverpflichtungen bestehen regelmäßig eine Vielzahl ähnlicher Verpflichtungen, sodass die Wahrscheinlichkeit eines Mittelabflusses nach IAS 37.24 durch eine Betrachtung der ähnlichen Verpflichtungen als Ganzes ermittelt wird. Folglich ist die Tatsache, dass die

Wahrscheinlichkeit eines Abflusses im Einzelfall unwahrscheinlich ist, unschädlich (vgl. Coenenberg et al., 2018, S. 471).

Gewährleistungen ohne rechtliche Verpflichtungen (Kulanzleistungen) entsprechen einer faktischen Verpflichtung, d.h. das Unternehmen hat durch sein bisher übliches Vorgehen öffentlich angekündigte Maßnahmen oder hinreichend spezifischen Aussagen eine gerechtfertigte Erwartung auf Erfüllung geweckt (vgl. IAS 37.10). Vorausgesetzt die weiteren Anforderungen des IAS 37.14 sind kumulativ erfüllt, besteht die Notwendigkeit zur Bildung einer Kulanzrückstellung.

Die Bewertung der Rückstellung erfolgt in Höhe der „bestmöglichen Schätzung", die zur Erfüllung der Garantie- oder Kulanzverpflichtung benötigt wird (vgl. IAS 37.36). Dies gelingt durch Schätzungen des Managements auf Grundlage der Erfahrungswerte aus vergleichbaren Transaktionen, mathematischer Wahrscheinlichkeitstheorien und Einschätzungen von Experten (vgl. IAS 37.38).

3.3.3. Auswirkungen von Covid-19 auf die Garantie- und Kulanzrückstellungen

Die Auswirkungen der Corona-Pandemie sind weitreichend und vielseitig. Viele Unternehmen sind wirtschaftlich angeschlagen und die allgemeine wirtschaftliche Unsicherheit ist stark erhöht. Diese Entwicklungen können unter Umständen Einfluss auf die Garantie- und Kulanzrückstellungen eines Unternehmens nehmen, welche regelmäßig auf Grundlage von Erfahrungen aus der Vergangenheit gebildet werden. Da Covid-19 eine beispiellose Herausforderung darstellt, werden die bisherigen Erfahrungswerte der aktuellen Situation nicht gerecht. Somit müssen die Rückstellungen für Verkäufe mit am Bilanzstichtag noch nicht abgelaufenen Garantiezusagen und Kulanzregelungen kritisch auf ihre Höhe hinterfragt werden (vgl. Müller, 2021).

Möglicherweise kommt es zu vermehrten Reklamationen oder Stornierungen, was zusätzlich zu Rechtsstreitigkeiten führen kann. Daher ist in diesem Zusammenhang auch die Bildung von Rückstellungen für konkrete Prozesse zu prüfen (vgl. Müller, 2021).

Die Corona-Pandemie muss aber nicht zwangsweise negative Auswirkungen entfalten, so verzeichnete BMW im vierten Quartal 2020 in Folge der Lockdown-Maßnahmen geringere Abflüsse aus Gewährleistungsrückstellungen (vgl. BMW Ad-hoc-Meldung).

Die Automobilindustrie ist hinsichtlich Garantie- und Kulanzleistungen als besonderes Beispiel anzuführen. Viele Hersteller zeigen ihren Kunden in Anbetracht der Corona-Pandemie eine besondere Wertschätzung und erkennen damit die einzigartigen Umstände an: Porsche verlängert für seine Kunden weltweit und bei allen Fahrzeugen, unter bestimmten Voraussetzungen, die Garantie um weitere drei Monate (vgl. Porsche-Mitteilung). Auch VW und Audi kündigten an, die Neuwagen- und Anschlussgarantien ihrer Kunden kurzfristig verlängern zu wollen (vgl. Mitteilung von VW und Audi).

3.4. Pensionsrückstellungen

3.4.1. Allgemeine Grundsätze nach HGB

Pensionsrückstellungen erfassen die Aufwendungen im Rahmen einer unmittelbaren Versorgungszusage seitens des Arbeitgebers. Aus dieser unmittelbaren Zusage resultieren Verbindlichkeiten, die bezüglich ihrer Fälligkeit und Höhe ungewiss sind, sodass grundsätzlich eine Passivierungspflicht gem. § 249 Abs. 1 S. 1 HGB besteht (vgl. Schwirkslies-Filler, 2019, S. 169).

Die Bewertung von Pensionsrückstellungen erfolgt gem. § 253 Abs. 1 S. 2 HGB „in Höhe des nach vernünftiger kaufmännischer Beurteilung notwendigen Erfüllungsbetrags". Dies impliziert eine Berücksichtigung von zukünftigen Preis- und Kostensteigerungen, vorausgesetzt es liegen ausreichend objektive Hinweise für diese Annahme vor. Im Rahmen der Pensionsrückstellungen entspricht dies insbesondere zukünftigen Gehalts- und Rentenentwicklungen. Zusätzlich werden Daten über Fluktuation von Arbeitnehmern sowie Sterbe- und Invaliditätswahrscheinlichkeiten in die Bewertung integriert (vgl. Grottel & Haußer, 2020, Rn. 195).

Nach § 253 Abs. 2 S. 1 HGB sind Rückstellungen mit einer Restlaufzeit von mehr als einem Jahr abzuzinsen. Für Pensionsrückstellungen besteht dabei ein Wahlrecht. Demnach kann eine Abzinsung der Pensionszusage mit dem entsprechenden laufzeitabhängigen durchschnittlichen Marktzinssatz der vergangen 10 Geschäftsjahre erfolgen oder sie erfolgt gem. § 253 Abs. 2 S. 2 HGB pauschal mit dem durchschnittlichen Marktzinssatz, der sich bei einer angenommenen Restlaufzeit von 15 Jahren ergibt (vgl. Grottel & Haußer, 2020, Rn. 196).

Zusätzlich wird gem. § 253 Abs. 6 S. 1 HGB der Unterschiedsbetrag zwischen einer auf einer 10- und einer 7-jährigen Abzinsungsbasis beruhenden Pensionsrückstellung berechnet. Dieser Unterschiedsbetrag stellt gem. § 253 Abs. 6 S. 2 und 3 HGB eine bedingte Ausschüttungssperre für Gewinne dar und muss in jedem Geschäftsjahr im Anhang oder unter der Bilanz ausgewiesen werden (vgl. Kirsch, 2020; Tanski, 2019, S. 238).

Weiterhin fordert § 246 Abs. 2 S. 2 HGB die Verrechnung der Vermögensgegenstände, die ausschließlich für die Erfüllung von Pensionsverpflichtungen bestimmt sind und somit dem Zugriff sämtlicher übrigen Gläubiger entzogen sind, mit der entsprechenden Verpflichtung (vgl. Schwirkslies-Filler, 2019, S. 169).

3.4.2. Allgemeine Grundsätze nach IFRS

Die Bilanzierung von Pensionsrückstellungen im Rahmen der IFRS erfolgt nach IAS 19. Dort wird zwischen beitragsorientierten Pensionszusagen und leistungsorientierten Pensionszusagen differenziert (vgl. IAS 19.8).

Bei leistungsorientierten Zusagen, also Verpflichtungen gegenüber dem Arbeitnehmer zur Zahlung künftiger Pensionsleistungen, bei denen der Arbeitgeber sowohl das versicherungsmathematische Risiko als auch das Kapitalanlagerisiko trägt, sind die Voraussetzungen nach IAS 19.1 erfüllt und eine langfristige Rückstellung ist zu bilanzieren (vgl. Schrimpf-Dörges, 2020, Rn. 161).

Die Bewertung von leistungsorientierten Plänen erfolgt mit dem Barwert und erfordert gem. IAS 19.55 die Berücksichtigung von versicherungsmathematischen Bewertungsmethoden und demografischen Annahmen, beispielsweise über Lebenserwartung und Invalidität sowie Arbeitnehmerfluktuationen (vgl. IAS 19.66; Kirsch, 2020). Außerdem werden, wie im Rahmen des HGB, finanzielle Annahmen über künftige Gehalts- und Kostentrends berücksichtigt (vgl. IAS 19.57a (i)). Anders als im HGB erfolgt die Abzinsung nach IAS19.78 mit dem stichtagsaktuellen Zinssatz. Der ermittelte Barwert muss anschließend, analog zu der Vorgehensweise des HGB, mit dem beizulegenden Zeitwert des Planvermögens saldiert werden (vgl. IAS 19.57a (iii)). Das Planvermögen umfasst Vermögen, welches die Erfüllung von Versorgungsansprüchen der Arbeitnehmer sicherstellt und daher im Insolvenzfall unberührt bleibt (vgl. IAS 19.8).

3.4.3. Auswirkungen von Covid-19 auf die Pensionsrückstellungen

Der Ausbruch der Corona-Pandemie im Frühjahr 2020 bewirkte einen deutlichen Anstieg der Renditesätze. Folglich wirkt sich diese Entwicklung insbesondere auf die Bewertung der Pensionsrückstellungen nach IFRS aus, da dort der stichtagsaktuelle Zinssatz zur Diskontierung verwendet wird. Die versicherungsmathematischen Effekte einer Zinserhöhung resultieren in einer Entlastung der Bilanzen, da sich der Wert der Verpflichtungen verringert. Gleichzeitig entstehen durch die Erhöhung des Zinssatzes Gewinne, welche im sonstigen Gesamtergebnis zu erfassen sind (vgl. Stamm & Owerdieck, 2021; Kirsch, 2020). Abbildung 2 der Aon Holding Deutschland GmbH zeigt die exemplarische Veränderung des Verpflichtungsumfangs leistungsorientierter Pensionszusagen durch den Anstieg der Renditesätze nach IFRS für eine gemischte Bestandsstruktur:

Abbildung 2 verdeutlicht die Effekte auf ein nach IFRS bilanzierendes Unternehmen durch einen angestiegenen Rechnungszins, welcher in einer deutlichen Verringerung des Verpflichtungsumfangs für Abschlussstichtage zum 31.03.2020 bzw. zu diesem Stichtag erstellte Zwischenabschlüsse resultiert. Da die Zinssätze im Verlaufe des Jahres wieder gesunken sind, ergeben sich für spätere Abschlussstichtage lediglich marginale Effekte (vgl. Stamm & Owerdieck, 2021). Für Unternehmen mit Bilanzstichtag 31.12. ergibt sich für den 2020er Jahresabschluss und deren Pensionsrückstellungsbewertung Folgendes: Die anhand von Musterbeständen ermittelten Zinssätze für die Bewertung von Pensionsverpflichtungen nach IAS 19.83 betragen zum 31.12.2020 0,75 % (31.12.2019: 1,05 %) (vgl. Heubeck Zins-Info, 2021).

Durch einen Anstieg der Renditesätze im März 2020 können sich auch in HGB-Abschlüssen zum Abschlussstichtag 31.03.2020 Entlastungen in den Bilanzen ergeben. Diese werden aufgrund der Diskontierung mit durchschnittlichen Marktzinssätzen der vergangenen Geschäftsjahre jedoch geringer ausfallen als bei Abschlüssen nach IFRS.

Für spätere Abschlussstichtage kehrt sich der im März konstatierte Anstieg auch hier um. Für handelsrechtliche Jahresabschlüsse zum 31.12.2020 gilt:

Der für die Ausschüttungsbemessung relevante 7-Jahresdurchschnitt ist von 1,97 % (Stand 12/2019) auf 1,60 % (Stand 12/2020) gesunken; der für die Bilanzierung relevante 10-Jahresdurchschnitt verzeichnete eine ähnliche Entwicklung von 2,71 % (Stand 12/2019) auf 2,30 % (Stand 12/2020) (vgl. Zinssätze Deutsche Bundesbank). Dabei wurde eine Restlaufzeit von 15 Jahren angenommen.

Für die übrigen relevanten Bewertungsparameter kann sowohl für IFRS wie für HGB Folgendes konstatiert werden: Während sich mit Blick auf den Impffortschritt die Sterblichkeitsrate nicht nachhaltig signifikant verändern sollte, könnten kurzfristig die bislang erwarteten Lohn-und Gehaltssteigerungen zu hoch erscheinen und eine Anpassung nach sich ziehen. Analog gilt dies für den Karriere- und Rententrend. Dies würde tendenziell zu geringeren Pensionsrückstellungen führen (vgl. Müller, 2021).

3.5. Weitere mögliche Rückstellungen, die von Covid-19 beeinflusst werden können

Rückbauverpflichtungen: Regelmäßig ist in Filialmietverträgen eine Rückbauverpflichtung vertraglich vereinbart, die grundsätzlich im Zeitpunkt des Vertragsabschlusses handelsrechtlich eine Ansammlungsrückstellung nach sich zieht. Eine Ansammlungsrückstellung ist eine Rückstellung i. S. d. § 249 Abs. 1 S. 1 HGB, die eine Außenverpflichtung voraussetzt. Außerdem muss der laufende Geschäftsbetrieb für die Entstehung der Rückstellung wirtschaftlich verantwortlich sein, damit die Rückstellung zeitanteilig in gleichen Raten angesammelt werden kann. Das hier typische Szenario stellt die mietvertragliche Verpflichtung zur Beseitigung von Mietereinbauten dar, denn die Verpflichtung zum Entfernen von Mietereinbauten ist wirtschaftlich durch den laufenden Geschäftsbetrieb verursacht worden (vgl. Atilgan, 2021). Nach den IFRS hat die Rückbauverpflichtung den Charakter von Anschaffungs- oder Herstellungskosten des Anlagevermögens (vgl. IAS 16.16c). Im Zeitpunkt der Kündigung des Mietvertrages sind die gesamten Kosten des Rückbaus als Barwert zu aktivieren. Korrespondierend erfolgt eine Passivierung einer Rückstellung für Rückbauverpflichtungen. Im Fall einer coronabedingten Schließung wären sowohl nach HGB als auch nach IFRS die der Rückstellung zugrunde liegenden Bewertungsparameter insbesondere die Restlaufzeiten anzupassen (vgl. Müller, 2021).

Patronatserklärungen: In der Praxis werden insbesondere in Konzernen regelmäßig harte Patronatserklärungen der Muttergesellschaft abgegeben. Harte Patronatserklärungen sind rechtsverbindliche Einstandspflichten, bei denen der „Patron“ (i. d. R. ein MU) dem Kreditgeber eines Dritten (i. d. R. seinem TU) zur Förderung oder Erhaltung der Kreditbereitschaft Maßnahmen oder Unterlassungen in Aussicht stellt oder zusagt. Es handelt sich somit um einseitig verpflichtende Rechtsgeschäfte, welche zunächst im Anhang oder unter der Bilanz vermerkt werden müssen (vgl. Grottel & Haußer, 2020, Rn. 41; Maier-Reimer & Etzbach, 2011).

Eine Rückstellung muss dann gebildet werden, wenn aufgrund der Auswirkungen des Coronavirus ernsthaft mit einer Inanspruchnahme aus der Haftung gerechnet werden muss

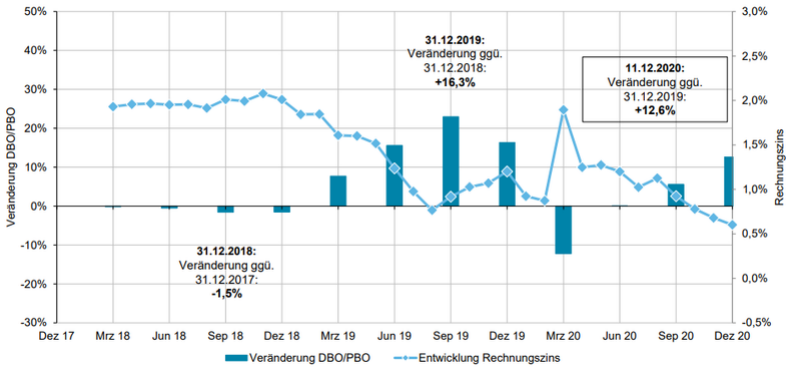

Abbildung 2: Veränderung des Verpflichtungsumfangs leistungsorientierter Pensionszusagen, entnommen von: Aon Holding Deutschland GmbH

(vgl. Oser & Wirtz, 2021; Rinker, 2020). Die obigen Ausführungen gelten unter Berücksichtigung des Grundsatzes „more likely than not“ gleichermaßen für IFRS (vgl. IAS 37.14).

Verlustübernahme: Die Corona-Pandemie hat weitreichende negative Auswirkungen und kann die Entwicklung eines Unternehmens nachhaltig beeinträchtigen und dadurch zu Verlusten führen (vgl. Rinker, 2020). Bei Beherrschungs- und/oder Gewinnabführungsverträgen, welche eine Verpflichtung zur Verlustübernahme begründen, erfolgt bei dem MU ggf. die handelsrechtliche Passivierung einer Rückstellung für ungewisse Verbindlichkeiten in Höhe des angefallenen Verlustes bei dem TU (vgl. Schubert, 2020, Rn. 100). Der angefallene Verlust muss unter Umständen geschätzt werden, beispielsweise wenn bei unterschiedlichen Abschlussstichtagen kein Zwischenergebnis des TU vorliegt. Dies stellt zu Zeiten einer Pandemie eine besondere Herausforderung dar.

Beherrschungs- und/oder Gewinnabführungsverträge sind gesellschaftsrechtliche Organisationsverträge, sodass die Voraussetzungen für die Bildung von Drohverlustrückstellungen nicht vorliegen (vgl. IDW RS HFA 4, Tz.4; Schubert, 2020, Rn. 100). Die Tatsache, dass zum Bilanzstichtag des MU nicht damit gerechnet werden kann, dass das TU innerhalb der voraussichtlichen Laufzeit der Verpflichtung jemals die Gewinnschwelle erreichen wird, entfaltet in diesem Zusammenhang keine Wirkung (vgl. Schubert, 2020, Rn. 100). Die oben genannten Ausführungen gelten analog für IFRS (vgl. IAS 37.14). Jedoch kommt obige Rückstellung nach IFRS in der Bilanzierungspraxis selten vor, da in der Regel nach IFRS Konzernabschlüsse erstellt werden und daher entsprechende Rückstellungen im Rahmen der Konsolidierung nicht anzusetzen sind.

Prozessrückstellungen: Es ist davon auszugehen, dass es in der Pandemie zu einem Anstieg der Rechtsstreitigkeiten kommen wird (vgl. Müller, 2021; Schütte & Götz, 2021). Hierbei ist zu beachten, dass Rückstellungen für Prozesskosten sowohl nach HGB als auch nach IFRS erst dann geboten sind, wenn ganz konkrete Prozesse angestrengt wurden. Im Zusammenhang mit Passivprozessen erfolgt die Berücksichtigung von den Prozesskosten, Bußgeldern und wahrscheinlichen Schadenersatzverpflichtungen, während bei Aktivprozessen lediglich das Kostenrisiko berücksichtigt wird (vgl. Schubert, 2020, Rn. 100). Es wird darauf hingewiesen, dass aus den Prozessen drohende Schadensersatzzahlungen beim Beklagten handelsrechtlich auf Basis des Vorsichtsprinzips zu Rückstellungen führen können, während nach IFRS die Passivierung nur bei einer Eintrittswahrscheinlichkeit von mehr als 50 % (more likely than not) geboten ist (vgl. IAS 37.14).

4. Fazit und Ausblick

In dieser Arbeit wurde der Einfluss von Covid-19 auf die Rückstellungen thematisiert und hinsichtlich der Auswirkungen auf Ansatz und Bewertung nach HGB sowie IFRS untersucht. Das Ziel war es insbesondere, das pandemische Umfeld der Unternehmen unter Einbeziehung der Branchen daraufhin zu untersuchen, welche Rückstellungen besonders stark von der Covid-19 Pandemie betroffen sein könnten und wie sich die Pandemie auf Ansatz und Bewertung der identifizierten Rückstellungen auswirkt.

Im Rahmen der vorliegenden Arbeit wurde deutlich, dass insbesondere für Jahresabschlüsse ab dem 31.03.2020 (sowohl nach HGB als auch IFRS) die Herausforderung besteht, Covid-19 als wertaufhellendes Ereignis durch Bildung und Anpassung von Rückstellungen angemessen zu erfassen. Es ist jedoch zu beachten, dass weiterhin eine einzelfallspezifische Betrachtung geboten ist.

Das Ergebnis dieser Arbeit verdeutlicht, dass sich im Zuge der Pandemie die Unternehmen, je nach Branche, verstärkt mit dem Ansatz von Drohverlustrückstellungen sowie mit Restrukturierungsrückstellungen auseinandersetzen müssen. Hier drohen den Unternehmen im Zeitpunkt einer Rückstellungspassivierung ggf. erhebliche Belastungen der Vermögens- und Ertragslage.

Weiterhin erfordert die Corona-Krise im Zusammenhang mit Rückstellungen eine kritische Überprüfung der zugrunde gelegten Annahmen und Einflussfaktoren:

Die Untersuchung im Rahmen der Garantierückstellungen hat gezeigt, dass die Krise sowohl zu einer Entlastung, als auch einer Belastung der Unternehmen führen kann.

Im Rahmen der vorliegenden Arbeit wurde im Hinblick auf den vorübergehenden starken Anstieg der Marktzinssätze zu Beginn der Covid-19 Pandemie außerdem der Frage nachgegangen, inwieweit Pensionsrückstellungen, die einen erheblichen Teil der in Deutschland bilanzierten Rückstellungen ausmachen, hiervon betroffen sind. Die Untersuchungen kommen zu dem Ergebnis, dass es insbesondere bei IFRS-Jahresabschlüssen mit Bilanzstichtag 31.3.2020 (oder entsprechenden Zwischenabschlüssen) aufgrund des Zinsanstiegs zu geringeren Pensionsrückstellungen kommen kann. Da die Zinssätze im Verlaufe des Jahres 2020 wieder gesunken sind, ergeben sich für spätere Abschlussstichtage lediglich marginale Effekte. Bei handelsrechtlichen Abschlüssen ist der Zinseffekt zum 31.3.2020 mit Hinweis auf eine Diskontierung mit dem durchschnittlichen Marktzinssatz der vergangenen Geschäftsjahre, deutlich geringer.

Abschließend verweist diese Arbeit auf weitere Rückstellungen, die unter Umständen und im Einzelfall von den Auswirkungen der Pandemie betroffen sein können.

In Anbetracht der obigen Ausführungen muss davon ausgegangen werden, dass die Auswirkungen der Corona-Krise die Bilanzen zahlreicher Unternehmen, insbesondere durch die weiterhin existierende wirtschaftliche Unsicherheit, nicht nur in 2020, sondern auch in den kommenden Jahren weiterhin belasten werden. Im Zusammenhang mit Rückstellungen bedarf es einer fortlaufenden unternehmensindividuellen Analyse, um die entsprechenden Auswirkungen auf den Ansatz und die Bewertung von Rückstellungen sachgerecht darzustellen und in Einklang mit den bestehenden Regelungen nach HGB und IFRS zu bringen.

Die obige Arbeit macht deutlich, dass die Pandemie als Krise je nach Branche schwer auf den Unternehmen lastet. Jedoch zeigt sich, dass Unternehmen hierauf insbesondere mit Anpassung der Geschäftstätigkeit, Organisations- und Ablaufstrukturen reagieren. Wenn die Unternehmen entsprechende Chancen nutzen, besteht die Möglichkeit gestärkt aus der Krise hervorzugehen.

Literatur

Atilgan, E. (2021). Drohverlust- und Ansammlungsrückstellungen bei vorzeitiger Filialschließung bedingt durch die Corona-Pandemie. *Der Betrieb*, *14*, 693–696.

Baetge, J., Kirsch, H.-J. & Thiele, S. (2017). *Die Bilanzierung der Rückstellungen*. IDW Verlag Düsseldorf.

Ballwieser, W. (2020). §249 HGB. In K. Schmidt (Hrsg.), *Münchener Kommentar zum HGB.* C.H.BECK.

Bardt, H. & Hürther, M. (2020). Corona stoppt die Volkswirtschaft von allen Seiten. *IW-Kurzbericht*, *31/2020*.

Böckem, H. (2020). Wie sind Veränderungen bei den Arbeitnehmervergütungen zu berücksichtigen? *FAQs zu Auswirkungen von Covid-19 auf die Rechnungslegung (KPGM)*. Zugriff am 2021-05-16 auf https://home.kpmg/de/de/home/themen/2020/04/auswirkungen-von-covid-19-auf-die-rechnungslegung/wie-sind-veraenderungen-bei-den-arbeitnehmerverguetungen-zu-beruecksichtigen.html

Böckem, H. & von Heynitz, B. (2010). Aktuelle Praxisfälle zu Restrukturierungsrückstellungen in der Handelsbilanz. *BC 2010- Zeitschrift für Bilanzierung, Rechnungswesen und Controlling*, 398–401.

Bertram, K. (2019). §249 HGB. In K. Bertram, R. Brinkmann, H. Kessler & S. Müller (Hrsg.), *Haufe hgb-kommentar* (10th Aufl.).

Bolik, A. S. & Schuhmann, S. (2016). Ansatz und Bewertung von Umstrukturierungs- und Sozialplanrückstellungen. *StuB*, *18/2016*, 684–686.

Bonnecke, T. (2021). Rückstellungen für Restrukturierungsmaßnahmen. *StuB*, *2021* (5), 185–190.

Burkhardt-Böck, H. (2020). Update IAS 37 – Belastende Verträge. *PiR*, *2020* (10), 315–317.

Cockpit. (2020). Lufthansa Group kündigt massive Restrukturierungen an. *Cockpit, Das Schweizer Luft- und Raumfahrt-Magazin*. Zugriff am 2021-05-16 auf https://www.cockpit.aero/rubriken/detailseite/news/lufthansa-group-kuendigt-massive-restrukturierungen-an/?no_cache=1

Coenenberg, A. G., Haller, A. & Schultze, W. (2018). *Jahresabschluss und Jahresabschlussanalyse* (25th Aufl.). Schäffer-Poeschel Verlag Stuttgart.

Deutsche Bundesbank. (2020). *Jahresabschlussstatistik (Verhältniszahlen)*. Zugriff am 2021-05-16 auf https://www.bundesbank.de/resource/blob/834640/a5039354898fdf3c59c958d5aed92ce9/mL/2020-06-12-08-15-52-jahresabschlussstatistik-verhaeltniszahlen-data.pdf

DIHK. (2020). *Auswirkungen des Coronavirus auf Beschäftigungspläne deutscher Unternehmen.* Zugriff am 2021-05-16 auf https://de.statista.com/statistik/daten/studie/1110554/umfrage/umfrage-unter-deutschen-unternehmen-zu-auswirkungen-des-coronavirus-auf-beschaeftigsplaene/

Fischer, D. T. (2020). Covid-19: Implikationen für die IFRS-Berichterstattung. *PiR*, *2020* (4), 141–142.

Freiberger, H. (2020). Grabesruhe in der Filiale. *Süddeutsche Zeitung vom 14. November 2020*. Zugriff am 2021-05-16 auf https://www.sueddeutsche.de/wirtschaft/banken-corona-filialschliessungen-krise-1.5114548

Gopinath, G. (2020). *Reopening from the Great Lockdown: Uneven and Uncertain Recovery.* Zugriff am 2021-05-16 auf https://blogs.imf.org/2020/06/24/reopening-from-the-great-lockdown-uneven-and-uncertain-recovery/

Grömling, M., Hüther, M., Beznoska, M. & Demary, M. (2020). Wirtschaftspolitische Antworten auf die Corona-Krise – Liquidität hat Vorrang! In *IW-Policy Paper* (S. 1–6). Institut der deutschen Wirtschaft (IW).

Grottel, B. & Haußer, J. (2020). §251 HGB Haftungsverhältnisse. In B. Grottel, S. F. Schmidt, W. J. Schubert & U. Störk (Hrsg.), *Beck´scher Bilanz-Kommentar* (12th Aufl.). C.H.BECK.

Henkel, K., Isensee, W. & Esquivel, S. M. (2021). Corona-Pandemie: Unternehmensradar zur Rechnungslegung. *WPg*, *2021* (5), 294.

Heubeck Zins-Info. (2021). Rechnungszins für internationale Rechnungslegung. *Rechnungszins*. Zugriff am 2021-05-16 auf https://www.heubeck.de/assets/Download/HI_Zinsinfo/2021/01/HI_IFRS_Rechnungszins_dt_2020-12-31.pdf

IDW. (2020a). *Auswirkungen der Ausbreitung des Coronavirus auf die Rechnungslegung und deren Prüfung (Teil 2)*.

IDW. (2020b). *Auswirkungen der Ausbreitung des Coronavirus auf die Rechnungslegung zum Stichtag 31.12.2019 und deren Prüfung (Teil 1)*.

IDW. (2021). *Zweifelsfragen zu den Auswirkungen der Ausbreitung des Coronavirus auf die Rechnungslegung und deren Prüfung (Teil 3, 5. Update)*.

ifo-Institut. (2020). *Corona-Krise: Umfrage unter Unternehmen zur Überlebensdauer 2020.* Zugriff am 2021-05-16 auf https://de.statista.com/statistik/daten/studie/1113650/umfrage/corona-krise-umfrage-unter-unternehmen-zur-ueberlebensdauer/

International Monetary Fund. (2020). *World Economic Outlook, April 2020: The Great Lockdown.* Zugriff am 2021-05-16 auf https://www.imf.org/en/Publications/WEO/Issues/2020/04/14/weo-april-2020

Jonas, M. (2021). Wertaufhellung durch enttäuschte Erwartungen? *Der Betrieb* (15), 18–20.

Kantar. (2020). *Betroffenheit deutscher Unternehmen durch die Corona-Pandemie, Eine Studie von KANTAR im Auftrag des Bundesministeriums für Wirtschaft und Energie.* Zugriff am 2021-05-16 auf https://www.bmwi.de/Redaktion/DE/Downloads/B/betroffenheit-deutscher-unternehmen-durch-die-corona-pandemie.pdf?__blob=publicationFile&v=4

Kirsch, H. (2020). Pensionsrückstellungen nach Handels- und Steuerrecht sowie IFRS. *PiR*, *2020* (12), 398–399.

Kirsch, H. (2021). Voraussichtliche Auswirkungen der Corona-Pandemie auf die IFRS-Abschlüsse 2020. *PiR*, *2021* (1), 16.

Krudewig, W. (2018). Rückstellungen im Jahresabschluss der GmbH. *GmbH-Steuerpraxis*, *2018* (4), 107–111.

Kunz, A. (2020). Corona-Krise beschleunigt Zweigstellenabbau deutscher Banken. *Welt am Sonntag*. Zugriff am 2021-05-16 auf https://www.welt.de/wirtschaft/article207033825/Investors-Marketing-Corona-Krise-fuehrt-zu-Filialabbau-bei-Banken.html

Lüdenbach, N. (2019). *IFRS, Training für Ausbildung und Praxis*.

Lüdenbach, N. (2020). Kurzarbeitergeld und Aufstockung. *StuB*, *2020* (18), 727.

Maier-Reimer, G. & Etzbach, P. (2011). Die Patronatserklärung, in: NJW–Neue Juristische Wochenschrift. *NJW– Neue Juristische Wochenschrift*, *2011* (16), 1110–1112.

Merkt, H. (2021). HGB §249 Rückstellungen. In A. Baumbach & K. J. Hopt (Hrsg.), *Handelsgesetzbuch* (40th Aufl.). C.H.BECK.

Müller, S. (2021). Coronavirus und Rückstellung: Ansatz und Bewertung im Jahresabschluss 2019/2020 und 2020 nach HGB. *Haufe*. Zugriff am 2021-05-16 auf https://www.haufe.de/finance/jahresabschluss-bilanzierung/coronavirus-und-rueckstellungen-ansatz-und-bewertung_188_517506.html

Müller, S. & Reinke, J. (2020). Corona-Krise: Handelsbilanzielle Auswirkungen. *BC - Zeitschrift für Bilanzierung, Rechnungswesen und Controlling*, 568–572.

Nagengast, K. & Boecker, C. (2016). Restrukturierungsrückstellungen nach IFRS im Vergleich zum HGB. *Zeitschrift für Internationale Rechnungslegung (IRZ)*, 306–308.

Oser, P. & Wirtz, H. (2021). Rückstellungsreport 2020. *StuB*, *2021* (1), 1–6.

Pellens, B., Fülbier, R. W., Gassen, J. & Selhorn, T. (2017). *Internationale Rechnungslegung: IFRS 1 bis 16, IAS 1 bis 41* (10th Aufl.).

Pütz, T. (2020). Sind alle Schulden vollständig erfasst? *FAQs zu Auswirkungen von Covid-19 auf die Rechnungslegung (KPMG)*. Zugriff am 2021-05-16 auf https://home.kpmg/de/de/home/themen/2020/04/auswirkungen-von-covid-19-auf-die-rechnungslegung/sind-alle-schulden-vollstaendig-erfasst.html

Raffelhüschen, B., Bahnsen, L., Kohlstruck, T., Seuffert, S. & Wimmesberger, F. (2021). Was kann sich der ehrbare Staat noch leisten? Update 2020/21: Corona, Schulden – und noch eine Pflegereform. In *Argumente zu Marktwirtschaft und Politik.* Stiftung Marktwirtschaft.

Ragnitz, J. (2020). Langfristige wirtschaftliche Auswirkungen der Corona-Pandemie. *ifo Institut – Leibniz- Institut für Wirtschaftsforschung an der Universität München*, *73* (1), 25–30.

Rahe, I. (2020). Für welche Stichtage ist Covid-19 relevant? *FAQs*

zu Auswirkungen von Covid-19 auf die Rechnungslegung (KPMG). Zugriff am 2021-05-16 auf https://home.kpmg/de/de/home/insights/2020/04/auswirkungen-von-covid-19-auf-die-rechnungslegung/fuer-welche-stichtage-ist-covid-19-relevant.html#wertaufhellend

Rimmelspacher, D. & Kliem, B. (2020). Auswirkungen des Coronavirus auf die handelsrechtliche Finanzberichterstattung. *WPg*, *2020* (7), 381–385.

Rinker, C. (2020). Auswirkungen des Corona-Virus auf den HGB-Jahresabschluss und Lagebericht. *StuB*, *2020* (7), 256–262.

Roos, B. (2016). Zur Passivierung von Rückstellungen nach IAS 37 – Überlagerung einer nicht vorhandenen faktischen durch eine rechtliche Verpflichtung? *DStR*, *2016* (37), 2174–2178.

Ruhnke, K. & Simons, D. (2018). *Rechnungslegung nach IFRS und HGB*. Schäffer-Poeschel Verlag Stuttgart.

Schellenberger, J. & Baggen, L. (2020). *Die „Covid-19-Pandemie" – rechtliche Auswirkungen auf werkvertragliche Rechtsbeziehungen.* Zugriff am 2021-05-16 auf https://www.goerg.de/sites/default/files/2020-03/200323_GÖRG_Legal_Update_RechtlicheAuswirkungenaufwerkvertraglicheRechtsbeziehungen_0.pdf

Schrimpf-Dörges, C. (2020). §12 Rückstellungen. In *Beck´sches IFRS-Handbuch.* Brune, J. and Driesch, D. and Schulz-Danso, M. and Senger, T.

Schütte, S. & Götz, C. (2021). Jahresabschluss 2020 im Lichte der Corona-Pandemie. *DStR*, *2021* (6), 371–372.

Schubert, W. J. (2020). §249 HGB. In B. Grottel, S. F. Schmidt, W. J. Schubert & U. Störk (Hrsg.), *Beck´scher Bilanz-Kommentar.*

Schwirkslies-Filler, M. (2019). *Rückstellungen bilden, auflösen und buchen* (1st Aufl.).

Stamm, A. & Owerdieck, B. (2021). *Die Auswirkungen der Corona-Pandemie auf die Rechnungslegung nach HGB und IFRS.* Zugriff am 2021-05-16 auf https://www.dhpg.de/de/newsroom/blog/corona-rechnungslegung/

Störk, U. & Büssow, T. (2020). §253 HGB. In B. Grottel, S. F. Schmidt, W. J. Schubert & U. Störk (Hrsg.), *Beck´scher Bilanz-Kommentar* (12th Aufl.).

Tanski, J. S. (2019). 6 Passiva: Bilanzierung und Bewertung. In *Jahresabschluss – inkl. Arbeitshilfen* (5th Aufl.).

Thommen, J. P., Achleitner, A. K., Gilbert, D. U., Hachmeister, D., Jarchow, S. & Kaiser, G. (2020a). Rechnungslegung nach HGB. In *Allgemeine Betriebswirtschaftslehre.* Springer Gabler Verlag Wiesbaden.

Thommen, J. P., Achleitner, A. K., Gilbert, D. U., Hachmeister, D., Jarchow, S. & Kaiser, G. (2020b). Rechnungslegung nach IFRS. In *Allgemeine Betriebswirtschaftslehre.* Springer Gabler Verlag.

Weaver, I. (2020). Störung der Lieferkette durch Covid-19 – Force Majeure? Es kommt darauf an! *ZVertriebsR*, 159–162.

Weber, J. & Weißenberger, B. E. (2015). Bilanzierung von Fremdkapital. In *Einführung in das Rechnungswesen.* Schäffer-Poeschel Verlag Stuttgart.

Winnefeld, R. (2015). Kapitel D: Ansatz der Aktiva und Passiva in Handels- und Steuerbilanz sowie IAS/IFRS-Rechnungslegung. In *Bilanz-Handbuch.*

Zugehör, P. (2021). Folgen des Coronavirus auf die Rechnungslegung. *StuB*, *2021* (5), 1.

Zwirner, C. & Boecker, C. (2021). Corona-Pandemie im Jahresabschluss: Auswirkungen auf die Rechnungslegung nach IFRS. *IRZ – Zeitschrift für Internationale Rechnungslegung* (2), 57–60.

Zwirner, C., Busch, J. & Krauß, F. (2020). IDW-Hinweise zu den Auswirkungen der Ausbreitung des Coronavirus auf die Rechnungslegung. *Der Betrieb*, *2020* (24), 1241–1244.

Junior Management Science 7(2) (2022) 429-454

Junior Management Science

journal homepage: www.jums.academy

Carbon Risk in European Equity Returns

Fabian Alexander Meyer

Universität Wien

Abstract

Investors perceive climate change and the volatility of asset prices caused by the ongoing low carbon transition of the economy, so-called carbon risk, to have an impact on their portfolio performance. However, the extent of carbon risk's impact on asset prices is still largely unknown. This paper provides a comprehensive quantification of carbon risk in European equity prices and examines whether it constitutes a systematic risk factor. I construct a carbon risk factor to determine the unique share of return attributable to differences in carbon intensity. During the sample period less (more) carbon intensive firms offer higher (lower) returns, which leads to a significant positive return of the carbon risk factor. Moreover, the carbon factor is significantly related to the sample covariance matrix of returns and offers a carbon risk premium in the cross-section of returns. In combination with the enhanced explanatory power relative to standard asset pricing models, this indicates that carbon risk constitutes a systematic risk factor. Consequently, investors can estimate carbon risk exposures based on widely available stock returns and include stocks without explicit carbon emission information in their risk management and investment process.

Keywords: Carbon risk; carbon risk factor; factor model; asset pricing.

1. Introduction

Compared to pre-industrial times, human activities are assumed to have led to a 1°C increase in global temperature (IPCC, 2018). Scientists agree that maintaining this development has potentially devastating consequences for natural and human systems. Hence, experts call for collective efforts to limit the global temperature rise to 1.5°C. Especially the financial system and its participants, who provide funding for proven yet costly cleaner technologies, are identified as crucial stakeholders to achieve this goal (IPCC, 2018). In this context, the European Union (EU) increases its efforts regarding sustainable finance to promote "the transition to a low-carbon, more resource-efficient and sustainable economy" (European Commission, 2021). In 2019 the EU announced the *European Green Deal* as the spearhead of its measures with the main goal of reaching climate neutrality by 2050. This transition process will be financed with an investment plan amounting to at least one trillion Euros of public and private investments. These political efforts are accompanied by heightened climate awareness in the private sector, where carbon divestment and green investment movements are entering the mainstream. For example, the Institutional Investors Group on Climate Change (IIGCC) has 270 members from 16 European countries with more than 35 trillion Euros in assets under management who pledge to support the low carbon transition of the economy (IIGCC, 2021). The enormous magnitude of the outlined financial streams has the potential to affect asset prices. As climate change's financial impact is not restricted to certain industries or firms but affects the entire economic system, the underlying carbon risk might explain systematic broad stock market movements.

Although the question, how carbon risk affects stock returns, generates widespread interest in the nascent climate finance literature, current evidence is mixed and primarily focused on the United States of America (US) (e.g., Bolton & Kacperczyk, 2020; In, Park, & Monk, 2019; Pedersen, Fitzgibbons, & Pomorski, 2021). In contrast, academic articles on carbon risk's effect on asset prices focusing solely on Europe are scarce. Moreover, ambiguity exists regarding the appropriate measures of carbon risk. This motivates authors to use complex, composite measures of carbon risk (e.g., Görgen et al., 2020), which raises questions regarding the comparability of existing results.

Based on common carbon intensity measures of the STOXX Europe Total Market Index constituents in the pe-

DOI: https://doi.org/10.5282/jums/v7i2pp429-454

riod from December 2006 to June 2020, the main goal of this thesis is to examine the following research questions:

1. Is carbon risk a systematic risk factor in European equity returns?
2. What is the magnitude and direction of the relationship between carbon risk and asset prices?

A further aspect within the scope of this thesis is to highlight the carbon risk factor's usefulness for risk management purposes.

Ambiguous empirical results on the relationship between carbon emission and stock returns and contradicting theoretical arguments emphasize that carbon risk could affect stock returns in different ways.[1] First, considering the currently observed divestment of carbon intensive assets, investors in brown firms might hold stakes larger than optimal in them and require higher compensation for non-optimal risk-sharing (Heinkel, Kraus, & Zechner, 2001). However, even in the absence of discriminatory tastes, holding carbon intensive stocks exposes shareholders to policy and technological risks, as regulatory changes target especially those firms accelerating global warming. Consequently, rational investors should require compensation for bearing this additional risk. Recent research, furthermore, suggests that green assets exert hedging properties against climate risk (Choi, Gao, & Jiang, 2020; Engle, Giglio, Kelly, Lee, & Stroebel, 2020), which should lead to lower returns when no climate event materializes. Thus, based on discriminatory tastes and risk consideration, one could expect high emission firms to outperform cleaner ones. I refer to this as the *dirty alpha hypothesis* in the following.

Alternatively, stocks with low carbon intensity potentially perform better if this attribute relates positively to future profits and this is not correctly priced yet. This argument is based on the empirical observation that low carbon intensity positively predicts accounting and market performance (e.g., Busch & Lewandowski, 2018; Pedersen et al., 2021). Moreover, a positive outperformance of low emission stocks could also be caused by higher demands due to recent shifts in tastes (e.g., Pástor, Stambaugh, & Taylor, 2021). Such effects could be driven by the rising awareness of market participants for climate change. However, in this case one would not expect a long lasting effect once investors' preferences are correctly reflected in asset prices. In general, mispricing and unexpected changes in preferences might cause an outperformance of low emission stocks which is referred to as the *clean alpha hypothesis*.

To examine my research questions, the thesis proceeds as follows. Chapters 2 and 3 briefly review the theoretical basis for this thesis. First, the principles of carbon accounting are outlined before different specifications of factor models are discussed. Chapter 4 reviews the related literature and chapter 5 describes the data. The construction of the carbon factor and its properties are included in chapter 6. Chapter 7 contains a rigorous test whether the carbon risk factor is indeed a priced risk factor and chapter 8 proposes fields of application. The results of the various analyses are aggregated and discussed in chapter 9 before a short conclusion is drawn in chapter 10.

2. Carbon Accounting

2.1. Measures of CO_2 and Global Warming Potential

The greenhouse effect is a phenomenon caused by so-called greenhouse gases (GHG) which occur naturally in the Earth's atmosphere and maintain the habitability of the planet. However, human activities have led to a concentration of GHGs in the atmosphere and give rise to the anthropogenic greenhouse effect caused by positive *radiative forcing*, i.e., more radiation is received than emitted by the Earth's climatic system (e.g., Brohé, 2017). Global awareness for the issue arose when the United Nations (UN) launched their first international framework with the aim to "prevent dangerous anthropogenic interference with the climate system" (UNFCCC, 1992, p. 9). As a consequence, efforts were undertaken to make GHG emissions traceable and quantifiable through so-called GHG inventories. The process of measuring CO_2 equivalents at the entity level is commonly referred to as carbon or GHG accounting.[2]

Although carbon dioxide is frequently equated with greenhouse gas, carbon accounting also covers other GHGs. More precisely, the *Kyoto Protocol* identifies five additional relevant GHGs: methane (CH_4), nitrous oxide (N_2O), sulphur hexafluoride (SF_6), hydrofluorocarbons, and perfluorocarbons (UNFCCC, 2008). Each GHG has a different global warming potential (GWP) depending on its radiative efficiency and residence time in the atmosphere, which characterizes its radiative forcing effect over a specific time span. I.e., GWP measures how much energy is absorbed by one tonne of a GHG over a given time period relative to one tonne of carbon dioxide (United States Environmental Protection Agency, 2021). Consequently, carbon dioxide's GWP is normalized to one. Organizations or states then record the emission of each GHG individually before converting it into carbon accounting's main unit of account, metric tonnes of CO_2 equivalent (tCO_2e), based on GWP values. Due to the great complexity involved in measuring the impact of individual greenhouse gases on radiative forcing and changing atmospheric conditions, GWP values are regularly updated and therefore constitute a source of uncertainty for carbon accounting. A commonly proposed best practice is the usage of the most recent GWP values for 100 years as provided by UN's Intergovernmental Panel on Climate Change (e.g., WBCSD & WRI, 2013). In the following, I use these CO_2

[1] In the following hypotheses on the direction of carbon intensity's impact on stock prices, I do not strictly differentiate between risk-based explanations and characteristics relating to investors' biases. For a differentiation compare, for instance, Pukthuanthong, Roll, and Subrahmanyam (2019).

[2] For a comprehensive literature review of the different meanings of carbon accounting in different disciplines, see Stechemesser and Guenther (2012).

equivalents reported following the GHG protocol to determine the greenhouse gas emission of a company.[3]

2.2. Boundaries of Carbon Accounting

The GHG protocol (WBCSD & WRI, 2004), which sets out rules for the emission quantification and reporting, was first published in 2001 and updated in 2004 based on practitioners' feedback. Since then the GHG protocol has developed into a widely applied standard for the preparation of emission inventories at the corporate level and is endorsed by major data providers (Busch, Johnson, Pioch, & Kopp, 2018). Like traditional financial accounting, carbon accounting also considers varying legal and organizational structures of companies. Hence, reporting firms can define the organizational boundaries according to which the emission inventory is compiled. The two approaches proposed by the GHG protocol are the *equity share approach* and the *control approach*. With the equity share approach the company reports emissions according to its economic interest in the company, which typically coincides with the equity stake in an operation (WBCSD & WRI, 2004). As an alternative, the control approach only considers CO_2 emissions from controlled operations, rather than those the company has interest in. This approach can be further distinguished based on whether financial or operational control over the entities of interest exists.

After identifying the organizational boundaries of considered GHG emissions, a further differentiation between direct and indirect emissions is necessary. This is referred to as the operational boundaries or scopes of carbon accounting. The GHG protocol proposes three scopes (WBCSD & WRI, 2004): The direct CO_2 emissions of a company are subsumed in the scope 1 measure. Depending on the organizational boundaries, this encompasses all GHG emissions from sources controlled or owned by the firm. Among others, it includes the GHG emissions set free by the combustion of fossil fuels, transportation of materials or the processing of chemicals. Scope 2 and 3 cover indirect carbon emissions, i.e., those emissions that are caused by company activities "but occur at sources owned or controlled by another company" (WBCSD & WRI, 2004, p. 25). More precisely, scope 2 refers to emissions from the purchase of externally generated electricity. As it represents one of the largest sources of a firm's GHG emission, it is individually reported and identified as one of the foremost sources of GHG emission reductions (e.g., Brohé, 2017; Busch & Lewandowski, 2018). The optional category of scope 3 emission covers the remaining indirect CO_2 emissions along the firm's value chain. This includes up- and downstream emissions, for instance, emissions from purchased input goods or the usage of sold products. Due to its optional character and the inherent ambiguity of determining emissions for the entire value chain, scope 3 reporting typically only covers a small fraction of the real value (Busch, 2011).

[3] For simplicity I use the terms GHG emission, carbon emission, and CO_2 emission interchangeably in the following.

3. Factor Models

3.1. General Structure of Factor Models

Nowadays the capital asset pricing model's (CAPM) inability to individually explain the cross-section of expected returns is commonly accepted and part of finance textbooks (e.g., Bodie, Kane, & Marcus, 2018; Campbell, Lo, & MacKinlay, 1997). Rather, returns of individual assets or portfolios of multiple assets are likely to depend on more than one determinant at the same time and even to a varying degree over time. Since return time series are observed to follow similar patterns and the joint analysis of multiple asset returns quickly results in complex or even inefficient multivariate statistical analyses (e.g., Tsay, 2014), a large stream of the finance literature is concerned with the identification of broad underlying factors for dimension reduction. In general, three main types of factor models emerged in the finance literature (Zivot & Wang, 2003). The first two types can be classified as theoretical approaches which either use macroeconomic variables as factors or construct factor portfolios relating to firm characteristics. These types are therefore commonly referred to as *macroeconomic factor model* and *fundamental factor model*, respectively. Thirdly, *statistical factor models* exist which extract the non-observable factors from the realized stock returns (Zivot & Wang, 2003).

Generally, the structure introduced below underlies each of the three types of factor models, but for clarity the particularities of statistical factor models are highlighted in the next section. It is assumed that an asset's return generating process is a linear function of a limited number of common factors. For notational simplicity, only the cross-sectional regression formulation of the model is presented which has the form

$$\underset{(N\times 1)}{\boldsymbol{r_t}} = \underset{(N\times 1)}{\boldsymbol{\alpha}} + \underset{(N\times K)(K\times 1)}{\boldsymbol{\beta}\ \boldsymbol{f_t}} + \underset{(N\times 1)}{\boldsymbol{\varepsilon_t}}, \forall t, \tag{1}$$

where $\boldsymbol{r_t} = \left(r_{1,t}, \ldots, r_{N,t}\right)'$ denotes the vector of either real or excess returns of assets i $(i = 1, \ldots, N)$ at time t $(t = 1, \ldots, T)$.[4] The vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)'$ represents the intercept, $\boldsymbol{\beta}$ is the matrix of factor loadings, and $\boldsymbol{f_t}$ refers to the common factor realizations at time t.[5] K denotes the number of common factors included in the model. The first two moments of $\boldsymbol{f_t}$ are described by $E[\boldsymbol{f_t}] = \overline{\boldsymbol{f}}$ and the covariance matrix $\boldsymbol{\Omega_f}$. Finally, $\boldsymbol{\varepsilon_t}$ denotes the vector of asset specific factors, or error terms. The covariance matrix of the error terms, $\boldsymbol{D}$, is assumed to be diagonal. Moreover, common factors and error terms are assumed to be uncorrelated among each other. In contrast, especially in the macroeconomic and fundamental factor model, common factors can

[4] If factors are observable, equation 1 can be interpreted as cross-sectional regression. This interpretation, however, is incorrect if factors are unobservable in statistical factor models (Tsay, 2010). For alternative representations of the general factor model, see Zivot and Wang (2003) or Tsay (2010).

[5] Pukthuanthong et al. (2019) further classify common factors into priced and unpriced common factors.

be correlated. The resulting covariance matrix of asset returns thus can be denoted by

$$\boldsymbol{\Omega_r} = \boldsymbol{\beta}\,\boldsymbol{\Omega_f}\,\boldsymbol{\beta}' + \boldsymbol{D}. \tag{2}$$

The model formulation presented in equation 1 is particularly common for macroeconomic and fundamental factor models. For instance, it is the basis for the well-known macroeconomic single factor model by Sharpe, the market model, and its notable later extensions to a fundamental factor model by Fama and French (1993), Carhart (1997), and Fama and French (2015). Especially the methodology by Fama and French (1993), where factor returns are derived from hedge portfolios, is mentioned as the archetype of a fundamental factor model (e.g., Tsay, 2010). Conceptually, the mentioned macroeconomic and fundamental factor models have two main differences compared to the statistical factor model. First, since returns and therefore factors are observable, factor loadings and idiosyncratic volatility can be estimated directly through time series regression (Zivot & Wang, 2003).[6] A second conceptual difference mentioned by Tsay (2010) lies in the indeterminacy of the number of factors in fundamental factor models because characteristics could be (dis-)aggregated to adjust the number of factors. Statistical models offer a remedy by dissociating the factors mathematically, but issues regarding the number of factors to be chosen and the lack of economic interpretability arise (Campbell et al., 1997). Linking fundamental and statistical factors through canonical correlation analysis in section 7.1 therefore closes the gap between economical intuition and computational rigor.

3.2. Statistical Factor Model and Asymptotic Principal Component Analysis

In statistical factor models both factor realizations and factor loadings of assets are latent and must be estimated based on the return vector of N assets, $\boldsymbol{r_t}$. In comparison to the fundamental factor model, statistical factor models therefore have the advantage of only requiring return data and the absence of multicollinearity (Alexander, 2001). The traditional statistical factor model is based on an orthogonal factor structure and assumes that $\boldsymbol{r_t}$ is generated by a stationary process with mean $\overline{\boldsymbol{r}}$ and a covariance matrix of $\boldsymbol{\Omega_r}$ (Tsay, 2010). As above, it is assumed that few linear combinations, the unobserved K common factors $\boldsymbol{f_t} = \left(f_{1,t}, \ldots, f_{K,t}\right)$, can be used to explain a large fraction of $\boldsymbol{\Omega_r}$. The remaining unexplained share of variance is linearly explained by the vector of error terms $\boldsymbol{\varepsilon_t}$, which is assumed to be independent of $\boldsymbol{f_t}$. However, for identifiability it is now assumed that $E[\boldsymbol{f_t}] = 0$ and the factor covariance matrix equals the $K \times K$ identity matrix $\boldsymbol{I_K}$. Since the common factors explain all cross-covariances of asset returns, the return covariance matrix simplifies to

$$\boldsymbol{\Omega_r} = \boldsymbol{\beta}\,\boldsymbol{I_K}\,\boldsymbol{\beta}' + \boldsymbol{D} = \boldsymbol{\beta}\boldsymbol{\beta}' + \boldsymbol{D}, \tag{4}$$

compared to equation 2. Moreover, now the matrix of factor loadings, $\boldsymbol{\beta}$, must be full column rank, otherwise one or more factors are redundant (Tsay, 2014). As shown in equation 5 below, the only formula-related change in the model relative to the general factor model in equation 1 occurs in the constant; i.e., $\boldsymbol{\alpha}$ is replaced by $\overline{\boldsymbol{r}}$. This is commonly done without loss of generality since the factors are computed from the covariance matrix (Campbell et al., 1997). Equation 5, however, uniquely identifies $\boldsymbol{f_t}$ and $\boldsymbol{\beta}$ only up to an orthogonal transformation which can hamper the interpretation of the factors (Zivot & Wang, 2003).

$$\underset{(N\times 1)}{\boldsymbol{r_t}} - \underset{(N\times 1)}{\overline{\boldsymbol{r}}} = \underset{(N\times K)}{\boldsymbol{\beta}}\,\underset{(K\times 1)}{\boldsymbol{f_t}} + \underset{(N\times 1)}{\boldsymbol{\varepsilon_t}}, \forall t \tag{5}$$

If the number of time periods exceeds the number of assets (i.e., $T > N$), a common method for factor estimation is principal component analysis (PCA). However, in many financial applications N exceeds T. In this case, the return covariance matrix of the sample cannot be inverted anymore, i.e., it becomes singular and restricts the usage of PCA (Zivot & Wang, 2003). Asymptotic principal component analysis (APCA), a method proposed by Connor and Korajczyk (1986, 1988), provides a remedy and is shown to be asymptotically ($N \to \infty$) equivalent to the traditional factor analysis under certain assumptions.[7] It applies traditional PCA to the sample cross-sectional covariance matrix of demeaned stock returns $\boldsymbol{R}$,

$$\underset{(T\times T)}{\boldsymbol{\Omega_R}} = \frac{1}{N}\,\underset{(T\times N)}{\boldsymbol{R}'}\,\underset{(N\times T)}{\boldsymbol{R}}. \tag{6}$$

The eigenvectors of the K largest eigenvalues of $\boldsymbol{\Omega_R}$ are then used to consistently estimate the common factors (Tsay, 2014).

4. Related Literature

A broad stream of literature studies the impact of environmental, social, and governance (ESG) ratings in financial markets. Due to the crucial importance of global warming, particularly the environmental perspective of ESG with focus on carbon emission has generated widespread interest. My thesis is primarily linked to this dynamically growing literature on carbon risk's effect on asset prices. Although the literature in the field is primarily empirical, also notable theoretical contributions have been made.

[6] This feature is extensively used in the following chapters to determine the relevance of the carbon factor. Refining equation 1 for individual assets, i, and explicitly incorporating the period's risk-free rate, $r_{f,t}$, the resulting model can be represented by

$$r_{i,t} - r_{f,t} = \alpha_i + \boldsymbol{\beta_i'}\boldsymbol{f_t} + \varepsilon_{i,t}. \tag{3}$$

[7] See Tsay (2014) for a discussion of the assumptions of APCA and the presented formula.

An early theoretical contribution regarding the effects of environmentally friendly investing on asset prices in equilibrium is made by Heinkel et al. (2001). The authors argue that if the market is segmented into green and non-green investors, the latter require compensation for limited risk-sharing when holding polluting firms shunned by green investors. Hence, Heinkel et al.'s model predicts higher expected returns for non-green firms. An alternative equilibrium model incorporating ESG considerations is proposed by Pástor et al. (2021). According to their model, green stocks command negative CAPM alphas, whereas brown stocks have positive abnormal returns. In the model's base setting, the underperformance of green stocks is explained by the investors' ESG preferences. The investors' equilibrium asset choice is described by three-fund separation where each portfolio consists of the risk-free asset, the market portfolio, and an ESG portfolio. As a first extension of their model, Pástor et al. introduce a priced ESG factor which explains the above mentioned CAPM alphas. If ESG concerns increase unexpectedly (i.e., positive realization of the ESG factor), green firms perform better due to higher demand for green products and greater utility derived by investors holding them. In this case, green firms outperform brown ones in spite of lower expected returns. In a second extension of their model for climate risk, Pástor et al. introduce an additional risk-based rationale for brown stocks' higher expected returns. Green stocks can be seen as hedges against climate shocks, which investors dislike, and therefore should offer lower returns than their browner counterparts.[8] This leads to the conclusion that in the presence of climate risk sensitivities investors should additionally hold a climate-hedging portfolio (four-fund separation). Finally, in a further recent theoretical contribution, Pedersen et al. (2021) extend the traditional CAPM for ESG preferences and present the ESG-Sharpe ratio frontier which conceptualizes the trade-off between risk, return, and ESG preferences. Similar to Pástor et al. (2021), they find a four-fund separation solution to be optimal. However, funds are allocated between the risk-free asset, the market portfolio, the minimum variance portfolio, and the optimal ESG portfolio in their model. Moreover, the authors show that higher ESG assets can either generate lower or higher expected returns relative to their conventional counterparts. The latter is the case if ESG characteristics predict higher future profits and a large fraction of investors is unaware of this relationship. That means prices of these profitable assets are not bid up yet. In an empirical examination of their predictions based on the S&P 500 index, Pedersen et al. measure the environmental dimension as the scope 1 and 2 carbon intensity provided by Trucost for the period from January 2009 to March 2019. They examine whether ESG measures help to generate abnormal returns by testing their relation to future fundamentals and investors' demands. Lower carbon intensity is shown to positively predict accounting returns (measured by return on net operating assets) and increase institutional ownership, which results in higher valuations for greener assets. Moreover, carbon intensity-sorted quintile hedge portfolios generate weakly significant excess returns and positive alphas relative to common factor models.

Besides the mentioned theoretical contributions, a broad empirical literature has emerged recently. Using a similar approach as Pedersen et al. (2021), In et al. (2019) study the performance of carbon intensity hedge portfolios. Their sample ranges from 2005 to 2015 and covers 736 US stocks. As main measure of carbon risk they use scope 1 to 3 carbon intensity provided by Trucost. Overall, they find economically and statistically significant returns when going long in low carbon intensity stocks and shorting high intensity stocks. Bolton and Kacperczyk (2020) examine whether carbon emissions have an impact on the cross-section of US stock returns. Their sample spans the period from 2005 to 2017 and includes more than 3,400 firms for which Trucost provides scope 1 to 3 CO_2 emission data. Controlling for known firm-specific return predictors (e.g., market value, book-to-market ratio, and profitability), Bolton and Kacperczyk do not find a significant cross-sectional relationship between emission intensity and stock return. However, both absolute amounts of emission and emission growth rates are significantly related to returns for all three scopes. This points to the existence of a carbon return premium by which investors are compensated for bearing carbon risk associated with high emission firms. Moreover, this premium cannot be explained by known risk factors such as the Fama-French factors. Testing whether the carbon premium is driven by institutional investors' withdrawal of funds from high emission firms and resulting limited risk-sharing (e.g., as in Heinkel et al. (2001)), Bolton and Kacperczyk only find weak evidence for systematic effects. While investors divest from companies with high scope 1 carbon intensity to a statistically significant degree, this effect is not observable for scope 2 and 3 carbon intensity. Consequently, they do not seem to avoid high emission companies per se.

Another stream of empirical literature does not only ask whether but also when carbon risks materialize. Recent research suggests that carbon risk affects asset prices when the impact of climate change becomes tangible reality. For a sample of 74 stock exchanges worldwide, Choi et al. (2020) find that in times of abnormally warm local temperatures the attention to global warming rises and low carbon stocks outperform. Moreover, they show that especially retail investors, who are prone to limited attention problems, react to climatic events by replacing stock ownership in high emission firms with low emission firms. Instead of looking at climate disasters directly, Engle et al. (2020) propose a dynamic equity hedging approach whose returns hedge against climate news innovations. Using environmental scores of US firms from MSCI and Sustainalytics in the time from 2009 to 2016, they find significant covariations between the hedge portfolios and the climate change news indices. In a related article, Ilhan, Sautner, and Vilkov (2020) add further evidence for the existence of priced carbon risk. Studying the impact of climate policy uncertainty in option prices of S&P 500 firms

[8] For empirical evidence regarding the hedging properties of green stocks, see Engle et al. (2020) and Choi et al. (2020).

in the time from 2009 to 2016, they show that protection against left-tail risk is more costly for carbon intensive firms. This effect tends to be even more systematic on the sector-level, when firm-specific risk is diversified away. Moreover, using proxies of climate change attention as in Engle et al. (2020) and Choi et al. (2020), Ilhan et al. also find a stronger positive effect of carbon intensities on the implied volatility slope and hence on the option price when climate change awareness is higher.

Most closely related to my thesis is the paper by Görgen et al. (2020). They study carbon risk in asset prices based on a global sample of 1,657 firms in the period from 2010 to 2017. Joining environmental data from four databases, they construct a carbon risk factor using the Fama and French (1993) methodology. Besides carbon intensity, which they classify as the *value chain* dimension of carbon risk, Görgen et al. also consider additional *public perception* and *adaptability* measures as firm characteristics in the calculation of their factor. Among others, the latter two dimensions include aggregate environmental scores and related subscores. Employing a panel regression approach, they show that brown firms provide higher returns on average. Although their carbon risk factor explains variation in stock returns, they do not find evidence of an associated risk premium. Lastly, Hübel and Scholz (2020) use the same factor construction methodology and create an environmental, social, and governance factor, respectively. Based on Datastream's aggregate ESG scores of the annual STOXX Europe Total Market Index constituents from 2003 to 2016, they observe an outperformance of firms with low environmental rating. Furthermore, they show how firms without ESG ratings can be incorporated in investors' risk management by computing exposures of individual stocks to their three factors.

5. Data

The thesis focuses on European equities which are proxied by the STOXX Europe Total Market Index in the time from December 2006 to June 2020. I download annual constituents as per year-end from Thomson Reuters Datastream to avoid potential survivorship bias over the sample period. The index covers at least 95% of the free-float market capitalization in each of the 17 European countries included in the index (Deutsche Börse Group, 2021).[9] Hence, this sample reflects the investable universe of most investors and ensures the practical relevance of my results. All financial data retrieved from Datastream are US dollar (USD) denominated to accord with other data sources. As the direct usage of financial data from Thomson Reuters can lead to incorrect inference compared to other data providers,[10] I follow the guidelines on static and dynamic screens by Ince and Porter (2006) and Schmidt, von Arx, Schrimpf, Wagner, and Ziegler (2019) for data cleaning. Static screens refer to time-invariable information, whereas dynamic screens address information which changes over time. Applied static screens comprise, for example, the exclusion of non-equity assets and foreign or non-primary listed stocks. The used dynamic screens aim at resolving potential liquidity biases arising from incorrectly padded zero-returns and Datastream's decimalization, which leads to distortions for pennystock returns. For a comprehensive overview of the applied data screens see table 19 in appendix A. Similar to related literature (e.g., Görgen et al., 2020; In et al., 2019), I exclude firms labeled as *financials* according to the Industry Classification Benchmark (ICB). Financials tend to have low levels and a small dispersion of carbon intensities, despite potentially high carbon risk exposures due to their business model. For instance, reinsurance companies are among the firms most vulnerable to climate change in spite of low carbon intensities. Furthermore, firms – especially financials – have considerable freedom regarding reported carbon emissions depending on their defined organizational boundaries. Restricting the sample to non-financial firms results in 1,710 unique stocks from ten industries. The financial data set is complemented by USD denominated European monthly factor return data provided by French (2020). As risk-free rate the one month US treasury bill rate included in French's data set is used.

In the following, I use carbon intensity as measure of carbon risk. Carbon intensity is defined as metric tonnes of scope 1 and 2 CO_2 emission equivalents per one million USD of net sales:

$$\text{Carbon Intensity} = \frac{\text{Scope 1 \& Scope 2 Emission}}{\text{Net Sales}} \left[\frac{tCO_2e}{\text{million USD}}\right]. \tag{7}$$

Hence, it reflects how much tonnes of CO_2 the company requires to generate one million USD of net sales. Due to its relative nature, carbon intensity is more robust towards firm size or process adjustments, and accounts for changes in emission from variations in production over the business cycle (Busch & Lewandowski, 2018). The used emission figures from Datastream comprise company reported and estimated carbon emissions. A study by Busch et al. (2018) finds that not only company reported emissions but also scope 1 and 2 estimates are mostly homogeneous across major data providers. Since scope 3 emissions are less widely available and inconsistencies in emission estimates exist (Busch et al., 2018), I do not include this emission scope in my analyses.

Table 1 shows the industry composition of my sample including stocks with and without emission information and the time series average summary statistics of carbon intensity. The largest industries in my sample are industrials (25.26%) and consumer discretionary (20.47%), whereas none of the remaining eight industries has a share above 10%. The mean of carbon intensity exceeds the median in each of the ten industries and the third quartile in seven industries. This indicates the presence of highly carbon intensive firms in all

[9] Countries included are: Austria, Belgium, Poland, Denmark, Finland, France, Germany, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the UK.

[10] For instance, Ince and Porter (2006) do not find the widely observed momentum effect in US returns using raw data from Datastream.

Table 1: Summary statistics of carbon intensity and industry composition.

	Count		Carbon Intensity			
Industry	Absolute	%	1st Q	Median	Mean	3rd Q
Basic Materials	128	7.49	165.50	415.56	648.91	798.16
Consumer Discretionary	350	20.47	8.17	19.08	106.22	47.22
Consumer Staples	127	7.43	28.87	48.85	73.94	80.41
Energy	101	5.91	63.43	224.01	475.04	427.66
Health Care	144	8.42	13.70	24.04	44.92	40.86
Industrials	432	25.26	17.51	32.86	304.00	82.70
Real Estate	133	7.78	21.36	44.04	66.43	69.33
Technology	144	8.42	7.47	12.43	28.44	22.51
Telecommunication	75	4.39	10.77	23.29	53.07	34.35
Utilities	76	4.44	331.96	866.59	1,753.90	1,509.55
Total	1,710	100				

industries. In particular, the striking difference between the median and the mean for industrials is primarily driven by the inclusion of large cement producers in this industry sector, e.g., LafargeHolcim. Among the ten industries, utilities, basic materials, and energy can be identified as especially carbon intensive industries. In contrast, less carbon intensive industries are consumer discretionary, telecommunication, and technology.

Table 2 shows the resulting annual sample used in the subsequent sections. As the emission data are compiled from company reports, I rebalance my sample at the end of June of each year to account for potential look-ahead bias arising from the lagged availability of company filings. The row *Rated* refers to stocks with emission information available at the end of the previous year, $t-1$, and price information in July of year t. In contrast, the *Unrated* row includes stocks without emission information. The main focus of my analyses is the last decade from July 2010 to June 2020, whose start is indicated by the vertical line in the table. The preceding period from July 2007 to June 2010 is used for auxiliary purposes in rolling regressions to generate estimates in later tests of the factors. While the sample's total number of stocks is relatively stable and ranges between 1,212 and 1,323, the number of rated stocks grows gradually from 364 in 2007 to 1,020 in 2019. Although the selected time span is relatively short compared to common asset pricing studies, the restriction to this time period appears meaningful for the following reasons. First, carbon accounting tends to be a relatively new phenomenon as discussed in section 2. While Thomson Reuters offers the longest time series of emission data among major data providers going back to 2002 (Busch et al., 2018), previous to the fiscal year 2006 the data availability drops sharply. Moreover, most other major data providers start their coverage during or only shortly before my sample period (Busch et al., 2018). Thus, including the earlier, additional years could introduce systematic differences and lead to false inference. Using the years after 2009 as main sample period is also consistent with other studies and ensures the comparability of my results (e.g., Engle et al., 2020; Görgen et al., 2020; Pedersen et al., 2021). Secondly, research by Engle et al. (2020) suggests that public awareness for climate change has gradually increased since the change of the millennium and intensified during the sample period. Key events mentioned in their study include the UN Climate Change Conferences in Copenhagen (12/2009), Doha (12/2012), and Paris (12/2015). The latter culminated in the notorious *Paris Agreement*. Especially in Europe, recently also grassroots movements such as *Fridays for Future* manifest this heightened awareness.

Prior to the subsequent main analyses I check for a selection bias caused by the exclusion of financials. For this purpose, I test the screened sample relative to the Fama and French (1993) 3-factor model extended for the momentum factor proposed by Carhart (1997).[11] Similar to the factor model outlined in section 3.1, the regression equation for an individual portfolio i is denoted by

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i^{MKT} f_t^{MKT} + \beta_i^{HML} f_t^{HML} + \beta_i^{SMB} f_t^{SMB} + \beta_i^{MOM} f_t^{MOM} + \varepsilon_{i,t}. \quad (8)$$

In the equation, the superscript MKT, HML, SMB, and MOM refer to the widely applied market, value, size, and momentum factor, respectively.[12] Table 3 shows the fit of my sample relative to this 4-factor model. Both the full sample and the sample of stocks with emission information available (*Rated*) are well explained by the 4-factor model with an adjusted R^2 close to one. Despite the exclusion of financials, the market betas are close to one and the alphas are close to zero and statistically not significant. I conclude that neither the exclusion of financials nor the restriction to firms with emission information introduce a systematic bias.

[11] In the literature this model constellation is frequently referred to as Carhart 4-factor model. For brevity, I frequently refer to this model as the 4-factor model in the following.

[12] For a comprehensive description of the factors and their construction see, e.g., French (2020).

Table 2: Number of stocks in the sample in July of year t.

t	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Rated	364	649	637	664	692	687	704	715	732	749	743	832	1,020
Unrated	936	632	575	569	561	547	535	571	583	567	580	489	264
Total	1,300	1,281	1,212	1,233	1,253	1,234	1,239	1,286	1,315	1,316	1,323	1,321	1,284

Note. The vertical line represents the start of the main sample period. The period from July 2007 to June 2010 is used for auxiliary purposes.

Table 3: 4-factor model for the value-weighted portfolio of all stocks and rated stocks from 07/2010 to 06/2020.

	MKT	HML	SMB	MOM	Alpha	Adj. R^2
All	1.006***	-0.192***	-0.094***	0.033	-0.011	98.88
	(93.02)	(-9.23)	(-3.55)	(1.55)	(-0.23)	
Rated	1.005***	-0.204***	-0.138***	0.025	-0.022	98.82
	(90.11)	(-8.78)	(-4.85)	(1.17)	(-0.45)	

Note. *p<0.1, **p<0.05, ***p<0.01; t-statistics in parentheses; heteroskedasticity robust standard errors are used; alpha is reported as monthly percentage.

6. Understanding Carbon Risk in Asset Prices

6.1. Return of Carbon Intensity-Sorted Portfolios

In order to establish an understanding of the relation between carbon intensity and returns, I construct carbon intensity-sorted quintile portfolios. The value-weighted portfolios are annually updated at the end of June based on the reported carbon intensity of the previous year end. The first quintile (*Clean*) contains the firms with the lowest carbon intensities, whereas the fifth quintile contains the firms with the highest emissions intensities (*Dirty*). Figure 1 plots the cumulative returns of the quintile portfolios adjusted for the value-weighted return of the market portfolio in the time from July 2007 to June 2020. The vertical red dotted line represents the start of the main sample period in mid 2010. While more carbon intensive portfolios performed better in the earlier years, the low emission portfolios outperformed clearly in the more recent years. Notable is also the inverse development of the top and bottom quintile starting in mid 2008. The cleanest 20% of stocks show a relatively linear positive development, whereas the opposite is true for the dirtiest quintile.

Over the sample period a monotonous relationship between excess returns and carbon-intensity quintiles emerges. Table 4 reports the risk and return profile of the quintiles. Panel A shows the results from the main sample period from 07/2010 to 06/2020 and panel B additionally includes the preceding three years for comparison. The latter sample includes the great financial crisis and is characterized by lower average excess returns and greater volatility. In both panels the extreme portfolios have a higher standard deviation. While in panel A this does not affect the monotonous relationship in Sharpe ratios, in panel B the second quintile offers the most favorable risk-return trade-off. The calculated Sharpe measures are also consistent with MacKinlay (1995) who finds that over long periods the annualized Sharpe ratios of diversified portfolios lie well below one. Overall, only the monthly excess return of the first quintile in panel A is statistically different from zero at the 5% significance level. At least weak significance is found for the two adjacent portfolios.

6.2. Construction of the SCOPE12 Factor

After observing a monotonous negative relationship between returns and carbon intensities, I proceed by constructing a carbon risk factor. Since the carbon intensity is calculated based on the scope 1 and 2 emissions, I refer to the constructed carbon factor as *SCOPE12*. I apply the Fama and French (1993) methodology for factor construction, where a characteristic-based return predictor is converted into a factor by calculating the return differential of a zero-investment portfolio. Comparable to Fama and French's HML factor, SCOPE12 controls for size and goes long (short) in stocks with low (high) carbon intensity. I annually double sort stocks into six portfolios based on their market capitalization in June of year t and their carbon intensity at the end of the previous year $t-1$. I apply the median size in my sample and the terciles of carbon intensity as breakpoints. The annual update at the end of June aims at reducing potential look-ahead bias compared to related studies (e.g., Görgen et al., 2020; Hübel & Scholz, 2020). I then calculate the value-weighted monthly simple returns, r_t, for the four relevant portfolios for factor construction: small size and high carbon intensity (SH), small size and low carbon intensity (SL), big size and high carbon intensity (BH), and big and low carbon intensity (BL). The SCOPE12 factor goes long in the equally-weighted portfolio of firms with the lowest carbon intensity and shorts the equally-weighted portfolio of firms with the highest carbon intensity. Hence, the return on the carbon risk factor in period t, $f_t^{SCOPE12}$, is defined as the return of the following portfolio,

$$f_t^{SCOPE12} = \frac{1}{2}(r_t^{SL} + r_t^{BL}) - \frac{1}{2}(r_t^{SH} + r_t^{BH}). \tag{9}$$

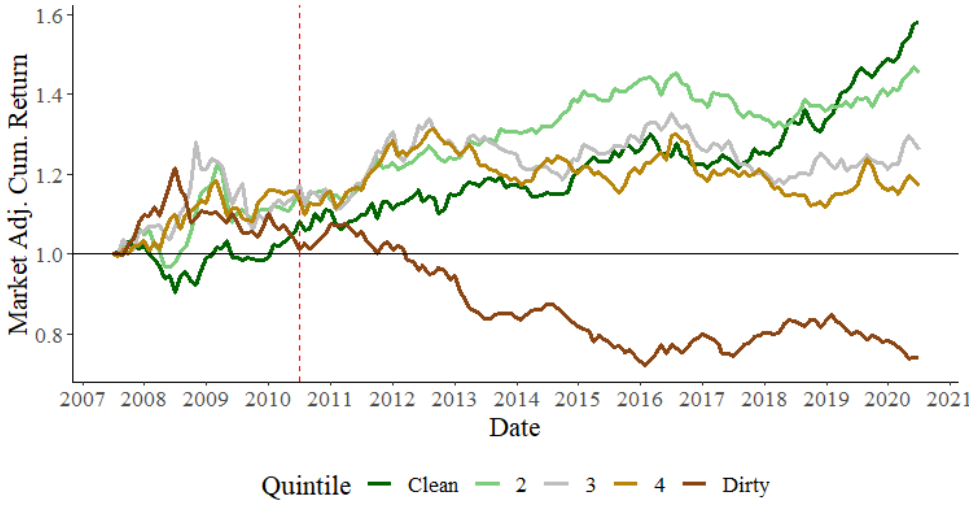

Figure 1: Market adjusted cumulative returns of the carbon intensity-sorted, value-weighted quintile portfolios in the period from 07/2007 to 06/2020.

Table 4: Risk and return of carbon intensity-sorted quintile portfolios.

Panel A. Period from 07/2010 to 06/2020.

Quintile	Clean	2	3	4	Dirty
Excess return	0.94**	0.81*	0.69*	0.65	0.37
Standard deviation	4.98	4.62	4.27	4.81	5.50
t-statistic	2.07	1.91	1.76	1.48	0.73
Sharpe ratio	0.65	0.61	0.56	0.47	0.23

Panel B. Period from 07/2007 to 06/2020.

Quintile	Clean	2	3	4	Dirty
Excess return	0.52	0.47	0.38	0.33	0.04
Standard deviation	5.84	5.06	4.80	5.55	6.31
t-statistic	1.12	1.15	1.00	0.74	0.07
Sharpe ratio	0.31	0.32	0.28	0.21	0.02

Note. *p<0.1, **p<0.05, ***p<0.01; excess return is reported as monthly percentage; the Sharpe ratio is annualized.

To ensure the robustness of my results, I also consider alternative specifications of the carbon factor. First, I address potential objections that the above factor simply reflects the differential between well-performing, low emission technology companies and low-performing, high emission industries. For this purpose, I replicate the SCOPE12 factor with industry adjusted carbon intensity terciles. That is, I keep the size sort unchanged and repeat the second sort into carbon intensity terciles for each industry individually. Then I allocate the stocks to the four relevant portfolios as above. The resulting factor, *IND_SCOPE12*, is significantly correlated ($\rho = 0.63$) with the SCOPE12 factor. In addition to this, I also consider the factor construction without an additional size control. The clean minus dirty (CMD) factor goes long in the value-weighted portfolio of the least carbon intensive tercile and short in the most carbon intensive tercile. The correlation between the CMD and SCOPE12 factor is 0.92, suggesting that SCOPE12's additional size sort has only limited impact on the performance of the factor. For comparison, figure 2 below depicts the cumulative returns of the three factors in the period from July 2007 to June 2020. The vertical red dotted line again represents the start of the main sample. While the SCOPE12 factor is more volatile and experienced a pronounced drawdown during the great financial crisis, it performed well since mid 2008 generating average monthly returns of 0.323% ($t = 2.137$, $\sigma = 1.885\%$). IND_SCOPE12, in contrast, generated a mean return of 0.201% ($t = 2.205$) with a lower standard deviation of 1.140% per month. The CMD factor, which neither adjusts for size nor for industry, offered an average monthly return of 0.395% ($t = 2.218$, $\sigma = 2.224\%$). As indicated by the t-statistics in parentheses, all three average monthly factor returns are statistically

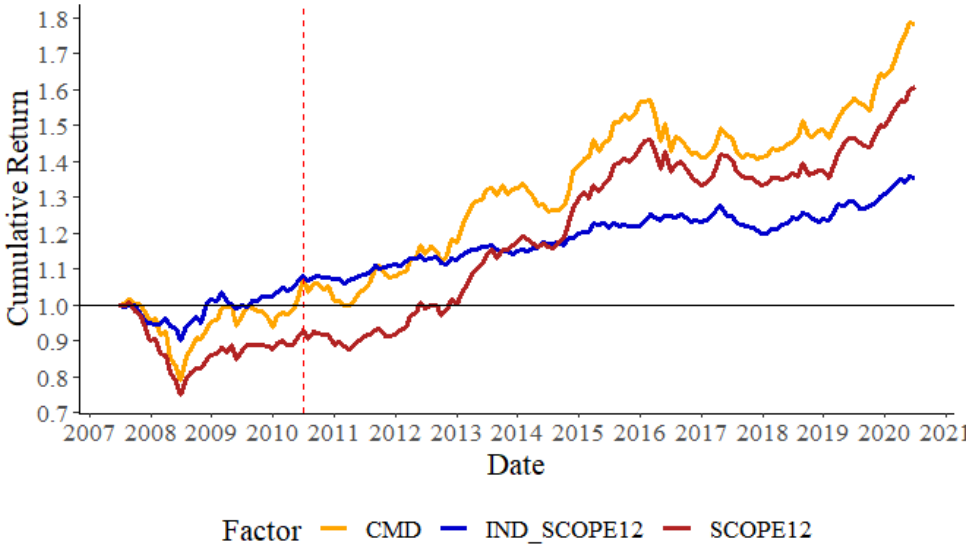

Figure 2: Cumulative returns of the CMD, SCOPE12, and IND_SCOPE12 factor in the period from 07/2007 to 06/2020.

significant at the 5% level.

Despite the close resemblance of SCOPE12 and the two alternative factors, I only consider the SCOPE12 factor in my subsequent analyses for the following reasons. First, carbon risks are more likely to materialize based on absolute carbon intensities rather than emission intensity differences within industries. Hence, the economic intuition appears weaker for the industry-sorted factor. Secondly, related research finds that ESG ratings positively depend on firm size (Kaiser, 2020) because larger firms might be able to invest in greener technologies. As a consequence, double sorting rules out potential size effects present in my data and the CMD factor.

Table 5 reports the descriptive statistics and Pearson correlation coefficients of the SCOPE12 factor with the Fama-French 3-factor and 5-factor model extended for momentum. The analysis is based on monthly returns in the period from 07/2007 to 06/2020. Panel A reports the results for the 4-factor model underlying my further analyses. Additionally, panel B shows the relation of the SCOPE12 factor to the Fama and French (2015) 5-factor model augmented with momentum.[13] Although I do not further consider the Fama-French 5-factor model in later analyses, I include it here to ensure the robustness of the carbon factor and rule out potentially strong interdependencies between the factors. The restriction to the 4-factor model as benchmark appears reasonable as CMA and RMW themselves are relatively new factors which still raise concerns regarding their robustness and economic rationale (e.g., Blitz, Hanauer, Vidojevic, & van Vliet, 2018). Moreover, inevitable additional dependencies between the factors would further complicate the interpretation of the results. Particularly, the highly negative correlation of -0.78 between HML and RMW raises concerns regarding potential redundancies.

Among the downloaded factors only the returns on HML, MOM, and RMW are at least weakly statistically significant. While common factors do not need to be uncorrelated in fundamental factor models (Tsay, 2010), problems of multicollinearity might arise if the SCOPE12 factor covaries strongly with one or more of the standard factors used in the asset pricing literature. Overall, the SCOPE12 factor seems to be only weakly correlated ($|\rho| < 0.3$) with the factors of either model. Of course, looking at the correlation matrix does not detect constellations in which collinearity exists between three or more variables and pairwise correlations are low (Kennedy, 2008). Therefore, I additionally orthogonalize the factors in section 6.4 and eliminate all linear relations between the factors. Nevertheless, the observed low correlations in both panels provide a first indication that the SCOPE12 factor is not merely a replication of known factors but might extend the explanatory power of standard factor models.

6.3. Explanatory Power of the SCOPE12 Factor

6.3.1. Carbon Intensity-Sorted Quintile Portfolios

A frequent criterion for a factor to be deemed relevant is that it enhances the explanatory power in a time series regression relative to previously proposed factors. I therefore test whether the carbon factor can improve the adjusted R^2 compared to the Carhart 4-factor model. The underlying model is the model presented in equation 8 extended with the SCOPE12 factor as additional explanatory variable. The resulting model used in the following subsections can be writ-

[13] While SMB refers to the original size factor constructed based on the book-market ratio sorted portfolios, I denote the Fama and French (2015) size factor as $SMB^{\dagger}$. $SMB^{\dagger}$ is constructed as the average of the three size factors based on the book-market, profitability, and investment sort. RMW and CMA refer to the Fama and French (2015) profitability and investment factor, respectively.

Table 5: Descriptive statistics and correlations of the factors from 07/2007 to 06/2020.

Panel A. Carhart 4-factor model and SCOPE12.

				Correlations			
	Mean	t-stat	σ	MKT	SMB	HML	MOM
MKT	0.215	0.47	5.677				
SMB	0.049	0.320	1.917	0.014			
HML	-0.340*	-1.689	2.512	0.495	-0.049		
MOM	0.815***	2.631	3.870	-0.477	-0.034	-0.532	
SCOPE12	0.323**	2.137	1.885	-0.230	0.072	-0.274	0.053

Panel B. Fama-French 5-factor + momentum model and SCOPE12.

				Correlations					
	Mean	t-stat	σ	MKT	SMB†	HML	MOM	RMW	CMA
MKT	0.215	0.47	5.677						
SMB†	0.072	0.48	1.876	0.028					
HML	-0.340*	-1.689	2.512	0.495	0.013				
MOM	0.815***	2.631	3.870	-0.477	-0.063	-0.532			
CMA	-0.047	-0.403	1.463	-0.213	-0.209	0.406	0.120		
RMW	0.431***	3.425	1.573	-0.362	-0.051	-0.784	0.401	-0.348	
SCOPE12	0.323**	2.137	1.885	-0.230	0.066	-0.274	0.053	-0.083	0.141

Note. *p<0.1, **p<0.05, ***p<0.01; mean is reported as monthly percentage.

ten as

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i^{MKT} f_t^{MKT} + \beta_i^{HML} f_t^{HML} + \beta_i^{SMB} f_t^{SMB} + \beta_i^{MOM} f_t^{MOM} + \beta_i^{SCOPE12} f_t^{SCOPE12} + \varepsilon_{i,t}. \quad (10)$$

Since it is lively debated in the finance literature whether asset pricing tests should be conducted with portfolios or single assets (e.g., Pukthuanthong et al., 2019), I first test the SCOPE12 factor with portfolio returns as left-hand-side variable. That is, I use the carbon intensity-sorted and value-weighted quintile portfolio returns from above. As before, the first quintile contains the clean companies, i.e., the 20% of stocks with the lowest carbon intensity. In contrast, the fifth quintile contains the most carbon intensive or dirty stocks. Table 6 reports the median carbon intensity of the five portfolios during the main sample period. As the carbon intensities tend to decrease in all quintiles during the sample period, I report the average of the annual median figures. While the median carbon intensities of the first three portfolios are relatively close to each other in absolute terms, the gap widens for the two upper quintiles. The median in the dirtiest quintile is even more than six times as high as in the fourth quintile and nearly 95 times larger than in the least carbon intensive portfolio.

Table 7 reports the results of the quintile portfolio regressions on the 4-factor model (panel A) and the extended 5-factor model (panel B). In all regressions the market betas are close to one and statistically significant indicating a broad exposure to the market. With respect to the HML factor there is a contrarian exposure between the fifth quintile and the other four quintiles in both panels. While all quintile portfolios load negatively on the SMB factor, only the betas of the three intermediate carbon intensity portfolios are statistically significant. Lastly, the momentum factor reliably explains returns only in the third portfolio and all regression intercepts are insignificant in both panels. Although the traditional 4-factor model cannot be rejected based on this observation, the alphas' absolute distance from zero decreases in all quintiles through the addition of SCOPE12. Panel B shows the corresponding SCOPE12 factor loadings. The second quintile has a marginally higher carbon factor beta than the first quintile. For the higher carbon intensity portfolios the factor loading decreases monotonically and becomes the largest in magnitude for the fifth quintile with -0.729. The third portfolio's SCOPE12 beta of close to zero indicates that no significant exposure to carbon risk exists in this quintile.

Looking at the change in adjusted R^2, ΔR^2, reported in the last column of panel B, the SCOPE12 factor adds explanatory power in each of the quintiles except for the third one. To check whether this difference is statistically significant, the last column in panel B also reports the result of the heteroskedasticity robust F-statistic for nested models (e.g., Wooldridge, 2016). Except for the third quintile, the F-test is highly significant and confirms the SCOPE12 factor's ability to enhance explanatory power. Especially for the high carbon intensity portfolios, the additional adjusted R^2 turns out larger with 1.82% in the fourth quintile and 3.80% in the fifth quintile.

To ensure the robustness of my results, the portfolio regression analysis is repeated for the IND_SCOPE12 and the

Table 6: Median carbon intensities in the carbon intensity- sorted quintile portfolios from 07/2010 to 06/2020.

Quintile	Clean	2	3	4	Dirty
Median	6.50	17.95	36.24	91.97	616.73

Table 7: 4-factor model and 5-factor model with SCOPE12 for carbon intensity-sorted quintile portfolios from 07/2010 to 06/2020.

Panel A. 4-factor model.

Quintile	MKT	HML	SMB	MOM	Alpha	Adj. R^2
Clean	1.072***	-0.338***	-0.018	0.066	0.094	93.68
	(31.34)	(-6.75)	(-0.25)	(1.34)	(0.72)	
2	0.985***	-0.293***	-0.096*	-0.032	0.138	95.89
	(40.87)	(-6.12)	(-1.81)	(-0.83)	(1.45)	
3	0.929***	-0.337***	-0.242***	0.076*	-0.058	94.13
	(33.10)	(-5.81)	(-3.79)	(1.70)	(-0.51)	
4	1.022***	-0.345***	-0.204***	0.003	-0.087	92.96
	(39.74)	(-5.17)	(-2.68)	(0.06)	(-0.69)	
Dirty	1.036***	0.229***	-0.091	0.014	-0.188	93.10
	(37.92)	(3.17)	(-1.23)	(0.25)	(-1.43)	

Panel B. 5-factor model.

Quintile	MKT	HML	SMB	MOM	SCOPE12	Alpha	Adj. R^2	ΔR^2
Clean	1.079***	-0.272***	-0.034	0.061	0.276***	-0.009	94.30	0.62***
	(31.78)	(-5.36)	(-0.54)	(1.31)	(4.15)	(-0.07)		
2	0.992***	-0.225***	-0.112**	-0.037	0.282***	0.032	96.67	0.78***
	(46.29)	(-5.41)	(-2.43)	(-0.94)	(4.42)	(0.36)		
3	0.929***	-0.345***	-0.240***	0.077*	-0.033	-0.046	94.09	-0.04
	(33.28)	(-5.66)	(-3.78)	(1.73)	(-0.60)	(-0.41)		
4	1.010***	-0.451***	-0.178***	0.012	-0.445***	0.079	94.78	1.82***
	(40.41)	(-6.25)	(-2.70)	(0.29)	(-5.46)	(0.66)		
Dirty	1.017***	0.055	-0.048	0.028	-0.729***	0.085	96.90	3.80***
	(48.81)	(1.00)	(-0.86)	(0.71)	(-11.82)	(0.89)		

Note. *p<0.1, **p<0.05, ***p<0.01; t-statistics in parentheses; heteroskedasticity robust standard errors used. The last column reports the significance of the heteroskedasticity consistent F-test for nested models.

CMD factor, respectively. Table 20 in appendix B contains the regression table for both factors as additional explanatory variables. In contrast to SCOPE12, the industry-sorted carbon factor adds more explanatory power for the least carbon intensive portfolio. The CMD factor performs similar to SCOPE12 for carbon intensive portfolios and explains more variance in the less carbon intensive portfolios. Overall, the results from above remain largely unchanged with significant carbon betas and F-tests in the same portfolios as before.

6.3.2. Impact of SCOPE12 on Single Stock Returns

The previous section shows that the SCOPE12 factor performs well when applied to carbon intensity-sorted portfolios. To be included in these portfolios, however, carbon emission information must be available for the respective stock. As shown in table 2, such information is not available for a large fraction of firms. Furthermore, a frequently mentioned concern with portfolio formation is that it renders stock-specific effects unobservable. Therefore, I repeat the time series regression on the 5-factor model (see equation 10) with individual stocks' excess returns as dependent variable. In comparison to the previous portfolio analysis, I break down the ten year period into two subperiods of five years from 07/2010 to 06/2015 with 868 observations and from 07/2015 to 06/2020 with 897 available complete time series. This procedure allows for approximately 30% larger subsamples in each of the periods because a complete ten year return time series is only available for 667 firms.

Table 8 reports the absolute and relative frequency of significant factor loadings for each of the conventional significance levels. Results for the period from 07/2010 to 06/2015 are shown in panel A and the subsequent period in panel B. The proportion of significant SCOPE12 regression coefficients is relatively similar in both panels, albeit slightly higher

in the first five year period. In the first five year subperiod more stocks have a significant exposure to SCOPE12 than to the momentum factor. Furthermore, SCOPE12 performs only slightly worse than the HML factor during this time span. In panel B, the carbon factor is only marginally less significant than the momentum factor. In sum, SCOPE12 helps to explain the variation of individual stock returns.

The above analysis jointly examines the full sample of emission rated and unrated stocks. However, concern exists that the carbon factor's significance is driven by the emission rated stocks, which are also included in the factor construction. Therefore, a further differentiation into *rated* and *unrated* subsamples appears reasonable. Since some assets' carbon emissions are not continuously available and the universe of stocks with CO_2 emission information is gradually extended by data providers, I classify stocks with less than three years (out of ten years) of emission information as unrated. Tables 21 and 22 in appendix C report the subperiod performances of the SCOPE12 factor for rated and unrated stocks, respectively. Looking at the significant factor betas, it can be seen that the SMB factor is of higher relevance in the unrated sample. This is consistent with the frequent observation that firms without ESG-related ratings tend to be smaller on average (e.g., Hübel & Scholz, 2020). Although the SCOPE12 factor is a bit less significant in the unrated sample, the patterns remain largely unchanged. It is to be noted, however, that the other factor candidates are also less significant in the unrated subsample, except for SMB and MOM in the first subperiod. Looking at panel B in table 22, the carbon factor's significance levels closely resembles those of MOM and HML. These solid results in the unrated sample suggest that the carbon factor can be generalized beyond stocks with available emission information, which creates a sound foundation for the more rigorous single stock analyses in chapter 7.

6.4. Orthogonal Factor Model

6.4.1. Orthogonalization of the Common Factors

Despite only low or moderate correlations between the factors, as shown in table 5, equation 2 indicates that the variance of asset returns not only depends on factor loadings but also on the covariances of factor returns. I therefore orthogonalize the common factors and reassess the sensitivities of asset returns to the unique variation of the underlying factor. This helps to determine how much of the systematic variance is indeed attributable to the respective factor. In order to disentangle the linear relationships in the 5-factor model, I use the democratic orthogonalization procedure suggested by Klein and Chow (2013). Their method builds on Löwdin's symmetric orthogonalization and assigns equal shares of common variance to the respective orthogonalized factors. This results in the identical R^2 in a multiple regression as if the original factors were used. Moreover, using symmetric orthogonalization ensures maximal similarity with the original factors and a unique result in comparison to sequential procedures.[14]

I apply the symmetric orthogonalization to the downloaded Carhart 4-factors and the self-calculated SCOPE12 factor.[15] Table 9 reports the distributional properties of the orthogonalized factors and their correlations with the original factors. The similarity of the orthogonal factors with the original factors is very high with most correlation coefficients close to unity. That is, correlations are 0.970, 0.998, 0.896, 0.941, and 0.976 for the MKT, SMB, HML, MOM, and SCOPE12 factor, respectively. Only HML, which is the factor with the highest cross-correlations in table 5, is slightly lower correlated with its orthogonal counterpart. With regard to the orthogonal factors' mean returns, the general direction and magnitude remains similar to the original factors. The largest change can be observed for the mean return of the market factor, which changes from 0.215% (MKT) to 0.621% ($MKT^{\perp}$) per month. Analogously to the original factors, only momentum and the carbon factor have mean returns different from zero at the 5% level. By construction, the standard deviations are unchanged compared to the original factors and the correlations among the orthogonally transformed factors themselves are zero.

I use the orthogonally transformed factor returns as explanatory variables for the value-weighted returns of the carbon intensity-sorted quintile portfolios in table 10. Since the variation of the explanatory variable is now unique to it, ordinary least squares (OLS) regression estimates tend to be more precise given a correct allocation of the jointly explained variance (Kennedy, 2008). In comparison with the results of the original factors in panel B of table 7, I find changes in the signs and magnitudes of the regression coefficients. The regression alphas and (adjusted) R^2, however, remain unchanged to the previous analysis by construction. First, market betas tend to be lower for the uncorrelated factors but are still close to one, except for the third quintile. The greatest changes can be observed with regard to $HML^{\perp}$ and $MOM^{\perp}$. In contrast to the original value factor, the loading on $HML^{\perp}$ is statistically and economically highly significant only for the most carbon intensive portfolio. Furthermore, the orthogonal momentum beta is negative and highly significant for most portfolios. The observed changes suggest that the original factors tend to overestimate the portfolios' exposure to systematic risk arising from the market and value factor, while underestimating the systematic risk from momentum. In contrast, factor loadings on $SMB^{\perp}$ and $SCOPE12^{\perp}$ closely resemble the original factors. The $SCOPE12^{\perp}$ betas are all highly significant and monotonically decrease from the lowest to the highest carbon intensity portfolio. Especially the dirty quintile loads even stronger on the orthogonal carbon factor than the original factor.

Focusing on the carbon factor, the similar results when using $SCOPE12^{\perp}$ instead of SCOPE12 as explanatory variable provide further evidence that the original carbon factor

[14] The uniqueness prove goes back to Aiken, Erdos, and Goldstein (1980).

[15] In the following the orthogonal factors are identified by the superscript perpendicular symbol, $\perp$.

Table 8: Significance levels of the single stock factor betas.

Panel A. Single stock regressions for the period 07/2010 to 06/2015.						
Significance	10%		5%		1%	
$N = 868$	Abs.	%	Abs.	%	Abs.	%
MKT	847	97.58	842	97.00	806	92.86
HML	251	28.92	179	20.62	82	9.45
SMB	317	36.52	230	26.50	115	13.25
MOM	110	12.67	61	7.03	15	1.73
SCOPE12	201	23.16	137	15.78	65	7.49

Panel B. Single stock regressions for the period 07/2015 to 06/2020.						
Significance	10%		5%		1%	
$N = 897$	Abs.	%	Abs.	%	Abs.	%
MKT	863	96.21	846	94.31	795	88.63
HML	220	24.53	156	17.39	71	7.92
SMB	373	41.58	278	30.99	154	17.17
MOM	198	22.07	128	14.27	47	5.24
SCOPE12	182	20.29	117	13.04	47	5.24

Note. Heteroskedasticity robust standard errors used.

Table 9: Descriptive statistics of orthogonal factors and correlation with original factors from 07/2007 to 06/2020.

				Correlations with original factors				
	Mean	t-statistic	σ	MKT	SMB	HML	MOM	SCOPE12
$MKT^{\perp}$	0.621	1.366	5.677	0.970				
$SMB^{\perp}$	0.041	0.264	1.917		0.998			
$HML^{\perp}$	-0.225	-1.120	2.512			0.896		
$MOM^{\perp}$	0.928***	2.995	3.870				0.941	
$SCOPE12^{\perp}$	0.346**	2.294	1.885					0.976

Note. *p<0.1, **p<0.05, ***p<0.01; mean is reported as monthly percentage.

Table 10: Orthogonalized 5-factor model for carbon intensity-sorted quintile portfolios from 07/2010 to 06/2020.

Quintile	$MKT^{\perp}$	$HML^{\perp}$	$SMB^{\perp}$	$MOM^{\perp}$	$SCOPE12^{\perp}$	Alpha	Adj. R^2
Clean	0.991***	0.014	-0.005	-0.117***	0.151**	-0.009	94.30
	(34.55)	(0.37)	(-0.08)	(-3.10)	(2.49)	(-0.07)	
2	0.926***	0.056*	-0.086*	-0.199***	0.147**	0.032	96.67
	(48.21)	(1.73)	(-1.85)	(-6.10)	(2.57)	(0.36)	
3	0.847***	-0.055	-0.223***	-0.071*	-0.107**	-0.046	94.09
	(33.80)	(-1.16)	(-3.51)	(-1.93)	(-2.17)	(-0.41)	
4	0.942***	-0.050	-0.173**	-0.154***	-0.495***	0.079	94.78
	(37.71)	(-0.84)	(-2.59)	(-4.48)	(-6.75)	(0.66)	
Dirty	1.038***	0.448***	-0.056	-0.272***	-0.874***	0.085	96.90
	(60.77)	(11.33)	(-1.01)	(-8.54)	(-16.63)	(0.89)	

Note. *p<0.1, **p<0.05, ***p<0.01; t-statistics in parentheses; heteroskedasticity robust standard errors used.

only has weak linear relationships with the other four factors and provides unique explanatory power. Since all factors are mutually uncorrelated, the presented analysis also rules out more complex linear relationships among three or more factors in contrast to the correlation analysis in table 5.

6.4.2. Decomposition of the Coefficient of Determination

While the previous section provides an overview of the portfolio returns' sensitivities towards the underlying factors, the risk can be closer examined decomposing the coefficient of determination. Due to the absence of covariance between the explanatory variables, fractions of R^2 can be attributed to the individual factors. Asset i's total coefficient of determination, R_i^2, is therefore the sum of the contributions of the K individual factors, $R_{i,k}^2$ (Klein & Chow, 2013). The resulting formula is

$$R_i^2 = \sum_{k=1}^{K} R_{i,k}^2 , \quad \text{where } R_{i,k}^2 = \left(\frac{\hat{\beta}_{k,i}^{\perp} \hat{\sigma}_k^{\perp}}{\hat{\sigma}_i} \right)^2 . \tag{11}$$

The sample estimate of asset i's factor loading on the kth orthogonalized factor is denoted by $\hat{\beta}_{k,i}^{\perp}$. The variables $\hat{\sigma}_k^{\perp}$ and $\hat{\sigma}_i$ denote the sample estimate of the standard deviation of orthogonal factor k's and asset i's returns, respectively.

I apply the systematic risk decomposition to the carbon intensity-sorted quintile portfolios and report the five factors' provision of additional explanatory power in table 11. As mentioned above, the cumulative portion of explained variance in portfolio returns is equivalent to the original factors. As expected for diversified portfolios, MKT$^{\perp}$ contributes most of the explanatory power for each of the quintile portfolios with about 90% of the explained variance. In contrast, the R^2 contributions of HML$^{\perp}$, SMB$^{\perp}$, and MOM$^{\perp}$ lie below 1% in most regressions. Only in the most carbon intensive portfolio, the quintile with the highest R^2, HML$^{\perp}$ and MOM$^{\perp}$ have higher explanatory power with 4.04% and 1.99%, respectively. Of even higher importance in this portfolio, however, is the SCOPE12$^{\perp}$ factor which explains 6.53% of the return's variance. It can be seen that SCOPE12$^{\perp}$ is particularly important in more carbon intensive portfolios, whereas in quintiles 1 to 3 its additional explanatory power is lower. The carbon factor therefore seems to primarily explain return variation of equities especially exposed to the low carbon transition of the European economy.

6.5. Maximum Squared Sharpe Ratios

An alternative to running time series regressions on the factor candidates is the right-hand-side approach in form of spanning regressions proposed by Fama and French (2018). In a spanning regression each factor is individually regressed on the model's remaining factors. In case the regression constant is significantly different from zero, the factor used as dependent variable enhances the right-hand-side model's explanation of average returns. Hence, this approach allows for an additional perspective on whether the carbon factor should be added to the Carhart 4-factor model.

Based on the original results by Gibbons, Ross, and Shanken (1989), Fama and French (2018) show that

$$\hat{\boldsymbol{\alpha}}' \boldsymbol{\Omega}_{\varepsilon}^{-1} \hat{\boldsymbol{\alpha}} = SR^2(\boldsymbol{\Pi}, \boldsymbol{f}) - SR^2(\boldsymbol{f}) = SR^2(\boldsymbol{\alpha}), \tag{12}$$

holds. Where $\boldsymbol{\Pi}$ is the combination of assets' excess returns and alternative factors' returns, which are currently not included in the model. $\boldsymbol{f}$ denotes the factors of the model under consideration. On the left-hand-side of the equation, $\hat{\boldsymbol{\alpha}}$ is a $N \times 1$ vector of estimated regression intercepts from the time series regression of $\boldsymbol{\Pi}$ on $\boldsymbol{f}$. The resulting estimate of the $N \times N$ residual covariance matrix is $\boldsymbol{\Omega}_{\varepsilon}$. Hence, the left-hand-side expression in equation 12 is solely determined by the choice of the factors and their resulting maximum squared Sharpe ratio (SR). Since the squared Sharpe ratio of the intercept, $SR^2(\boldsymbol{\alpha})$, has to be minimized, the factor model with the highest squared Sharpe ratio, $SR^2(\boldsymbol{f})$, is chosen. $SR^2(\boldsymbol{\Pi}, \boldsymbol{f})$ denotes the maximum squared Sharpe ratio which can be achieved by the combination of $\boldsymbol{\Pi}$ and $\boldsymbol{f}$.

The formulation in equation 12 also provides the basis for the estimation of the marginal contribution an additional factor k has to the model's maximum squared Sharpe ratio (Fama & French, 2018):

$$\frac{\hat{\alpha}_k^2}{\hat{\sigma}_{\varepsilon,k}^2} = SR^2(\boldsymbol{f}, k) - SR^2(\boldsymbol{f}) \tag{13}$$

In this equation, $\hat{\alpha}_k$ denotes the constant and $\hat{\sigma}_{\varepsilon,k}$ the residual standard deviation resulting from the time series regression of factor k on the remaining factors of the model. Their squared ratio equates the difference between the maximum squared Sharpe ratio of the model of $\boldsymbol{f}$ and k combined and $\boldsymbol{f}$ individually. The left side of equation 13 indicates that the contribution of factor k to the model's squared Sharpe ratio depends on two factors. First, if $\hat{\alpha}_k$ is close to zero, the returns of factor k are well explained by the other factors and its contribution is small. Secondly, the contribution to the squared Sharpe ratio is small if the other factors only explain little variation in factor k's return ($\hat{\sigma}_{\varepsilon,k}$ is large). As mentioned above, the statistical significance of a factor's marginal contribution is measured by the size of the t-statistic for the intercept in a spanning regression.

Table 12 reports the results from the spanning regressions of the Carhart 4-factor model (panel A) and the extended 5-factor model (panel B) in the time from July 2010 to June 2020. The last column reports the marginal contribution of each factor to the maximum squared Sharpe ratio when added to the model's other factors. The marginal contributions and regression coefficients of the traditional four factors are similar across both models, indicating that the SCOPE12 factor adds a unique portion to the 4-factor maximum squared Sharpe ratio. Among the traditional four factors, only MKT and MOM have significant intercepts and therefore reliably contribute to the Sharpe measure in both models. Although the regression constant is larger for MKT than MOM in both panels, MKT's larger residual variance leads to a smaller Sharpe ratio contribution in comparison

Table 11: R^2 decomposition of carbon intensity-sorted quintile portfolios from 07/2010 to 06/2020.

	Decomposed R^2						
Quintile	$MKT^{\perp}$	$HML^{\perp}$	$SMB^{\perp}$	$MOM^{\perp}$	$SCOPE12^{\perp}$	R^2	$1-R^2$
Clean	93.85	0.00	0.00	0.45	0.24	94.54	5.46
2	94.85	0.09	0.10	1.51	0.26	96.81	3.19
3	93.08	0.10	0.77	0.23	0.16	94.34	5.66
4	90.99	0.07	0.36	0.84	2.75	95.00	5.00
Dirty	84.44	4.04	0.03	1.99	6.53	97.03	2.97

Note. (Decomposed) R^2 values are given as percentages.

Table 12: Spanning regressions and $SR^2(\boldsymbol{f})$ decomposition for the 4-factor and 5-factor model from 07/2010 to 06/2020.

Panel A. 4-factor model.

	α	MKT	HML	SMB	MOM	R^2	σ_ε	$\alpha^2/\sigma_\varepsilon^2$
MKT	1.223***		0.704***	0.071	-0.356**	24.90	4.217	0.084
	(2.949)		(3.766)	(0.303)	(-2.202)			
HML	-0.112	0.155***		-0.019	-0.350***	35.60	1.977	0.003
	(-0.556)	(3.766)		(-0.175)	(-4.996)			
SMB	0.090	0.011	-0.014		0.009	0.10	1.675	0.003
	(0.530)	(0.303)	(-0.175)		(0.138)			
MOM	0.883***	-0.113**	-0.506***	0.018		30.63	2.376	0.138
	(3.875)	(-2.202)	(-4.996)	(0.138)				

Panel B. 5-factor model.

	α	MKT	HML	SMB	MOM	SCOPE12	R^2	σ_ε	$\alpha^2/\sigma_\varepsilon^2$
MKT	1.299***		0.646***	0.084	-0.349**	-0.224	25.34	4.205	0.095
	(3.056)		(3.236)	(0.357)	(-2.157)	(-0.831)			
HML	0.049	0.129***		0.006	-0.309***	-0.402***	41.80	1.879	0.001
	(0.248)	(3.236)		(0.060)	(-4.547)	(-3.497)			
SMB	0.061	0.013	0.005		0.008	0.078	0.56	1.671	0.001
	(0.346)	(0.357)	(0.060)		(0.114)	(0.729)			
MOM	0.863***	-0.111**	-0.493***	0.015		0.050	30.70	2.374	0.132
	(3.653)	(-2.157)	(-4.547)	(0.114)		(0.329)			
SCOPE12	0.373**	-0.027	-0.239***	0.059	0.019		18.83	1.448	0.066
	(2.518)	(-0.831)	(-3.497)	(0.729)	(0.329)				

Note. *p<0.1, **p<0.05, ***p<0.01; α and R^2 are reported as a percentage.

to MOM. In both models, the HML and SMB factor's intercept is statistically non-distinguishable from zero. That is, HML's returns can be explained through highly significant loadings on the MKT, MOM, and SCOPE12 factor. Although SMB cannot be explained by the other factors, SMB's average returns are close to zero during the sample period and therefore cannot be reliably distinguished from zero either. Focusing on the SCOPE12 factor's spanning regression in the last row of panel B, a statistically significant intercept of 0.373% per month can be found. Among the explanatory variables only HML reliably explains the carbon factor's returns. Besides, SCOPE12's spanning regression has the lowest residual standard deviation. This results in the carbon factor having the third largest marginal contribution to the 5-factor model's squared Sharpe ratio with $\frac{\hat{\alpha}^2}{\hat{\sigma}_\varepsilon^2} = 0.066$ behind MOM and MKT.

Overall, the spanning regression analysis suggests that the SCOPE12 factor should be included in the 5-factor model as it provides a statistically significant increase in the model's $SR^2(\boldsymbol{f})$. Although SMB and HML appear redundant in the above analysis, the results of spanning regressions tend to be sensitive to the chosen time period (Fama & French, 2018). Hence, I provide a more rigorous examination of which factors should be considered in a factor model in section 7.1.

7. Factor Identification

7.1. Covariance Matrix of Returns and Factor Returns

The analyses presented so far provide a first indication for the relevance of a carbon factor in order to explain the cross-

section of European stock returns. Being aware of potential data-snooping objections against additionally introduced factors (e.g., Campbell et al., 1997), I conduct additional rigorous tests of my factor. I follow the protocol for the identification of genuine risk factors proposed by Pukthuanthong et al. (2019). According to the authors, a genuine risk factor must fulfill the following requirements: a) it must relate to the covariance matrix of stock returns, b) it must be priced in the cross-section of stock-returns, and c) it should offer a risk-reward ratio consistent with risk pricing.

This section constitutes the first stage in Pukthuanthong et al.'s protocol and aims to determine factors that systematically move asset prices. This is achieved by extracting statistical factors from the sample covariance matrix and relating them to the fundamental factors[16] via canonical correlation analysis. Facing the problem of a large number of observations N and a smaller number of time periods T, I extract principal components (PC) from the transformed $T \times T$ sample covariance matrix of European stock returns using the asymptotic principal component analysis as described in section 3.2. For the full ten year period from July 2010 to June 2020 ($T = 120$), I have complete return time series for $N = 667$ stocks. I construct the $\mathbf{\Omega_R}$ matrix (compare equation 6) based on this data and retain the first ten PCs ($L = 10$), which account for approximately 90% of the volatility in the covariance matrix. The threshold is in line with Pukthuanthong et al. (2019) and ensures that most of the covariances are explained by the calculated statistical factors. Consequently, equation 4 holds approximately.

The next step is to link those ten PCs with the five factor candidates via canonical correlations. The rationale for the usage of canonical correlations is the rotational indeterminacy of statistical factor models mentioned in section 3.2, which restricts their direct interpretation. Hence, checking for correlations between linear combinations of the extracted PCs and linear combinations of the factor candidates provides a reasonable proxy whether a factor is linked to the sample covariance matrix (Pukthuanthong et al., 2019). If a factor candidate is not significantly canonically correlated with the eigenvectors maximizing the explained variance, it is rejected as a genuine risk factor and does not enter the cross-sectional regression in the second stage of the protocol. The canonical correlation analysis between the ten eigenvectors extracted from the $\mathbf{\Omega_R}$ matrix and the five factor candidates ($K = 5$) results in $\min(K, L)$ five canonical variate pairs and hence five canonical correlations. For each canonical variate pair, the weights of the linear combinations of both the eigenvectors and the factor candidates are chosen such that their resulting canonical correlation is maximized. Moreover, additional constraints ensure the uncorrelatedness of the canonical variate pairs. Panel A of table 13 reports the canonical correlations and their respective F-statistics obtained from Wilks' lambda test. The test's null hypothesis is that the correlations in the current and all following rows are equal to zero. It can be rejected for the first four canonical correlations. Especially the first canonical correlation is large and relatively close to one. Hence, the hypothesis of independence of the two multivariate sets of variables can be rejected overall. Panel B builds on this result and reports the significance level for the factor candidates. The reported t-statistics are obtained from the regression of each canonical variate for the set of principal components on all five factor candidates over the whole sample period. Since there are five canonical variate pairs, I run five regressions in total and calculate the arithmetic average of the absolute t-statistics for each factor. The last row in panel B repeats the analysis and calculates only the mean absolute t-statistics of canonical correlations which are statistically significant at the 5% level. Therefore, it only considers the t-statistics of the first four regressions. Since an insignificant canonical correlation indicates that no relationship between the two sets of variables exists, only significant correlations are considered in the following. Moreover, as absolute t-statistics are reported, one-tailed cut-off levels are used in the following discussion. Pukthuanthong et al. (2019) argue that a genuine risk factor should have an average absolute t-statistic of statistically significant canonical correlations in excess of the one-tailed 2.5% cut-off level ($t \approx 1.96$). Focusing on the average t-statistics from significant correlations, only the MOM factor is below the approximate critical rejection level of 1.96. However, its t-statistic ($t = 1.705$) exceeds the threshold for the one-tailed 5% significance level ($t \approx 1.65$). Besides the market risk factor ($t = 5.669$), the SCOPE12 factor has the second highest average absolute t-statistic ($t = 3.537$), which indicates that the carbon factor should be considered in the cross-sectional regressions.

Pukthuanthong et al. (2019) mention the potential presence of nonstationarity as a relevant restriction to the factor extraction from the sample covariance matrix. In order to avoid excluding a relevant risk factor and to account for fluctuations in the relative importance of the five factor candidates over time, I also conduct APCA and calculate canonical correlations for the two five year subperiods from 07/2010 to 06/2015 and 07/2015 to 06/2020, respectively. The results from the additional analyses are reported in tables 23 and 24 in appendix D. Considering the shorter time periods offers larger samples of complete return time series, which again allows to check for robustness of the asymptotic PCs from the full sample period. In general, the canonical correlations and significance levels of the subperiods are close to the full ten year period. However, as expected and in line with Pukthuanthong et al. (2019), I find that the factor candidates vary in relative importance between the subsamples. Particularly, the momentum factor is highly significant with an average absolute t-statistic of 2.225 in the more recent subperiod. Since my time series is rather short and due to the weak significance of MOM for the full time period, I do not want to incorrectly exclude the momentum factor. Based on my results, I conclude that each of the five factor candidates is sufficiently related to the sample covariance matrix. Hence, all five factors are incorporated in the second step

[16] Strictly speaking, MKT is a macroeconomic factor. However, for simplicity I refer to all five factor candidates (MKT, HML, SMB, MOM, and SCOPE12) as fundamental factors in the following.

Table 13: Canonical correlations with asymptotic principal components and significance of factor candidates for the period from 07/2010 to 06/2020 ($N = 667$).

Panel A. Canonical correlations.

Canonical variate	Canonical correlation	F-stat	df1	df2	p-value
1	0.847***	7.423	50	482.2	0
2	0.715***	4.757	36	399	0
3	0.557***	3.095	24	310.9	0
4	0.440**	2.060	14	216	0.02
5	0.186	0.653	6	109	0.69

Panel B. Significance of factor candidates.

	MKT	SMB	HML	MOM	SCOPE12
Mean absolute t-stat	4.659	1.936	2.569	1.529	2.849
Mean absolute t-stat (significant corr.)	5.669	2.344	2.843	1.705	3.537

Note. *p<0.1; **p<0.05; ***p<0.01; t-statistics in panel B are calculated with heteroskedasticity robust standard errors.

cross-sectional regressions.

7.2. Cross-Sectional Regressions and the Carbon Risk Premium

After ensuring that the factor candidates are related to the covariance matrix of stock returns, I conduct the second stage of Pukthuanthong et al.'s protocol for factor identification and check whether the factor candidates command risk premiums by running a variant of Fama and MacBeth (1973) cross-sectional regressions. Fama and MacBeth's methodology offers a widely used way to generate standard errors adjusted for cross-sectional correlations (Cochrane, 2005). Although there is an ongoing debate on whether the cross-sectional regressions should use single stock or portfolio returns as dependent variable, I use individual returns for the following analyses.[17] Among the foremost arguments against the usage of portfolio returns "is that diversification into portfolios can mask cross-sectional phenomena in individual stocks that are unrelated to the portfolio grouping procedure" (Jegadeesh, Noh, Pukthuanthong, Roll, & Wang, 2019, p. 274).

In the first step of the Fama-MacBeth procedure, the first-pass regression, I estimate the betas by running the time series regression shown in equation 10 for each stock i individually. For the time series regression either a rolling window regression approach or a complete time series regression with full sample betas can be used (Cochrane, 2005). While Pukthuanthong et al. compute betas based on all available return observations, I follow the rolling regression approach (e.g., Fama & MacBeth, 1973; Görgen et al., 2020). As the available historical return time series varies across the sample assets and the cross-section is smaller compared to other asset pricing papers, I use 36-month rolling window regressions starting in July 2007 to estimate the five beta coefficients. This approach ensures the comparability of betas across assets at each point in time and retains a larger cross-section than longer regression windows.

In the second step, I test whether individual stocks' risky excess returns linearly depend on their covariances with the five factor candidates (Jegadeesh et al., 2019). For this analysis, the beta estimates are treated as explanatory variables and the following second-pass or cross-sectional regression is performed for each month in the time from July 2010 to June 2020 ($t = 1, \ldots, 120$):

$$r_{i,t} - r_{f,t} = \alpha_t + \gamma_t^{MKT}\hat{\beta}_{i,t}^{MKT} + \gamma_t^{HML}\hat{\beta}_{i,t}^{HML} + \gamma_t^{SMB}\hat{\beta}_{i,t}^{SMB} + \gamma_t^{MOM}\hat{\beta}_{i,t}^{MOM} + \gamma_t^{SCOPE12}\hat{\beta}_{i,t}^{SCOPE12} + \varepsilon_{i,t}, \quad \text{where } i = 1, \ldots, N. \tag{14}$$

The hat on the betas denotes that the estimates from the first-pass regression are used and the subscript t indicates the time-variability of the beta coefficients. The regression coefficients, γ_t, are interpreted as the risk premiums per month for each of the five factor candidates. Due to the rolling regression approach, the size of the cross-section varies over time. On average, I have beta coefficients of 1,146 stocks in the cross-sectional regressions. The risk premium for each factor is then calculated as the mean over all cross-sectional regression estimates as

$$\overline{\gamma} = \frac{1}{T}\sum_{t=1}^{T}\hat{\gamma}_t. \tag{15}$$

The Fama-MacBeth procedure then determines the variance

[17] See also Pukthuanthong et al. (2019) for a rationale for the usage of individual assets as left-hand-side variables.

of this estimate as

$$\sigma^2_{\overline{\gamma}} = \frac{1}{T}\sum_{t=1}^{T}\frac{(\hat{\gamma}_t - \overline{\gamma})^2}{T-1}. \quad (16)$$

On the basis of the estimates from equations 15 and 16, I compute the t-statistics according to the usual formula to determine the risk premiums' statistical significance. However, as the betas from the first-pass regression are estimates rather than exact values, the Fama-MacBeth approach is prone to an errors-in-variables (EIV) problem. In line with Pukthuanthong et al. (2019), I therefore repeat the cross-sectional regression with double-sorted portfolio betas as explanatory variables to alleviate the EIV problem. Annually at the end of June, I sort stocks into decile portfolios based on market value. Then, within each of the deciles, I conduct a second sort based on beta percentiles. Pukthuanthong et al. (2019) suggest the usage of deciles in the second step but I additionally report the result for betas sorted into quintiles to account for the smaller size of my sample. Note that this second sort is repeated for each factor individually. Within each of the 10×10 or 10×5 portfolios, I assign the arithmetic average beta to the stocks in the respective portfolio. The average betas are updated each month based on my rolling 36-month estimates to account for retired stocks.

The results from the computations outlined above are reported in table 14. The risk premiums are relatively stable in sign and magnitude across all three regressions. However, the explanatory power is reduced through the EIV correction as the adjusted R^2 drops from 19.14% in the first column to about 4% in the regressions with the double-sorted betas. Across all three models only the intercept is highly significant. In contrast, the other factors lack significant risk premiums, except for SCOPE12 in the EIV corrected models. SCOPE12's monthly risk premium of 0.155% in the 10×5 sorted model is weakly significant, whereas the risk premium of 0.195% in the 10×10 sorted model is highly significant. Overall, the results from the EIV corrected regressions point to the existence of a positive risk premium for holding low carbon intensity firms in my sample. Nevertheless, the results should be interpreted cautiously as the inevitable EIV problem constitutes a serious violation of the OLS assumptions and alternative approaches to treat this problem are discussed in the literature (e.g., Jegadeesh et al., 2019).

7.3. Hedge Portfolio Returns

The previous two sections of this chapter examined one necessary and one sufficient condition of factor testing. However, Pukthuanthong et al. (2019) suggest an additional test to determine whether the risk-reward trade-off is within reasonable limits for the respective risk factor candidate. This test rules out the case that the SCOPE12 factor reflects systematic empirical regularities which are in fact driven by market inefficiency or behavioral biases (Campbell et al., 1997). If deviations from the CAPM or alternative multifactor models are driven by a genuine risk factor, they should be bounded by the factor's relation to variance (MacKinlay, 1995). For non-risk explanations no such relation exists and the Sharpe ratio can be unbounded in theory. Based on theoretical and empirical arguments, MacKinlay (1995) proposes an annual Sharpe ratio of 0.6 as upper bound for a risk factor. Although this threshold is motivated by the historical excess return and standard deviation of the US stock market, I use it as a reasonable proxy for my European sample. If a risk factor delivers a Sharpe ratio significantly exceeding MacKinlay's threshold, this would provide an indication that the underlying factor might not be consistent with a risk-based pricing explanation (Pukthuanthong et al., 2019).

In contrast to the previous section, this robustness check is performed based on hedge portfolios. I construct an equally-weighted hedge portfolio for each of my five factor candidates by going long (short) in the stock quintile with the highest (lowest) factor beta. The betas used for monthly sorting are estimated as in the first-pass time series regression in section 7.2. Panel A in table 15 reports the mean return of the five quintile hedge portfolios. With 0.557% per month the SCOPE12 hedge portfolio offers the largest return, and – except for HML – all hedge portfolios generate positive returns. For none of the portfolios, however, returns are statistically different from zero.

The threshold proposed by MacKinlay (1995) refers to a long-only portfolio. Therefore, the zero-investment hedge portfolios must be combined with a representative long-only portfolio to test whether they exceed the 0.6 bound (Pukthuanthong et al., 2019). For this purpose, the market excess return is added to the hedge portfolio returns, which is referred to as *combined returns* in the following. The resulting mean and volatility of the combined returns are reported in panel B of table 15. Combining the momentum and SCOPE12 hedge portfolio returns with the market excess return offers significant returns at the 5% and 10% level, respectively. The last column reports the statistics of the market excess return for comparison. Based on the combined returns the annual Sharpe ratio is calculated for each factor, which ranges from 0.21 (HML) to 0.68 (MOM). The t-statistic in the last row of panel B indicates whether the Sharpe ratio is statistically different from 0.6. Sharpe ratios for MKT, HML, and SMB are significantly smaller than the threshold, whereas MOM's and SCOPE12's are not statistically different from it. This indicates that the traded versions of the factor candidates yield Sharpe ratios consistent with risk-based explanations. I therefore proceed by retaining all five factor candidates.

8. Carbon Risk Management with Exposure Portfolios

8.1. Construction of Industry Adjusted Exposure Portfolios

Based on the finding that the SCOPE12 factor can be considered a genuine equity risk factor, I address two key issues socially responsible investors face: First, investors are aware of potential negative financial implications carbon risks pose to their portfolio and actively want to manage those instead of withdrawing from problematic stocks (Krueger, Sautner,

Table 14: Estimated risk-premiums for factor candidates from 07/2010 to 06/2020.

	No EIV Correction	EIV Correction	
		10×5	10×10
MKT	-0.070	-0.286	-0.262
	(-0.147)	(-0.913)	(-0.819)
HML	-0.248	-0.174	-0.195
	(-1.050)	(-1.279)	(-1.358)
SMB	0.066	0.112	0.119
	(0.408)	(1.387)	(1.433)
MOM	0.384	0.243	0.251
	(1.551)	(1.344)	(1.430)
SCOPE12	0.222	0.155*	0.195**
	(1.516)	(1.814)	(2.240)
Intercept	0.910***	1.093***	1.070***
	(5.341)	(3.798)	(3.830)
Adj. R^2	19.14	3.52	4.15

Note. *p<0.1, **p<0.05, ***p<0.01; t-statistics in parentheses; heteroskedasticity robust standard errors are used; risk premiums are expressed as percentages per month.

Table 15: Risk and return of beta-sorted hedge portfolios from 07/2010 to 06/2020.

Panel A. Quintile hedge portfolios.

	MKT	HML	SMB	MOM	SCOPE12
Mean	0.134	-0.130	0.115	0.493	0.557
σ	5.413	5.098	4.663	4.260	4.574
t (mean)	0.27	-0.28	0.27	1.27	1.33

Panel B. Quintile hedge portfolios combined with market excess return.

	Combined portfolio returns					
	MKT	HML	SMB	MOM	SCOPE12	$r_{MKT} - r_f$
Mean	0.747	0.483	0.728	1.106**	1.170*	0.613
σ	9.825	7.902	7.688	5.671	6.723	4.866
t (mean)	0.83	0.67	1.04	2.14	1.91	1.38
SR	0.26***	0.21***	0.33***	0.68	0.60	0.44*
t (SR)	-3.69	-4.25	-2.98	0.83	0.03	-1.79

Note. *p<0.1, **p<0.05, ***p<0.01; mean is reported as monthly percentage; the Sharpe ratio (SR) is annualized; the last row in panel B reports the t-statistic of the Sharpe ratio against the MacKinlay (1995) threshold of 0.6.

& Starks, 2020). Hence, a main goal for risk management must be to help investors identify stocks exposed to carbon risk while maintaining a broad market exposure. Secondly, critics correctly point out that carbon emission information is mainly available for large firms because emission disclosure is still limited. This restricts investors who are aware of carbon risks in their investable universe. To mitigate this problem and to allow investors to identify carbon risks, widely available stock returns can be used. Based on the estimated exposure to the SCOPE12 factor, which equals the regression beta, both firms with and without CO_2 emission information could be considered in the management of carbon risk. For comparability with my previous analyses, I consider quintile exposure portfolios in the following.[18]

For the estimation of return-based exposures I run single stock, 36-month rolling window regressions on the 5-factor model (equation 10) starting in July 2007. I proceed by sorting stocks into industry adjusted and value-weighted quintile portfolios based on beta. For instance, the clean portfolio

[18]Note that by construction a positive (negative) SCOPE12 beta corresponds to less (more) carbon intensive stocks. To avoid confusion and ensure consistency with previous analyses, I therefore classify stocks in the highest (lowest) exposure quintile as clean (dirty).

Table 16: Average industry composition of exposure portfolios.

	Quintile						
Industry	Clean	2	3	4	Dirty	Total	%
Basic Materials	19	19	19	19	18	93	8.1
Consumer Discretionary	45	45	45	45	44	223	19.5
Consumer Staples	18	18	18	18	17	88	7.7
Energy	13	13	13	13	13	64	5.6
Health Care	20	19	19	19	19	96	8.4
Industrials	63	63	63	63	62	313	27.3
Real Estate	17	17	17	17	16	84	7.3
Technology	18	17	17	17	17	85	7.4
Telecommunication	10	10	10	9	9	47	4.1
Utilities	11	11	11	11	10	53	4.6
Total	233	229	229	229	226	1,146	100

Note. Values are rounded.

contains the 20% of stocks with the largest SCOPE12 beta in each industry. As in previous analyses, the portfolios are rebalanced annually at the end of June. The chosen approach offers diversification across industries whilst maintaining an equal industry composition of the quintile portfolios. It can therefore be interpreted as a variant of the best-in-class approach widely employed by ESG funds. In contrast, sorting portfolios merely based on exposure to the SCOPE12 factor would result in an unequal weighting of industries and contradict the investors' wish for a broad market exposure. Analogous to the first-pass regression in section 7.2, I have on average beta estimates of 1,146 stocks in my sample. 414 of these 1,146 companies lack explicit carbon intensity information provided by Datastream for the respective year, which equals more than one-third of the overall sample size. Hence, the investable universe is considerably extended through the consideration of SCOPE12 exposures in comparison to relying on reported carbon emissions alone. Table 16 shows the average industry compositions of the quintile portfolios. The percentage distribution across industries reported in the last column is similar to the complete sample presented in table 1. Consequently, the quintile portfolios offer a similar industry exposure as holding the market. As above, consumer discretionary and industrials constitute the largest industries and jointly account for nearly half of the sample.

To check for the resemblance of the exposure-sorted portfolios and the previously considered carbon intensity-sorted portfolios, I compare the inverse percentile ranks of the estimated exposures[19] with the actual carbon intensities. I find a significant positive Spearman rank correlation of approximately 0.3. This is comparable to the results of Hübel and Scholz (2020) who follow a similar approach for their environmental, social, and governance factor. Differences between the carbon intensities and exposures might arise from the additional industry sort, the change of a firm's carbon risk during the three year estimation period, and carbon risks potentially captured by beta but not by the reported emissions. For instance, sensitivities to SCOPE12 might reflect that certain firms got bad press for environmental reasons and therefore behave like high carbon stocks, which is not captured by the reported carbon intensities. This aspect also suggests that the usage of exposures could be more robust to greenwashing and strategic reporting of environmental risks (Hübel & Scholz, 2020).

8.2. Performance of Exposure Portfolios

Financial and risk considerations rank highly among investors' core motives to incorporate carbon risks in their investment process (Krueger et al., 2020). The risk-return profiles of the industry adjusted exposure portfolios for all stocks are reported in panel A of table 17. Panel B shows the exposure-sorted portfolios consisting of the above mentioned stocks without explicit emission information on Datastream. However, due to the smaller size of the unrated subsample these portfolios are *not* industry adjusted.

Both panels show a similar risk-return profile as the portfolios based on emission intensities in table 4. In both panels the excess return decreases monotonically from the cleanest to the dirtiest quintile. Additionally, I find that the extreme portfolios have the highest standard deviations. As investors are particularly concerned to reduce both portfolio and tail risks associated with carbon risk (Krueger et al., 2020), looking at exposures can identify particularly volatile and hence undesirable stocks. In both panels, the higher volatility in the clean portfolio results in a marginally smaller Sharpe ratio compared to the second portfolio. Although unrated stocks in panel B offer higher average excess returns in all quintiles, they are also more volatile. This results in Sharpe ratios of similar magnitude in both panels. To illustrate the performance of the portfolios over time, appendix E contains a plot of the cumulative returns of the portfolios from panel A and B in figure 3 and 4, respectively.

[19] As noted, high carbon stocks have low SCOPE12 betas.

Table 17: Risk and return of industry- and SCOPE12 exposure-sorted quintile portfolios from 07/2010 to 06/2020.

Panel A. Performance of quintile portfolios of all stocks.

	Clean	2	3	4	Dirty
Excess Return	0.85*	0.80*	0.67*	0.60	0.49
Standard deviation	5.17	4.53	4.44	4.72	5.52
t-statistic	1.81	1.92	1.65	1.39	0.98
Sharpe ratio	0.57	0.61	0.52	0.44	0.31

Panel B. Performance of quintile portfolios of unrated stocks.

	Clean	2	3	4	Dirty
Excess Return	1.01**	0.97**	0.84*	0.83*	0.55
Standard deviation	5.31	4.92	4.68	5.08	5.15
t-statistic	2.08	2.15	1.96	1.78	1.18
Sharpe ratio	0.66	0.68	0.62	0.56	0.37

Note. *p<0.1; **p<0.05; ***p<0.01; excess return is reported as monthly percentage; the Sharpe ratio is annualized.

Table 18: 4-factor and 5-factor model for industry- and SCOPE12 exposure-sorted quintile portfolios.

Panel A. 4-factor model.

Quintile	MKT	HML	SMB	MOM	SCOPE12	Alpha	Adj. R^2
Clean	1.093***	-0.217***	0.028	0.038		0.061	95.64
	(39.82)	(-4.27)	(0.55)	(0.92)		(0.58)	
2	0.970***	-0.242***	-0.233***	0.027		0.111	96.48
	(53.72)	(-6.08)	(-4.65)	(0.87)		(1.37)	
3	0.937***	-0.175***	-0.167**	0.048		0.003	94.80
	(31.39)	(-3.23)	(-2.60)	(0.95)		(0.03)	
4	0.996***	-0.195***	-0.129**	0.017		-0.086	95.95
	(51.39)	(-4.45)	(-2.34)	(0.42)		(-1.02)	
Dirty	1.103***	-0.047	0.188**	0.029		-0.252*	91.90
	(31.41)	(-0.72)	(2.28)	(0.44)		(-1.74)	

Panel B. 5-factor model.

Quintile	MKT	HML	SMB	MOM	SCOPE12	Alpha	Adj. R^2	ΔR^2
Clean	1.097***	-0.180***	0.019	0.035	0.152**	0.004	95.79	0.15**
	(40.86)	(-3.41)	(0.39)	(0.88)	(2.45)	(0.04)		
2	0.970***	-0.241***	-0.233***	0.027	0.005	0.109	96.45	-0.03
	(53.97)	(-5.66)	(-4.67)	(0.87)	(0.11)	(1.36)		
3	0.933***	-0.208***	-0.159**	0.051	-0.139**	0.055	94.96	0.16**
	(31.39)	(-3.58)	(-2.50)	(1.04)	(-2.55)	(0.50)		
4	0.989***	-0.263***	-0.112**	0.022	-0.283***	0.020	96.70	0.75***
	(53.33)	(-5.62)	(-2.54)	(0.68)	(-4.04)	(0.25)		
Dirty	1.086***	-0.198***	0.225***	0.041	-0.634***	-0.015	94.72	2.82***
	(38.32)	(-3.13)	(3.32)	(0.73)	(-9.00)	-0.12		

Note. *p<0.1, **p<0.05, ***p<0.01; t-statistics in parentheses; heteroskedasticity robust standard errors used. The last column reports the significance of heteroskedasticity consistent F-test for nested models.

Finally, I check whether the exposure-sorted portfolios generate abnormal returns and whether the SCOPE12 factor adds explanatory power. For this purpose, I regress the value-weighted and industry adjusted portfolio returns from panel A in table 17 on the 4-factor model and 5-factor model, respectively. Again, I report the results of the heteroskedasticity consistent F-test for nested models in the last column. The magnitude and the significance of the regression coefficients in table 18 is similar to the carbon intensity-sorted quintile portfolios in table 7. The results of the 4-factor re-

gression in panel A indicate that all portfolios have market betas close to one and a tilt towards growth stocks. The latter, however, is not significant for the dirty quintile. While exposure to the size factor is mixed across the quintiles, none of the portfolios loads significantly on the MOM factor. With regard to abnormal returns, only for the dirty portfolio a weakly significant monthly alpha of -0.25% is found. This suggests that the portfolio of high carbon stocks underperforms after adjusting for common risk factors. One explanation for this finding could be that stocks in this quintile are overvalued relative to low carbon stocks and therefore offer lower returns. For example, investors perceive carbon intensive sectors such as the oil, the automotive, and the electric utilities industry to be overvalued considering climate change (Krueger et al., 2020). Although the presented quintiles have the same industry composition, the observed negative alpha in the dirty quintile might reflect the investors' perception of incorrect carbon risk pricing. Panel B shows that the SCOPE12 factor captures this abnormal return and leads to smaller alphas in most portfolios. In comparison to panel A, the other four regression coefficients remain largely unchanged. As expected, the additional SCOPE12 factor loadings decrease monotonically from the clean to the dirty portfolio. Comparable to table 7, the additional explanatory power is higher and more statistically significant for the more dirty portfolios. The change in adjusted R^2 is only non-significant for the second quintile.

Overall, the results indicate that the SCOPE12 factor adds explanatory power after adjusting the portfolios for a broad industry exposure. My findings also suggest that the SCOPE12 factor indeed captures the fraction of systematic risk attributable to priced carbon risk and therefore should be considered by investors to avoid negative financial outcomes. Lastly, using exposures rather than actual emission information seems to yield comparable risk-return results while extending the investment universe. These aspects emphasize the potential usefulness of the SCOPE12 factor for investors' carbon risk management.

9. Discussion

9.1. SCOPE12 as Genuine Risk Factor

The guiding research question of this thesis is to determine whether carbon risk is a systematic equity risk factor in European stocks. Previous research primarily uses time series regressions and the additional explanatory power of newly introduced factors as criterion for their relevance (e.g., Hübel & Scholz, 2020). However, publishing factors which merely add R^2 or reduce alpha relative to a benchmark model has proven to be too little of a hurdle in the factor model literature and resulted in "a zoo of new factors" (Cochrane, 2011, p. 1047). These supposed improvements can have several causes (e.g., data-snooping) and do not necessarily reflect progress in the identification of superior factor models (Campbell et al., 1997). Hence, recently new methodologies were proposed to identify systematic risk factors which should be considered in factor models (e.g., Fama & French, 2018; Klein & Chow, 2013; Pukthuanthong et al., 2019). Based on the combination of various methodologies for testing the carbon factor, SCOPE12 can be identified as a systematic risk factor.

First, the carbon factor is shown to be only weakly related to other accepted factors and explains returns when tested with both portfolios and individual stocks. The magnitude of the regression coefficients and additional R^2 is consistent with related research by Hübel and Scholz (2020) and Görgen et al. (2020). Similar to Görgen et al. (2020), I find that especially the variance in more carbon intensive portfolios – irrespective of whether actual ratings or carbon factor exposures are considered – is well captured by adding the SCOPE12 factor. This observation becomes particularly striking when factors are democratically orthogonalized. The unique R^2 from adding SCOPE12 is about 25 times higher for the most carbon intensive quintile than the clean one. Overall, this points to the existence of a large common variation of highly carbon intensive stocks due to the low carbon transition of the European economy.

The necessary condition for SCOPE12 to be considered as a risk factor is shown through the significance of the relation between SCOPE12's returns and the principal components of the sample covariance matrix (Pukthuanthong et al., 2019). This result is reliable because the ten extracted PCs in my analysis explain approximately 90% of the variance in stock returns and therefore provide a good proxy. Then I test for factor risk premiums using the methodology of Fama and MacBeth (1973) with single stock returns as dependent variable. There is evidence for the existence of a positive risk premium related to the carbon factor in the EIV corrected cross-sectional regressions. That means, investors are compensated for holding cleaner stocks. Although the chosen methodology is consistent with related literature, the result that only SCOPE12 offers a risk premium should be treated with certain caution. The treatment of EIV issues is complex and the t-statistics are shown to be sensitive to the beta sort chosen. Furthermore, alternative EIV correction approaches for cross-sectional regressions based on individual stocks and portfolios as dependent variable exist (e.g., Fama & MacBeth, 1973; Jegadeesh et al., 2019). Those methods, however, come with their own restrictions and tend to have data requirements exceeding the natural limitations of the presented sample. Apart from that, the observed weak significance of other regression coefficients than the intercepts is consistent with related papers (e.g., Görgen et al., 2020; Pukthuanthong et al., 2019). In contrast to my results, however, Görgen et al. (2020) do not find statistically significant evidence that carbon risk is priced in the cross-section of returns. A potential cause for this difference might be their consideration of additional return predictors as control variables. Lastly, SCOPE12 offers a reward-to-risk ratio that appears to be consistent with risk pricing limits suggested by the finance literature (e.g., MacKinlay, 1995).

As a genuine risk factor, SCOPE12 also becomes interesting for investors who seek to manage the carbon risk of their

holdings even in the absence of explicit emission information. The factor construction through double sorting based on size and carbon intensities allows to extract the unique portion of return caused by emission differences irrespective of potential size effects. Considering that stocks without ESG information tend to be significantly smaller on average (e.g., Hübel & Scholz, 2020), this feature appears particularly desirable for beta estimation. Measuring stocks' exposure to the SCOPE12 factor through time series regression therefore helps socially responsible investors in two ways. First, it allows them to considerably extend their universe of investable equities (i.e., small caps), and secondly, they can identify assets to be included in carbon risk management. With regard to risk, I find that the most extreme quintiles also have the highest standard deviations. This makes these portfolios also interesting from a risk management perspective as investors are particularly concerned about both portfolio and tail risks related to climate change (Krueger et al., 2020). Regarding the dirty exposure portfolio, I find weakly significant evidence that the Carhart 4-factor model cannot sufficiently explain its returns. The identified negative 4-factor alpha, however, can be attributed to the SCOPE12 factor and emphasizes the carbon factors relevance. Overall, my analyses suggest that investors can achieve similar results by using a return-based exposure measure instead of directly reported carbon intensities.

9.2. Outperformance of Low Carbon Stocks

The second research question addresses the direction and magnitude of the price effect of carbon risk and relates to a strand of literature finding mixed results. In the European sample less carbon intensive portfolios provide higher returns than more carbon intensive portfolios. This finding supports the clean alpha hypothesis mentioned in the introduction. Sorting stocks into both carbon intensity and SCOPE12 exposure portfolios, I find a monotonous decline of excess returns from the clean to the dirty portfolio in all my analyses. Especially the cleaner portfolios generate excess returns significantly different from zero in the last decade, whereas the carbon intensive portfolios do not. The observed relationship between excess returns and carbon risk on the portfolio level is also consistent with SCOPE12's positive risk premium found in the cross-sectional regressions.

Moreover, I generally do not find significant abnormal returns for the mentioned long-only quintile portfolios based on the Carhart 4-factor model. As stated above, only the dirty exposure portfolio has a weakly significant negative alpha relative to this model and provides some indication for a risk adjusted underperformance of high carbon stocks. In contrast, long-short strategies seem to generate returns which cannot be explained by the common risk factors. In the time from July 2007 to June 2020, the outperformance of cleaner assets results in a highly significant monthly average return of 0.32% on the SCOPE12 factor, which is constructed as the return on a size adjusted hedge portfolio going long (short) in less (more) carbon intensive firms. As shown in the spanning regressions, this return cannot be explained by the 4-factor model and generates a significant monthly alpha of 0.37% during the main sample period. This result is consistent with Pedersen et al. (2021) and In et al. (2019) who both find significant average excess returns on carbon intensity-sorted hedge portfolios for their US samples. Moreover, both studies report (weakly) significant positive alphas when regressing the hedge portfolio returns on standard factor models.

While the average returns are higher on less carbon intensive stocks during my sample period, the outperformance is not constant over time. Rather, the plot of the portfolios' market adjusted cumulative returns indicates that cleaner stocks started to outperform their higher emission peers in mid 2008. The performance differential between the clean and the dirty portfolio gradually expanded further during the main sample period. Interestingly, the onset of the outperformance of clean stocks corresponds closely to the period in which most of the spikes in the climate change news index by Engle et al. (2020) occur. If the observed outperformance is driven by unexpected changes in tastes for greener assets due to higher climate risk awareness, the observed pattern could be consistent with the ESG factor proposed by Pástor et al. (2021). That is, Pástor et al. (2021, p. 8) state that "if one computes average returns over a sample period when ESG concerns strengthen more than investors expected,(. . .) then green stocks outperform brown stocks, contrary to what is expected". Choi et al. (2020) find global evidence for such effects. They show that greener assets perform better when awareness for climate change is high during times of abnormally hot weather. Similarly, Pedersen et al. (2021) find higher institutional holdings and valuations for less carbon intensive assets based on their US sample. Lastly, institutional investors perceive certain carbon intensive sectors to be overvalued, as prices do not correctly reflect climate change risks yet. Vice versa, they see potentially cleaner sectors (e.g., battery or renewable energy producers) as undervalued (Krueger et al., 2020). A promising avenue for future research would therefore be to test whether the observed pattern during my sample period is due to a learning period in which the prices of stocks with less (more) carbon intensity are adjusted upwards (downwards). If the sample indeed reflects an adjustment period, one would expect a reversal in the performance pattern in a longer time series. Such effects would then be consistent with the dirty alpha hypothesis.

9.3. Limitations

Considering the complexity of capturing the extent of GHG emissions, current climate finance research is limited by the availability of accurate proxies. Although scope 1 and 2 emissions tend to be the most reliable and consistent measures of carbon emission across data providers, they only provide an incomplete picture of actual carbon emissions. Relevant additional information such as a firm's other indirect emissions along the value chain, scope 3 emissions, "are rarely reported by companies and are at best noisily estimated and inconsistent across different data providers" (Pedersen et al., 2021, pp. 12-13). Especially the option to

choose between different organizational boundaries for carbon accounting bears the potential that the true emissions caused by a company are misrepresented. Moreover, the current accounting standards do not properly reflect the business model of financial firms (Görgen et al., 2020). The decision to exclude financials based on this observation, nevertheless, poses a main limitation of my research as it restricts the generalizability of the presented results.

In addition to this, Campbell et al. (1997) point out the fundamental factor literature's proneness to data-snooping issues. Fundamental factors are usually constructed based on characteristics which are empirically found to be relevant in explaining the cross-section of returns. Hence, the relevance of a particular measure might be overstated and not indicative for future periods. Carbon intensity measures – as the one used in this thesis – are widely used in the climate finance literature and are frequently found to positively relate to accounting and market performance (e.g., Pedersen et al., 2021). Moreover, relative emission measures tend to be more likely to yield statistically significant results (Busch & Lewandowski, 2018). Although the presented factor has a sound theoretical motivation and therefore the likelihood of data dredging is reduced (Campbell et al., 1997), such objections usually can only be overcome with very long time series.

With the ongoing development of the carbon accounting standards and reporting best practices, repeating the presented analyses with more comprehensive measures of carbon emission and a longer time series appears promising. Doing so would equally address both key limitations of my thesis.

10. Conclusion

Large-scale reforms of the European financial system initiated by the EU and additional commitments by institutional investors shift the capital market's focus onto climate change considerations. Especially the aspect of climate change mitigation through carbon emission reduction has entered mainstream financial decision making. This heightened awareness in the past decade bears the potential to systematically influence the cross-section of asset prices as global warming affects the entire economy, not just individual industries. In this context, the question of carbon risk's impact on asset prices generated some interest in the nascent climate finance literature. However, the currently available mixed evidence is primarily restricted to US or global samples. This thesis contributes to the literature by quantifying the carbon risk in European equity returns, determining whether it constitutes a systematic risk factor, and highlighting the practical application of the carbon factor for risk management.

Using carbon intensity as a firm characteristic, I calculate the SCOPE12 carbon factor. The factor is constructed as a double-sorted hedge portfolio which is long in cleaner stocks and short in dirty stocks. It provides a significantly positive average return during the sample period and the onset of its positive performance corresponds closely with heightened awareness for climate change in the media. In combination with the observed monotonous negative relationship between return and both carbon intensity and SCOPE12 exposure across quintile portfolios, this provides evidence for the clean alpha hypothesis.

I also find that SCOPE12 can be considered as a genuine risk factor because it is related to the sample covariance matrix of returns, commands a (weakly) significant positive risk premium, and offers a risk-reward trade-off within reasonable limits. This finding is corroborated by complementary analyses which confirm that the carbon factor cannot be explained by alternative factor candidates. SCOPE12 therefore explains a unique portion of systematic variation in European stock returns.

Lastly, current research emphasizes that investors want to manage carbon risks in their portfolio but lack appropriate tools and best practices. Especially the insufficient disclosure of GHG emissions poses a serious restriction for the identification of carbon risks. Using return-based SCOPE12 factor exposures instead addresses this issue and considerably expands the universe of stocks which can be considered for investment and risk management. With the gradually growing importance of climate change considerations in financial markets and institutional investors' awareness for its financial implications, strategically managing carbon risk with the SCOPE12 factor could therefore bear positive effects.

References

Aiken, J. G., Erdos, J. A., & Goldstein, J. A. (1980). On Löwdin Orthogonalization. *International Journal of Quantum Chemistry*, *18*, 1101–1108.

Alexander, C. (2001). *Market Models: A Guide To Financial Data Analysis*. Hoboken, New Jersey: Wiley.

Blitz, D., Hanauer, M. X., Vidojevic, M., & van Vliet, P. (2018). Five Concerns with the Five-Factor Model. *Journal of Portfolio Management*, *44*(4), 71–78.

Bodie, Z., Kane, A., & Marcus, A. J. (2018). *Investments* (11th ed.). New York: McGraw-Hill Education.

Bolton, P., & Kacperczyk, M. (2020). Do Investors Care about Carbon Risk? *Journal of Financial Economics (forthcoming)*.

Brohé, A. (2017). *The Handbook of Carbon Accounting*. New York: Routledge.

Busch, T. (2011). Which emissions do we need to account for in corporate carbon performance? Response to Murray and colleagues. *Journal of Industrial Ecology*, *15*(1), 160–163.

Busch, T., Johnson, M., Pioch, T., & Kopp, M. (2018). Consistency of Corporate Carbon Emission Data. *University of Hamburg / WWF Deutschland, Hamburg*.

Busch, T., & Lewandowski, S. (2018). Corporate Carbon and Financial Performance: A Meta-analysis. *Journal of Industrial Ecology*, *22*(4), 745–759.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton, New Jersey: Princeton University Press.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, *52*(1), 57–82.

Choi, D., Gao, Z., & Jiang, W. (2020). Attention to Global Warming. *The Review of Financial Studies*, *33*(3), 1112–1144.

Cochrane, J. H. (2005). *Asset Pricing* (2nd ed.). Princeton, New Jersey: Princeton University Press.

Cochrane, J. H. (2011). Presidential Address: Discount Rates. *The Journal of Finance*, *66*(4), 1047-1108.

Connor, G., & Korajczyk, R. A. (1986). Performance Measurement with the Arbitrage Pricing Theory: A New Framework For Analysis. *Journal of Financial Economics*, *15*, 373–394.

Connor, G., & Korajczyk, R. A. (1988). Risk and Return in an equilibrium APT: Application of a New Test Methodology. *Journal of Financial Economics*, *21*, 255–289.

Deutsche Börse Group. (2021). *STOXX® Index Methodology Guide (Portfolio Based Indices)*. Retrieved from https://www.stoxx.com/document/Indices/Common/Indexguide/stoxx_index_guide.pdf

Engle, R. F., Giglio, S., Kelly, B., Lee, H., & Stroebel, J. (2020). Hedging Climate Change News. *The Review of Financial Studies*, *33*(3), 1184–1216.

European Commission. (2021). *Overview of sustainable finance.* Retrieved from https://ec.europa.eu/info/business-economy-euro/banking-and-finance/sustainable-finance/overview-sustainable-finance_en

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*, 3–56.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*, 1–22.

Fama, E. F., & French, K. R. (2018). Choosing factors. *Journal of Financial Economics*, *128*, 234–252.

Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, *81*(3), 607–636.

French, K. R. (2020). *Current Research Returns.* Retrieved from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#International

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A Test of the Efficiency of a Given Portfolio. *Econometrica*, *57*(5), 1121–1152.

Görgen, M., Jacob, A., Nerlinger, M., Riordan, R., Rohleder, M., & Wilkens, M. (2020). Carbon Risk. *Working Paper Augsburg University*.

Hübel, B., & Scholz, H. (2020). Integrating sustainability risks in asset management: the role of ESG exposures and ESG ratings. *Journal of Asset Management*, *21*, 52–69.

Heinkel, R., Kraus, A., & Zechner, J. (2001). The effect of green investment on corporate behavior. *Journal of Financial and Quantitative Analysis*, *36*(4), 431–449.

IIGCC. (2021). *About us.* Retrieved from https://www.iigcc.org/about-us/

Ilhan, E., Sautner, Z., & Vilkov, G. (2020). Carbon Tail Risk. *The Review of Financial Studies (forthcoming)*.

In, S. Y., Park, K. Y., & Monk, A. (2019). Is 'Being Green' Rewarded in the Market?: An Empirical Investigation of Decarbonization and Stock Returns. *Stanford Global Project Center Working Paper*.

Ince, O. S., & Porter, R. B. (2006). Individual Equity Return Data from Thomson Datastream: Handle With Care! *The Journal of Financial Research*, *29*(4), 463–479.

IPCC. (2018). *Intergovernmental Panel on Climate Change. Special report: Global warming of 1.5°C - Summary for Policymakers.* Retrieved from https://report.ipcc.ch/sr15/pdf/sr15_spm_final.pdf

Jegadeesh, N., Noh, J., Pukthuanthong, K., Roll, R., & Wang, J. (2019). Empirical tests of asset pricing models with individual assets: Resolving the errors-in-variables bias in risk premium estimation. *Journal of Financial Economics*, *133*(2), 273 – 298.

Kaiser, L. (2020). ESG Integration: value, growth and momentum. *Journal of Asset Management*, *21*, 32–51.

Kennedy, P. (2008). *A Guide to Econometrics* (6th ed.). Malden, MA: Wiley-Blackwell.

Klein, R. F., & Chow, V. K. (2013). Orthogonalized factors and systematic risk decomposition. *The Quarterly Review of Economics and Finance*, *53*, 175–187.

Krueger, P., Sautner, Z., & Starks, L. T. (2020). The Importance of Climate Risks for Institutional Investors. *The Review of Financial Studies*, *33*(3), 1067–1111.

MacKinlay, A. C. (1995). Multifactor models do not explain deviations from the CAPM. *Journal of Financial Economics*, *38*(1), 3 – 28.

Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics (forthcoming)*.

Pástor, L., Stambaugh, R. F., & Taylor, L. A. (2021). Sustainable Investing in Equilibrium. *Journal of Financial Economics (forthcoming)*.

Pukthuanthong, K., Roll, R., & Subrahmanyam, A. (2019). A Protocol for Factor Identification. *The Review of Financial Studies*, *32*(4), 1573–1607.

Schmidt, P. S., von Arx, U., Schrimpf, A., Wagner, A. F., & Ziegler, A. (2019). Common risk factors in international stock markets. *Financial Markets and Portfolio Management*, *33*(3), 213–241.

Stechemesser, K., & Guenther, E. (2012). Carbon accounting: a systematic literature review. *Journal of Cleaner Production*, *36*, 17–36.

Tsay, R. S. (2010). *Analysis of Financial Time Series* (3rd ed.). Hoboken, New Jersey: Wiley.

Tsay, R. S. (2014). *Multivariate Time Series Analysis - With R and Financial Applications*. Hoboken, New Jersey: Wiley.

UNFCCC. (1992). United Nations Framework Convention on Climate Change. New York.

UNFCCC. (2008). *Kyoto Protocol Reference Manual - On Accounting of Emissions and Assigned Amount.* Retrieved from https://unfccc.int/resource/docs/publications/08_unfccc_kp_ref_manual.pdf

United States Environmental Protection Agency. (2021). *Understanding Global Warming Potentials.* Retrieved from https://www.epa.gov/ghgemissions/understanding-global-warming-potentials

WBCSD, & WRI. (2004). *The Greenhouse Gas Protocol - A Corporate Accounting and Reporting Standard (Revised Edition).* Retrieved from https://ghgprotocol.org/sites/default/files/standards/ghg-protocol-revised.pdf

WBCSD, & WRI. (2013). *Required Greenhouse Gases in Inventories - Accounting and Reporting Standard Amendment.* Retrieved from https://ghgprotocol.org/sites/default/files/standards_supporting/RequiredgasesandGWPvalues_1.pdf

Wooldridge, J. M. (2016). *Introductory Econometrics - A Modern Approach* (6th ed.). Boston: Cengage Learning.

Zivot, E., & Wang, J. (2003). *Modeling Financial Time Series with S-PLUS*. New York: Springer.

Junior Management Science 7(2) (2022) 455-503

Junior Management Science

journal homepage: www.jums.academy

Conceptualising the Value-in-Use of Services for Researchers

Konzeptionierung des Value-in-Use von Dienstleistungen für Forschende

Theresa Kroschewski

Technische Universität Braunschweig

Abstract

The research area of value-centred and customer-centred service design transforms customers' needs in marketable offers. This field is dominated by a focus on services that are individually oriented towards people and their values and gives providers a competition advantage. Even outside the market, in the practical field of research, there has been an increasing demand for individual and collaborative support for several years. However, it has been barely investigated so far, how offers for research-related services can benefit from market approaches. This paper is dedicated to addressing this issue and draws on the "value-in-use" as a basis for design described in service-dominant logic. Therefore, a theoretical model is used that transfers goal dimensions and so-called activated work resources as conceptualisation levels to the professional context of researchers. In an empirical study these levels are specified with the subjective perception of value from 14 service experiences. The resulting findings build a deep understanding of the various levels of possible value and how this value can be created for researchers when using services.

Zusammenfassung

Im Forschungsbereich der wert- und kundenzentrierten Dienstleistungsgestaltung werden Bedürfnisse von Kund:innen in marktfähige Angebote überführt. Dabei trifft die Fokussierung auf individuell am Menschen und dessen Werte ausgerichtete Dienstleistungen auf hohe Nachfrage und schafft Anbietenden einen Vorteil im unternehmerischen Wettbewerb. Auch außerhalb des Marktes, im Praxisgebiet der Forschung, zeichnet sich seit einigen Jahren ein steigender Bedarf an individueller und kollaborativer Unterstützung ab. Wie Angebote für forschungsnahe Dienstleistungen von marktlichen Ansätzen profitieren können, ist bislang wenig untersucht. Die vorliegende Arbeit widmet sich ebendieser Betrachtung und zieht den in der Service-dominierten Logik beschriebenen „Value-in-Use" als Gestaltungsgrundlage heran. Über ein theoretisches Modell, das Zieldimensionen und aktivierte Arbeitsressourcen als Konzeptionierungsebenen in den beruflichen Kontext von Forschenden überführt, wird in einer empirischen Untersuchung die subjektive Wertwahrnehmung aus 14 Dienstleistungserfahrungen untersucht. Die hierbei entstandenen Ergebnisse bauen ein tiefgreifendes Verständnis dafür auf, wie und auf welchen Ebenen ein individueller Wert für Forschende bei der Inanspruchnahme von Dienstleistungen geschaffen werden kann.

Keywords: Value-in-Use; forschungsnahe Dienstleistungen; Service-dominierte Logik.

1. Einleitung

Die Rahmenbedingungen der Arbeit von Forschenden haben sich maßgeblich verändert. Individuell geförderte, wissenschaftliche Arbeit innerhalb der eigenen Disziplin hat an Bedeutung verloren (Böhm, 2006; Defila, Di Giulio & Scheuermann, 2006; Laudel, 2001). Stattdessen fordert die zunehmende Komplexität heutiger Problemstellungen fachübergreifende Forschungskooperationen, um ein integriertes Systemverständnis zu erreichen. Derartige Kooperationen werden vornehmlich über zeitlich begrenzte Forschungsprojekte in Form von Verbünden aus disziplinübergreifenden Einzelprojekten realisiert (DFG, 2018; Vowe & Meißner, 2020). Diese zunehmend kollaborative und projektgebundene Organisation der Forschungstätigkeiten stellt neue Anforderungen an den Arbeitsalltag von Forschenden – unter anderem durch erhöhte Akquise-, Integrations- oder Ab-

DOI: https://doi.org/10.5282/jums/v7i2pp455-503

stimmungsaufwände (Brandstäfter & Sonntag, 2016; König, Diehl, Tscherning & Helming, 2013; Locker-Grütjen, Ehmann & Jongmanns, 2012).

1.1. Motivation

Aufgrund dieser neuen Anforderungen an die Forschungsarbeit wurden über die vergangenen Jahre Möglichkeiten zur Unterstützung von Forschenden diskutiert, die zur Verbesserung der Forschungsleistung beitragen sollen (Auer & Herlitschka, 2008; Locker-Grütjen, 2011).

In diesem Zusammenhang haben sich zahlreiche Entwicklungen innerhalb der Disziplin des *Forschungsmanagements* ergeben (Boden, 2016a, 2016c; Locker-Grütjen et al., 2012). Die Aktivitäten des Managements beziehen sich nicht auf den expliziten Forschungsprozess, sondern auf „die Organisation aller die Forschung begleitender und unterstützender Prozesse“ (Hendrichs, 2017, S. 274). Hierbei sollen Dienstleistungen bereitgestellt werden, welche die Forschenden in ihren Tätigkeiten unterstützen und von administrativen Aufgaben entlasten (Hendrichs, 2017). Unter dem Begriff des Forschungsmanagements werden insbesondere die *institutionellen* Möglichkeiten zur Unterstützung von Forschenden untersucht (Locker-Grütjen et al., 2012). Neben einer technologiebasierten Unterstützung (Auth, Czarnecki, Bensberg & Thor, 2019) werden auch personalintensive Angebote über eine neue Profession innerhalb des Hochschulpersonals realisiert. Sogenannte *New Professionals* (Heuer, 2017) oder auch *HoPros* (Schneijderberg, Merkator, Teichler & Kehm, 2013) verstehen sich selbst als Dienstleister*innen für Forschende und bieten im Zuge dessen diverse administrative und koordinative Managementtätigkeiten an. Doch die Möglichkeiten der Unterstützung innerhalb der zentralen Strukturen, wie etwa der Hochschule, sind begrenzt.

Inzwischen ist ein Großteil der wissenschaftlichen Tätigkeiten innerhalb komplexer, zeitlich begrenzter Forschungsprojekte eingebettet (Aljets & Lettkemann, 2012; DFG, 2018; Vowe & Meißner, 2020), welche dezentrale Management- und Koordinationsstrukturen für die Dauer des Vorhabens projektintern aufbauen (Lee, 2015). Die Ausgestaltung dieser Strukturen wurde bereits als wichtige Determinante des Projekterfolgs kollaborativer Forschung anerkannt (König et al., 2013) und zahlreich diskutiert (Defila, Di Giulio & Scheuermann, 2008; König et al., 2013; Lippe & vom Brocke, 2016; Schneider, Buser, Keller, Tribaldos & Rist, 2019). Die Diskussion begrenzt sich bislang allerdings auf die übergeordnete und organisationstheoretische Perspektive von Forschungsprojekten, aus welcher Anforderungen an die Tätigkeiten des Managements und der Koordination ableitet werden. Hier wird zwar auch die Relevanz der einzelnen Forschenden anerkannt (Defila et al., 2008), es wird jedoch kein Rahmen geschaffen, der die *individuellen* Bedürfnisse der Forschenden bei der Anspruchsformulierung an das Management berücksichtigt.

1.2. Problemstellung und Zielsetzung

Wie die einzelnen Forschenden bei der Gestaltung von Management- und Koordinationsangeboten berücksichtigt werden könnten, wird von Lux und Robra-Bissantz (2020) vorgeschlagen. Demnach sollen derartige Tätigkeiten in Forschungsverbünden als Dienstleistungen für den Forschenden verstanden werden. Mithilfe der sogenannten *service-dominierte Logik* (SD-Logik) als theoretischem Rahmenwerk, setzten die Autorinnen die Forschenden und den für sie geschaffenen Wert durch die Dienstleistungen in den Fokus.

Die SD-Logik entstammt ursprünglich dem Marketing und unterstellt, dass der Wert einer Dienstleistung sich erst während der Inanspruchnahme ergibt und allein von den Kund*innen bestimmt werden kann. Im Zuge dessen wird der sogenannte *Value-in-Use* (ViU) definiert. Dieser bezeichnet das Wertergebnis, welches sich während der Nutzung eines Angebotes ergibt und subjektiv von den Kund*innen wahrgenommen wird (Vargo & Lusch, 2004). Eine solche wertzentrierte Sichtweise ermöglicht ein tieferes Verständnis dafür, welche Aspekte einer Dienstleistung seitens der Anspruchsgruppe als wertvoll empfunden werden und von welchen Faktoren diese Wertschätzung abhängt. Hieraus leitet sich die übergeordnete Fragestellung der vorliegenden Arbeit ab:

> *Wie kann der Value-in-Use von Forschenden abgebildet werden und so zur Gestaltung von Dienstleistungen für Forschende beitragen?*

Solche wertorientierten Einblicke sind bereits in diversen Disziplinen Teil der Forschungsarbeit zur Gestaltung von Dienstleistungen für den Markt. Davon motiviert betonen Lux und Robra-Bissantz (2020, S. 697) die Diskrepanz zwischen „innovativen Entwicklungen aus der Dienstleistungsforschung für den Markt und dem Stand der Dienstleistungsentwicklung im Praxisfeld der Wissenschaft“ selbst. Dieses Ungleichgewicht soll überwunden werden, indem die Erkenntnisse aus der SD-Logik auf die Entwicklung von Dienstleistungen für Forschende übertragen werden. Hierzu formuliert die vorliegende Arbeit einen konzeptionellen Rahmen, welcher den ViU von Forschenden abbildet und als Grundlage für die Gestaltung der Management- und Koordinationsangebote eingesetzt werden kann.

1.3. Aufbau der Arbeit

Um der Zielsetzung der vorliegenden Arbeit zu begegnen, erfolgt zunächst in Kapitel 2 eine Vorstellung der theoretischen Grundlagen. Hier wird der Begriff der *forschungsnahen Dienstleistungen* eingeführt (siehe Kapitel 2.1) und anhand des so gewonnenen Dienstleistungsverständnisses die SD-Logik als theoretisches Rahmenwerk und deren Wertverständnis im Sinne des ViU eingeführt (siehe Kapitel 2.2). Hieraus präzisiert sich die übergeordnete Fragestellung der vorliegenden Arbeit (siehe Kapitel 1.2) in zwei Forschungsfragen (siehe Kapitel 2.3). Es sollen zum einen die Ziele, welche Forschende an die Dienstleistungen herantragen, untersucht werden. Zum anderen werden die wertschöpfenden Faktoren entlang der Dienstleistungsinanspruchnahme in Form von Ressourcen untersucht. Die Forschungsfragen repräsentieren zwei Konzeptionierungsebenen des ViU, die in

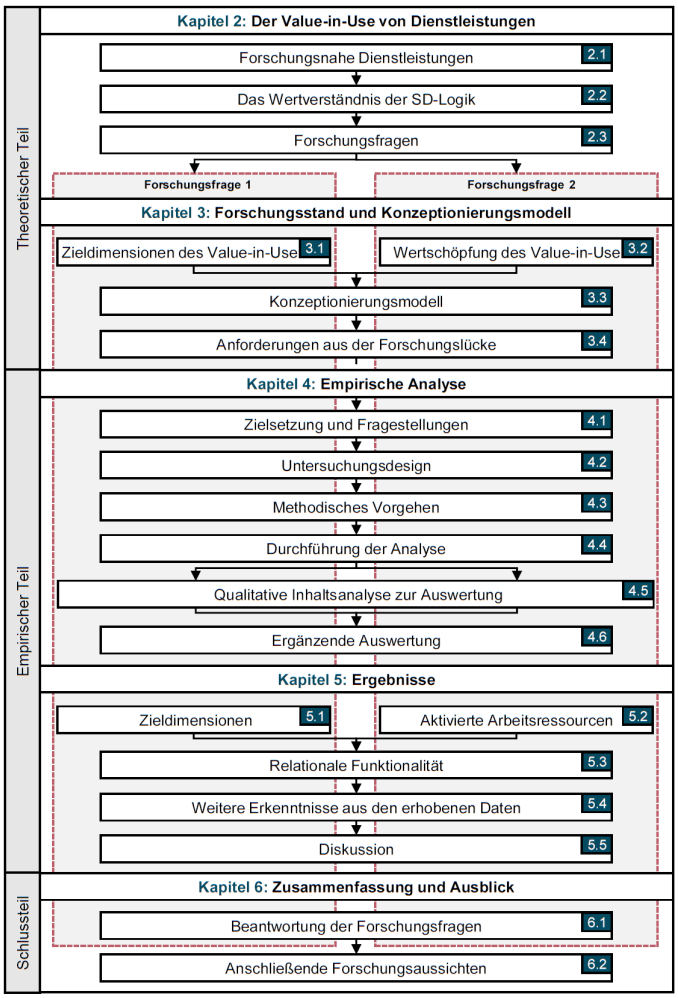

Abbildung 1: Aufbau der Arbeit

der Arbeit angestrebt werden und deren Unterteilung sich durch den Aufbau der Arbeit zieht (siehe Abbildung 1).

In Kapitel 3 erfolgt eine Darstellung des Forschungsstandes zu den beiden Forschungsfragen. Insbesondere für die erste Frage (Zieldimensionen) kann auf eine breite theoretische Basis zurückgegriffen werden, da der Konzeptionierungsansatz in der Literatur bereits umfassend Beachtung findet (siehe Kapitel 3.1). Die zweite Forschungsfrage zu den wertschöpfenden Ressourcen findet bislang weniger Aufmerksamkeit. Aus diesem Grund wird in Kapitel 3.2 anhand von Literatur aus dem Bereich der Arbeitspsychologie und der SD-Logik ein eigener Konzeptionierungsansatz zu deren Untersuchung abgeleitet. Anschließend wird ein zusammenführendes Konzeptionierungsmodell mit den beiden Ebenen für den ViU vorgestellt (siehe Kapitel 3.3), welches als Ausgangpunkt für eine anschließende empirische Untersuchung dient (siehe Kapitel 3.4).

In Kapitel 4 ist das Vorgehen der empirischen Analyse beschrieben. Deren Zielsetzung und Fragestellungen (siehe Kapitel 4.1) sowie das Untersuchungsdesign (siehe Kapitel 4.2) ergeben sich aus den abgeleiteten Forschungslücken des vorhergehenden theoretischen Kapitels. Kapitel 4.3 befasst sich mit dem methodischen Vorgehen zur empirischen Analyse der beiden Konzeptionierungsebenen. Konkret sollen Einzelinterviews geführt werden, welche anschließend mittels einer qualitativen Inhaltsanalyse ausgewertet werden. In Kapitel 4.4 und Kapitel 4.5 werden zunächst die Studiendurchführung und anschließend das Vorgehen zur Auswertung der Daten beschrieben. Die Auswertung erfolgt einzeln für die beiden Konzeptionierungsebenen. Um die Ergebnisse der beiden Ebenen abschließend zu verbinden, wird in Kapitel 4.6 eine Ergänzung des methodischen Rahmens zur Auswertung beschrieben, welcher ebenfalls eine quantitative Analyse vorsieht.

Die so gewonnen Ergebnisse sind in Kapitel 5 vorgestellt. Zunächst werden die qualitativen und quantitativen Ergebnisse der beiden Konzeptionierungsebenen separat beschrieben (siehe Kapitel 5.1 und Kapitel 5.2). Die zusammenführende Betrachtung der Ergebnisse ist Gegenstand des darauffolgenden Kapitels 5.3. Darüber hinaus konnten im Zuge der Interviews weitere Erkenntnisse gewonnen werden, welche aufgrund des zuvor gesetzten methodischen Auswertungsrahmens zunächst keine Beachtung finden. Diese sind in Kapitel 5.4 exemplarisch beschrieben und auf deren Grundla-

ge mögliche Erweiterungen des Konzeptionierungsmodells in Aussicht gestellt. Anschließend werden die Ergebnisse sowie das methodische Vorgehen zur Datenerhebung und Auswertung diskutiert (siehe Kapitel 5.5).

Kapitel 6 fasst die vorliegende Arbeit zusammen. Kapitel 6.1 beschreibt zunächst abschließend die gewonnen theoretischen und empirischen Erkenntnisse und stellt deren zukünftige Verwendung in der Praxis in Aussicht. Welche weiteren Forschungsrichtungen im Zuge der vorliegenden Arbeit geschaffen wurden, wird in Kapitel 6.2 beschrieben.

Der beschriebene Aufbau der Arbeit ist bis zur zweiten Gliederungsebene in Abbildung 1 zusammenfassend dargestellt.

2. Der Value-in-Use von Dienstleistungen

Die Tätigkeiten der Koordination und des Managements von Forschungsverbünden können als Dienstleistungen für Forschende verstanden werden (Lux & Robra-Bissantz, 2020). Dazu wird zunächst der Begriff der *forschungsnahen Dienstleistungen* eingeführt (siehe Kapitel 2.1). Anschließend wird die dienstleistungszentrierte Perspektive der SD-Logik als metatheoretisches Rahmenwerk vorgestellt (siehe Kapitel 2.2). Die SD-Logik findet Anwendung für alle Bereiche der Dienstleistungsforschung (Vargo, Koskela-Huotari & Vink, 2020) und bietet mit dem Konzept des *Value-in-Use* die Chance, Forschende als Anspruchsgruppe in den Fokus der Entwicklung und Gestaltung von forschungsnahen Dienstleistungen zu setzen. Um diese Chance nutzbar zu machen, stellt die vorliegende Arbeit zwei Forschungsfragen, welche in Kapitel 2.3 beschrieben werden.

2.1. Forschungsnahe Dienstleistungen

Aktuell werden die Management- und Koordinationstätigkeiten nicht als Dienstleistungsangebote für Forschende begriffen, sondern insbesondere unter dem Begriff des *Managements* von Forschungsverbünden untersucht. Dabei stehen Aufgaben zur internen sowie externen Kommunikation, Vernetzung, Synthesebildung und Organisation der Forschungsarbeit im Mittelpunkt (Defila et al., 2008). Bei der Anforderungsformulierung dieser Tätigkeiten wird die Einzigartigkeit von Verbundforschungsprojekten und die Diversität der daraus resultierenden Ansprüche an das Management betont (König et al., 2013; Lippe & vom Brocke, 2016). Die Untersuchungen nehmen jedoch eine organisationstheoretische Metaebene des Projekt- und Koordinationsmanagements ein, welche die Mikroperspektive auf die Forschenden selbst und deren individuelle Bedürfnisse nicht berücksichtigt (John, 2019).

Diese Mikroperspektive wird bei der Untersuchung der institutionellen bzw. universitären Möglichkeiten zur Unterstützung von Forschenden bereits vereinzelt eingenommen. Hier werden die Angebote für Forschende als Dienstleistungen verstanden und im Zuge dessen auch die Notwendigkeit einer nutzerzentrierten Entwicklung der Dienstleistungen betont, welche ein tiefgreifendes Verständnis und eine Integration der Anspruchsgruppe benötigt (Auth et al., 2019; Boden, 2016b; Locker-Grütjen, 2011). Die Dienstleistungen werden dort als *Forschungsunterstützung* bezeichnet und als „ein Komplex von Dienstleistungen für Wissenschaftler verstanden, der es, ohne selbst Forschung zu sein, ermöglicht, Forschung zu unterstützen“ (Boden, 2016b, S. 24).

Diese Definition kann für die vorliegende Arbeit auf den Begriff der *forschungsnahen Dienstleistungen* übertragen werden. Als Anspruchsgruppe werden im Folgenden *Forschende* zusammengefasst, welche in Anlehnung an Meschke (2015) alle Akteure bezeichnet, die wissenschaftliche Tätigkeiten innerhalb eines Forschungsverbundes ausüben. Betrachtet werden dabei insbesondere interdisziplinäre Forschungsverbünde[1]. Diese werden in der vorliegenden Arbeit als eine Kooperation von mehreren Teilprojekten unterschiedlicher Herkunftsdisziplinen verstanden (Defila et al., 2008; John, 2019; König et al., 2013). Die Forschungstätigkeiten, welche innerhalb dieser Verbünde ausgeübt werden, zeichnen sich durch *kooperatives Forschungshandeln* aus. Dieses kooperative Forschungshandeln ist nach Laudel (1999) nur ein Teil der Handlungen innerhalb von Forschungskooperationen. Darüber hinaus werden Tätigkeiten benötigt, welche das kooperative Forschungshandeln koordinieren (Laudel, 1999). Diese Aufgaben werden seitens der koordinierenden Projektstrukturen sowie einzelnen Akteuren innerhalb von Forschungsverbünden geleistet und unter anderem von König et al. (2013) zusammengefasst[2]. Beispielhaft können aus der Studie folgende Aufgaben genannt werden:

- Beratung und Begleitung hinsichtlich der Einhaltung von Zeitplänen, Vorschriften und finanziellen Fragen
- Angebote zur offenen Diskussion und Partizipation der Projektmitglieder
- Aufbau einer gemeinschaftlichen Kommunikationsstruktur (z.B. Website)
- Erleichterung der interdisziplinären Kommunikation mithilfe eines projektinternen Wissensmanagement (z.B. Glossar)

Diese und vergleichbare Tätigkeiten sollen folgend als *forschungsnahe Dienstleistungen* begriffen werden. Dafür wird in Anlehnung an Boden (2016b) und Laudel (1999) das Dienstleistungsverständnis auf die Management- und Koordinationstätigkeiten übertragen.

[1] Anhand welcher konkreten Merkmale sich derartige Forschungsverbünde auszeichnen fassen Defila et al. (2006) und Defila et al. (2008) zusammen.

[2] König et al. (2013) definieren vier einzelne Quadranten und ordnen diesen Rollen und dort inhärente Tätigkeiten zu, welche durch das Management von Forschungsverbünden abgebildet werden müssen. Die vier Quadranten klassifizieren das interdisziplinäre Forschungsmanagement unter anderen nach dem gesetzten Fokus (intern – extern), dem Zeithorizont und den Differenzierungsansprüchen der Tätigkeiten.

Definition: Forschungsnahe Dienstleistungen

Forschungsnahe Dienstleistungen umfassen einen Komplex von Dienstleistungen, welche in Forschungsverbünden das kooperative Forschungshandeln der Forschenden koordinieren und unterstützen, ohne selbst Forschung zu sein.

Durch die Übernahme eines solchen Dienstleistungsverständnisses werden die Forschenden als primäre Anspruchsgruppe der Angebote definiert. Um deren individuellen Bedürfnisse und Präferenzen bei der Entwicklung und Gestaltung von Dienstleistungen zu berücksichtigen, bietet die Marketing- und Dienstleistungsforschung unterschiedliche Ansätze. Ein populäres Konzept stellt hierbei die *wertzentrierte* Perspektive dar, welche es ermöglicht, die Anspruchsgruppe besser zu verstehen anhand dessen *wertvolle* Angebote zu gestalten (Durgee, O'Connor & Veryzer, 1996; Morar, 2013).

Eine derartige wert- und nutzerzentrierte Perspektive wird auch innheralb des theoretischen Rahmenwerks der SD-Logik eingenommen.

2.2. Das Wertverständnis der SD-Logik

Die SD-Logik charakterisiert den Wert von Dienstleistungen über eine Steigerung des Wohlbefindens der begünstigten Akteure (Vargo, Maglio & Akaka, 2008). Demnach besitzen Dienstleistungen[3] keinen inhärenten Wert, sondern stellen zunächst nur ein potenziell wertvolles Angebot dar. Der Wert ergibt sich erst aus dessen Nutzung und wird phänomenologisch und individuell vom begünstigten Akteur bestimmt (Vargo & Lusch, 2004, 2008a; Vargo et al., 2008). Diesen zentralen Gedanken fassen die Begründer der SD-Logik mit dem Satz zusammen: „there is no value until an offering is used" (Vargo & Lusch, 2006, S. 44) und charakterisieren damit das Wertkonstrukt des *Value-in-Use (ViU)*. Die Betrachtung von Wert im Sinne des ViU nimmt in der Marketing- und Dienstleistungsforschung eine zentrale Rolle bei der Gestaltung von Dienstleistungen ein (Grönroos & Voima, 2013; Ostrom, Parasuraman, Bowen, Patrício & Voss, 2015; Vargo & Lusch, 2004).

2.2.1. Beitrag zur Dienstleistungsgestaltung

Im Zuge dessen werden unter anderem zahlreiche Ansätze zur Messung des Wertes einer Dienstleistung aus der Perspektive der SD-Logik vorgestellt (Helkkula, Kelleher & Pihlström, 2012; Löbler & Hahn, 2013; Ranjan & Read, 2016). Diese ermöglichen eine nachgelagerte Betrachtung des Wertes, der sich durch eine subjektive Wahrnehmung für den begünstigten Akteur ergibt. Eine solche Untersuchung kann für die Entwicklung von Dienstleistungen ein erstes Verständnis für die Wertbeurteilung aufbauen. In der Literatur wird jedoch diskutiert, ob die Messung von Wert (nach der SD-Logik) überhaupt möglich sein kann (Gallan & Jefferies, 2020). Darüber hinaus wird ebenfalls in Frage gestellt, ob solche statischen Ansätze Implikationen für die Entwicklung von Dienstleistungen bieten können (Capra & Luisi, 2014; Vargo & Lusch, 2017), welche aufgrund sozialer Interaktionen von einer hohen Komplexität und Unsicherheit geprägt sind. Um dieser Dynamik zu begegnen kann in Anlehnung an Morelli (2009) ein offenes System von Komponenten formuliert werden, das als „Infrastruktur" für die Dienstleistungsentwicklung dient. Als eine solche Infrastruktur können innerhalb der Dienstleistungsforschung unter anderem Ansätze zur Konzeptionierung des ViU angesehen werden, welche die implizite Natur des Wertes sichtbar und nutzbar machen (u.a. Macdonald, Wilson, Martinez & Toossi, 2011; Bruns & Jacob, 2014; Hartwig & Jacob, 2018).

Für die vorliegende Arbeit kann demnach eine Konzeptionierung des ViU ein Verständnis dafür aufbauen, wie Wert für Forschende über forschungsnahe Dienstleistungen entsteht. Diese Einblicke können als Gestaltungsgrundlage für die Entwicklung und Verbesserung forschungsnaher Dienstleistungen angewandt werden (Helkkula, Kowalkowski & Tronvoll, 2018; Ostrom et al., 2010, 2015) und zwei Arten der Dienstleistungsentwicklung unterstützen. Zum einen baut die Konzeptionierung des ViU für die anbietenden Akteure ein Verständnis dafür auf, wie ein spezifisches Angebot ausgestaltet werden muss, um Wert für die Anspruchsgruppe zu schaffen. Zum anderen können die Ergebnisse auch als Verständnis- und Verständigungsgrundlage für eine gemeinsame Entwicklung[4] von Dienstleistungsangeboten zwischen Forschenden und den anbietenden Akteuren dienen. Die Chance dabei ist, dass die begünstigten Akteuren ihre eigenen Präferenzen reflektieren und anschließend in die Gestaltung eines Angebotes mit einbringen (Etgar, 2008). Dadurch können Dienstleistungen angeboten werden, welche den individuellen Präferenzen und Bedürfnissen der Forschenden entsprechen und somit wiederum den entstehenden ViU positiv beeinflussen (Lusch et al., 2007).

Welche Implikationen sich aus der SD-Logik für die Konzeptionierung des ViU forschungsnaher Dienstleistungen ergeben, wird folgend vorgestellt.

2.2.2. Value-in-Use aus der Zielerreichung

Über die vergangenen 20 Jahre haben sich eine Vielzahl von Definitionsansätzen für den ViU ergeben[5]. Neben

[3] In der vorliegenden Arbeit wird zum allgemeinen Verständnis der deutschsprachige Begriff der Dienstleistung(en) übernommen. Dadurch geht eine Differenzierung verloren, welche explizit in der englischsprachigen Primärliteratur eingeführt wird. Dort werden services (plural) und service (singular) betrachtet. Vargo und Lusch (2006) differenzieren diese Begriffe nach Plural und Singular, um den Grundgedanken der SD-Logik zu verdeutlichen. Die SD-Logik unterscheidet nicht nach Produkten und Dienstleistungen (services) wie es in der güterzentrierten Logik üblich ist. Stattdessen wird unterstellt, dass alles als Service im Sinne des Dienens verstanden werden muss. Demnach wird nach der SD-Logik durch die Anwendung von Kompetenzen den begünstigten Akteuren ein Vorteil verschafft – sowohl bei materiellen als auch bei immateriellen Angeboten. Hierfür wird im folgenden Verlauf der Arbeit der Begriff Dienstleistung(en) verwendet.

[4] Diese gemeinsame Gestaltung eines Angebotes wird von der SD-Logik als Co-Production bezeichnet und unter anderem von Lusch, Vargo und O'Brien (2007) und Etgar (2008) näher untersucht.

[5] Ein Überblick über die unterschiedlichen Definitionsansätze finden sich unter anderem in Macdonald, Kleinaltenkamp und Wilson (2016) und Med-

der generischen und abstrakten Definition des ViU im Sinne der Steigerung des Wohlbefindens, wird für die vorliegende Arbeit die Definition nach Macdonald et al. (2016, S. 98) übernommen, die den ViU als „all customer-perceived consequences arising from a solution that facilitate or hinder achievement of the customer's goals" beschreiben. Diese Definition wird angepasst und lediglich die *Erleichterung der Zielerreichung* betrachtet. Eine Berücksichtigung der Konsequenzen, welche die Zielerreichung behindern, impliziert, dass der ViU nicht zwangsläufig ein positives Wertergebnis darstellt. Dieser Aspekt wird unter anderem von Plé und Chumpitaz Cáceres (2010) im Sinne der Möglichkeit einer negativen Wertschöpfung zur Diskussion gestellt. Da die vorliegende Arbeit sich aber auf die positive Wertbetrachtung beschränkt, wird der ViU forschungsnaher Dienstleistungen demnach in Anlehnung an Macdonald et al. (2016) wie folgt definiert:

Definition: Value-in-Use forschungsnaher Dienstleistungen
Der **Value-in-Use forschungsnaher Dienstleistungen** beschreibt alle vom begünstigten Akteur wahrgenommenen Konsequenzen, die sich aus der Inanspruchnahme einer forschungsnahen Dienstleistung ergeben und die das Erreichen der Ziele des Akteures ermöglichen oder erleichtern.

Dieser Definition folgend ergibt sich der ViU aus dem Erreichen der Ziele der Forschenden. Der ViU kann somit als mehrdimensionales Wertergebnis beschrieben werden, dessen einzelne Dimensionen erreichte Ziele repräsentieren. Diese Verbindung von Zielen und Werten stellt erstmals die *Zieltheorie* nach Woodruff (1997) her. Nach Woodruff (1997) ergibt sich der Wert durch die Beurteilung des begünstigten Akteurs, inwieweit eine Dienstleistung dazu beigetragen hat, die persönlichen Ziele zu befriedigen. Wie in Abbildung 2 dargestellt, können somit die spezifischen Attribute einer Dienstleistung oder eines Produktes über deren Konsequenzen aus der Anwendung zu den übergeordneten abstrakten Zielen in Verbindung gesetzt werden.

Dieser Zusammenhang wird auch als *Attribution* bezeichnet (Raghubir & Corfman, 1999). Es lässt sich eine Art Kausalkette (Reynolds & Gutmann, 1988) aufstellen, zwischen Merkmalen eines Angebotes und den übergeordneten Zielen, welche die begünstigten Akteure bei dessen Nutzung verfolgen. Trägt das Angebot während der Inanspruchnahme zur Zielerreichung bei, wird ein Wert für den Akteur geschaffen. Die Ziele im Sinne der Zieltheorie sind dabei auf der gleichen Abstraktionsebene wie die Wertdimensionen, wie sie zur Differenzierung des ViU eingesetzt werden können. Selbstverwirklichung als ViU-Dimension wird bspw. geschaffen, wenn das Ziel nach Selbstverwirklichung durch die Dienstleistung erfüllt wurde. Demzufolge kann das mehrdimensionale Wertkonstrukt des ViU über die einzelnen Ziele operationalisiert werden.

berg und Grönroos (2020).

2.2.3. Value-in-Use aus dem Wertschöpfungsprozess

Ob innerhalb einer solchen Dimensionen Wert geschaffen wird, bestimmen die *Konsequenzen*, welche die Zielerreichung erleichtern (siehe Definition ViU).

Die Konsequenzen ergeben sich aus der Inanspruchnahme der forschungsnahen Dienstleistungen und somit aus dem Wertschöpfungsprozess. Dieser Prozess wird von der SD-Logik als Aktivitäten der *Value Co-Creation* verstanden (Vargo et al., 2020). Die Value Co-Creation besagt, dass Wert immer gemeinsam über Prozesse des Austausches von Ressourcen zwischen Akteuren geschaffen wird, wobei der begünstigte Akteur immer mit eingeschlossen ist[6] (Vargo et al., 2020). Der Wert, der sich daraus für den begünstigten Akteur ergibt, ist das „emergente Ergebnis der Integration von Ressourcen", welches dessen Wohlbefinden aufrechterhält oder steigert [übersetzt nach Vargo et al., 2020, S. 10].

Mit dem Konzept der Value Co-Creation zur Beschreibung des Wertschöpfungsprozesses und der mehrdimensionalen Differenzierung des ViU als Wertergebnis nimmt die SD-Logik zwei unterschiedliche, aber miteinander verbundene, Ebenen für die Wertbetrachtung ein. Der ViU beschreibt den Wert als ein Ergebnis, das zu einem bestimmten Zeitpunkt vom begünstigen Akteur wahrgenommen wird und über einen gemeinsamen Wertschöpfungs-prozess durch den Austausch von Ressourcen entsteht (Chandler & Vargo, 2011; Vargo et al., 2020; Vargo & Lusch, 2008b).

2.3. Forschungsfragen

In der Problemstellung der vorliegenden Arbeit (siehe Kapitel 1.2) wird folgende übergeordnete Fragestellung aufgeworfen:

> *Wie kann der Value-in-Use von Forschenden abgebildet werden und so zur Gestaltung von Dienstleistungen für Forschende beitragen?*

Da die Bezeichnung *Dienstleistungen für Forschende* ein umfassenderes Verständnis für mögliche Angebote impliziert (insb. zentrale Angebote durch das Forschungsmanagement) wird im folgenden Verlauf der Arbeit der Begriff *forschungsnahe Dienstleistung* verwendet (siehe Kapitel 2.1). Der ViU als Wertkonstrukt der SD-Logik dient der vorliegenden Arbeit als theoretische Grundlage einer eine wert- und nutzerzentrierte Untersuchung solcher Dienstleistungsangebote in Forschungsverbünden. Da der ViU ein phänomenologisch wahrgenommenes Wertergebnis darstellt, bietet ein Ansatz zur nachgelagerten Messung dieser subjektiven und spezifischen Erfahrung keine ausreichenden Implikationen für die Gestaltung von Dienstleistungsangeboten (siehe Kapitel 2.2.1).

Stattdessen strebt die vorliegende Arbeit eine Konzeptionierung des ViU an, welche auf abstrahierter Ebene und losgelöst von spezifischen Dienstleistungsattributen die Komponenten des ViU darstellt. Diese ermöglichen zum einen ein

[6] Dessen Beteiligung am Wertschöpfungsprozess spiegelt den Grundgedanken des ViU wider, nach dem sich der Wert für den begünstigten Akteur erst aus der Inanspruchnahme eines Angebotes ergibt.

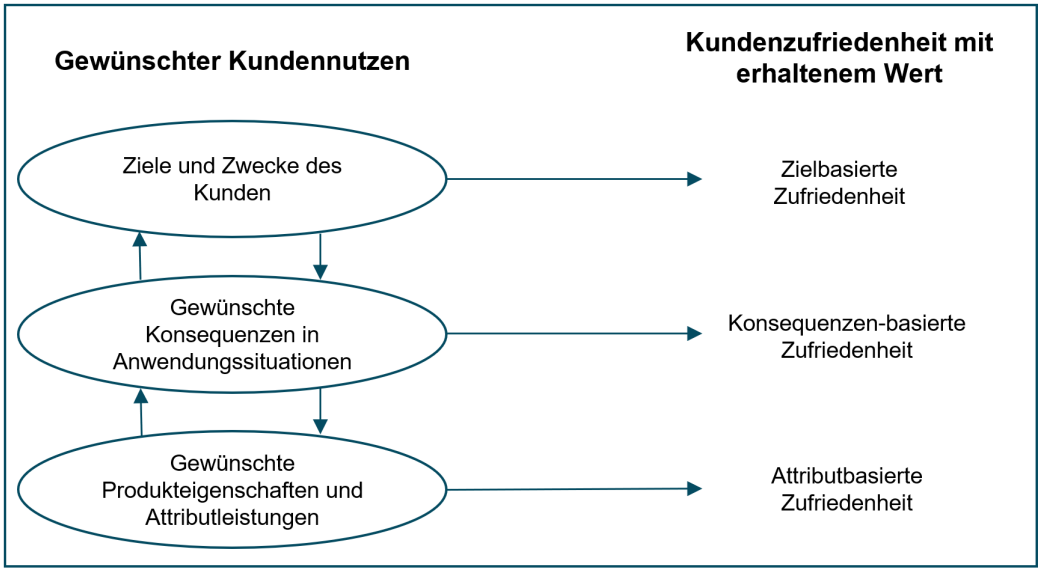

Abbildung 2: Kundenwerthierarchie

Quelle: Eigene Darstellung nach Woodruff (1997)

Verständnis für die implizite und inhärente Wertbeurteilung und zum anderen eine Verständigungsgrundlage für die gemeinsame Gestaltung von Angeboten.

Als Grundlage der Konzeptionierung werden die beiden zuvor beschrieben Betrachtungsebenen eingenommen (siehe Kapitel 2.2.2 und 2.2.3) und der ViU aus der Perspektive der Zielerreichung und des Wertschöpfungsprozesses untersucht (siehe Abbildung 3).

Dadurch löst sich die Arbeit davon, den ViU als ein nachgelagertes Ergebnis darzustellen. Stattdessen wird das *angestrebte* Wertergebnis mithilfe von Zieldimensionen operationalisiert, welche die Bedürfnisse und Bedarfe der Forschenden darstellen und somit den mehrdimensionalen ViU determinieren. Trägt die Dienstleistung zu dem Erreichen der Ziele bei, wird Wert innerhalb dieser Dimensionen für die Forschenden geschaffen[7]. Für die Konzeptionierung aus der Perspektive der potenziellen ViU-Dimensionen (in Form von Zieldimensionen[8]) wird die folgende Forschungsfrage gestellt:

> *Forschungsfrage 1: Welche Ziele verfolgen Forschende bei der Inanspruchnahme forschungsnaher Dienstleistungen?*

Nach der Definition des ViU wird Wert innerhalb der einzelnen Zieldimensionen infolge der Konsequenzen aus dem Wertschöpfungsprozess geschaffen. Aus diesem Grund soll die Konzeptionierung des ViU ebenfalls die relevanten Einflussfaktoren des Wertschöpfungsprozesses berücksichtigen. Somit wird ein besseres Verständnis der Wertwahrnehmung ermöglicht, als es die alleinige Darstellung des Wertergebnisses über die einzelnen Dimensionen leisten kann (Eggert, Ulaga, Frow & Payne, 2018; Payne, Storbacka & Frow, 2008).

Der Wertschöpfungsprozess im Sinne der SD-Logik stellt einen Austausch von Ressourcen dar. Das theoretische Konzept von Ressourcen und deren Wert findet nicht nur in der Dienstleistungsliteratur (Peters et al., 2014), sondern auch innerhalb der Arbeits- und Sozialpsychologie Beachtung (Demerouti, Bakker, Nachreiner & Schaufeli, 2001; Knecht & Schubert, 2012). Aus diesem Grund wird ein ressourcenorientierter Ansatz für die Untersuchung der Wertbeurteilung von forschungsnahen Dienstleistungen eingenommen und folgende Forschungsfrage für die zweite Konzeptionierungsebene gestellt:

> *Forschungsfrage 2:Welche Ressourcen aus dem Wertschöpfungsprozess der forschungsnahen Dienstleistungen determinieren den ViU für Forschende?*

Die beiden Forschungsfragen sollen somit ein Verständnis für die Entstehung des ViU aufbauen. Es wird nicht der Anspruch formuliert den ViU als einmaliges Ergebnis darzustellen oder zu messen, was allgemein aufgrund des phänomenologischen Charakters des ViU nicht als zielführend erachtet wird. Stattdessen ermöglicht die angestrebte Konzeptionierung über die beiden Forschungsfragen Einblicke in den Prozess der Wertschöpfung zu gelangen und zu verstehen, auf welchen Ebenen durch forschungsnahe Dienstleistungen Wert für die Forschenden geschaffen werden kann.

[7] Wie in Kapitel 2.2.3 dargestellt, wird im Sinne der SD-Logik Wert immer gemeinsam geschaffen. Demnach müsste im Verlauf der Arbeit stets davon die Rede sein, dass Wert für und von den Forschenden geschaffen wird. Zur sprachlichen Vereinfachung im Verlauf der Arbeit wird aufgrund der unterstützenden Intention der Dienstleistungen lediglich davon gesprochen, dass Wert für Forschende geschaffen wird.

[8] Da sich die Ziel- und ViU-Dimensionen auf der gleichen Abstraktionsebene befinden, repräsentieren die begrifflichen Klassifikationen der Ziele und Werte gleichbedeutende Dimensionen.

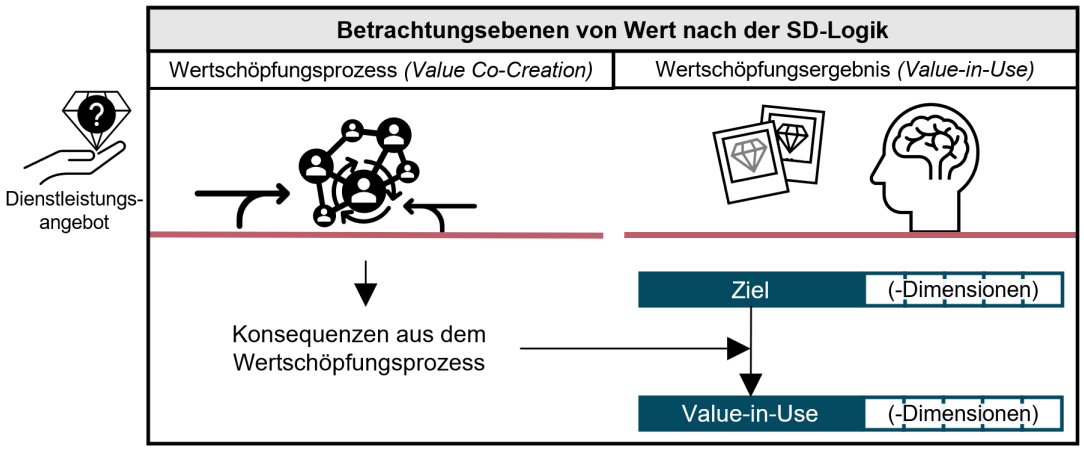

Abbildung 3: Betrachtungsebenen von Wert im Sinne der SD-Logik

Quelle: Eigene Darstellung

3. Forschungsstand und Konzeptionierungsmodell

Zur Beantwortung der beiden Forschungsfragen wird zunächst der aktuelle Stand der Forschung untersucht. Anhand dessen können erste Erkenntnisse für die beiden Forschungsfragen gewonnen (siehe Kapitel 3.1 und 3.2), und ein Konzeptionierungsmodell für den ViU abgeleitet werden (siehe Kapitel 3.3). Das so aufgestellte Modell und die Erkenntnisse aus dem Stand der Forschung dienen anschließend der Anforderungs-formulierung an eine empirische Untersuchung der Forschungsfragen (siehe Kapitel 3.4).

3.1. Zieldimensionen des Value-in-Use

Die erste Konzeptionierungsebene stellen die Zieldimensionen dar, welche Forschende an die forschungsnahen Dienstleistungen herantragen.

Die Zieldimensionen entsprechen nach Woodruff (1997) operationalisierten Wertdimensionen. In Form von ViU-Dimensionen wurden solche Wertdimensionen, als Ergebnis der Zielerreichung, bereits umfangreich in der Marketing- und Dienstleistungsforschung untersucht (Bruns & Jacob, 2014, 2016; Kleinaltenkamp, Löbler, Eßer & Fennert, 2018; Kleinaltenkamp, Storck, Gumprecht & Li, 2018; Macdonald et al., 2016). Im Zuge dessen wurde anhand diverser Dienstleistungsangebote ein umfassendes Portfolio der möglichen ViU-Dimensionen empirisch ermittelt (Kleinaltenkamp & Dekanozishvili, 2019). Die Ergebnisse unterscheiden sich zwischen den untersuchten Dienstleistungen und Nutzungssituationen.

Im Bereich der (Unterstützungs-)Dienstleistungen für Forschende – oder Mitarbeiter*innen allgemein – existiert nach Kenntnis der Autorin keine vergleichbare Klassifikation von Wert- oder Zieldimensionen. Aus diesem Grund wird folgend geprüft, inwieweit der bestehende Stand der Forschung bzgl. der ViU-Dimensionen auf die Konzeptionierung der Zieldimensionen forschungsnaher Dienstleistungen übertragen werden kann. Übergeordnet ist anzumerken, dass nach der Zieltheorie die einzelnen Dimensionen des ViU erneut als Ziele operationalisiert werden können, nachdem die abstrakten Ziele den angestrebten Werten eines Akteurs entsprechen.

3.1.1. Allgemein anerkannte Dimensionen

Einer zusammenfassenden Literaturstudie von Kleinaltenkamp und Dekanozishvili (2019) zufolge, kann bestimmten ViU-Dimensionen des aktuellen Forschungsstandes eine allgemeine Gültigkeit zugeschrieben werden und sind demnach für unterschiedliche Nutzungssituationen diverser Produkt- und Dienstleistungsangebote relevant (siehe Tabelle 1).

Diese Liste allgemeingültiger ViU-Dimensionen ergibt sich aus unterschiedlichen marktbezogenen Untersuchungen. Die Studien basieren sowohl auf Nutzungssituationen im Business-to-Business-Bereich (Macdonald et al., 2016) als auch auf unterschiedlichen Business-to-Consumer-Anwendungen. Zu nennen sind hier bspw. die Nutzung von Smartphones (Bruns & Jacob, 2014), Plattformen (Kleinaltenkamp, Storck et al., 2018) und Carsharing (Kleinaltenkamp, Löbler et al., 2018).

Die Untersuchungsumgebungen unterscheiden sich zu dem vorliegenden Forschungskontext insbesondere hinsichtlich des betrachteten Marktumfeldes (Geigenmüller & Lenk, 2014). Forschungsnahe Dienstleistungen können auch als sogenannte *interne Dienstleistungen* verstanden werden (Bruhn, 2012; Strauss, 2010), welche Dienstleistungen bezeichnen, die Mitarbeitenden innerhalb deren Organisation angeboten werden[9]. Dadurch ändert sich das Verständnis

[9] Die Literatur beschäftigt sich teilweise bereits mit der Qualität und dem Wert von internen Dienstleistungen. Bei Veröffentlichungen wie bspw. Kang, Jame und Alexandris (2002) oder Johnston (2008) wird jedoch insbesondere die interne Servicequalität hinsichtlich deren Wirkung auf die externe Servicequalität untersucht. Der Beitrag von Geigenmüller und Lenk (2014) untersucht explizit den Wert für Mitarbeitende. Dieser wird allerdings mit-

Tabelle 1: ViU-Dimensionen mit allgemeiner Relevanz

ViU-Dimension	Charakterisierung
Produktivität	Effizienz, Prozessverbesserung, Optimierung und Steigerung des Outputs, schnelle Problemlösung, Einfachheit der Aufgaben, Kontinuität, einfacher Zugriff auf relevante Informationen, Kostensenkung und Leistung
Sozialer Wert	Der soziale Nutzen, die Interaktion mit anderen, die angenehme Atmosphäre, persönliche Beziehungen, die Fähigkeit, ein soziales Netzwerk aufzubauen und Beziehungen zu pflegen, das Erreichen von Prestige und der Aufbau eines Images; all das beeinflusst das Wohlbefinden des Akteurs
Unsicherheitsvermeidung	Minimierung und Reduzierung der unbekannten und verbundenen Risiken; Erhöhung der wahrgenommenen Kontrolle, Anpassungsfähigkeit und Sicherheit
Selbstdarstellung	Persönliche Reputation, Selbstpräsentation, Selbstdarstellung, Selbstbestätigung, Ausdruck eigener Kompetenzen, Gedanken, Ideen und Leistungen
Selbstverwirklichung	Verwirklichung der eigenen Ambitionen, Wünsche, Interessen und eigenen Kompetenzen
Hedonistischer Wert	Gefühl der unmittelbaren affektiven Freude, Zufriedenheit, Bequemlichkeit und Spaß
Vermeidung von Abhängigkeiten	Unabhängigkeit und Freiheit gewinnen, Autonomie erhalten, verbindliche Pflichten und Verantwortlichkeiten reduzieren und das Risiko, kontrolliert zu werden, minimieren
Sachverständnis	Hohes Maß an Kompetenz, Leistungs- und Erfolgsstreben, Urteilsvermögen, kontinuierliches Lernen, Einsichten schaffen und Professionalität

Quelle: Kleinaltenkamp und Dekanozishvili (2019), eigene Übersetzung

der Anspruchsgruppe und deren Einbindung in das Dienstleistungsumfeld. Insbesondere die Nachfragebedingungen können sich bei internen Dienstleistungen von marktbasierten Dienstleistungen unterscheiden. Bspw. bei sogenannten obligatorischen internen Dienstleistungen ist die Inanspruchnahme seitens der Institution vorgeschrieben und somit keine aktive Nachfrageentscheidung, geschweige denn ein wettbewerbliches Angebot, erforderlich (Geigenmüller & Lenk, 2014; Strauss, 2010). Allgemein zeichnet sich jedoch der arbeitsorganisatorische Kontext von Wissensarbeit durch eine hohe Autonomie in der Aufgabenausführung aus (Gläser, 2006; Schaper, 2014). Diese Autonomie erlaubt den Forschenden in den meisten Fällen darüber zu entscheiden, ob sie ein Unterstützungsangebot in Anspruch nehmen. Dadurch ergibt sich für die forschungsnahen Dienstleistungen ein marktähnlicher Charakter. Diese Annahme wird von aktuellen Untersuchungen zu Dienstleistungen für Forschende im Zuge des universitären Forschungsmanagements gestützt (Locker-Grütjen, 2011). Aus diesem Grund wird angenommen, dass das Portfolio der allgemeingültigen ViU-Dimensionen, welche innerhalb eines wettbewerblichen Umfelds ermittelt wurden, auch als Grundlage für das Wertverständnis forschungsnaher Dienstleistungen dienen können.

hilfe des Wertverständnisses nach Zeithaml (1988) untersucht und die wahrgenommen Kosten zu den Nutzen in Relation gesetzt. Dieser Ansatz unterscheidet sich von dem der vorliegenden Arbeit, welcher eine positive Wirkrichtung für Wert unterstellt [vgl. weiterführend zu den unterschiedlichen Wertverständnissen: Gummerus (2013)].

3.1.2. Kontextgebundenheit der Dimensionen

Neben den allgemeingültigen ViU-Dimensionen bestimmt der Kontext, in welchem die Dienstleistung erbracht wird, weitere *kontextspezifische* Dimensionen (Kleinaltenkamp & Dekanozishvili, 2019). Der Kontext[10] wird im Sinne der SD-Logik als „set of unique actors with unique reciprocal links among them“ definiert (Chandler & Vargo, 2011, S. 40).

Mit dem Anspruch ein Wertverständnis für forschungsnahe Dienstleistungen aufzubauen, kann dieser Detailebene der Definition des Kontextes, bei der Konzeptionierung des ViU nicht entsprochen werden. Aus diesem Grund wird folgend der Kontext der forschungsnahen Dienstleistungen (im Sinne der vorliegenden Arbeit, siehe Kapitel 2.1) anhand der Forschenden und den Akteuren im Umfeld des Forschungsverbundes als „set of actors“ charakterisiert. Das kooperative Forschungshandeln im Verbund wird von Akteuren unterstützt, welche Dienstleistungen für die Forschenden anbieten und somit den Rahmen für eine gemeinsame Wertschöpfungskooperation „reciprocal links among them“ bilden.

Der Forschungsverbund als soziales System [vgl. weiterführend Krohn & Küppers, 1989] entspricht demnach dem (sozialen) Umfeld der Forschenden. Eine solche Einbindung der Anspruchsgruppe in ein soziales Umfeld muss bei der

[10] Die theoretische Diskussion bei der Konzeptionierung von Wert geht bzgl. der Relevanz des Kontextes weit über den hier betrachteten Einfluss auf die ViU-Dimensionen hinaus. Im Sinne der SD-Logik werden alle Wertschöpfungsprozesse von dem Kontext, in welchem diese stattfinden, determiniert (Chandler & Vargo, 2011). Dieses Wertverständnis hat insbesondere Einfluss auf Ansätze einer prozessorientierten Konzeptionierung und wird in Kapitel 3.2 erneut aufgegriffen.

Konzeptionierung des ViU berücksichtigt werden (Chandler & Vargo, 2011; Edvardsson, Tronvoll & Gruber, 2011). Edvardsson et al. (2011) betonen den Einfluss des sozialen Systems auf die Wertwahrnehmung und stellen dabei unter anderem die These auf, dass der Wert eine kollektive und intersubjektive (bzw. individuelle) Dimension hat[11] (Edvardsson et al., 2011). Dieser These zufolge sind die Ziele, welche Forschende bei der Inanspruchnahme von forschungsnahen Dienstleistungen verfolgen, nicht nur auf die persönlichen Interessen ausgerichtet, sondern ebenfalls auf die Ziele des Verbundprojektes. Die Annahme wird durch die Ergebnisse empirischer Studien bestätigt, welche innerhalb vergleichbarer Nutzungskontexte durchgeführt wurden. Die Studien untersuchen Dienstleistungsangebote für Tätigkeiten im Arbeitsumfeld und berücksichtigen dabei die Einbindung der Anspruchsgruppe in die Interessensgemeinschaft der Arbeit (Kleinaltenkamp, Storck et al., 2018; Macdonald et al., 2016). Demnach können die beiden Studien[12] in zweierlei Hinsicht Orientierung für die Spezifikation möglicher ViU-Dimensionen forschungsnaher Dienstleistungen bieten. Zum einen werden die bekannten allgemeingültigen ViU-Dimensionen hinsichtlich wertvoller Aspekte im Arbeitsumfeld definiert und inhaltlich charakterisiert. Darüber hinaus werden die ViU-Dimensionen nach ihrer Ausrichtung strukturiert und somit auch Wertebenen identifiziert, die sich auf die Ziele des Kollektivs beziehen, in welches das Individuum eingebettet ist.

Demnach wird folgend in Anlehnung an die Ergebnisse von Kleinaltenkamp, Storck et al. (2018) und Macdonald et al. (2016) das Portfolio der allgemeingültigen ViU-Dimensionen (siehe Tabelle 1) für den Kontext forschungsnaher Dienstleistungen angepasst und in Tabelle 2 (S. 20) zusammenfassend dargestellt.

Die allgemeingültigen ViU-Dimensionen werden für den Anwendungsfall spezifiziert. Bspw. gewinnt durch die Betrachtung von Dienstleistungen im Arbeitskontext die allgemeingültige Wertebene der Produktivität eine höhere Relevanz und wird in fünf Wertebenen ausdifferenziert (*Produktivität, Aufgabenvereinfachung, Prozessverbesserung, Schnelle Problemlösung* und *Vermeiden von Stillstandzeiten*). Darüber hinaus wird das Portfolio um kontextspezifische ViU-Dimensionen erweitert. Diese sind sowohl auf die persönlichen Interessen als auch auf die Interessen des Verbundprojektes ausgerichtet[13]. Auf der individuellen Wertebene wird das Portfolio durch die Wertdimensionen der *Druckreduzierung* und *Motivation* erweitert. Eine Aufnahme dieser Dimensionen kann auch durch den Betrachtungsgegenstand der forschungsnahen Dienstleistungen begründet werden. Der Aspekt der *Druckreduzierung* ergibt sich durch die primäre Intension der Dienstleistungsangebote die Forschenden bei ihren Tätigkeiten zu unterstützen und zu entlasten (vgl. Kapitel 2.1). Die ViU-Dimension *Motivation* wird definiert als „grundlegende Antriebskraft und kognitive Zielorientierung bei der täglichen Arbeit“ (Kleinaltenkamp, Storck et al., 2018, S. 104). Dieses Ziel wird auch als persönliches Interesse von Forschenden innerhalb von Verbundprojekten beschrieben. Nach John (2019, S. 23) kann durch die Einbettung der Forschungstätigkeiten in Verbundprojekten „die eigene Arbeitsmotivation durch Selbstverpflichtungen aufrechterhalten“ werden. Darüber hinaus wird die Dimension der *Autonomie* ergänzt, da sich die Wissensarbeit durch eine hohe Entscheidungsautonomie der Forschenden bzgl. der Aufgabenausführung auszeichnet (Gläser, 2006; May, Korezynski & Frenkel, 2002; Schaper, 2014). Diese kann zwar teilweise durch die institutionellen Strukturen oder Vorgaben der finanzierenden Instanzen begrenzt werden, Entscheidungen über das Vorgehen bei der Forschungsarbeit bleiben dennoch weitgehend autonom (Defila et al., 2008).

Auf der Wertebene, welche eher auf das Kollektiv ausgerichtet ist, werden von Kleinaltenkamp, Storck et al. (2018) und Macdonald et al. (2016) die Dimensionen *Zugang zu Finanzmitteln*, *Innovationskraft* und *Konkurrenzfähigkeit* übernommen. Diese Dimensionen spiegeln sich auch in der Intension und den Ansprüchen kooperativer Forschung wider. Das übergeordnete Ziel der kooperativen Forschung ist eine verbesserte Wissensproduktion (Katz & Martin, 1997). Die Wissensproduktion stellt dabei den Output der Forschungstätigkeiten dar (Laudel, 2001). Durch die Einbettung der Forschungstätigkeiten im Verbund, haben die Forschenden Zugang zu Ressourcen in Form von Wissen, technischen Geräten oder Disseminationsnetzen (*Zugang zur Scientific Community*), welche die eigenen Möglichkeiten übersteigen (Fleßa, John & Mahnke, 2011; John, 2019; Katz & Martin, 1997). Durch diese Möglichkeiten kann durch die Organisation der Forschenden in einem Verbund das Ziel einer erhöhten *Wettbewerbsfähigkeit bzw. Konkurrenzfähigkeit* der Individuen als auch des Kollektivs gegenüber der wissenschaftlichen Gemeinschaft verfolgt werden. Im Zuge dessen wird auch der *Zugang zu Finanzmitteln* erleichtert, welche zunehmend wettbewerbsorientiert und an kooperative, interdisziplinäre Forschungsprojekte vergeben werden (DFG, 2018; Hinze, 2016; Vowe & Meißner, 2020; Winterhager, 2015). Darüber hinaus werden durch die Forschung im Verbund im Optimalfall Synergien geschaffen, welche sich in Form von gegenseitiger Stimulation und Kreativität äußern (John, 2019; Katz & Martin, 1997; Kratzer, Leenders & van Engelen, 2005; Krücken, 2009; Laudel, 2001). Diese syn-

[11] Der Ansatz zur sozialen Konstruktion von Wert von Edvardsson et al. (2011) umfasst neben der genannten These weitere Aspekte, welche sich nicht auf das Wertergebnis sondern den Wertschöpfungsprozess beziehen und in Kapitel 3.2 aufgegriffen werden

[12] Kleinaltenkamp, Storck et al. (2018) zeigen die ViU-Dimensionen der Nutzung eines E-Learning-System für Studierende und Lehrerende auf, während Macdonald et al. (2016) innerhalb eines Business-to-Business-Marktumfeldes unterschiedlicher Leistungsangebote für Mitarbeitende untersuchen.

[13] An dieser Stelle sei darauf hingewiesen, dass ein wesentlicher Teil der Aufgaben des Managements von Forschungsverbünden auf die Umsetzung der strategischen Ziele des Kollektivs entfällt. Eine Balance zwischen den Zielen des Verbundes und den Zielen der einzelnen Forschenden zu finden, wird als große Herausforderung beschrieben (Defila et al., 2008). Die vorliegende Studie beschränkt sich jedoch nach dem ViU-Verständnis auf die persönlichen Ziele der Forschenden, welche neben individuellen Interessen auch kollektive Ziele im Sinne des Verbundes darstellen können.

ergetischen Effekte finden sich in der ViU-Dimension *Innovationskraft*, welche definiert wird ist als „Generierung von neuen und nützlichen Ideen und Umsetzung neuer kreativer Lösungen“ (Kleinaltenkamp & Dekanozishvili, 2019, S. 229).

Das so abgeleitete Portfolio möglicher ViU-Dimensionen stellt zunächst die theoretische Grundlage für die erste Konzeptionierungsebene des ViU dar. Welche der ViU-Dimensionen in der spezifischen Nutzungssituationen und aus Sicht des Individuums relevant sind, können die begünstigten Akteure artikulieren (Hartwig & Jacob, 2018; Macdonald et al., 2011).

Abschließend formuliert, stellt die Konzeptionierungsebene dieses Kapitels, den Wert einer Dienstleistung in Abhängigkeit zu den Zielen, welche die Forschenden bei der Inanspruchnahme verfolgen. Der wahrgenommene Beitrag der Dienstleistung zur Zielerreichung determiniert damit den ViU als Wertergebnis. Die Zieldimensionen beschreiben somit *was* der Forschende an einer Dienstleistung bzw. an deren Nutzung wertschätzt (Hartwig & Jacob, 2018; Kleinaltenkamp & Dekanozishvili, 2019; Morar, 2013).

Doch *wie* diese Wertbeurteilung durch die Erreichung der Zieldimensionen über den Wertschöpfungsprozess seitens des Forschenden zustande kommt, können derartige Klassifikationen nicht abbilden (Hartwig & Jacob, 2018; Woodruff & Flint, 2006).

3.2. Wertschöpfung des Value-in-Use

Die Wertschöpfung wird nach der SD-Logik als ein Prozess der Ressourcenintegration verstanden (Kleinaltenkamp et al., 2012; Peters et al., 2014). Der Wert der sich daraus für den begünstigten Akteur ergibt, ist das „emergente Ergebnis der Integration von Ressourcen“, welche dessen Wohlbefinden[14] aufrechterhalten oder steigern [übersetzt nach Vargo et al., 2020, S. 10]. Demnach stellt aus der Wertschöpfungsperspektive der ViU forschungsnaher Dienstleistungen ein Ergebnis aus einem Prozess der Integration von Ressourcen dar. Welche Ressourcen dabei den ViU determinieren wird folgend untersucht.

3.2.1. Ressourcen im Arbeitsumfeld

Hierfür stellt sich zunächst die Frage, welche Ressourcen aus dem Arbeitsumfeld im Zuge des Wertschöpfungsprozesses von forschungsnahen Dienstleistungen Wert für die Forschenden schaffen können.

Allgemein ist der Ressourcenbegriff weit gefasst. Im Sinne der SD-Logik werden Ressourcen immer integriert zum Nutzen eines Akteurs (Vargo & Lusch, 2004). Dabei sind nach der SD-Logik Ressourcen „alles, greifbar oder nicht greifbar, intern oder extern, operand oder operant [15]“ [übersetzt nach Lusch & Vargo, 2014, S. 121], das diesen Nutzen für einen Akteur schafft. Somit können Ressourcen anhand ihres spezifischen Zwecks für deren Einsatz abgeleitet werden (Peters et al., 2014). Dieser *Zweck* stellt für forschungsnahe Dienstleistungen die Unterstützung von Forschenden dar (siehe Kapitel 2.1).

Eine Orientierung welche Ressourcen aus dem Arbeitsumfeld wertgenerierend sein können, findet sich in Modellen aus dem Bereich der Arbeitspsychologie.

Das heuristische *Job Demands-Resources Model* (JD-R-Modell) nach Demerouti et al. (2001) weist entscheidende Parallelen zu den Überlegungen der SD-Logik auf. Es gibt zum einen an wie das „*Wohlbefinden* von Mitarbeitern durch zwei spezifische Sätze von Arbeitsbedingungen entsteht“ [übersetzt nach Bakker & Demerouti, 2007, S. 275, eigene Hervorhebung]. Zum anderen wird die Vielfalt an Einflussfaktoren des Arbeitsumfeldes eines Mitarbeitenden in Arbeitsanforderungen und *Arbeitsressourcen* unterteilt (siehe Abbildung 4). Die Balance zwischen Anforderungen und Ressourcen der Arbeitsumgebung stehen nach dem Modell in wechselwirkender oder direkter Beziehung zu positiven und negativen arbeitsbezogenen Erfahrungen wie Arbeitsengagement oder Burnout (Demerouti & Nachreiner, 2019). Für die vorliegende Arbeit werden diese Auswirkungen aus dem Gleichgewicht und die nachgelagerten psychologische Prozesse nicht näher betrachtet[16] (in Abbildung 4 ausgegraut). Mit dem Ziel einer Klassifizierung von (wertvollen) Ressourcen für Forschende wird die zugrundeliegende Prämisse der positiven Wirkung von Arbeitsressourcen sowie deren Definition aus dem JD-R-Modell übernommen.

Ein Kernaspekt von Arbeitsressourcen im Sinne des JD-R-Modells ist die moderierende und abpuffernde Wirkung auf den Einfluss von Arbeitsanforderungen[17] (Bakker & Demerouti, 2007; Bakker, Hakanen, Demerouti & Xanthopoulou, 2007; Demerouti & Nachreiner, 2019). Darüber hinaus werden in der Definition von Arbeitsressourcen zwei weitere

[14] Für das Konstrukt des „Wohlbefindens“ existieren zahlreiche Definitionen und Ansätze zu dessen Messung, wie unter anderem von Dodge, Daly, Huyton und Sanders (2012) aufgezeigt wird. Die vorliegende Arbeit bezieht sich, den Ansätzen der SD-Logik folgend, auf das subjektive Wohlbefinden von Forschenden.

[15] Innerhalb der SD-Logik wird vornehmlich zwischen operanden und operanten Ressourcen unterschieden. Nach Madhavaram und Hunt (2008) sind operante Ressourcen in der Regel physische Ressourcen, während operande Ressourcen insbesondere menschlicher, organisatorischer, informationeller oder relationaler Natur sind

[16] Eine Übersicht der aktuellen Entwicklungen und Weiterführungen des Modells zeigen Bakker und Demerouti (2007). Das Modell bietet Ansätze zur Steigerung des Wohlbefindens von Mitarbeitenden und findet über ein breites Spektrum an Beschäftigungsverhältnissen empirische Evidenz (Bakker & Demerouti, 2007). Für das vorliegende Forschungsobjekt kann bspw. die Studie von Bakker, Demerouti und Euwema (2005) genannt werden, bei welcher Angestellte einer Hochschule als Stichprobe dienten. Die Studie bewies unter anderem, dass der Einfluss von Überlastung durch die untersuchten Arbeitsressourcen (soziale Unterstützung, Autonomie, Beziehung zum Vorgesetzten und Leistungsfeedback) abgepuffert werden kann.

[17] Dieser Aspekt scheint in Konflikt mit dem zuvor gesetzten Fokus der vorliegenden Arbeit zu stehen, da die positive Wirkung von Arbeitsressourcen (Steigerung des Wohlbefindens) in Zusammenhang zu Arbeitsanforderungen gestellt werden, welche folgend aus der Betrachtung ausgeschlossen werden. Die Annahme, dass Arbeitsressourcen lediglich den Einfluss spezifischer Anforderungen abmildern, wenn diese übereinstimmen, wird als Matching-Hypothese bezeichnet. Diese wurde im Zuge einer empirischen Untersuchung von Bakker und Demerouti (2007) für das JD-R Modell widerlegt, welche unabhängig von deren Übereinstimmung zu den Arbeitsanforderungen, moderierende Effekte von Arbeitsressourcen für das Wohlbefinden feststellten.

Tabelle 2: Portfolio möglicher ViU-Dimensionen forschungsnaher Dienstleistungen

Allgemeine ViU-Dimensionen [Quelle]	ViU-Dimensionen im Kontext [Quelle]	Beschreibung	Quelle der Beschreibung
Produktivität [3]	**Produktivität** [1]	Verbesserung des Input/Output-Verhältnisses	[1]
	Aufgabenvereinfachung [1;2]	Einfachheit und Zeiteffizienz der Abläufe, die für die eigene Arbeit notwendig sind	[1], [2]
	Prozessverbesserung [1;2]	Vereinfachung oder Verbesserung der Forschungsprozesse auf Verbundebene	In Anlehnung an [1] und [2]
	Schnelle Problemlösung [1;2]	Schnelles Lösen von operativen und inhaltlichen Problemen	[1]
	Vermeiden von Stillstandzeiten [1;2]	Minimierung von unproduktiven Phasen während der Forschungsarbeit im Projekt	In Anlehnung an [1] und [2]
Sozialer Wert [3]	**Sozialer Wert** [1;2]	Soziale Bedingungen, die das Wohlbefinden steigern	[1]
Hedonistischer Wert [3]	**Hedonistischer Wert** [1]	Befriedigung durch unmittelbares gefühlsmäßiges Vergnügen bei der eigenen täglichen Arbeit	[1]
Selbstdarstellung [3]	**Selbstdarstellung** [1]	Ausdruck eigener Kompetenzen, Gedanken und Ideen bei der täglichen Arbeit	[1]
Selbstverwirklichung [3]	**Selbstverwirklichung** [1]	Verwirklichung der eigenen Ambitionen und Wünsche während der täglichen Arbeit	[1]
Unsicherheitsvermeidung [3]	**Unsicherheitsvermeidung** [1;2]	Minimierung von Unsicherheiten in Bezug auf Prozesse, Vorgaben und Entscheidungen, die mit der eigenen Arbeit zusammenhängen	[1], [2]
	Wahrgenommene Kontrolle [1;2]	Wahrnehmung der Kontrolle von Prozessen und Entscheidungen mit der eigenen Arbeit verbunden sind	In Anlehnung an [1] und [2]
	Flexibilität [1]	Möglichkeit, das eigene Arbeitsverhalten schnell an veränderte Bedingungen anzupassen	[1]
Vermeidung von Abhängigkeiten [3]	**Vermeidung von Abhängigkeiten** [2]	Reduzierung der Abhängigkeit eigener Aufgaben von Forschungsergebnissen anderer	In Anlehnung an [2]
Sachverständnis [3]	**Sachverständnis** [1]	Kompetenz, Urteilsvermögen und Streben nach Leistung und Erfolg durch die eigene Arbeit	In Anlehnung an [1]
Kontext-spezifische ViU-Dimensionen	**Zugang zu Finanzmitteln** [1]	Schaffung zusätzlicher Einnahmequellen	[1], [3]
	Druckreduzierung [1;2]	Reduktion von Stress und Belastung während den täglichen Aufgaben	[2]

(Continued)

Table 2—continued

Kreativität und Innovationskraft [angepasst nach 2]	Generierung von neuen und nützlichen Ideen und Umsetzung von kreativen Lösungen	[3]
Wettbewerbs-fähigkeit [angepasst nach 1 und 2]	Gesteigerte Wettbewerbsfähigkeit der eigenen und kollektiven Forschung	In Anlehnung an [1] und [2]
Motivation [1]	Grundlegende Antriebskraft und kognitive Zielorientierung bei der täglichen Arbeit	[1]
Zugang zur Scientific Community	Schaffen von Vernetzungsmöglichkeiten innerhalb der Scientific Community	
Autonomie	Freiheit und Unabhängigkeit, was die Entscheidungen bzgl. der eigenen Forschung anbelangt	

Quellen: [1] Kleinaltenkamp, Storck et al. (2018); [2] Macdonald et al. (2016); [3] Kleinaltenkamp und Dekanozishvili (2019), eigene Übersetzung mit Abgleich der Dimensionsbezeichnungen anhand der deutschsprachigen Ergebnisse der Studien Kleinaltenkamp, Storck et al. (2018) und Kleinaltenkamp, Löbler et al. (2018), welche von Prof. Dr. Dr. Michael Kleinaltenkamp auf Anfrage zur Verfügung gestellt wurden.

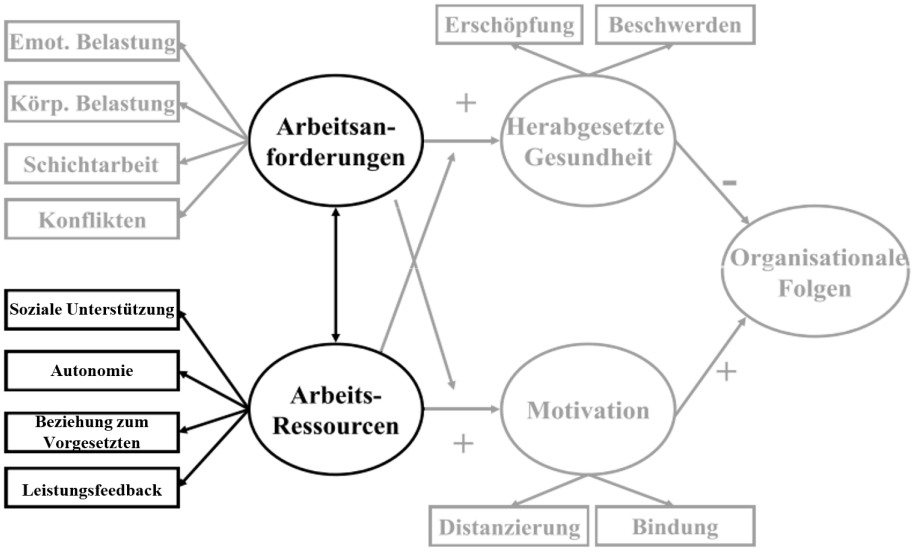

Abbildung 4: Das JD-R-Modell

Quelle: angepasst nach Demerouti und Nachreiner (2019) mit den Ergebnissen bzgl. der Arbeitsressourcen von Bakker et al. (2005)

Wirkrichtungen definiert: „Arbeitsressourcen sind die physischen, psychischen, sozialen und organisatorischen Arbeitsbedingungen, die

(1) funktional für das Erreichen der arbeitsbezogenen Ziele sind,
(2) Arbeitsanforderungen und damit zusammenhängende physische und psychische Kosten reduzieren und
(3) persönliches Wachstum und persönliche Entwicklung stimulieren“ [Bakker & Demerouti, 2007, S. 312; Demerouti & Nachreiner, 2019, S. 122].

Die möglichen Arbeitsressourcen nach dieser Definition sind weiter gefasst, als es im Zuge der Konzeptionierung von Ressourcen forschungsnaher Dienstleistungen relevant und zielführend ist, da sie sich auf die Arbeitsumgebung im Allgemeinen beziehen. Dadurch werden auch Aspekte wie die monetäre Entlohnung oder die Arbeitsplatzsicherheit zu Ar-

beitsressourcen. Welche Ziele Forschende explizit bei der Inanspruchnahme forschungsnaher Dienstleistungen verfolgen wird in Kapitel 3.1 aus empirischen Untersuchungen nach der SD-Logik abgeleitet. Innerhalb dieses Portfolios (siehe Tabelle 2) lassen sich Parallelen zu den beschriebenen Wirkrichtungen der Arbeitsressourcen nach Demerouti und Nachreiner (2019) ausmachen. Bspw. findet sich der *soziale Wert* oder die *Selbstverwirklichung* auch in den allgemeinen Arbeitszielen (1) nach Super (1970). Die Reduktion der Arbeitsanforderungen (2) werden durch Dimensionen wie *Aufgabenvereinfachung* oder *Unsicherheitsvermeidung* repräsentiert und dem persönlichen Wachstum (3) lassen sich bspw. die Dimensionen *Kreativität und Innovationskraft* oder *Sachverständnis* zuordnen. Aus diesem Grund wird zur Eingrenzung möglicher Ressourcen, welche im Wertschöpfungsprozess forschungsnaher Dienstleistungen integriert werden, der definitorische Ansatz der Arbeitsressourcen mit der zweckausgerichteten Charakterisierung von Ressourcen im Sinne der SD-Logik vereint. Daraus ergibt sich in Anlehnung an das JD-R-Modell (Bakker & Demerouti, 2007; Demerouti & Nachreiner, 2019) und der SD-Logik (Lusch & Vargo, 2014; Peters et al., 2014; Vargo & Lusch, 2004) folgende Definition für die Arbeitsressourcen:

Definition: Arbeitsressourcen

Arbeitsressourcen sind die physischen, psychischen, sozialen und organisatorischen Arbeitsbedingungen, welche während der Inanspruchnahme forschungsnaher Dienstleistungen den Forschenden zur persönlichen Zielerreichung zur Verfügung gestellt werden.

Eine erste Orientierung der möglichen Arbeitsressourcen bieten Demerouti und Nachreiner (2019). Hier werden unter anderem Autonomie, die Beziehung zum Vorgesetzten, Leistungsfeedback, Kooperations- und Kommunikationsmöglichkeiten sowie soziale Unterstützung genannt. Aufgrund der gruppenbasierten Arbeitsbedingungen sowie dem Konzept der gemeinsamen Wertschöpfung (Value Co-Creation) kann angenommen werden, dass das soziale Umfeld als Ressource im Wertschöpfungsprozess forschungsnaher Dienstleistungen von hoher Relevanz ist. Welche Ebenen diese soziale Unterstützung im Arbeitskontext einnehmen kann, strukturiert unter anderem House (1981). Er differenziert mit seiner *Taxonomie sozialer Unterstützung* nach emotionaler, evaluativer, informativer und instrumenteller Unterstützung (House, 1981).

Diese exemplarisch genannten Ressourcen stellen sogenannte *externe* oder auch *Umweltressourcen* dar. D.h. Ressourcen, welche im Zuge von forschungsnahen Dienstleistungen den Forschenden zur Verfügung gestellt werden, um diese zur persönlichen Wertschöpfung zu integrieren. Welche Ressourcen integriert werden können und inwieweit diese zur Zielerreichung von Forschenden funktional sind, wird entlang des Wertschöpfungsprozesses bestimmt, welcher von der SD-Logik genauer beschrieben wird.

3.2.2. Determinanten des Wertschöpfungsprozesses

Die SD-Logik schreibt Ressourcen einen dynamischen Charakter zu (Lusch & Vargo, 2014). Dabei bezieht sich SD-Logik auf die Sichtweise von Baum und Zimmermann (1951) und die Aussage nach de Gregori (1987) „resources are not, they become". Demnach werden Ressourcen erst entlang deren Integration bzw. Einsatz wertvoll (Chandler & Vargo, 2011) und können auch ihren übergeordneten Status als Ressource entsprechend ihrer Nutzbarkeit gewinnen oder verlieren (Löbler & Hahn, 2013; Peters et al., 2014). In diesem Zusammenhang wird auch der Begriff der *Ressourceness* eingeführt. Er beschreibt „die Fähigkeit potenzieller Ressourcen, die Verwirklichung von etwas Wünschenswertem zu erleichtern – die durch die Verfügbarkeit anderer, sich ergänzender und hemmender potenzieller Ressourcen bestimmt wird, einschließlich der Fähigkeit der Akteure, diese Ressourcen zu integrieren und anzuwenden" [übersetzt nach Koskela-Huotari & Vargo, 2016, S. 166].

Die Berücksichtigung der Fähigkeiten eines Akteurs steht in Zusammenhang mit dem Konzept der Value Co-Creation, in dem Sinne, dass der Wert über Austauschprozesse von Ressourcen zwischen Akteuren gemeinsam geschaffen wird, wobei der begünstigte Akteur immer mit eingeschlossen ist (Vargo et al., 2020). Dieser Aspekt findet sich ebenfalls in der Human- und Sozialökologie (Antonovsky, 1997; Becker, 2006). Hier wird betont, dass die Funktionalität und damit der Nutzen, den Individuen aus Umweltressourcen generieren können, davon abhängt, ob ausreichende und angemessene persönliche Ressourcen[18] vorhanden sind (Schubert & Knecht, 2012).

Neben den persönlichen Ressourcen wird auch auf die „Verfügbarkeit anderer, sich ergänzender und hemmender potenzieller Ressourcen" verwiesen [übersetzt nach Koskela-Huotari & Vargo, 2016, S. 165 f.]. Da prinzipiell alle Akteure als mögliche Ressourcenintegrator*innen zu sehen sind (Vargo et al., 2020), bestimmt sich diese Verfügbarkeit durch die weiteren Akteure, die an der Wertschöpfungskooperation beteiligt sind. Dadurch bilden die Akteure im Zuge dieser Kooperation den *(sozialen) Kontext* der Ressourcenintegration (Chandler & Vargo, 2011; Edvardsson et al., 2011). Bei der Analyse des Kontextes und der beteiligten Akteure können verschiedene Aggregationsebenen eingenommen werden (Vargo & Lusch, 2017). Eine gängige Klassifizierung der Analyseebenen stellt die Unterteilung in die Makro-, Meso- und Mikroebene dar (Chandler & Vargo, 2011). Neben der Formation der beteiligten Akteure definiert der soziale Kontext auch Aspekte[19], welche die gemeinsamen Aktivitäten der Ressourcenintegration beeinflussen. In diesem Zusammen-

[18] Unter persönlichen Ressourcen versteht man individuelle Ressourcen, welche einer Person inhärent sind. Diese können differenziert werden nach physischen Ressourcen (z.B. Gesundheit und Fitness), psychischen Ressourcen (z.B. Selbstwertgefühl), interaktionellen Ressourcen (z.B. Beziehungsfähigkeit) und ökonomischen Ressourcen (z.B. Kapitalbesitz) (Schubert & Knecht, 2012).

[19] Im Arbeitsumfeld umfasst der Kontext nach Bamberger (2008, S. 840) diverse Aspekte wie „physische Arbeitsplatzbedingungen, breitere soziale oder normative Umgebungen (einschließlich verschiedener Dimensionen

hang wird auch der Begriff der *Institutionen*[20] und institutionellen Logiken eingeführt. Dabei wird angenommen, dass Institutionen die „Art und Weise prägen, in der die Akteure Ressourcen integrieren und Werte schaffen und bewerten“ [übersetzt nach Koskela-Huotari und Vargo (2016), S. 169].

Zusammenfassend unterliegt der Prozess der Ressourcenintegration im Sinne der SD-Logik also einer sehr hohen Dynamik und Komplexität, welche in dem Umfang in der vorliegenden Arbeit nicht abgebildet werden kann. Neben den Fähigkeiten in Form von persönlichen Ressourcen des begünstigten Akteurs[21], könnte auch der Forschungsverbund als soziales System (siehe Fußnote 19) oder Institutionen (siehe Fußnote 20) untersucht werden, in welche die Forschenden eingebunden sind.

Neben der hohen Dynamik des Wertschöpfungsprozesses müsste aufgrund des Aspektes der Value Co-Creation mindestens eine dyadische (zwei Akteure) Austauschbeziehung für die Konzeptionierung des Prozesses analysiert werden (Chandler & Vargo, 2011). Während der dabei geschaffene Wert jedoch individuell und subjektiv von den Akteuren in ihrem (sozialen) Umfeld wahrgenommen und bestimmt wird (Chandler & Vargo, 2011; Edvardsson et al., 2011; Vargo et al., 2008). Somit nimmt die prozessorientierte Konzeptionierung eine andere Perspektive (Makroebene) ein als die Analyse des individuell wahrgenommenen Wertergebnisses (Gummerus, 2013). Um diese Asymmetrie zu überwinden, wird vorgeschlagen anzunehmen, dass sich diverse Akteure an der Value Co-Creation beteiligen und dadurch eine *Co-Creation-Erfahrung* für das Individuum schaffen: „Experience is the missing link and the common denominator of value creation processes and value outcomes” (Gummerus, 2013, S. 30).

Dieser Vorschlag wird für die Konzeptionierung wertvoller Ressourcen aus dem Wertschöpfungsprozess forschungsnaher Dienstleistungen übernommen. Durch die Untersuchung von Erfahrungen werden die theoretischen Einflussfaktoren, die die Ressourcenintegration und damit die Wertschöpfung bestimmen, implizit mitberücksichtigt, ohne die dahinterstehenden determinierenden Prozesse explizit zu strukturieren. Hierfür wird der Begriff der *aktivierten* Arbeitsressourcen eingeführt, dessen Verortung im Wertschöpfungsprozess folgend dargestellt wird.

Während der Inanspruchnahme findet ein gemeinsamer Wertschöpfungsprozess statt, in welchem Forschende ihrer eigenen Ressourcen, die angebotsseitigen Ressourcen sowie Ressourcen aus dem (sozialen) Umfeld integrieren. Die kundenseitige Wertschöpfung kann also als ein Prozess diverser Aktivitäten zur Ressourcenintegration verstanden werden, welche seitens des begünstigten Akteurs ausgeübt werden, *um ein bestimmtes Ziel zu erreichen* (Payne et al., 2008; Peters et al., 2014). Der Prozess der Ressourcenintegration kann nicht abgebildet werden, aber die anschließende subjektive Beurteilung des Wertschöpfungsprozesses hinsichtlich dessen Beitrag zur Zielerreichung (Kleinaltenkamp et al., 2012; Lemke, Clark & Wilson, 2011). Wie der Nutzen der Ressourcen beurteilt wird, wird von der Funktionalität der Ressourcen geprägt zu den Motiven, Interessen und Zielen einer Person beizutragen (Willutzki, 2003). Diese Motive, Interessen und Ziele stellen in der vorliegenden Arbeit die Zieldimensionen der forschungsnahen Dienstleistungen dar, welche in Kapitel 3.1 theoretisch abgeleitet werden. Die Arbeitsressourcen, die aus der Erfahrung der Value Co-Creation seitens der Forschenden als wertvoll beschrieben werden, entsprechen sogenannten *aktivierten* Arbeitsressourcen, welche „zur Anforderungsbewältigung oder Zielerreichung als brauchbar erkannt und entsprechend eingesetzt werden“ (Schubert & Knecht, 2012).

Definition: Aktivierte Arbeitsressourcen

Aktivierte Arbeitsressourcen sind die physischen, psychischen, sozialen und organisatorischen Arbeitsbedingungen, welche während der Inanspruchnahme forschungsnaher Dienstleistungen den Forschenden zur persönlichen Zielerreichung zur Verfügung gestellt, von ihnen als brauchbar erkannt und entsprechend eingesetzt werden.

3.3. Konzeptionierungsmodell

Aus den theoretischen Erkenntnissen lässt sich demnach das in Abbildung 5 dargestellte Konzeptionierungsmodell für den ViU forschungsnaher Dienstleistungen aufstellen.

Forschungsnahe Dienstleistungen müssen zunächst als potenziell wertvolles *Angebot* von Ressourcen verstanden werden. Dieses Angebot hat keinen Wert, bis die Forschenden es nutzen. Der Wert für die Forschenden ergibt sich über das Erreichen der Zieldimensionen, welche die Dimensionen des ViU als Wertergebnis repräsentierten (siehe Kapitel 3.1). Ob auf diesen Ebenen Wert für die Forschenden geschaffen

der organisatorischen oder nationalen Kultur oder des Klimas in der Einheit oder Organisation) oder sogar externe Arbeitsmärkte“ [aus dem Englischen übersetzt]. Mit dem Kontext interdisziplinärer Forschung befassen sich bspw. Stokols, Misra, Moser, Hall und Taylor (2008) oder Siedlok und Hibbert (2014).

[20] Unter anderem in den Veröffentlichungen von Edvardsson, Kleinaltenkamp, Tronvoll, McHugh und Windahl (2014), Koskela-Huotari und Vargo (2016) und Koskela-Huotari, Vink und Edvardsson (2020) setzen die Autoren die Ressourcen und deren dynamischen Charakter in Verbindung zu den Institutionen, in welchen diese integriert werden. Dabei lehnen sie sich an die Theorie von North (2006) an und diskutieren, inwiefern Institutionen als kognitive, normative und regulative Regelsysteme auch die „Regeln“ der Wertschöpfung abbilden. Bei der Untersuchung forschungsnaher Dienstleistungen könnte ein tiefgreifendes Verständnis der Auswirkungen der Verbundprojekte im Sinne von Institutionen wertvolle Erkenntnisse bzgl. des ViU liefern. Bspw. von Schophaus, Dienel und von Braun (2004) wurde interdisziplinäre Forschungsarbeit bereits in Zusammenhang zu Institutionen gesetzt.

[21] Die Forschenden als Akteure haben auf zwei Ebenen Einfluss auf die Wertschöpfung. Zum einen beeinflussen die persönlichen Ressourcen den Bedarf spezifischer weiterer Ressourcen sowie die Fähigkeit die angebotenen externen Ressourcen zu integrieren. Welche Ressourcen in Form von Kompetenzen im Zuge von interdisziplinärer Forschung relevant sein können untersucht bspw. Brandstäfter und Sonntag (2016). Theoretisch könnte für die Gestaltung von forschungsnahen Dienstleistungen anhand der persönlichen Ressourcen ein bedarfsorientiertes Angebot abgeleitet werden, indem die Dienstleistung explizit bei der Unterstützung der defizitären Ressourcen der Forschenden ansetzt. Als Rahmen hierfür könnte eineTypisierung der Akteure dienen (vgl. weiterführend Jung (2014)). John (2019) bietet bspw. einen Ansatz hierfür, indem er vier Kooperationstypen interdisziplinärer Verbundforschung definiert.

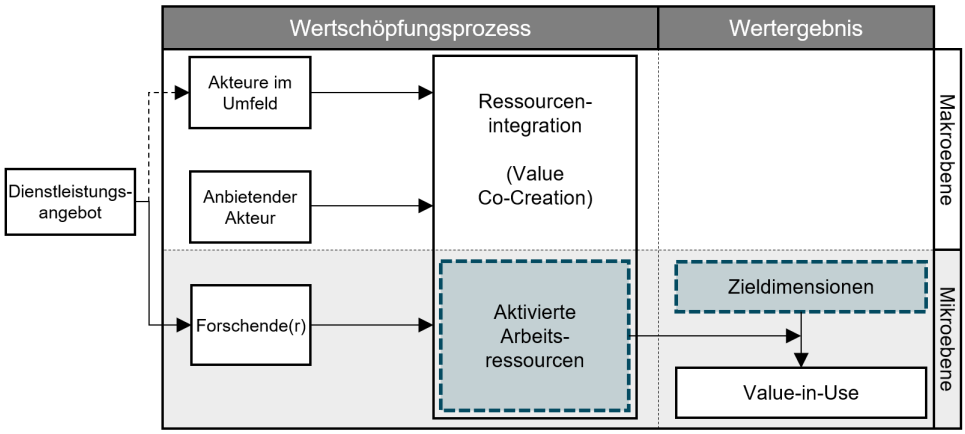

Abbildung 5: Konzeptionierungsmodell des ViU

Quelle: Eigene Darstellung

wird, bestimmt sich aus einem gemeinschaftlichen Wertschöpfungsprozess, welcher über die Mikroperspektive der einzelnen Forschenden hinausgeht. Hierbei integrieren die anbietenden Akteure, der oder die forschende Person sowie weitere Akteure aus dem Umfeld Ressourcen. Welche Ressourcen zur Zielerreichung seitens der einzelnen Forschenden auf Mikroebene als funktional erkannt und integriert werden, kann anhand von aktivierten Arbeitsressourcen dargestellt werden (siehe Kapitel 3.2).

3.4. Anforderungen aus der Forschungslücke

Die aktivierten Arbeitsressourcen und Zieldimensionen stellen die beiden Betrachtungsebenen des ViU forschungsnaher Dienstleistungen dar.

Für die Konzeptionierungsebene der Zieldimensionen kann auf ein breiten Forschungsstand zurückgegriffen werden, welcher sich mit der empirischen Analyse von ViU-Dimensionen innerhalb unterschiedlicher Nutzungssituationen befasst. Welche Wert-bzw. Zieldimensionen für forschungsnahe Dienstleistungen als konkreter Anwendungsfall relevant sind, wird bislang nicht untersucht. Aus diesem Grund wird aus dem aktuellen Forschungsstand und theoretischen Vorüberlegungen bzgl. des Einflusses des Anwendungsfalls (Dienstleistungen in Forschungsverbünden) ein Portfolio möglicher Zieldimensionen abgeleitet. Welche dieser potenziellen Dimensionen für forschungsnahe Dienstleistungen relevant sind, muss spezifiziert werden. Nach Lusch und Vargo (2006, S. 44) sind „Erfahrung und Wahrnehmung [...] wesentlich für die Wertbestimmung“ [aus dem Englischen übersetzt]. Aus diesem Grund kann eine solche Spezifikation nur mithilfe einer empirischen Untersuchung gelingen, in welcher Forschende als begünstigte Akteure anhand konkreter *erlebter* Nutzungssituationen artikulieren, welche Zieldimensionen für sie relevant sind (Hartwig & Jacob, 2018; Macdonald et al., 2011).

Die Relevanz von Erfahrung und Wahrnehmung (Lusch & Vargo, 2006) gilt auch für die Konzeptionierungsebene der aktivierten Arbeitsressourcen, weshalb auch hier eine empirische Untersuchung angestrebt wird. Eine Darstellung von spezifischen Ressourcen, welche sich als Konsequenzen auf den ViU in Form einer Zielerreichung auswirken, wird nach Kenntnis der Autorin bisweilen nicht untersucht. Aus diesem Grund werden anhand der Forschung aus dem Bereich der Arbeitspsychologie Einblicke gewonnen, welche Arbeitsressourcen für Forschende wertvoll sein können. Eine vergleichbare Klassifizierung, wie bei den Zieldimensionen, ist für die Arbeitsressourcen nicht aus dem Forschungsstand ableitbar. Außerdem wird der dynamische Wertschöpfungsprozess und damit der tatsächliche Wert von Ressourcen aus der Perspektive der SD-Logik dargestellt. Hieraus wird hergeleitet, dass mit dem Anspruch der Konzeptionierung des ViU aus der Sicht von Forschenden dieser Prozess lediglich implizit mitberücksichtigt werden kann. Dies gelingt durch die Untersuchung von Erfahrungen mit forschungsnahen Dienstleistungen, auf welche sich die Forschenden beziehen können (Gummerus, 2013; Kleinaltenkamp, Storck et al., 2018; Lusch & Vargo, 2006).

Durch die somit legitimierte angestrebte empirische Analyse, folgt die Arbeit ebenfalls Forschungsansprüchen aus dem Bereich der Dienstleistungsinnovationen, welche empirische Ergänzungen der bestehenden theoretischen Konzeptionierungsansätze fordern (Helkkula et al., 2018).

4. Empirische Analyse

Aus dem aktuellen Forschungstand kann ein Modell abgeleitet werden (siehe Kapitel 3.3), welches zunächst eine theoretische Konzeptionierung des ViU forschungsnaher Dienstleistungen darstellt. Dieses soll nun für den konkreten Anwendungsfall empirisch untersucht werden (siehe Kapitel 3.4). Das hierfür entwickelte Untersuchungsmodell sowie das methodische Vorgehen der Erhebung und Auswertung wird nachfolgend beschrieben.

4.1. Zielsetzung und Fragestellungen

Die empirische Untersuchung verfolgt das übergeordnete Ziel, den ViU als Wertergebnis forschungsnaher Dienstleistungen aus der Sicht der Forschenden zu verstehen. Wie zuvor erläutert, kann der ViU anhand unterschiedlicher Dimensionen spezifiziert werden (siehe Kapitel 3.1). Darüber hinaus wird in Kapitel 3.2 der ViU als prozessbasiertes Wertkonstrukt beschrieben, dessen Ergebnis von der Bereitstellung aktivierter Arbeitsressourcen innerhalb des Wertschöpfungsprozesses determiniert wird. Mithilfe der empirischen Untersuchung sollen diese beiden Konzeptionsebenen des ViU für das vorliegende Forschungsobjekt inhaltlich spezifiziert werden. Daraus ergeben sich die ersten beiden übergeordneten Fragestellungen (FS) für die empirische Untersuchung.

FS 1: Welche Zieldimensionen sind für Forschende bei der Inanspruchnahme forschungsnaher Dienstleistungen relevant und wie lassen sich diese charakterisieren?

FS 2: Welche aktivierten Arbeitsressourcen ermöglichen im Forschungskontext das Erreichen der Zieldimensionen forschungsnaher Dienstleistungen?

Über diese inhaltliche Spezifikation der Zieldimensionen und Exploration der aktivierten Arbeitsressourcen hinaus, unterstellt das Konzeptionierungsmodell eine funktionale Verbindung zwischen den beiden Betrachtungsebenen. Diese soll anhand der empirischen Ergebnisse überprüft werden, wofür die folgende Fragestellung ergänzend formuliert wird:

FS 3: Welche aktivierten Arbeitsressourcen aus dem Dienstleistungsprozess tragen aus Sicht der Forschenden zum Erreichen der Zieldimensionen und somit zur Wertschöpfung innerhalb der ViU-Dimensionen bei?

4.2. Untersuchungsdesign

Aus diesen Fragestellungen ergibt sich sowohl die theoretische Verortung der empirischen Analyse als auch die Ausrichtung des Forschungsdesigns, welche nachfolgend beschrieben wird.

Die Untersuchung bezieht sich auf die *methodisch-theoretischen Grundlagen der Sozialwissenschaften*. „Ihr [Sozialwissenschaften] Analysegegenstand ist der in einem sozialen Kontext lebende und handelnde Mensch, das soziale Individuum, dessen Handeln mit Sinn, mit Bezug auf andere versehen ist“ (Lamnek, 2010, S. 13). Diese Ausrichtung der Analyse auf das soziale Individuum deckt sich mit dem ViU-Verständnis im Sinne der SD-Logik. Nach dieser stellt der ViU eine individuelle und phänomenologische Wertbeurteilung aus der Sicht von Forschenden dar, welche in einen komplexen (sozialen) Kontext eingebettet ist.

Der Schwerpunk der Untersuchung auf die Wertwahrnehmung durch die Forschenden entspricht dem übergeordneten Anspruch der *qualitativen Sozialforschung*, die „Lebenswelten ’von innen heraus’ aus Sicht der handelnden Menschen zu beschreiben“ (Flick, von Kardorff & Steinke, 2017, S. 14). Demzufolge ergibt sich für die vorliegende Untersuchung ein qualitatives Forschungsdesign, wodurch sich der vorliegende Ansatz in die aktuelle qualitative Forschungspraxis bzgl. der Wertforschung im Rahmen des ViU einreiht.

Das übergeordnete Untersuchungsdesign ist in Abbildung 6 dargestellt.

Der aktuelle Stand der Forschung zu den ViU-Dimensionen bzw. Zieldimensionen (siehe Kapitel 3.1) wird als Inputdaten für die Untersuchung verwendet. Diese sollen im Zuge der Erhebung hinsichtlich ihrer Relevanz für den Forschungskontext geprüft sowie inhaltlich charakterisiert werden (FS 1), wozu konkrete Erfahrungen von Forschenden mit forschungsnahen Dienstleistungen untersucht werden. Dieses Vorgehen stützt sich auf Macdonald et al. (2011) und Hartwig und Jacob (2018), welche feststellen, dass Individuen die persönlich relevanten Dimensionen anhand von konkreten Dienstleistungssituationen artikulieren können. Auf den so ermittelten Zieldimensionen forschungsnaher Dienstleistungen aufbauend, soll zusätzlich erhoben werden, anhand welcher Arbeitsressourcen Forschende beurteilen, ob ein Wert für sie innerhalb dieser Dimensionen geschaffen wurde (FS 2). Dieser Forschungsansatz für die zweite Fragestellung und das dahinterstehende methodische Vorgehen wird von Hartwig und Jacob (2018)[22] übernommen. Die erhobenen Daten zu den ersten beiden Fragestellungen der empirischen Analyse sollen anschließend mittels einer qualitativen Inhaltsanalyse ausgewertet werden. Eine zusammenführende Betrachtung der Ergebnisse zu FS 1 und FS 2 ermöglicht die Beantwortung der dritten Fragestellung.

4.3. Methodisches Vorgehen

Im Zentrum des Untersuchungsdesigns steht eine qualitative Erhebung über Einzelinterviews mit einer anschließenden Kategorisierung der Daten mittels qualitativer Inhaltsanalyse. Im Folgenden wird das methodische Vorgehen zur Erhebung und Auswertung dargestellt. Anschließend folgt eine Beschreibung der angestellten Datenerfassung (siehe Kapitel 4.4) und des Auswertungsprozesses (siehe Kapitel 4.5).

4.3.1. Erhebungsinstrument

Als Erhebungsinstrument werden strukturierte, leitfadengestützte Interviews eingesetzt. Die genaue Befragungsform lässt sich anhand einer Vielzahl von Gesichtspunkten spezifizieren. Die Auswahl des Erhebungsinstruments wird in Anlehnung an Lamnek (2010, S. 303) beschrieben und in Tabelle 3 charakterisiert.

Die Wahl von Interviews begründet sich aus den zugrundliegenden Fragestellungen der empirischen Analyse (siehe Kapitel 4.1). Diese erfordern Erkenntnisse aus der subjektiven Erfahrungswelt von Forschenden im Umgang

[22] Hartwig und Jacob (2018) setzen sich bei einer empirischen Untersuchung ein vergleichbares Forschungsziel. Unter dem Begriff Beurteilungsmodi wird analysiert, wie und anhand welcher Kriterien die Befragten entscheiden, ob ein Ziel für sie erreicht wurde. Der Ansatz ist aufgrund der impliziten Berücksichtigung des Wertschöpfungsprozesses mit dem vorliegenden Konzeptionierungsansatz des Wertschöpfungsprozesses vergleichbar, nimmt hierzu jedoch nicht die Perspektive der Ressourcen ein.

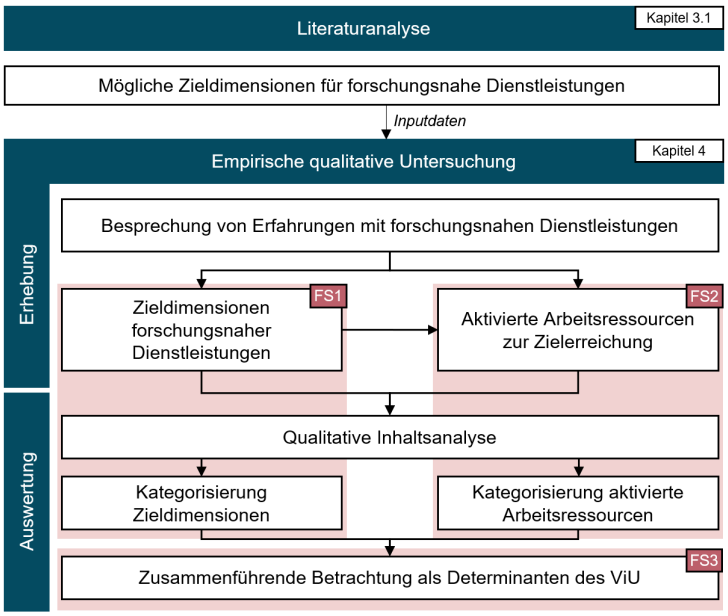

Abbildung 6: Untersuchungsdesign

Tabelle 3: Spezifikation der Befragung

Dimension der Differenzierung	Form des vorliegenden Interviews
Struktur der zu Befragenden	Einzelinterview
Stil der Kommunikation, Interviewverhalten	Weich
Form der Kommunikation	Mündlich
Intention des Interviews	Ermittelnd (informatorisch)
Standardisierung	Halb-standardisiert
Kommunikationsmedium	Videotelefonisch

mit forschungsnahen Dienstleistungen. Ein derartiger Zugang kann mithilfe von *Einzelinterviews* geschaffen werden, in welchen sich die Betroffenen auf vergangene Erlebnisse und Erfahrungen mit forschungsnahen Dienstleistungen beziehen und davon berichten. Eine weitere Möglichkeit hierfür stellen auch Gruppeninterviews dar. Die empirische Untersuchung bezieht sich jedoch auf Forschende, welche innerhalb ihres Arbeitsumfeldes befragt werden. Um hier die Auswirkungen gegenseitiger Beeinflussung und sozial erwünschter Antworttendenzen zu minimieren, stellen Einzelinterviews die präferierte Struktur der Befragten dar.

Im Zuge der Einzelinterviews soll ein *weicher Kommunikationsstil* eingesetzt werden, um eine Sympathie- und Vertrauensebene zu der interviewten Person aufzubauen. Dabei folgt die *mündliche* Befragung einer Struktur, welche bereits im Zuge der Fragestellungen und dem Untersuchungsdesign angedeutet wird und auf den theoretischen Vorannahmen bzgl. der ViU-Konzeption beruht. Konkret sollen die beiden zentralen Fragen FS 1 und FS 2 beantwortet werden. Dabei hat die Erhebung insbesondere eine *ermittelnde Intention,* da ein konzeptioneller Rahmen für den ViU bereits besteht. Mithilfe der Befragung sollen für diesen Rahmen Wissensbestände der Interviewten *informatorische* Erkenntnisse liefern, welche sich aus dem vorliegenden Forschungskontext ergeben (Lamnek, 2010).

Die Befragung wird mithilfe eines Leitfadens durchgeführt, welcher *standardisierte* und *nicht-standardisierte* Elemente enthält. Die inhaltlichen Elemente operationalisieren die ersten beiden Fragestellungen der Erhebung (siehe Kapitel 4.1) für die Interviews. Demnach hat der Interviewleitfaden zwei Befragungsebenen, welche sich auf die Zieldimensionen (FS 1) und Arbeitsressourcen (FS 2) beziehen. Die Befragung ist in diese zwei thematischen Einheiten unterteilt. Zunächst soll anhand des bestehenden Portfolios der Zieldimensionen (siehe Tabelle 2) eine gemeinsame Verständnisgrundlage geschaffen werden. Dabei werden die einzelnen Zieldimensionen mithilfe von niederschwellig formulierten Aussagen auf einem digitalen Whiteboard dargestellt (siehe Anhang A – Interviewleitfaden) und anschließend gemeinsam anhand der subjektiven Dienstleistungserfahrungen der Interviewten besprochen. Der zweite Teil der Befragung ist vornehmlich narrativ gestaltet. Die Befragten werden dazu aufgefordert zu beschreiben, wie sie beurteilen, ob eine spezifische Zieldimension für sie persönlich erreicht wurde. Hier wird bewusst auf die Verwendung des Begriffes der Ressource verzichtet und anstatt dessen offen nach den Arbeitsbedin-

gungen gefragt. Dabei werden die einzelnen Dienstleistungserfahrungen getrennt voneinander besprochen und lediglich die zuvor als relevant identifizierten Zieldimensionen untersucht. Eine genaue Beschreibung des Ablaufs ist dem Interviewleitfaden in Anhang A zu entnehmen.

Allgemein soll der Leitfaden als struktureller Rahmen für das Interview gesehen werden. Somit wird die Vergleichbarkeit der Interviews erhöht und sichergestellt, dass alle relevanten Themengebiete angesprochen werden. Die konkrete Abfolge sollte von der interviewenden Person jedoch flexibel an die Gesprächssituation angepasst werden. Neben den standardisierten Elementen des Interviews, sind die Fragen im Leitfaden *offen* formuliert und größtenteils als Erzählaufforderung zu verstehen. Diese offene Formulierung soll dem Interviewten eine autonome Priorisierung persönlich relevanter Aspekte zu dem betrachteten Gesprächsgegenstand erlauben.

4.3.2. Stichprobe

Wie bei qualitativen Erhebungen üblich (Lamnek, 2010), verfolgt auch die vorliegende Erhebung bei der Wahl der Stichprobe keine Ansprüche der Repräsentativität im statistischen Sinn. Vielmehr soll die gewählte Stichprobe im inhaltlichen Sinne für das Erhebungsziel repräsentativ sein, weshalb ein qualitativer Stichprobenplan verfolgt wird (Kelle & Kluge, 1999; Lamnek, 2010). Dieser sieht vor, die relevanten Merkmale der Stichprobe im Voraus zu definieren und darauf aufbauend die Stichprobe auszuwählen (Lamnek, 2010).

Das Erhebungsziel beläuft sich auf eine inhaltliche Spezifikation des ViU von forschungsnahen Dienstleistungen. Da der ViU als ein individuell wahrgenommenes Wertkonstrukt verstanden wird, müssen demnach Forschende befragt werden, welche derartige Dienstleistungen (siehe Kapitel 2.1) bereits in Anspruch genommen haben. Darüber hinaus stellt der Begriff forschungsnahe Dienstleistungen ein Überbegriff verschiedener Dienstleistungsoptionen dar. Um sich in der Untersuchung dieser Variabilität zu nähern und zusätzlich die Datendichte innerhalb einer einzelnen Befragung zu erhöhen, sollten die Interviewten von zwei Erfahrungen mit unterschiedlichen Dienstleistungen berichten können. Des Weiteren werden die Dienstleistungen immer im Kontext eines Forschungsverbundes erbracht. Dieses (soziale) Umfeld hat nach der SD-Logik Einfluss auf die Wahrnehmung des ViU. Diesem Einflussfaktor kann insofern Rechnung getragen werden, dass die Stichprobe Proband*innen aus unterschiedlichen Projektumfeldern beinhaltet. Um dennoch eine adäquate Datenmenge innerhalb eines (sozialen) Umfeldes zu generieren, wird eine proportional verteilte Stichprobe aus zwei unterschiedlichen Verbundprojekten angestrebt.

Für die so festgelegte Stichprobe der empirischen Untersuchung können zehn Forschende aus zwei unterschiedlichen Forschungsverbünden gewonnen werden. Der erste Forschungsverbund (Forschungsverbund A) setzt sich aus acht einzelnen Verbundprojekten zusammen mit einem dazugehörigen Metaprojekt. Das Metaprojekt stellt ein wissenschaftliches Begleitforschungsprojekt dar, welches die Erkenntnisse der Förderlinie wissenschaftlich weiterentwickelt und verbreitet. Für diese Tätigkeiten steht das Metaprojekt im engen Austausch mit den Projekten und initiiert für sie in dem Zusammenhang auch koordinative sowie begleitende Angebote. Beispielsweise werden so forschungsnahe Dienstleistungen in Form von Austauschterminen zwischen den einzelnen Verbundprojekten angeboten sowie individuelle Workshops zu methodischen und theoretischen Grundlagen. Der zweite Forschungsverbund (Forschungsverbund B) ist eine interdisziplinäre Forschungspartnerschaft zwischen sechs Kooperationsmitgliedern, hinter welchen größere Organisationen wie (Fach-)Hochschulen und außeruniversitäre Forschungseinrichtungen stehen. Hier werden seitens der Projektkoordination und -leitung forschungsnahe Dienstleistungen in Form von Austauschtreffen, einer organisatorischen und verwaltungstechnischen Begleitung der Projektinstitutionen sowie partnerübergreifende Vernetzungs- und Austauschtreffen angeboten. Von diesen beiden Verbünden werden insgesamt zehn Forschende befragt. Sechs dieser Forschenden haben eine Professur inne während drei als wissenschaftliche Mitarbeiter*innen in den Projekten arbeiten, von denen zwei derzeitig promovieren. Die befragten Forschenden stammen aus insgesamt acht unterschiedlichen Forschungsinstitutionen. Auf weitere personenbezogenen statistischen Angaben wird an dieser Stelle bewusst verzichtet, um die Anonymität der Befragten zu wahren.

4.3.3. Auswertungsstrategie

Die Auswertung der Interviews wird mithilfe einer qualitativen Inhaltsanalyse nach Mayring (1983) vorgenommen. Dabei wird das Kommunikationsmaterial hinsichtlich der explizit getroffenen Aussagen der Befragten, sogenannten *manifesten Kommunikationsinhalten*, analysiert (Lamnek, 2010). Das Vorgehen der Auswertung folgt einem festgelegten Ablaufmodell. Nachfolgend ist das inhaltsanalytische Ablaufmodell abstrahiert dargestellt (siehe Abbildung 7). Im Detail untergliedert sich dieses in weitere Vorgehensstufen, welche im Zuge der Durchführungsbeschreibung in Kapitel 4.4 näher erläutert werden.

Als Ausgangsmaterial werden die Transkripte der Interviews verwendet (siehe Kapitel 4.4.1). Die Analyse des Materials folgt den zwei Fragestellungen FS 1 und FS 2 (siehe Kapitel 4.4.2), aus welchen sich unterschiedliche Analysetechniken ergeben. Zum einen sollen Erkenntnisse bzgl. der Zieldimensionen von forschungsnahen Dienstleistungen gewonnen werden. Ein Kategoriensystem hierfür konnte im Vorfeld aus dem Stand der Forschung abgeleitet werden (siehe Kapitel 3.1). Demnach wird hier eine inhaltliche Strukturierung vorgenommen, welche eine deduktive inhaltliche Spezifikation der Kategorien anhand des Materials vorsieht (siehe Kapitel 4.5.1). Zum anderen sollen aktivierte Arbeitsressourcen kategorisiert werden, welche für die Forschenden zur Zielerreichung während der Inanspruchnahme forschungsnaher Dienstleistungen bereitgestellt werden. Diese Analyse wird induktiv vorgenommen und folgt einem festgelegten Schema (siehe Kapitel 4.5.2).

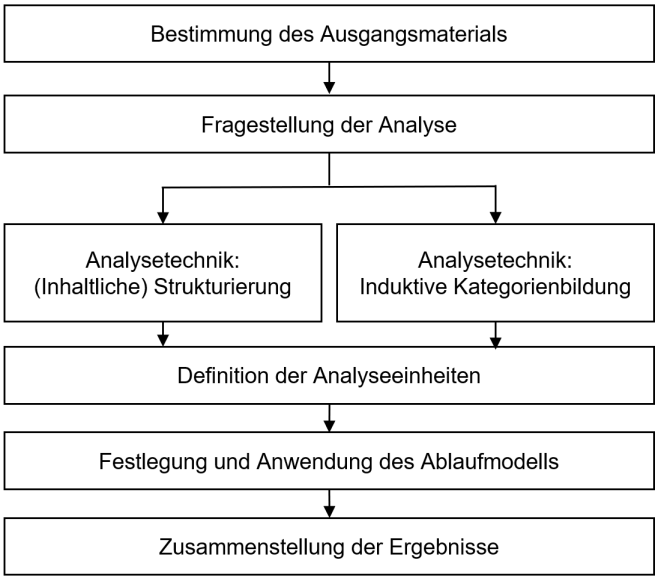

Abbildung 7: Inhaltsanalytisches Ablaufmodell

Quelle: Eigene Darstellung, in Anlehnung an Mayring (2015)

4.3.4. Gütekriterien

Um die Qualität der Erkenntnisgewinnung im Zuge der empirischen Untersuchung zu gewährleisten, wurde das methodische Vorgehen unter Berücksichtigung der Gütekriterien der sozialwissenschaftlichen Methodenlehre gestaltet und erweitert. Es wird nach den Maßen der *Reliabilität* und *Validität* unterschieden (Mayring, 2015, S. 123).

Die Zuverlässigkeit bzw. *Reliabilität* gilt als das Maß der „Stabilität und Genauigkeit der Messung sowie die Konstanz der Meßbedingungen“ (Friedrichs, 1973, S. 102). Die Reliabilität wird während der empirischen Datenerhebung sowie deren Auswertung über die Einhaltung verschiedener Zielvorgaben erreicht. Mayring (1996) formuliert im Zuge dessen drei allgemeine Prüfaspekte. Das Verfahren der Datenerhebung und Analyse soll detailliert dargestellt (*Verfahrensdokumentation*) und der interpretative Vorgang festgehalten (*argumentative Interpretationsabsicherung*) werden. Darüber hinaus soll die Auswertung systematisch vorgehen (*Regelgeleitetheit*). Durch die Übernahme des von Mayring (2015) beschrieben Auswertungskonzepts der qualitativen Inhaltsanalyse (Kapitel 4.3.3) werden die drei Prüfaspekte sichergestellt. Die genaue Beschreibung des Auswertungsprozesses erfolgt in den beiden folgenden Kapiteln 4.4 und 4.5.

Die *Validität* gibt an, „ob das gemessen wird, was gemessen werden sollte“ (Friedrichs, 1973, S. 100). Dabei ist nicht die Validität das Ziel qualitativer Forschung, sondern eine Validierung als Prozess (Lamnek, 2010). Gemäß diesem Ziel wird die vorliegende empirische Analyse auf zwei Ebenen mithilfe von Methoden der Triangulation validiert. Reichertz (2014) verweist darauf, dass die Validität der Datenanalyse erhöht werden kann, wenn das Material von mehreren Forschenden unabhängig interpretiert und die Ergebnisse anschließend diskutiert werden. Aus diesem Grund wird die qualitative Inhaltsanalyse in Kooperation und Rücksprache mit Anna Lux, der Betreuerin der vorliegenden Arbeit, durchgeführt. Darüber hinaus werden die theoretische Initiierung der empirischen Analyse, deren Durchführung sowie die Ergebnisse mit der Erstprüferin, Prof. Dr. Susanne Robra-Bissantz[23], abschließend besprochen und validiert.

4.4. Durchführung der Analyse

Folgend wird die Durchführung der Datenerhebung und Auswertung beschrieben. Dabei dient das Ablaufmodell der qualitativen Inhaltsanalyse zur Dokumentation des Verfahrens (siehe Abbildung 8, detaillierend zu Abbildung 7). In Kapitel 4.4.1 wird das Ausgangsmaterial bestimmt. Dabei wird unter anderem der Datenerhebungsprozess (Interviewdurchführung) beschrieben und das so gesammelte Material spezifiziert. Darauf werden die Fragestellungen der empirischen Analyse, welche sich aus den theoretischen Vorüberlegungen in Kapitel 3 ergeben, erneut kurz vorgestellt (siehe Kapitel 4.4.2). Aus den Fragestellungen leiten sich zwei unterschiedliche Analysetechniken zur Auswertung des Materials ab, deren Ablauf in Kapitel 4.5 beschrieben wird.

4.4.1. Bestimmung des Ausgangsmaterials

Zur Bestimmung des Ausgangsmaterials wird folgend die Festlegung des Materials, die Analyse der Entstehungssituation sowie formale Charakteristika des Materials beschrieben (Mayring, 2015).

[23] Als Forschende entspricht sie den Anforderungen der Strichprobenkriterien und ist damit im Sinne der vorliegenden Arbeit auch als Anspruchsgruppe forschungsnaher Dienstleistungen zu verstehen. Darüber hinaus kann sie auf umfassende Kenntnisse und Forschungserfahrung in dem Bereich der SD-Logik und der Theorie des ViU zurückgreifen.

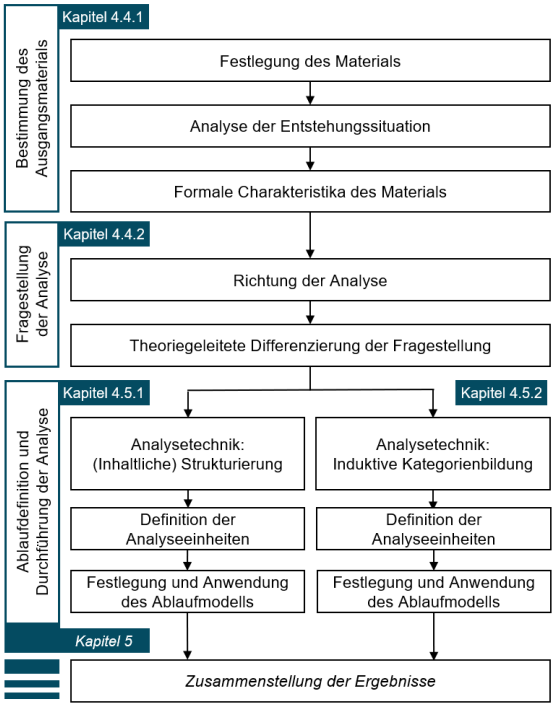

Abbildung 8: Durchführungsbeschreibung

Quelle: Eigene Darstellung, in Anlehnung an Mayring (2015)

Tabelle 4: Ausgangsmaterial der Analyse

Kennung der Interviewten	Anzahl der besprochenen Dienstleistungen	Gesamte Dauer des Interviews [Stunden]	Dauer des transkribierten Teils[Stunden]	Umfang der Transkription [Seiten]
A_1	2	01:48:42	01:29:57	28
A_2	2	01:14:50	01:11:08	24
A_3	2	01:20:16	00:59:52	20
Summe A	**6**	**04:23:48**	**03:40:57**	**72**
B_1	1	01:06:12	00:54:25	19
B_2	2	00:59:12	00:55:46	20
B_3	2	01:29:30	01:17:54	30
B_4	1	01:39:50	01:10:54	20
B_5	1	01:40:48	01:16:55	25
B_6	1	00:51:02	00:46:45	20
Summe B	**8**	**07:47:32**	**06:28:40**	**134**
Summe	**14**	**12:11:20**	**10:09:37**	**206**

Festlegung des Materials

Für die Auswertung werden die Transkriptionen von neun Einzelinterviews verwendet[24]. Zusätzlich wird das Portfolio der ViU-Dimensionen (Tabelle 2), welche als Verständnisgrundlage in den Interviews vorgestellt wurden, in die Auswertung impliziert. Tabelle 4 listet das Ausgangsmaterial hinsichtlich der Dauer der Interviews sowie des Umfangs der schriftlichen Transkription auf. Im Zuge der Interviews werden insgesamt 14 Dienstleistungserfahrungen besprochen. Innerhalb der Interviews können insbesondere aus zeitlichen Gründen nicht immer zwei Erfahrungen mit forschungsnahen Dienstleistungen besprochen werden. Aus diesem Grund

[24]Insgesamt wurden zehn Interviews durchgeführt. Ein Interview konnte nicht in die Auswertung impliziert werden, da die dort besprochene Dienstleistungserfahrung nicht dem, in der vorliegenden Arbeit definierten, Dienstleistungsverständnis für forschungsnahe Dienstleistungen entspricht.

umfasst die Stichprobe des zweiten Forschungsverbundes (B) mehr befragte Personen. Die beschriebenen Dienstleistungserfahrungen sind mit sechs (Forschungsverbund A) zu acht (Forschungsverbund B) annähernd gleich verteilt (siehe Kapitel 4.3.2).

Es werden jeweils die Textpassagen untersucht, innerhalb derer Zieldimensionen und Beurteilungsaspekte dieser Dimensionen im Sinne von aktivierten Arbeitsressourcen anhand von Beispielen forschungsnaher Dienstleistungen besprochen wurden.

Analyse der Entstehungssituation

Die Interviewteilnehmer wurden durch ein direktes Anschreiben über E-Mail kontaktiert und dabei von der Autorin direkt angesprochen (siehe Anhang B – Einladungsschreiben).

Die Teilnahme der interviewten Personen war freiwillig. Das Interview erfolgte bilateral zwischen der befragten Person und der Autorin der vorliegenden Arbeit. Dabei wurde ein Leitfaden eingesetzt (siehe Anhang A – Interviewleitfaden), welcher die inhaltlichen Aspekte der Interviews festlegt. Die konkrete Formulierung der Fragen und der Ablauf der Befragung wurde nach Ermessen der Interviewerin angepasst. Die Interviews wurden über eine Videokonferenz innerhalb des Zeitraums vom 08.03.2021 bis 24.03.2021 durchgeführt.

Eine kurze Einschätzung der Interviewsituation seitens der Interviewerin wird im Nachgang des Termins formuliert.

Formale Charakteristika des Materials

Die Interviews wurden aufgezeichnet und anschließend semantisch transkribiert. Dafür wurden im Vorfeld Hinweise zur Transkription festgelegt. Diese orientieren sich an einer Formatvorlage zur Interviewtranskription von Mayring (2015) (siehe Anhang C – Transkriptionshinweise).

4.4.2. Fragestellung der Analyse

Richtung der Analyse

Durch die Interviews sollen die Befragten dazu angeregt werden, ihre persönlichen Bedürfnisse und Präferenzen hinsichtlich der Unterstützungsangebote in Form von forschungsnahen Dienstleistungen zu berichten. Dabei dient ein zuvor definiertes Portfolio der Zieldimensionen einer gemeinsamen Verständigungsgrundlage. Anhand dessen können die Interviewten Ihre subjektiven Erfahrungen mit derartigen Dienstleistungsangeboten zielgerichtet innerhalb der theoretischen Vorkenntnisse artikulieren.

Theoriegeleitete Differenzierung der Fragestellung

Das Interviewmaterial beinhaltet Aussagen zu den Zielen, welche bei der Inanspruchnahme von forschungsnahen Dienstleistungen verfolgt werden. Darüber hinaus formulieren die Befragten, anhand welcher Aspekte sie für sich beurteilt haben, ob ein bestimmtes Ziel für sie persönlich erreicht wurde. Aus diesen Aussagen können die wertschaffenden aktivierten Arbeitsressourcen abgeleitet werden.

Die Fragestellungen der Analyse leiten sich aus dem theoretischen Konzeptionierungsmodell ab (siehe Kapitel 3.3 und Kapitel 3.4) ab, und werden bereits im Zuge der Zielsetzung der empirischen Analyse (siehe Kapitel 4.1) definiert:

FS 1: Welche Zieldimensionen sind für Forschende bei der Inanspruchnahme forschungsnaher Dienstleistungen relevant und wie lassen sich diese charakterisieren?

FS 2: Welche aktivierten Arbeitsressourcen ermöglichen im Forschungskontext das Erreichen der Zieldimensionen forschungsnaher Dienstleistungen?

Die zusammenführende Analyse, welche durch die dritte Fragestellung (siehe Kapitel 4.1) gefordert wird, stellt die Ergebnisse der Fragestellungen FS 1 und FS 2 zueinander in Relation. Dadurch ist keine systematische Auswertung im Sinne einer qualitativen Inhaltsanalyse notwendig, wie sie für die ersten beiden Fragestellungen folgend beschrieben wird.

4.5. Qualitative Inhaltsanalyse zur Auswertung

Die Auswertung der Interviews soll mithilfe einer qualitativen Inhaltsanalyse erfolgen. Hierfür existieren zahlreiche Analyseformen. Mayring (2015) unterscheidet drei Grundformen des Interpretierens. Für die vorliegenden Fragestellungen (siehe Kapitel 4.4.2) wird das Material (siehe Kapitel 4.4.1) *zusammenfassend* interpretiert. Dabei werden zwei unterschiedlichen Analysetechniken eingesetzt, im Zuge derer separate Ablaufmodelle für das systematische Vorgehen definiert werden (siehe Abbildung 9).

4.5.1. Zieldimensionen: Ablauf der inhaltlichen Strukturierung

Die erste Fragestellung (FS 1) bezieht sich auf die Konkretisierung der Zieldimensionen hinsichtlich deren Relevanz und Inhalt für forschungsnahe Dienstleistungen. Dabei wird das zuvor entwickelte Portfolio möglicher Zieldimensionen als deduktives Kategoriensystem verwendet und anschließend eine inhaltliche Strukturierung der Dimensionen nach Mayring (2015) durchgeführt. Dabei werden Textstellen, welche der inhaltlichen Charakterisierung der ViU-Dimensionen forschungsnaher Dienstleistungen dienen, extrahiert, paraphrasiert und zusammenfassend dargestellt. Für dieses Vorgehen werden folgende Analyseeinheiten festgelegt (Mayring, 2015, S. 88):

- Kodiereinheit: Bedeutungstragende Elemente im Transkriptionsmaterial
- Kontexteinheit: Interviewpassagen, in welcher die Zieldimensionen einer Person beschrieben werden, in Verbindung mit den theoretisch abgeleiteten ViU-Dimensionen
- Auswertungseinheit: Alle Interviewtranskriptionen (siehe Kapitel 4.4.1)

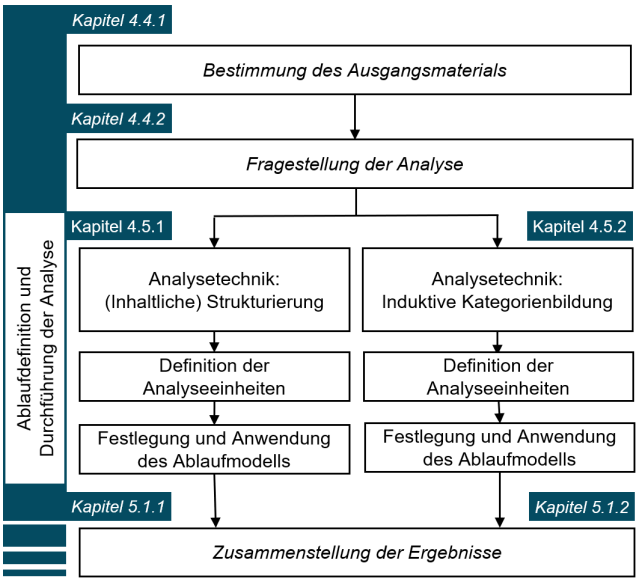

Abbildung 9: Ablaufdefinition und Durchführung der Analyse

Quelle: Eigene Darstellung, in Anlehnung an Mayring (2015)

Das analytische Vorgehen ist nach dem Ablaufmodell zur inhaltlichen Strukturierung vorgegeben (siehe Abbildung 10).

Die theoriegeleitete Festlegung der Kategorien wird innerhalb der Vorüberlegungen hinsichtlich des aktuellen Forschungsstandes beschrieben (siehe Kapitel 3.1). Die ViU-Dimensionen (siehe Tabelle 2) stellen für die vorliegende Analyse die inhaltlichen Hauptkategorien dar und können mithilfe von Zielen für die Befragten operationalisiert werden. Aus diesem Grund wird im Interview danach gefragt, welche Ziele die Interviewten bei der Inanspruchnahme forschungsnaher Dienstleistungen verfolgen und wie sie für sich ein solches Ziel charakterisieren. Dazu werden die Dimensionen des zuvor festgelegten Kategoriensystems als Verständigungsgrundlage während der Interviews den Befragten vorgestellt (siehe 4.3.1 und Anhang A – Interviewleitfaden).

Das Material wird nach Textpassagen durchlaufen, welche sich inhaltlich auf die Zieldimensionen beziehen. Die Fundstellen werden dabei fortlaufend im Transkriptionsmaterial hervorgehoben und nach dem Schema „Interviewtenkennung – Dienstleistungsnummer – fortlaufende Nummer“ nummeriert. Anschließend werden die Fundstellen extrahiert und den Kategorien (Zieldimensionen) zugeordnet. Hierbei wird zunächst die intuitive Zuordnung der Befragten befolgt, die ihre Beschreibungen innerhalb des Interviews auf eine konkrete Dimension bezogen haben. Bei definitorischen Abweichungen der Fundstellen zu den verbundenen Dimensionen wird diese erste Zuordnung angepasst.

Anhand der Fundstellen wird anschließend das Kategoriensystem für forschungsnahe Dienstleistungen entwickelt. Dabei werden die Zieldimensionen, welche von den Befragten als nicht relevant (siehe Kapitel 4.6.1, Bildung von Schlüsselkategorien) eingeordnet wurden aus dem Portfolio expliziert und Dimensionen bei hohen inhaltlichen Überscheidungen zusammengeführt. Darüber hinaus werden die Definitionen für den Anwendungsfall forschungsnahe Dienstleistungen weiter angepasst. Um einen praktischen Bezug zu den abstrakten Dimensionen herzustellen, werden außerdem repräsentative Aussagen für die einzelnen Dimensionen aus den Fundstellen paraphrasiert (Lamnek, 2010). Dabei wird den Regeln der Paraphrasierung (Z1-Regeln) nach Mayring (2015, S. 72) gefolgt:

> „Z1.1: Streiche alle nicht (oder wenig) inhaltstragenden Textbestandteile wie ausschmückende, wiederholende, verdeutlichende Wendungen!
>
> Z1.2: Übersetze die inhaltstragenden Textstellen auf eine einheitliche Sprachebene!
>
> Z1.3: Transformiere sie auf eine grammatikalische Kurzform!“

4.5.2. Arbeitsressourcen: Ablauf der induktiven Kategorienbildung

Zur Beantwortung der zweiten Fragestellung wird das Material mithilfe einer induktiven Kategorienentwicklung analysiert und der Entwicklungsprozess deduktiv vereinfacht durch die Verwendung von Kategorienbezeichnungen aus der Literatur (Gioia, Corley & Hamilton, 2013).

Die zweite Fragestellung (FS 2) ergibt sich aus den theoretischen Vorüberlegungen hinsichtlich der wertdeterminierenden Einflüsse entlang des Wertschöpfungsprozesses in Form von aktivierten Arbeitsressourcen (siehe Kapitel 3.2) und orientiert sich an dem methodischen Ansatz von Hartwig und Jacob (2018). Um die aktivierten Arbeitsressourcen forschungsnaher Dienstleistungen aus dem Material zu

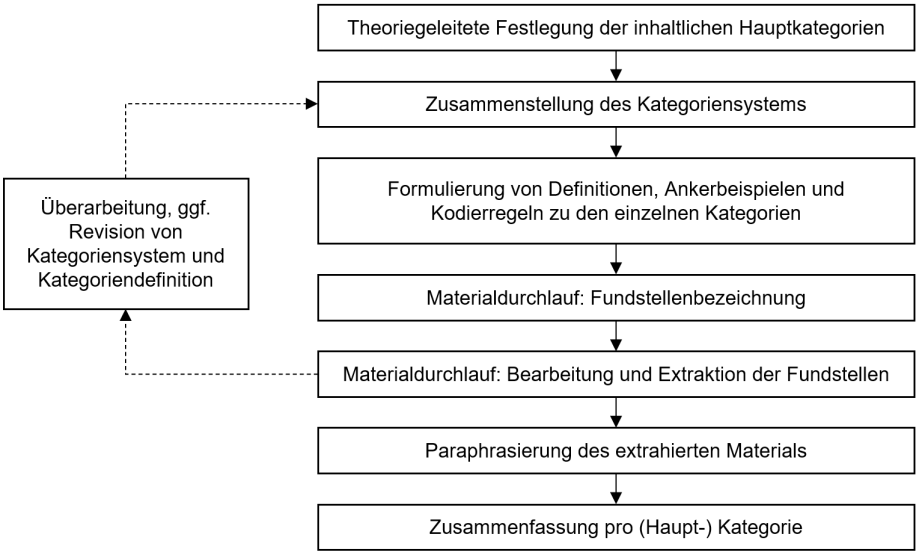

Abbildung 10: Ablaufmodell inhaltlicher Strukturierung

Quelle: Eigene Darstellung, in Anlehnung an Mayring (2015)

extrahieren, werden folgende Analyseeinheiten festgelegt (Mayring, 2015):

- Kodiereinheit: Bedeutungstragende Elemente im Transkriptionsmaterial
- Kontexteinheit: (Zweite) Interviewpassage, in welcher die Beurteilung der Zielerreichung in Form von Arbeitsressourcen beschrieben werden, plus zusätzliche Notizen der Interviewerin (Einschätzung der Interviewsituation)
- Auswertungseinheit: Alle Interviewtranskriptionen (siehe Kapitel 4.4.1)

Die Entwicklung des Kategoriensystems zu den, über die Inanspruchnahme der forschungsnahen Dienstleistungen aktivierten, Arbeitsressourcen folgt dem Prozessmodell zur induktiven Kategorienentwicklung (siehe Abbildung 11).

Das Material, der Gegenstand und das Ziel der Analyse sowie die dahinterstehende Theorie werden in Kapitel 4.4 beschrieben. Das Interviewmaterial wird nach Aussagen selektiert, in welchen die Befragten bedeutungstragende Elemente ihrer Beurteilung für einzelne Ziele beschreiben. Anhand dieser Aussagen sollen Kategorien gebildet werden, welche wertschaffende aktivierte Arbeitsressourcen in dem Kontext der Inanspruchnahme forschungsnaher Dienstleistungen beschreiben.

Die Fundstellen werden dabei fortlaufend im Transkriptionsmaterial hervorgehoben und, wie bei den Zieldimensionen, nach dem Schema „Interviewtenkennung – Dienstleistungsnummer – fortlaufende Nummer“ nummeriert.

Die zu entwickelnden Kategorien sollen ein Abstraktionsniveau aufweisen, welches von den spezifischen und persönlichen Einschätzungen gelöst ist und dadurch personen- sowie situationsübergreifend angewandt werden kann. Dabei werden die Kategorien in Anlehnung an (Klemenz, 2009) hinsichtlich deren Funktionalität klassifiziert. Eine erste Klassifizierung der Fundstellen folgt den bestehenden Kategorien, welche aus der Literatur abgeleitet werden konnten (siehe Kapitel 3.2.1). Fundstellen, welche nicht diesen Kategorien zugeordnet werden können, werden klassifiziert und die Literatur zielorientiert nach passenden Begriffen untersucht. Dieser deduktive Ansatz zur theoriebasierten Unterstützung der induktiven Klassifizierung folgt unter anderem Gioia et al. (2013). Sofern keine übereinstimmenden oder repräsentative Titel aus der Literatur abgeleitet werden können, erfolgt eine induktive Benennung. Das so entwickelte Kategoriensystem wird während des Auswertungsprozess mit der Betreuerin der vorliegenden Arbeit (Anna Lux) abgestimmt und abschließend mit der Erstprüferin (Prof. Dr. Susanne Robra-Bissantz) validiert.

Wie auch bei den Zieldimensionen, werden zu den Kategorien der aktivierten Arbeitsressourcen repräsentative Aussagen aus den Fundstellen nach den Z1-Regeln nach (Mayring, 2015) paraphrasiert um einen praktischen Bezug zu den abstrakten Kategorien herzustellen (Lamnek, 2010).

4.6. Ergänzende Auswertung

Im Anschluss an die qualitative Inhaltsanalyse wird, über das Ablaufmodell hinaus (siehe Kapitel 4.5), eine quantitative Analyse der Zieldimensionen und aktivierten Arbeitsressourcen vorgenommen sowie eine zusammenführende Betrachtung dieser beiden Auswertungsebenen.

4.6.1. Quantitative Auswertung

Die quantitative Auswertung soll die Ergebnisse auf relevante Kategorien für den Anwendungsfall reduzieren und

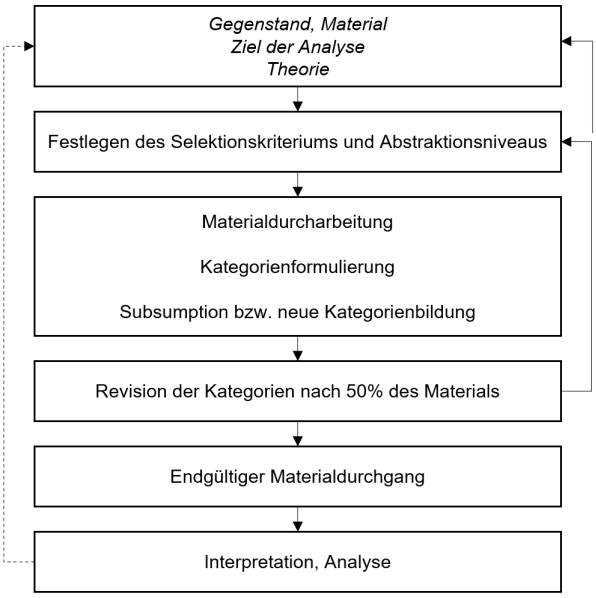

Abbildung 11: Prozessmodell induktiver Kategorienentwicklung

Quelle: Eigene Darstellung nach (Mayring, 2015)

eine erste Orientierung für mögliche Schwerpunktsetzungen von forschungsnahen Dienstleistungen liefern.

Hierfür wird die Nennungshäufigkeit der Zieldimensionen und Kategorien der Arbeitsressourcen untersucht. Innerhalb der Interviews haben die Befragten unterschiedlich stark Bezug zu den Begrifflichkeiten gefunden. Dadurch unterscheidet sich die Anzahl der inhaltstragenden Aussagen innerhalb der einzelnen Dimensionen und Kategorien. Aus diesem Grund wird für die quantitative Analyse die Anzahl der Fundstellen auf eine Nennung pro Dimension/ Kategorie und Dienstleistungserfahrung reduziert (Pfisterer, 2017).

Methodische Hinweise für das Vorgehen zur quantitativen Auswertung bieten Studien zur empirischen Analyse von persönlichen Konstrukten, welche die Repertory-Grid-Methode anwenden (Goffin, Lemke & Szwejczewski, 2006; Goffin, Raja, Claes, Szwejczewski & Martinez, 2012; Pfisterer, 2017). Diese Methode wird häufig in den empirischen Studien zu den ViU-Dimensionen verwendet (siehe Kapitel 3.1). Zur Analyse sogenannter Schlüsselkategorien führt Goffin et al. (2006) eine Entscheidungsregel ein, welche festlegt dass eine Kategorie als relevant für den Anwendungsfall zu werten ist, wenn diese von einem bestimmten Prozentsatz der Befragten genannt wird. Innerhalb der Repertory-Grid Literatur wird dieser prozentuale Grenzwert auf unterschiedlichen Niveaus festgelegt (Goffin et al., 2006). Ein typischer Wert hierfür liegt bei 25 % Nennungshäufigkeit (Goffin et al., 2006, 2012; Lemke et al., 2011; Pfisterer, 2017). Die vorliegende Arbeit orientiert sich an diesem Grenzwert. Die Zählung der Nennungshäufigkeit wird jedoch insofern angepasst, dass diese sich auf die einzelnen Dienstleistungserfahrungen ($\sum$ 14) bezieht und nicht auf eine interviewte Person ($\sum$ 9). Für die quantitative Analyse der aktivierten Arbeitsressourcen wird der Grenzwert von 25 % übernommen, während er für die Analyse der Zieldimensionen angepasst werden muss. Das begründet sich daraus, dass die Fundstellen, welche zur Bildung eines Kategoriensystems für Arbeitsressourcen dienen, explorativ ermittelt werden, indem die Befragten implizit erzählen, anhand welcher Arbeitsbedingungen sie das Erreichen eines Ziels bewerten. Für die Zieldimensionen werden jedoch innerhalb der Interviews bereits die möglichen Dimensionen vorgegeben. Aus diesem Grund ist anzunehmen, dass innerhalb der Untersuchungsebene der Zieldimensionen die relative Dichte der Fundstellen innerhalb des Kategoriensystems höher ist als bei einem explorativen Vorgehen wie bei den Arbeitsressourcen (oder persönlichen Konstrukten im Zuge von Repertory-Grid Studien). Um diesen Einfluss zu berücksichtigen, wird der Grenzwert auf ein Drittel (mindestens fünf Nennungen) für die Zieldimensionen angepasst.

4.6.2. Zusammenführende Betrachtung

Abschließend sollen die Ergebnisse der qualitativen Inhaltsanalyse beider Untersuchungsebenen zusammenführend dargestellt werden und somit die dritte Fragestellung der empirischen Analyse beantwortet werden (siehe Kapitel 4.1).

FS 3: Welche aktivierten Arbeitsressourcen aus dem Dienstleistungsprozess tragen aus Sicht der Forschenden zum Erreichen der Zieldimensionen und somit zur Wertschöpfung innerhalb der ViU-Dimensionen bei?

Da der ViU sich aus dem Erreichen der Zieldimensionen ergibt, welches durch die Verfügbarkeit aktivierter Arbeitsressourcen moderiert wird, sollen die Ergebnisse ergänzend zueinander in Relation gesetzt werden. Dadurch,

dass innerhalb der Interviews die aktivierten Arbeitsressourcen anhand der Kriterien zum Erreichen der einzelnen Zieldimensionen erfragt werden, sind die Fundstellen der Arbeitsressourcen stets einer Zieldimension zugeordnet. Diese Verbindung der Ergebnisse bietet eine erste Orientierung dafür, welche Ressourcen seitens der anbietenden Akteure oder aus dem Umfeld integriert werden müssen, um die angestrebten Ziele der Forschenden zu erreichen und somit für sie Wert zu schaffen.

5. Ergebnisse

Durch die Interviews können 14 Erfahrungen mit forschungsnahen Dienstleistungen untersucht werden. Zwölf der Erfahrungen lassen sich in vier vergleichbare Dienstleistungsangebote einordnen (siehe Tabelle 5). Das gewonnene Datenmaterial beinhaltet insgesamt 264 Fundstellen (durchschnittlich 18,8 pro Dienstleistungserfahrung) für die inhaltliche Strukturierung der Zieldimensionen und 234 (durchschnittlich 16,7 pro Dienstleistungserfahrung) für die induktive Kategorienentwicklung der Arbeitsressourcen. Die Fundstellen der Arbeitsressourcen sind stets einer Zieldimension zugeordnet.

5.1. Zieldimensionen

Im Fokus der Auswertung der Zieldimensionen steht die inhaltliche Strukturierung der Fundstellen (siehe Kapitel 4.5.1). Hierfür wird das theoretisch abgeleitete Portfolio (siehe Tabelle 2) möglicher Zieldimensionen für forschungsnahe Dienstleistungen anhand der empirischen Ergebnisse spezifiziert (siehe Kapitel 5.1.1). Darüber hinaus werden die so abgeleiteten Dimensionen hinsichtlich deren Nennungshäufigkeit untersucht (siehe Kapitel 5.1.2)

5.1.1. Inhaltlich strukturierte Zieldimensionen

Welche Ziele die Forschenden bei der Inanspruchnahme der einzelnen Dienstleistungen anstreben, wurde intuitiv von den Befragten innerhalb der Interviews formuliert. Die zu den einzelnen Dimensionen getätigten Aussagen werden zur inhaltlichen Charakterisierung ausgewertet. Im Zuge dessen werden fünf Dimensionen mit anderen Dimensionen aufgrund hoher inhaltlicher Überschneidungen zusammengeführt. Drei Dimensionen werden aus dem Kategoriensystem expliziert[25]. Das theoretisch abgeleitete Kategoriensystem kann somit von 21 Dimensionen auf zwölf Dimensionen für forschungsnahe Dienstleistungen konkretisiert werden. Aus Gründen der Übersichtlichkeit ist in Tabelle 6 das Ergebnis der inhaltlichen Strukturierung zusammengefasst. Der Prozess der inhaltlichen Strukturierung wird anschließend näher erläutert und legitimiert.

Tabelle 6 umfasst das abgeleitete Portfolio möglicher Zieldimensionen forschungsnaher Dienstleistungen. Für die einzelnen (zusammengeführten) Dimensionen werden die zuvor abgeleiteten Definitionen (siehe Tabelle 2) anhand der Erkenntnisse aus den Interviews angepasst. Darüber hinaus werden die Dimensionen anhand häufig thematisierter Aspekte und mithilfe von exemplarischen Aussagen aus den Interviews für forschungsnahe Dienstleistungen charakterisiert.

Prozessverbesserung

Die Dimensionen *Vermeiden von Stillstandzeiten, Produktivität* und *Prozessverbesserung* werden zusammengefasst. Die *Prozessverbesserung* bezieht sich auf kollektive Prozesse auf der Ebene des Verbundes oder der einzelnen Projekte. Genannt werden hier insbesondere die Ermöglichung und Vereinfachung der kollektiven Prozesse (Austausch, Abstimmung) sowie deren Sichtbarmachung. In diesem Zusammenhang verweisen die Befragten auch auf die Zweckdienlichkeit einer solchen Transparenz für die *Vermeidung von Stillstandzeiten*, weshalb die Dimension der *Prozessverbesserung* zugeordnet wird.

> *„. . . , dass Arbeitsprozesse transparenter und vernetzter werden [. . .]. Das ist ein geeignetes Mittel, um Unproduktivität oder auch Leerphasen in Projekten sichtbar zu machen.“*

Wie die zitierte Aussage andeutet, wird in den Interviews an dieser Stelle auch häufig der Bezug zur Dimension der *Produktivität* aufgebaut. Allgemein finden sich innerhalb der Dimension der *Produktivität* insbesondere Aussagen, welche sich auf eine Verbesserung der kooperativen (Austausch-) Prozesse beziehen.

> *„Wenn man so einen Austausch hat, dann kommt ja auch immer wieder Bewegung rein und es bleibt produktiv.“*

Aus diesem Grund wird die Dimension *Produktivität* ebenfalls der *Prozessverbesserung* zugeordnet. Vereinzelte Aussagen, welche sich auf die Steigerung der Produktivität individueller Arbeitsprozesse beziehen, können der Dimension der *Aufgabenvereinfachung* zugeordnet werden, welche den Aspekt der persönlichen Aufwandsreduktion für die Forschenden beinhaltet.

Aufgabenvereinfachung

Die Dimension der *Aufgabenvereinfachung* und deren Definition kann durch die Aussagen der Interviews für forschungsnahe Dienstleistungen bestätigt werden. Die Dimension wird insbesondere auf die individuelle Ebene bezogen. Hierbei steht der Aspekt der persönlichen Aufwandsreduktion im Vordergrund. Dies kann durch diverse Angebote erreicht werden, welche bspw. den Zugang zu neuen fachlichen und methodischen Aufgaben erleichtern, oder durch eine zentrale und übersichtliche Informationsbereitstellung,

[25] Der Ausschluss erfolgt insbesondere aufgrund unzureichender Relevanz (siehe Kapitel 4.6.1). Eine Dimension (Motivation) wurde aufgrund der Inhalte in den erhobenen Daten und neuen theoretischen Bezügen aus dem Kategoriensystem nachträglich ausgeschlossen (vgl. S. 60).

Tabelle 5: Datenmaterial nach Dienstleistungen

Dienstleistungen	Anzahl Dienstleistungs-erfahrungen	Anzahl Fundstellen Zieldimensionen	Anzahl Fundstellen Arbeitsressourcen
Angebot eines digitalen schwarzen Bretts	2	20 23	18 23
Regelmäßige Austausch-angebote (für Projekt-institutionen oder Projektpartner)	3	17 20 14	17 19 12
Unregelmäßige Austauschangebote (verbundübergreifend, teilweise mit Externen)	4	13 26 22 20	9 20 20 17
Workshopangebote (fachliche und methodische Einführung)	3	16 18 17	13 16 16
Sonstige	2	16 22	16 18
Summe	**14**	**264**	**234**
Durchschnitt	**1**	**18,8**	**16,7**

die zur individuellen Strukturierung von Arbeitsabläufen verhelfen kann.

Schnelle Problemlösung

Auch die theoretisch abgeleitete Dimension *schnelle Problemlösung* kann für forschungsnahe Dienstleistungen bestätigt und übernommen werden. Innerhalb dieser Dimension wird insbesondere auf die übergeordnete Notwendigkeit des Aufdeckens und der Sichtbarmachung von Problemen sowie das Potential einer gemeinsamen Problemlösung verwiesen[26].

Sozialer Wert

Die Dimension des *sozialen Wertes* kann ebenfalls durch die Schilderung diverser Begebenheiten aus dem sozialen Umfeld für forschungsnahe Dienstleistungen bestätigt werden. Eine befragte Person stellt einige Aspekte, welche ebenfalls von anderen genannt werden, zusammenfassend dar:

> *„... eine angenehme Atmosphäre, also ob man sich wertgeschätzt fühlt, ob da eine positive Grundhaltung ist oder auch ein verbindendes Element - meinetwegen auch thematisch."*

Selbstdarstellung

Die Dimension *Selbstdarstellung* kann aus dem theoretischen Portfolio übernommen werden. Aus den Daten der Interviews sollte hervorgehoben werden, dass die Befragten sich jedoch explizit von dem Erreichen persönlicher Anerkennung oder beruflicher Reputation distanzieren.

> *„Das wäre zu selbstvermarktungsmäßig."*

Lediglich im Zuge der Besprechung einer forschungsnahen Dienstleistung wird erwähnt, dass man persönlich Präsenz zeigen will. In den anderen Interviews wird betont, dass man während den Dienstleistungen gegenüber Dritten nicht als Repräsentant der eigenen Person auftritt, sondern die eigene Forschung präsentiert.

> *„Wobei ich das ganz bewusst nicht als Selbstpräsentation sehe, sondern als Austausch von Erfahrungen und Ergebnissen."*

Selbstverwirklichung

Die Dimension *Selbstverwirklichung* wird ebenfalls im Zuge der Interviews für forschungsnahe Dienstleistungen bestätigt. Während der Inanspruchnahme wird erwartet, dass den Forschenden Raum zur Einbringung und Umsetzung persönlicher Wünsche sowie Vorstellungen gegeben wird.

Unsicherheitsbewältigung

Die Dimension *Unsicherheitsvermeidung* wird mit den Dimensionen *Druckreduzierung* und *Vermeidung von Abhängigkeiten* zusammengeführt. Außerdem wird die Benennung in

[26] Anzumerken ist hier, dass es sich nicht um die allgemeine Analyse der Problemsituation und Problemformulierung der Verbundforschung handelt, sondern auf individuelle oder projektbezogene Heraus-forderungen und Problemstellungen, während der Ausübung von Forschungstätigkeiten.

Tabelle 6: Zieldimensionen forschungsnaher Dienstleistungen

Dimensionen (Input)	Neue Dimension	Definition *[Quellen]*	Charakterisierende Aspekte	Beispielhafte (paraphrasierte) Aussagen
Produktivität Vermeiden von Stillstandzeiten Prozessverbesserung	Prozessverbesserung	Verbesserung der kooperativen Arbeitsabläufe auf Projekt- und Verbundebene *[In Anlehnung an 1 und 2]*	Transparenz, Rahmensetzung und Routinen, Vereinfachung des Austauschprozesses (u.a. durch Leitung)	*„Das ist so angelegt, dass wir uns schnell und unbürokratisch [...] abstimmen können."* *„Ansonsten kann man sich zwei Stunden lang festreden, ohne dass was dabei herauskommt. Also das Geleitete finde ich immer wichtig."*
Aufgabenvereinfachung		Einfachheit und Zeiteffizienz der Abläufe, die für die eigene Arbeit notwendig sind *[1;2]*	Unterstützung bei eigenen Aufgaben (Aufwandsreduktion), Strukturierung und Übersichtlichkeit	*„Wenn ein Thema aufgearbeitet werden muss, von dem man vorher einfach keine Ahnung hatte. Dann ist es schön eine Expertenmeinung zu hören und dann auch zu wissen, wie man das Ganze angehen muss."* *„Ich weiß, wo die Sachen zu finden sind."*
Schnelle Problemlösung		Sichtbarmachung und Lösung operativer sowie inhaltlicher Probleme *[In Anlehnung an 1]*	Offene und direkte Problemkommunikation zur kooperativen Lösung, Problembewusstsein aufbauen	*„Die Probleme wurden nicht gelöst. Aber es hat uns dabei geholfen die Probleme besser selbst zu lösen, auch dadurch, dass neue Probleme aufgedeckt wurden."* *„Probleme werden schnell gelöst, wenn man dort zusammenfindet."*
Sozialer Wert		Soziale Bedingungen, die das Wohlbefinden steigern *[1]*	Gleichberechtigung, Verbundenheit und gemeinsame Identität, Engagement Dritter, Egalitäre Gemeinschaft	*„Ich finde das schön, dass man sich auch sozial trifft, wenn wir eine fachliche Aufgabe zu beantworten haben. Es war sehr schön, diese nette Atmosphäre zu erleben."* *„Da erleben wir uns als eingeschworene Gemeinschaft."*
Selbstdarstellung		Darstellung der eigenen Ideen, Gedanken und Forschungstätigkeiten *[In Anlehnung an 1]*	Sichtbarkeit eigener Beiträge, Präsentieren der Forschung	*„Ich finde das sehr relevant, weil es auch eine große Wertschätzung bedeutet, wenn es dafür Raum gibt."* *„Wenn mir was wichtig ist, kann ich es da posten und dann sehen es hoffentlich die anderen."*
Selbstverwirklichung		Verwirklichung der eigenen Ambitionen und Wünsche während der Arbeit *[1]*	Sich selbst einbringen, Partizipative Bearbeitung	*„Ich will da gerne äußern, wie ich mir dieses Projekt vorstelle."* *„Es ist mir wichtig, dass das ich auch die Gestaltungsmöglichkeiten spüre, die ich in so einem [Verbund] habe."*

(Continued)

Table 6—continued

Unsicherheits-vermeidung Vermeidung von Abhängigkeiten Druck-reduzierung	Unsicherheits-bewältigung	Reduktion von oder besserer Umgang mit Unsicherheiten in Bezug auf Anforderungen, Verantwortlichkeiten und Entscheidungen, die mit der eigenen Arbeit zusammenhängen *[In Anlehnung an 1 und 2]*	Sicherheit durch Kommunikation (Reflektion), Transparente Abhängigkeiten, Selbstwirksamkeitsüberzeugung aufbauen	*„Aufdecken von Abhängigkeiten: Wann muss was fertig sein, wann gibt wer wo Feedback? Das war gut."* *„Dass man sich mit anderen austauscht, abgleicht und dadurch natürlich auch Unsicherheiten verschwinden können."* *„Ich habe das jetzt verstanden. Ich kann den Punkt in meinem Projekt den ich bearbeiten muss auch vernünftig bearbeiten."*
Flexibilität Autonomie	Autonomie	Freiheit und Unabhängigkeit des eigenen Arbeitsverhaltens	Orts- oder Zeitunabhängigkeit, Autonomie bzgl. Entscheidungen und Umsetzung, Freiwillige Nutzung, Herrschaftsfreiheit	*„Das ist ein Riesenvorteil, dass man immer den Zugang hat."* *„. . . , dass es keine Pflichtveranstaltungen waren. Man konnte teilnehmen, wenn man wollte."*
Sachverständnis		Steigerung des persönlichen und kollektiven Sachverständnisses *[In Anlehnung an 1]*	Persönlicher Wissensgewinn, Interdisziplinäres Verständnis, Gemeinsame theoretische Basis	*„Da will ich lernen."* *„Ich empfinde es als bereichernd, wenn ich die Chance habe [bei solchen Formaten] mehr Verständnis für die anderen Fachdisziplinen zu entwickeln."*
Kreativität und Innovationskraft	Kreativität und Stimulation	Persönliche Inspiration und kollektives Generieren neuer Ideen	Impulse, Perspektivenerweiterung, gemeinsames Brainstorming und neue Ideen	*„Da sind nochmal andere Perspektiven aufgegangen."* *„Der Austausch liefert mir persönlich immer wieder neue Impulse."*
Wettbewerbs-fähigkeit	Impact und Wettbewerbs-fähigkeit	Erfolg, Relevanz und Reichweite der persönlichen und kooperativen Forschungsergebnisse	Verwertungserfolg, Erhöhte Reichweite und Reputation, Weiterführung von Forschungsthemen, Akquise neuer Projekte	*„Je mehr man mit anderen in Austausch tritt desto merkt man auch wo es vielleicht Schwachstellen gibt oder wo es noch Potenzial gibt."* *„Man wird dort vielleicht auf aktuelle Entwicklungen hingewiesen oder auch auf Methoden oder Treffen, die einen dann auch besser machen."*
Zugang zur Scientific Community	Netzwerkerweiterung und -vertiefung	Erweiterung und Vertiefung des bestehenden persönlichen Netzwerks	Disziplinenübergreifende Vernetzung, Neue Leute (außerhalb der bekannten Strukturen) kennenlernen, Personelle Zuordnung	*„Da ergeben sich auch Kontakte jenseits der etablierten und seit Jahren festen Scientific Community."* *„. . . , weil man sich dort auch näher kennenlernt und miteinander in Verbindung geht."*

Quellen: [1] Kleinaltenkamp, Storck et al. (2018); [2] Macdonald et al. (2016)

Unsicherheitsbewältigung[27] geändert, da sich die Aussagen innerhalb der Interviews auf die Chancen zur Reduktion von Unsicherheiten, aber auch auf eine Erhöhung der Toleranz gegenüber Unsicherheiten, beziehen.

Die von den Interviewten genannten Unsicherheiten haben unterschiedliche Quellen, wie bspw. ein möglicher Verlust der Kontrolle oder des Überblicks von kollektiven Prozessen und Entscheidungen. Aber auch in Bezug auf Herausforderungen im Zuge der eigenen Arbeit werden Unsicherheiten geschildert. Hier stellen die Forschenden sich unter anderem die Frage, ob nur sie selbst spezifische Probleme mit gewissen Tätigkeiten und Herausforderungen haben, oder ob es anderen Forschenden auch so geht. Darüber hinaus beschreiben die Interviewten auch ein Gefühl von Verunsicherung aufgrund von neuen fachlichen Herausforderungen.

> *„Es wurde ein Workshop zu einem Thema angeboten, wo ich mir unsicher war: Kriege ich das auf der gleichen Sachverstandsebene hin?“*

Anlässlich der Unsicherheiten schildern die Forschenden teilweise ein Druckgefühl. Diese Verbindung zwischen *Unsicherheitsbewältigung* und *Druckreduzierung* weist jede Aussage innerhalb der Dimension *Druckreduzierung* auf, weshalb diese Dimension der *Unsicherheitsbewältigung* zugeordnet wird.

Die Dimension der *Vermeidung von Abhängigkeiten* wird in ihrer ursprünglichen Definition von keiner der befragten Personen explizit verfolgt. Es wird lediglich angemerkt, dass es insbesondere um die Sichtbarmachung und Aufdeckung potenzieller Abhängigkeiten ginge, um diese kontrollieren zu können. In dieser Auslegung ordnet sich die Dimension ebenfalls der *Unsicherheitsvermeidung* zu.

Autonomie

Die Dimension *Flexibilität* wird mit *Autonomie* zusammengefasst. Die Aspekte der Möglichkeit zum flexiblen Arbeitsverhalten werden in einigen Fällen von den Befragten der übergeordneten wahrgenommenen Autonomie zugeordnet.

> *„Formate [Veranstaltungen] arbeiten per se gegen Flexibilität. Es ist immer ein Termin, wo du eine Verpflichtung spürst, selbst wenn alle sagen: Nein, du musst nicht hinkommen.“*

Sachverständnis

Die Dimension *Sachverständnis* und deren Definition kann durch die Aussagen der Interviewten für forschungsnahe Dienstleistungen aus dem theoretischen Portfolio übernommen werden. Auf individueller Ebene schildern die Forschenden persönliche Wissensgewinne innerhalb der eigenen Disziplin und auch über die Grenzen der eigenen Disziplin hinaus.

> *„Ich empfinde es als bereichernd, wenn ich die Chance habe [bei solchen Formaten] mehr Verständnis für die anderen Fachdisziplin zu entwickeln.“*

Eine Steigerung des kollektiven Sachverständnisses wird insbesondere zum Aufbau einer gemeinsamen theoretischen Verständigungsbasis angestrebt.

> *„…, dass alle unsere Projektbeteiligten verstehen, was das ist, dass uns das alle voranbringt das Thema, was wir da behandeln.“*

Kreativität und Stimulation

Die Dimension *Kreativität und Innovationskraft* wird für forschungsnahe Dienstleistungen von einem Großteil der Befragten als relevant erachtet. Die Forschenden differenzieren in ihren Aussagen nach einer verbesserten persönlichen und kollektiven Kreativität. Auffällig ist, dass in dem Zusammenhang selten von Innovationen oder Lösungen[28] als Ergebnis aus der Kreativität gesprochen wird. Die Forschenden benutzten eher Begriffe wie Impulse, Inspiration und Ideen[29]. Der Fokus liegt eher darauf im Austausch gemeinsam neue Ideen zu entwickeln oder neue Perspektiven und Stimulation für die eigene Forschungsarbeit zu gewinnen, auf denen aufbauend im Anschluss Lösungen entwickelt werden können. Aus diesem Grund wird die Dimension in Anlehnung an John (2019) in *Kreativität und Stimulation* umbenannt.

Impact und Wettbewerbsfähigkeit

Die Dimension *Wettbewerbsfähigkeit* wird von den Forschenden unterschiedlich ausgelegt. Vereinzelt lehnen die Befragten die Dimension strikt ab, mit dem Verweis darauf, dass Forschung kein Wettbewerb sei. Befragte mit praxisnaher Forschung beziehen die Dimension auf das Verstetigungsprodukt ihrer Forschungsprojekte.

[27] Die Begriffsfindung ist im Zuge eines Validierungsgesprächs mit der Erstprüferin, Prof. Dr. Susanne Robra-Bissantz, entstanden. Sie verweist auf die allgemein vorherrschende hohe Unsicherheit und Vagheit der Forschungsarbeit und wirft in dem Zusammenhang die Frage auf, ob eine tatsächliche Reduktion von Unsicherheiten durch forschungsnahe Dienstleistungen erfolgen kann. Ihrer Meinung nach kann insbesondere die Fähigkeit eines Forschenden, die ohnehin vorherrschenden Unsicherheiten zu bewältigen und mit ihnen umzugehen, verbessert werden.

[28] „Konkrete Lösungen“ im Zuge von forschungsnahen Dienstleistungen werden eher innerhalb der Dimension schnelle Problemlösung angesprochen, dort aber nicht explizit in Verbindung zu kreativen Arbeiten gebracht.

[29] Eine befragte Person kritisierte diesen Punkt explizit: „Wir untersuchen nach Schema F und wollen was Neues rausfinden. Aber man könnte ja auch kreativ überlegen, was man finden müsste und danach suchen. Und das ist tatsächlich etwas, was Wissenschaftler irgendwie so gar nicht auf dem Schirm haben.“

> *„Wir haben uns der Frage gestellt, wie man dann am Ende da ein Ergebnis rauskriegt, das man vermarkten oder weiterverwenden kann."*

Von zwei interviewten Personen wird vorgeschlagen, die Bezeichnung der Dimension, um den Begriff des *Impacts* zu erweitern und somit das allgemeine Verständnis der Dimension weiter zu fassen.

> *„Natürlich steht jede Forschung im Wettbewerb. Aber nicht auf die Art und Weise: Ich muss das Produkt loswerden, sondern ich will die Forschung, die von Steuergeldern finanziert wird, so gut machen, dass sie auch wirklich etwas bringt."*

Netzwerkerweiterung und -vertiefung

Die Dimension *Zugang zur Scientific Community* wird aufgrund der Rückmeldungen von den Forschenden in *Netzwerkerweiterung und -vertiefung* umbenannt. Die Benennung soll implizieren, dass die Forschenden eine Erweiterung des Netzwerks anstreben, welche jedoch nicht auf die Scientific Community begrenzt sein muss.

> *„Da ergeben sich ja auch Kontakte jenseits der etablierten und seit Jahren festen Scientific Community."*

Außerdem bildet der ursprüngliche Dimensionstitel den Aspekt der Vertiefung des Netzwerks nicht ab, welcher von den Forschenden als eine besondere Chance bei forschungsnahen Dienstleistungen betont wird:

> *„Austausch im Sinne von Netzwerken. Ich weiß was du machst. Du weißt, was ich mache. Vielleicht können wir das gemeinsam gebrauchen und wir haben schon mal eine Beziehungsebene herstellen können."*

Neu zugeordnete und verworfene Zieldimensionen

Die Dimension *Zugang zu Finanzmitteln* wird unter anderem im Zusammenhang mit der verbundinternen Vergabe von Finanzmitteln thematisiert. Hier wird der individuelle Aufwand für die Beantragung der internen Mittelvergabe genannt, welcher thematisch eher in der Dimension *Aufgabenvereinfachung* verortet werden kann. Ein neuer Zugang zu Finanzmitteln thematisiert eine befragte Person. Der Bezug liegt hier jedoch eher auf der Vernetzung mit neuen Akteuren und weniger auf der konkreten Akquise neuer Drittmittel.

> *„. . . , dass man Kollegen und Kolleginnen kennenlernt, die ähnliche Ideen haben und wo man neue Forschungskooperationen hinbekommen kann."*

Aus diesem Grund wird die Dimension *Zugang zu Finanzmittel* nicht mit aufgenommen und die Aussagen entsprechend zu *Aufgabenvereinfachung* sowie *Netzwerkerweiterung und -vertiefung* zugeordnet.

Die Dimension *Motivation* wird seitens der Befragten häufig als weniger relevant bezeichnet, mit dem Verweis auf eine ohnehin vorhandene hohe (intrinsische) Motivation.

> *„Ich glaube, da bringe ich einfach schon selbst eine intrinsische Motivation mit, sodass ich das eigentlich nicht brauche."*

Die Interviewten, welche die Dimension *Motivation* als relevant beschreiben, nennen als (extrinsische) Quellen der Motivation erreichte Ziele, welche in dem vorliegenden Portfolio repräsentiert sind. Bspw. werden die Ziele der *Netzwerkerweiterung und -vertiefung* sowie *Kreativität und Stimulation* genannt:

> *„Der Austausch und das Kennenlernen interessanter Menschen, das motiviert mich eigentlich immer."*

> *„. . . , dass ich dort inspiriert werde und Dinge mitnehme, die meinen Horizont erweitern."*

Diese motivierende Wirkung als Ergebnis einer persönlichen Zielerreichung, wird auch durch das JD-R Modell (siehe Kapitel 3.2.1) beschrieben (Demerouti & Nachreiner, 2019) und durch Aussagen der Interviewten bestätigt. Vereinzelt wird auch ein Gewinn an Motivation aus einer abschließenden Evaluation in Richtung der wahrgenommenen Sinnhaftigkeit einer forschungsnahen Dienstleistung beschrieben.

> *„Wenn [die Dienstleistung] einem was bringt, dann wird man auch motiviert."*

Mit den Daten aus den Interviews und deren theoretische Bestätigung, wird die Motivation[30] als nachgelagerter psychologischer Prozess aus dem Portfolio der Zieldimensionen expliziert.

Die Dimension des *hedonistischen Wertes* wird von den Forschenden kritisch diskutiert. Einige äußern sich ablehnend oder verweisen auf eine untergeordnete Rolle dieser Dimension.

> *„Es ist ein nützlicher Nebeneffekt, aber sicherlich erstmal nicht das Ziel, mit dem man da hineingeht."*

Im Kontrast dazu wird von einer befragten Person die These aufgestellt, dass die Dimension das Fundament sei, welches durch forschungsnahe Dienstleistungen[31] geschaffen werden muss.

[30] Weiterführende Literatur zur Arbeitsmotivation findet sich unter anderem in Kleinbeck (1996). Motivation in Verbindung mit sozialem Handeln wird thematisiert von Fischer und Wiswede (2002).

[31] In diesem Fall wurde von einem Austauschtermin auf Verbundebene gesprochen.

> *„Meiner Ansicht nach ist das die Basis für gute Arbeit, dass man die Leute entertaint. Und das muss jeder immer machen, wenn er gute Leistungen haben will."*

Insgesamt wird die Dimension jedoch lediglich innerhalb von vier Dienstleistungen thematisiert und somit aufgrund der Entscheidungsregel (siehe Kapitel 4.6.1) aus dem Portfolio ausgeschlossen.

Die Dimension *wahrgenommene Kontrolle* wird von keiner befragten Person als relevant erachtet und somit auch aus dem Portfolio der Zieldimensionen forschungsnaher Dienstleistungen ausgeschlossen.

5.1.2. Nennungshäufigkeit

Neben der inhaltlichen Strukturierung wird eine quantitative Analyse der Dimensionen vorgenommen. Einleitend ist anzumerken, dass die folgende quantitative Auswertung der qualitativen Daten nur bedingt als Handlungs- oder Gestaltungsempfehlung für forschungsnahe Dienstleistungen im Allgemeinen zu sehen ist. Welche Zieldimensionen im Sinne der Konzeptionierung des ViU tatsächlich relevant sind unterschiedet sich nach den Dienstleistungsangeboten und persönlichen Präferenzen der einzelnen Forschenden.

Die Analyse soll eine erste Orientierung hinsichtlich möglicher Schwerpunktsetzungen der Dimensionen für forschungsnahe Dienstleistungen liefern. Hierfür wird zunächst die Nennungshäufigkeit der Dimensionen dargestellt. Innerhalb der Interviews haben die Befragten unterschiedlich stark Bezug zu den Begrifflichkeiten gefunden. Dadurch unterscheidet sich die Anzahl der inhaltstragenden Aussagen innerhalb der einzelnen Dimensionen erheblich. Aus diesem Grund wird für die quantitative Analyse die Anzahl der Fundstellen auf eine Nennung pro Dimension und Dienstleistungserfahrung reduziert. Tabelle 7 stellt die Ergebnisse zu den Zieldimensionen über die gesamte Stichprobe (14 Dienstleistungserfahrungen) dar.

Nach der Nennungshäufigkeit der Zieldimensionen kann für sechs Dimensionen eine übergeordnete Relevanz (>70 %) für die *untersuchten* forschungsnahen Dienstleistungen identifiziert werden. Das primäre Ziel der Dienstleistungen stellt die Unterstützung von Forschenden bei deren kooperativen Forschungshandeln dar (siehe Kapitel 2.2.1). Diesem Motiv können insbesondere die Dimensionen *Prozessverbesserung* (79 %) und *Aufgabenvereinfachung* (86 %) zugeordnet werden, innerhalb derer durch eine Reduktion der Arbeitsanforderungen Wert für die Forschenden geschaffen werden kann. Darüber hinaus können die forschungsnahen Dienstleistungen auch bspw. eine Reduzierung oder Toleranzerhöhung von Unsicherheiten (*Unsicherheitsbewältigung*, 79 %) erwirken und somit durch eine Verringerung der (kognitiven) Arbeitsbelastungen wertvoll für die Forschenden werden. Daneben entsprechen die Dimensionen *Kreativität und Stimulation* (86 %), *Sachverständnis* (86 %) und *sozialer Wert* (79 %) eher übergeordneten Arbeitszielen (Demerouti & Nachreiner, 2019; Super, 1970), die auch innerhalb von forschungsnahen Dienstleistungen relevante Bedürfnisse darstellen.

Diese Ergebnisse ergeben sich aus der gesamten Stichprobe über 14 Dienstleistungserfahrungen. Da der Begriff der forschungsnahen Dienstleistung eine Vielzahl möglicher Angebote umfasst, unterscheiden sich auch die Ziele, welche an die Dienstleistungen seitens der Forschenden herangetragen werden. Ob sich in diesem Sinne bei vergleichbaren Angeboten (trotz unterschiedlicher individueller Präferenzen) Schwerpunkte bzgl. der Zielsetzungen identifizieren lassen, soll folgend anhand der Interviewdaten dargestellt werden.

Im Zuge der Interviews werden drei vergleichbare (oder gleiche) Dienstleistungen mehrfach besprochen[32]. Hierbei handelt es sich konkret um ein verbundübergreifendes Austauschangebot zwischen den Projektpartnern (teilweise mit Externen), welches von vier Befragten beschrieben wird. Drei Interviewte beziehen sich auf fachliche sowie methodische Workshopangebote und zwei schildern ihre Erfahrungen im Zusammenhang mit einem digitalen schwarzen Brett zur asynchronen und verbundübergreifenden Kommunikation. Abbildung 12 zeigt die Nennungshäufigkeit der Zieldimensionen innerhalb dieser drei Dienstleistungen.

Nach diesen Ergebnissen scheinen bestimmte Zieldimensionen eine allgemeingültige Relevanz für unterschiedliche Angebote forschungsnaher Dienstleistungen zu haben. Exemplarisch zu nennen sind hier *Aufgabenvereinfachung*, *Sachverständnis* oder *Kreativität und Stimulation*. Darüber hinaus bestätigen die Ergebnisse intuitive Zuschreibungen von relevanten Zieldimensionen für einzelne Angebote. Bspw. ist eine Vernetzung der Akteure eine primäre Intension der verbundübergreifenden Austauschangebote. Das spiegelt sich auch in der Nennung der Zieldimension *Netzwerkerweiterung und -vertiefung* von allen Befragten für diese Dienstleistung, während das Ziel für die anderen beiden Dienstleistungen (innerhalb der vorliegenden Stichprobe) keine Relevanz zu haben scheint.

5.2. Aktivierte Arbeitsressourcen

Die Ziele, welche Forschende an die Dienstleistungen herantragen, erfordern die Bereitstellung und Nutzung bestimmter Arbeitsressourcen, damit für die Forschenden innerhalb der Zieldimensionen Wert geschaffen werden kann.

5.2.1. Kategorien

Die Entwicklung des Kategoriensystems zu den, über die Inanspruchnahme der forschungsnahen Dienstleistungen aktivierten, Arbeitsressourcen erfolgt induktiv und wird teilweise von deduktiven Begriffen aus der Literatur gestützt. Dabei werden die Ressourcen in Anlehnung an Klemenz (2009) anhand deren Funktion kategorisiert. Das so entwickelte Kategoriensystem orientiert sich an der Abstraktionsebene der Arbeitsressourcen von Bakker et al. (2005) und beinhaltet

[32] In Kapitel 5 wird einleitend eine Unterteilung in vier vergleichbare Dienstleistungen vorgenommen. Für die folgende Analyse werden nur drei Dienstleistungen vergleichend ausgewertet, da die vierte Unterscheidung eine zu große Diversität aufweist (bspw. digitales Treffen vs. Treffen in Präsenz).

Tabelle 7: Nennungshäufigkeit der Zieldimensionen

Dimensionen	Anzahl Fundstellen [∑14 Erfahrungen]	Nennungshäufigkeit [1 Zählung pro pro Erfahrung]	Nennungshäufigkeit [% von ∑14]
Prozessverbesserung	36	11	79%
Aufgabenvereinfachung	36	12	86%
Schnelle Problemlösung	10	6	43%
Sozialer Wert	29	11	79%
Selbstdarstellung	13	9	64%
Selbstverwirklichung	12	9	64%
Unsicherheitsbewältigung	27	11	79%
Sachverständnis	24	12	86%
Kreativität und Stimulation	32	12	86%
Impact und Wettbewerbsfähigkeit	15	9	64%
Vernetzung	13	8	57%
Autonomie	11	8	57%
Gesamt	**258***	**119**	*70%*

**Insgesamt sechs der Fundstellen zu den Zieldimensionen (vgl. Tabelle 5) sind den ausgeschlossen Dimensionen Motivation und hedonistischer Wert zugeordnet und somit hier nicht dargestellt.*

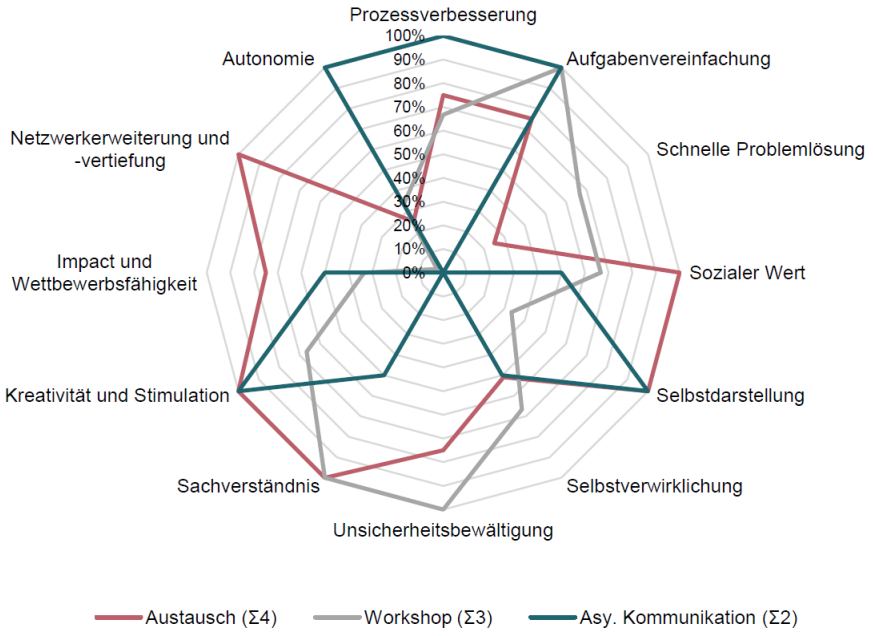

Abbildung 12: Vergleich der Zieldimensionen

sechs Kategorien. Diese werden durch insgesamt 20 zugehörige Aspekte charakterisiert, welche die Funktionalität der übergeordneten Kategorien gestalten. Von den insgesamt 234 Fundstellen zu den aktivierten Arbeitsressourcen können 224 diesen Aspekten zugeordnet werden, zehn übrige Fundstellen (4,3 %) können innerhalb des Kategoriensystems nicht verortet werden.

Zentrale Unterstützung

Die Kategorie *zentrale Unterstützung* fasst als aktivierte Arbeitsressource vornehmlich die Unterstützungswirkungen zusammen, welche seitens der anbietenden Instanz der forschungsnahen Dienstleistungen getätigt werden. Somit repräsentieren diese insbesondere Aspekte, welche direkt mit dem Dienstleistungsangebot assoziiert werden. Bspw. bietet ein digitales schwarzes Bett insbesondere eine zentrale und übersichtliche Sammlung von Informationen. In vielen Fällen

stellen die zentralen Unterstützungsangebote den Rahmen für die Bereitstellung weiterer Arbeitsressourcen, wie bspw. *soziale Unterstützung* , dar.

Unter *Rahmensetzung und Routinen (Z1)* werden Aspekte wie bspw. die Organisation und eine zielführende Moderation von kooperativen Austauschterminen zusammengefasst. Aber auch auf verwaltungstechnischer und administrativer Ebene können Mechanismen und Routinen die Prozesse und Aufgaben von Forschenden erleichtern. Die Ressource *Informatorische Unterstützung (Z2)* beschreibt eine zentrale Informationsbereitstellung. Die Befragten wertschätzen insbesondere eine transparente und zusammenführende Sammlung (relevanter) Informationen. Dabei werden Vorteile der individuellen Strukturierung von Arbeitsabläufen sowie der Transparenz von kooperativen Prozessen genannt. *Fachliche und methodische Unterstützung* (Z3) auf zentraler Ebene wird insbesondere durch das Angebot von Workshops erreicht. Unterstützt werden die Forschenden hier durch die methodische und fachliche Aufarbeitung neuer Themengebiete. Über die konkreten Unterstützungsangebote als Arbeitsressourcen hinaus, beschreiben die Interviewten außerdem eine notwendige *Anforderungsklarheit* im Zuge der Nutzung der Unterstützungsangebote. Die Aussagen hierzu beziehen sich auf Anforderungen, welche seitens der Dienstleistungsanbieter*innen an die Forschenden gerichtet werden, wie bspw. die Vorbereitung eines Austauschtreffens.

Autonomie

Die *Autonomie* stellt eine Arbeitsressource nach dem JD-R Modell dar (Bakker & Demerouti, 2007) und wird durch die Aussagen der Befragten bestätigt. Diese Selbstbestimmung ist für die einzelnen Mitarbeitenden unterschiedlich wertvoll (Carver & Scheier, 2000; Sauer & Weibel, 2012), dennoch wird der Arbeitsbedingung insbesondere im wissenschaftlichen Arbeitskontext ein hoher Stellenwert zugesprochen (Gläser, 2006; Schaper, 2014).

Die Befragten differenzieren nach vier Aspekten, welche der Hauptkategorie *Autonomie* zugeordnet werden (siehe Tabelle 9). Neben der Möglichkeit zur Selbstbestimmung bzgl. der *Teilnahme (A1)*, der *Arbeitszeit und -ort (A2)* sowie der konkreten *Umsetzung (A3)* wird übergeordnet auf die Relevanz einer *Herrschaftsfreiheit (A4)* im Zuge von forschungsnahen Dienstleistungen verwiesen.

Soziale Unterstützung

Neben der *zentralen Unterstützung* und *Autonomie* nimmt die Unterstützung durch Ressourcen aus dem sozialen Umfeld bei den besprochenen Dienstleistungserfahrungen einen hohen Stellenwert ein. Die Ressourcen ergeben sich durch die Interaktion mit den anbietenden Akteuren und weiteren Akteuren im Umfeld der Dienstleistungsinanspruchnahme. Die aktivierten Arbeitsressourcen in Form der Unterstützung aus dem sozialen Umfeld werden nach der Taxonomie sozialer Unterstützung nach House (1981) klassifiziert. Wodurch die Aussagen der Befragten nach den Begriffen der *emotionalen sozialen Untersetzung*, *evaluativen sozialen Unterstützung* sowie *informativen sozialen Unterstützung* gegliedert werden (siehe Tabelle 10).

Emotionale soziale Unterstützung wird unter anderem durch *Wertschätzung und Respekt (E1)* erreicht. Diese Arbeitsbedingung entspricht dem Konstrukt der sozialen Anerkennung, das es dem Menschen ermöglicht, sich als gleichberechtigtes sowie einmaliges Mitglied einer Gruppe zu sehen und das somit eine sozial-emotionale Ressource[33] darstellt (Schubert & Knecht, 2012). Ein weiterer Aspekt der emotionalen Unterstützung stellt, nach den Aussagen der Befragten, das Gefühl der *Zugehörigkeit und Kohäsion (E2)* dar (Fischer & Wiswede, 2002; Schubert & Knecht, 2012). Der Begriff der Kohäsion beschreibt den „Gruppenzusammenhalt bzw. die durchschnittliche Bindekraft im Rahmen einer Gruppe“ (Fischer & Wiswede, 2002, S. 595).

Innerhalb der *evaluativen sozialen Unterstützung* können die beiden Aspekte *Feedback und Legitimation (EV1)* sowie *Vergleich (EV2)* von House (1981) übernommen werden. Die beiden Aspekte werden bei den Erfahrungen mit forschungsnahen Dienstleistungen hauptsächlich im Zusammenhang mit einem Zugewinn an Sicherheit kommuniziert. *Feedback und Legitimation* werden sowohl direkt von den dienstleistungsanbietenden Akteuren als auch von anderen Forschenden bezogen, während für den *Vergleich* immer ein Austausch mit Akteuren in ähnlichen Situationen oder mit vergleichbaren Herausforderungen notwendig ist.

Informative soziale Unterstützung zeichnet sich nach House (1981) durch die Bereitstellung von Informationen seitens Dritter aus, welche nicht in sich hilfreich sind, sondern den Menschen helfen sich selbst zu helfen. Aus den Daten der Interviews können hierzu drei Aspekte identifiziert werden. Aufgrund der Informationsbereitstellung durch Dritte, wird die notwendige *Beteiligung (I1)* an forschungsnahen Dienstleistungen betont, um den Informationsfluss zu gewährleisten. Darüber hinaus empfinden die Befragten die Zusammenführung von *heterogenen Sichtweisen (I2)* als zielführend. Durch die differenzierten Perspektiven wird ein größeres Ressourcenpotential wahrgenommen. Dieses erhöhte Ressourcenpotential wird auch innerhalb der Sozialpsychologie bzgl. Gruppenstrukturen und deren Leistung diskutiert (Fischer & Wiswede, 2002). Während die *heterogenen Sichtweisen* eher übergeordnet sind, und bzgl. der beteiligten Akteure auch über die Grenzen der Forschungsverbünde hinausgehen können, wird von den Befragten auch spezifischer und auf Verbundebene ein Gewinn an *Informationen aus anderen Projekten (I3)* geschildert.

Synergieaufbau und -nutzung

Die aktivierte Arbeitsressource *Synergieaufbau und -nutzung* charakterisiert sich durch insgesamt fünf Aspekte

[33] Darüber hinaus beschreiben unter anderem Semmer, Jacobshagen, Meier und Elfering (2012) soziale Anerkennung und Wertschätzung aufgrund der daraus folgenden Steigerung des persönlichen Selbstwertgefühls als starke Ressourcen (zur Stressbewältigung).

Tabelle 8: Zentrale Unterstützung als aktivierte Arbeitsressource

Arbeitsressource		Exemplarische (paraphrasierte) Aussage
Zentrale Unterstützung		
Z1	Rahmensetzung und Routinen	*„Man braucht einen, der das moderiert und uns da durchführt."*
Z2	Informatorische Unterstützung	*„Die direkte Kommunikation, Jour Fixe oder auch Protokolle reichen nicht aus, weil das immer prozessartig ist und hier hat man es auf einen Blick."*
Z3	Fachliche und methodische Unterstützung	*„Wir haben eine Methodik erarbeitet bekommen von einem Experten."*
Z4	Anforderungsklarheit	*„Ich [möchte mir nicht] tausend Gedanken machen müssen, wie ich die Aufgabe richtig erfüllen könnte."*

Tabelle 9: Autonomie als aktivierte Arbeitsressource

Arbeitsressource		Exemplarische (paraphrasierte) Aussage
Autonomie		
A1	Freiwillige Nutzung	*„Man konnte teilnehmen, wenn man wollte und es gab keine negativen Punkte, wenn man mal nicht teilnehmen konnte."*
A2	Orts- oder zeitunabhängige Verfügbarkeit	*„Man kann schauen, wann und wo man Zeit hat."*
A3	Umsetzungsautonomie	*„Ich will immer noch derjenige sein, der dann weiter über das Vorgehen im Projekt entscheidet."*
A4	Herrschaftsfreiheit	*„Wissenschaft funktioniert nicht mit Macht. Wenn man mit Macht etwas durchdrücken will, kommt schlechte Wissenschaft raus."*

Tabelle 10: Soziale Unterstützung als aktivierte Arbeitsressourcen

Arbeitsressource		Exemplarische (paraphrasierte) Aussage
Emotionale soziale Unterstützung		
E1	Wertschätzung und Respekt	*„Du kannst dich überall in dem Verbund einbringen und jeder hört dir zu."*
E3	Zugehörigkeit und Kohäsion	*„Wir gehen oft in die Inhalte rein und merken, dass [. . .] wir uns nicht fern sind, egal an welchen Positionen wir sitzen, in welchen Institutionen wir sitzen und egal wie viel Mitarbeiterinnen und Mitarbeiter wir haben."*
Evaluative soziale Unterstützung		
EV1	Feedback und Legitimation	*„Eine dritte Legitimation der Sachen, die man schon vorhatte."*
EV2	Vergleich	*„. . . , dass man sieht, man ist da erstmal nicht allein. Es ist nicht mein persönliches Problem, sondern das ist ein Problem, was andere auch nicht so gut gelöst bekommen."*
Informative soziale Unterstützung		
I1	Beteiligung	*„Das lebt auch davon, dass alle mitmachen und somit auch der Informationsfluss entsprechend gewährleistet ist."*
I2	Heterogene Sichtweisen	*„Das zu hören ist sehr produktiv, weil sie einen ganz anderen Blick darauf haben."*
I3	Informationen aus anderen Projekten	*„Das war ganz schön, weil man relativ viele Informationen aus anderen Projekten bekommen hat."*

Tabelle 11: Synergieaufbau und -nutzung als aktivierte Arbeitsressource

Arbeitsressource		Exemplarische (paraphrasierte) Aussage
Synergieaufbau und -nutzung		
S1	Gemeinsame Bearbeitung	*„Man kann etwas gemeinsam angehen und dadurch in der Gruppe auch die eigene Arbeit vereinfachen."*
S2	Partizipation	*„Schon allein durch die Mitarbeit bringt man ja immer die eigenen Wünsche und Ambitionen ein."*
S3	Vernetzung	*„. . . , dass man sich kennt, dass man irgendwo schon mal miteinander gesprochen hat und dass man auch weiß, was die einzelnen Menschen bewegt."*
S4	Zusammenführung der Perspektiven	*„Sie machen was ähnliches wie wir oder das sind Synergien die man nutzen kann. Da entstehen echt tolle neue Ideen"*
S5	Informeller Austausch	*„Es wird die Möglichkeit gegeben, sich und das Projekt zu präsentieren und sich auch für andere und andere Projekte zu interessieren [. . .]. Oder ist das nur eine Bühne, um sich groß aufzuspielen¿'*

Tabelle 12: Nennungshäufigkeit der aktivierten Arbeitsressourcen

Aktivierte Arbeitsressourcen		Anzahl Fundstellen [$\sum$14 Erfahrungen]	Nennungs-häufigkeit [1 Zählung pro Erfahrung]	Nennungs-häufigkeit [% von $\sum$14]
Zentrale Unterstützung		**80**	**22**	**39%**
Z1	Rahmensetzung und Routinen	19	7	50%
Z2	Informatorische Unterstützung	23	5	36%
Z3	Fachliche und methodische Unterstützung	31	5	36%
Z4	Anforderungsklarheit	7	5	36%
Autonomie		**15**	**14**	**25%**
A1	Freiwillige Nutzung	3	3	21%
A2	Orts- oder zeitunabhängige Verfügbarkeit	5	4	29%
A3	Umsetzungsautonomie	4	4	29%
A4	Herrschaftsfreiheit	3	3	21%
Emotionale soziale Unterstützung		**17**	**14**	**50%**
E1	Wertschätzung und Respekt	8	8	57%
E2	Zugehörigkeit und Kohäsion	9	6	43%
Evaluative soziale Unterstützung		**22**	**12**	**43%**
EV1	Feedback und Legitimation	12	6	43%
EV2	Vergleich	10	6	43%
Informative soziale Unterstützung		**28**	**17**	**40%**
I1	Beteiligung	8	5	36%
I2	Heterogene Sichtweisen	9	5	36%
I3	Informationen aus anderen Projekten	11	7	50%
Synergieaufbau und -nutzung		**62**	**34**	**49%**
S1	Gemeinsame Bearbeitung	7	4	29%
S2	Partizipation	12	8	57%
S3	Vernetzung	16	8	57%
S4	Zusammenführung der Perspektiven	17	9	64%
S5	Informeller Austausch	10	5	36%
Gesamt		**224**	**113**	***41%***

(siehe Tabelle 11) und nimmt innerhalb von forschungsnahen Dienstleistungen einen hohen Stellenwert ein.

Hier wird von den Forschenden weniger eine direkte, auf die Einzelperson ausgerichtete Unterstützungswirkung beschrieben. Stattdessen ergeben sich aus der Sicht der Befragten insbesondere Vorteile in Form von kumulativen Effekten (Synergien) durch die Zusammenführung von kollektiven Ressourcen und Kompetenzen[34]. Die zuvor beschriebene Ressource *informative soziale Unterstützung* weist Parallelen zu der vorliegenden aktivierten Arbeitsressource auf, unterscheidet sich jedoch in einem wesentlichen Punkt. Die genannten kumulativen Effekte werden innerhalb der Ressource *Synergieaufbau und -nutzung* im Kollektiv aktiv wahrgenommen oder die Möglichkeit einer zukünftigen kollaborativen Nutzung aufgebaut. Bei der *informativen sozialen Unterstützung* werden bspw. durch heterogene Sichtweisen den einzelnen Forschenden lediglich Ressourcen (in Form von Kompetenzen und Informationen) zur Verfügung gestellt, welche persönlich genutzt werden können. Die aktivierte Arbeitsressource *Synergieaufbau und -nutzung* findet unter anderem über diese, auf die Einzelpersonen ausgerichtete, Perspektive hinaus Anwendung.

Genannt werden hier bspw. forschungsnahe Dienstleistungen, welche eine *gemeinsame Bearbeitung (S1)* von Thematiken ermöglichen. Dadurch wird aktiv auf den kollektiven Ressourcenpool (intellektueller) Fähigkeiten zurückgegriffen und Synergien, die sich daraus ergeben, genutzt (Katz & Martin, 1997). In diesem Zusammenhang beschreiben die Forschenden nicht nur einen reduzierten Aufwand aufgrund der kumulativen Effekte, sondern auch die Möglichkeit zur *Partizipation (S2)* und sich selbst einzubringen. Des Weiteren können die Dienstleistungen auch die Grundlage für derartige Kooperationen und Synergienutzungen über den Zeithorizont der Dienstleistung hinaus bilden. Von der Dienstleistung initiiert, wird hier auf den Aufbau von Synergien verwiesen, welche zukünftig genutzt werden können. Dieser beinhaltet zweierlei: Eine Erweiterung, aber auch eine Vertiefung und Intensivierung, bestehender Netzwerke (*Vernetzung, S3*) sowie die *Zusammenführung der Perspektiven (S4)*. Diese beide Aspekte (S3 und S4) stellen im Zusammenhang mit kooperativen Forschungshandeln, im Sinne der Integration und Zusammenführung disziplinärer Teilsichten, eine wesentliche Bedingung für den Erfolg der Forschung dar (Katz & Martin, 1997; Krott, 1996). Darüber hinaus wird auch *informeller Austausch (S5)* als notwendige Arbeitsbedingung beschrieben. Neben informellen Gesprächen, welche sich während der Inanspruchnahme der Dienstleistungen ergeben, wird auch darauf verwiesen, dass man selbst und die weiteren Akteure keine expressive Funktion einnehmen dürfen. Dieser Aspekt findet sich ebenfalls in der Sozialpsychologie bzw. der Betrachtung von Gruppenleistungen, welche die Forcierung einer positiven Selbstdarstellung als ein Hemmnis für Synergie-Effekte aufführen (Fischer & Wiswede, 2002). Die Befragten nennen diesen Aspekt im Zusammenhang mit der Ermöglichung einer offenen Kommunikation, welche frei ist von einer „künstlich aufgebauten Struktur“ oder wahrgenommenen Statusunterschieden, die beide eine offene Kommunikation behindern können.

[34]Die Intensivierung dieser Ressource durch die Identifizierung und Nutzung von Synergien wurde ebenfalls von zwei Befragten explizit genannt und auch als dringender Wunsch an forschungsnahe Dienstleistungen formuliert.

5.2.2. Nennungshäufigkeit

Wie auch für die Zieldimensionen, wird für die Arbeitsressourcen ebenfalls eine quantitative Analyse der Nennungshäufigkeit nach den 14 Dienstleistungserfahrungen durchgeführt. Auch hier wird die Nennungshäufigkeit pro Ressource und Dienstleistungserfahrung (diskrete Anzahl) ausgewertet (siehe Tabelle 12).

Tabelle 12 gibt einen ersten Einblick in eine mögliche Gewichtung der aktivierten Arbeitsressourcen bzgl. deren Relevanz für forschungsnahe Dienstleistungen. Die Autonomie weist mit 25 % die geringste Nennungshäufigkeit auf, während die restlichen fünf Kategorien zwischen 39 % und 50 % liegen.

Die ressourcenorientierte Betrachtung ergibt sich aus dem Wertschöpfungsprozess und damit aus dem konkreten Dienstleistungsprozess bzw. der Interaktion zwischen den Forschenden und weiteren Akteuren. Demnach ist bei einer quantitativen Untersuchung der aktivierten Arbeitsressourcen eine Differenzierung der Ergebnisse nach vergleichbaren Dienstleistungen aussagekräftiger als eine Betrachtung der gesamten Stichprobe mit einer breiten Varianz an Dienstleistungserfahrungen. Wie bereits bei der Analyse der Zieldimensionen in Kapitel 5.1.2 werden auch hier neun Dienstleistungserfahrungen miteinander verglichen, welche sich aus der gleichen oder einer vergleichbaren Dienstleistung ergeben haben. Hierbei handelt es sich um ein verbundübergreifendes Austauschangebot zwischen den Projektpartnern (teilweise mit Externen), fachliche sowie methodische Workshopangebote und das Angebot eines digitalen schwarzen Brettes zur asynchronen und verbundübergreifenden Kommunikation. Abbildung 13 zeigt die Verteilung der Zieldimensionen innerhalb dieser drei Dienstleistungen.

Für eine erste Übersicht werden die sechs Kategorien der aktivierten Arbeitsressourcen dargestellt. Eine detaillierte Einsicht in die einzelnen Ressourcenaspekte wird in Abbildung 14 geboten, deren Codierung in Tabelle 12 zu finden ist.

Bei der zusammenführenden Differenzierung nach den unterschiedlichen Dienstleistungen (siehe Abbildung 13) zeigen sich deutliche Unterscheide bzgl. der Arbeitsressourcen, welche zur Wertschöpfung aktiviert werden. Im Zuge des Austauschangebotes sind insbesondere Ressourcen aus dem sozialen Umfeld in Form von *sozialer Unterstützung* und *Synergieaufbau- und nutzung* relevant. Bei dem Angebot einer verbundübergreifenden asynchronen Austauschmöglichkeit in Form eines schwarzen Brettes sind die *emotionale und evaluative soziale Unterstützung* weniger bedeutsam. Stattdessen werden die *zentrale Unterstützung* und die *informative soziale Unterstützung* als aktivierte Arbeitsressourcen

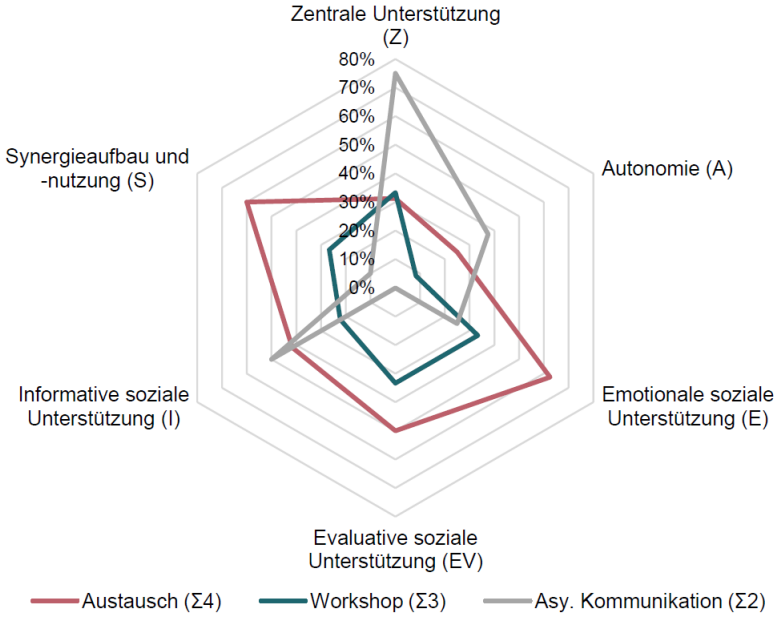

Abbildung 13: Vergleich der aktivierten Arbeitsressourcen

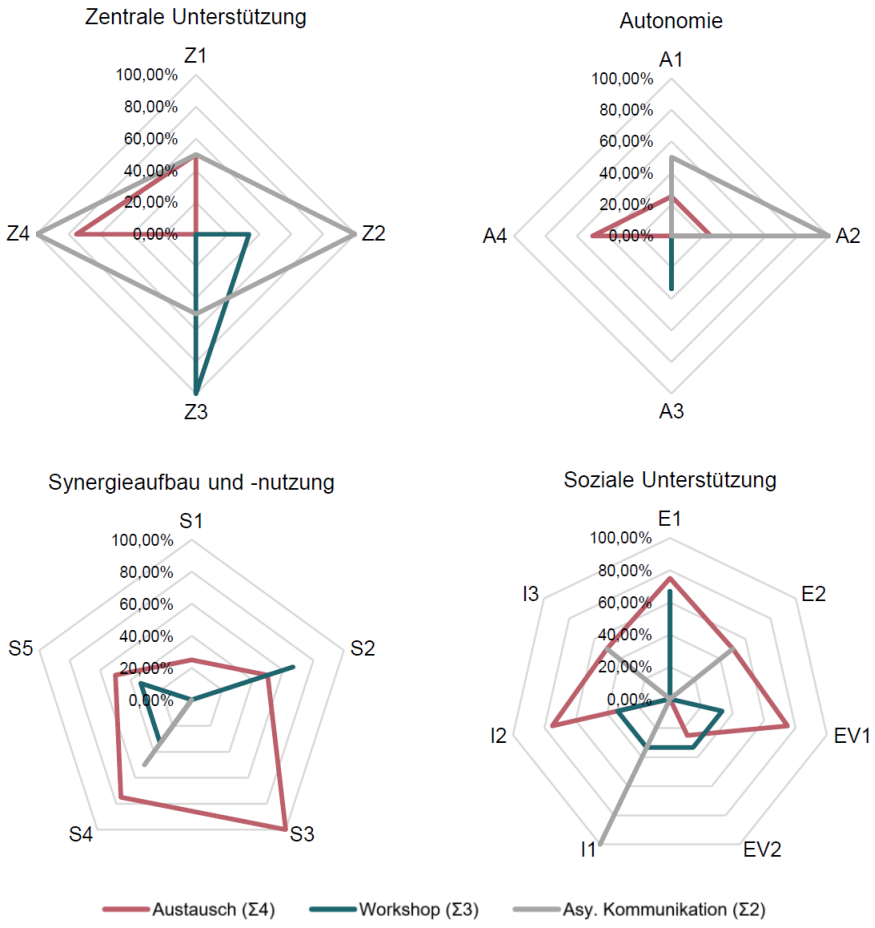

Abbildung 14: Vergleich der Aspekte aktivierter Arbeitsressourcen

genannt. Diese Gewichtung der Ergebnisse bestätigt die Ausgestaltung der Angebote. Während die Austauschangebote insbesondere einer Kooperation zwischen den Akteuren dienen, stellt das schwarze Brett eine technische Infrastruktur für die Informationsbereitstellung (insb. *informative soziale Unterstützung*) dar, welche von der forschungsnahen Dienstleistung (*zentrale Unterstützung*) initiiert wird. Ein Kernaspekt dieser technischen Infrastruktur ist die Möglichkeit zur asynchronen Nutzung, weshalb die Nennungshäufigkeit der *Autonomie* vergleichsweise gering erscheint. Das begründet sich jedoch durch die weiteren Aspekte innerhalb der aktivierten Arbeitsressource, welche, den Ergebnissen zufolge, für das digitale schwarze Brett weniger Relevanz haben (vgl. Abbildung 14). Die Ergebnisse in Abbildung 13 des Workshops sind in Vergleich zu den anderen Angeboten annähernd gleichmäßig verteilt[35]. Hier können allerdings aus der Auswertung nach den einzelnen Aspekten der aktivierten Arbeitsressourcen (vgl. Abbildung 14) relative Schwerpunktsetzungen ausgemacht werden. Bspw. die *fachliche und methodische Unterstützung (Z3)* sowie die Möglichkeit der *Partizipation (S2)* und der *informelle Austausch (S5)* weisen dabei gegenüber den anderen Dienstleistungen eine besondere Relevanz auf.

5.3. Relationale Funktionalität

Nach der Definition des ViU und dem daraus abgeleiteten Konzeptionierungsmodell sind die zuvor präsentierten Ergebnisse (siehe Kapitel 5.1 und 5.2) miteinander verbunden. Die beschriebenen aktivierten Arbeitsressourcen, welche während der Inanspruchnahme von forschungsnahen Dienstleistungen bereitgestellt werden, schaffen Wert für die Forschenden innerhalb der Zieldimensionen (siehe Abbildung 15).

Der sogenannten *Mehrfachbestimmtheit von Ressourcen* (Klemenz, 2009) zufolge, können Ressourcen mehrere Bedürfnisse und unterschiedliche Ziele gleichzeitig erfüllen. Demnach lassen die einzelnen Ressourcen (siehe Kapitel 5.2) sich nicht spezifisch den Zieldimensionen (siehe Kapitel 5.1) zuordnen. Eine Matrixbetrachtung der Ergebnisse ermöglicht jedoch Einsicht in relevante Ressourcen-Ziel-Verbindungen. Dazu wird die Nennungshäufigkeit der einzelnen Aspekte der aktivierten Arbeitsressourcen pro Zieldimension in Tabelle 13 dargestellt. Innerhalb einer Ziel-Ressourcen-Kombination werden die Fundstellen von Dopplungen innerhalb einer Dienstleistungserfahrung bereinigt.

Die Matrix zeigt teilweise deutliche Zusammenhänge zwischen den beiden Konzept-ionierungsebenen des ViU. So lassen sich bspw. klare Interdepedenzen zwischen der *emotionalen sozialen Unterstützung (E)* als aktivierte Arbeitsressource und dem *sozialen Wert* ausmachen. Auch die Ressource *evaluative soziale Unterstützung (EV)* erreicht insbesondere das Ziel *Unsicherheitsbewältigung*.

Dagegen haben die aktivierten Arbeitsressourcen *zentrale Unterstützung (Z)* sowie *Synergieaufbau und -nutzung (S)* Zielbeziehungen auf vielen Ebenen. Dadurch zeigt sich, dass die Ressource *zentrale Unterstützung* für das Erreichen vielfältiger Zielebenen funktional ist. Diese Multifunktionalität der *zentralen Unterstützung*, welche den Kern der forschungsnahen Dienstleistungen repräsentiert, stellt ein wünschenswertes Ergebnis dar. *Synergieaufbau und -nutzung* entspricht allgemein einer wichtigen Arbeitsbedingung bei der Forschung in Verbünden. Demnach sind die Ergebnisse für diese Ressource ebenfalls positiv hervorzuheben, welche durch die Unterstützung und Initiierung des kooperativen Forschungshandelns, funktional für das Erreichen einer Vielzahl von Zieldimensionen forschungsnaher Dienstleistungen ist.

Die Matrix bietet darüber hinaus auch ein Verständnis für die Ziel-Ressourcen-Verbindungen. Neben der Mehrfachbestimmtheit der Ressourcen, erfordert nach Klemenz (2009) das Erreichen spezifischer Ziele wiederum häufig bestimmte Ressourcen. Demnach können den Zieldimensionen notwendige Ressourcen zugeordnet werden. Dies geschieht unter der Einschränkung, dass aufgrund der Stichprobengröße und der Diversität der untersuchten Dienstleistungen die Ergebnisse nicht als allgemeingültig für forschungsnahe Dienstleistungen angenommen werden dürfen. Aus welchen Ressourcen welche Zieldimensionen erreicht, und somit Wert für die Forschenden geschaffen wird, lässt sich ebenfalls Tabelle 13 entnehmen. Darüber hinaus stellt diese Verbindung ein Teil der folgenden Auswertungsgrafik (siehe Abbildung 16) dar, welche die Ergebnisse der empirischen Analyse zusammenfasst.

In Abbildung 16 sind die entwickelten Kategoriensysteme zu den Zieldimensionen (siehe Kapitel 5.1.1) und aktivierten Arbeitsressourcen (siehe Kapitel 5.1.2) dargestellt sowie deren soeben beschriebene Ziel-Ressourcen-Verbindung. Außerdem wird auch die Relevanz der einzelnen Kategorien über deren relative Nennungshäufigkeit innerhalb der gesamten Stichprobe abgebildet (siehe Kapitel 5.1.2 und 5.2.2).

5.4. Weitere Erkenntnisse aus den erhobenen Daten

Neben den Fundstellen in den Interviews, welche nach der Auswertungsstrategie in den Ergebnissen dargestellt werden, liefern die Interviews weitere Daten zum Verständnis des ViU. Hieraus ergeben sich mögliche Erweiterungen des Konzeptionierungsmodells, welche folgend kurz in Aussicht gestellt werden.

Zum einen wird von einer befragten Person explizit auf den Zeithorizont der Wertschöpfung eingegangen.

> *„Der Wert, der wirklich entsteht, passiert danach, was damit [Ergebnis eines offenen Austausch- und Vernetzungsangebots] passiert."*

Die vorliegende Arbeit und empirische Analyse beschränkt sich auf den ViU, welcher sich im Zuge des Dienstleistungsprozesses ergibt. Bei der Erhebung thematisieren

[35] Darüber hinaus scheinen für die Workshops weniger Arbeitsressourcen relevant zu sein als bei den anderen beiden Dienstleistungen. Für die vorliegende Auswertung lässt sich das dadurch erklären, dass die Fundstellen von den Interviews bzgl. der Workshops insbesondere dem primären Angebot, der fachlichen und methodischen Unterstützung zugeordnet werden, welche für die Auswertung auf eine Nennung pro Dienstleistungserfahrung und Aspekt reduziert wird.

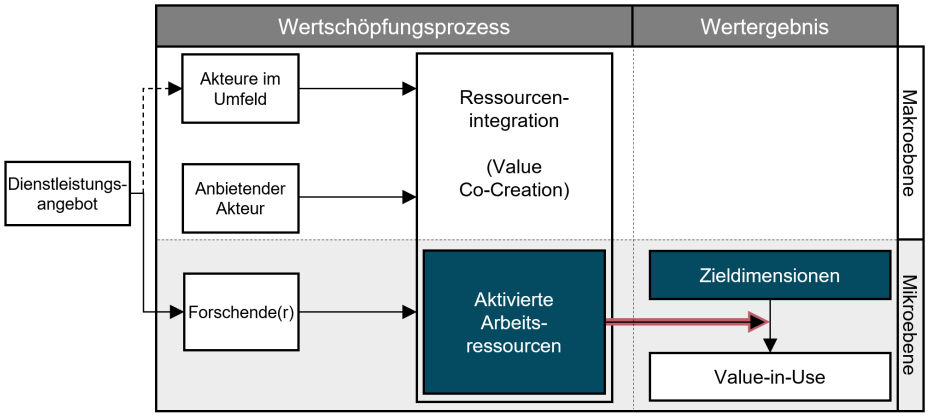

Abbildung 15: Relationale Funktionalität innerhalb des Modells

Tabelle 13: Ziel-Ressourcen-Matrix

Aktivierte Arbeitsressourcen			Zieldimension	Prozessverbesserung	Aufgabenvereinfachung	Schnelle Problemlösung	Sozialer Wert	Selbstdarstellung	Selbstverwirklichung	Unsicherheitsbewältigung	Sachverständnis	Kreativität und Stimulation	Impact und Wettbewerbsfähigkeit	Netzwerkerweiterung und -vertiefung	Autonomie	**Summe**
Zentrale Unterstützung	Z1	Rahmensetzung und Routinen		4	3	0	0	1	0	1	0	0	1	0	0	**10**
	Z2	Informatorische Unterstützung		3	2	0	0	2	1	2	0	0	0	0	0	**10**
	Z3	Fachliche und methodische Unterstützung		3	3	2	0	0	1	2	4	1	0	0	0	**16**
	Z4	Anforderungsklarheit		0	5	0	0	0	0	0	0	0	0	0	0	**5**
Autonomie	A1	Freiwillige Nutzung		0	0	0	0	0	0	0	0	0	0	0	3	**3**
	A2	Orts- oder zeitunabhängige Verfügbarkeit		0	0	0	0	0	0	0	0	0	0	0	4	**4**
	A3	Umsetzungsautonomie		0	0	0	0	0	0	0	0	0	0	0	4	**4**
	A4	Herrschaftsfreiheit		0	0	0	0	0	0	0	0	0	0	0	3	**3**
Soziale Unterstützung	E1	Wertschätzung und Respekt		0	0	0	6	2	0	0	0	0	0	0	0	**8**
	E2	Zugehörigkeit und Kohäsion		0	0	0	5	0	0	1	0	0	0	0	0	**6**
	EV1	Feedback und Legitimation		0	0	0	0	1	0	5	2	0	0	0	0	**8**
	EV2	Vergleich		0	0	0	0	0	0	6	0	0	1	0	0	**7**
	I1	Beteiligung		5	0	0	0	0	0	0	1	0	0	0	0	**6**
	I2	Heterogene Sichtweisen		0	0	0	0	0	0	0	1	5	0	0	0	**6**
	I3	Informationen aus anderen Projekten		3	1	0	0	0	0	0	4	1	1	0	0	**10**
Synergieaufbau	S1	Gemeinsame Bearbeitung		1	1	4	0	0	0	0	0	1	0	0	0	**7**
	S2	Partizipation		0	0	0	0	0	7	0	0	2	0	0	0	**9**
	S3	Vernetzung		1	1	0	0	0	0	0	0	0	2	7	0	**11**
	S4	Zusammenführung der Perspektiven		0	0	0	0	0	0	0	1	6	6	2	0	**15**
	S5	Informeller Austausch		0	0	0	3	3	0	0	0	1	0	0	0	**7**
	Summe			**20**	**16**	**6**	**14**	**9**	**9**	**17**	**13**	**17**	**11**	**9**	**14**	**155**

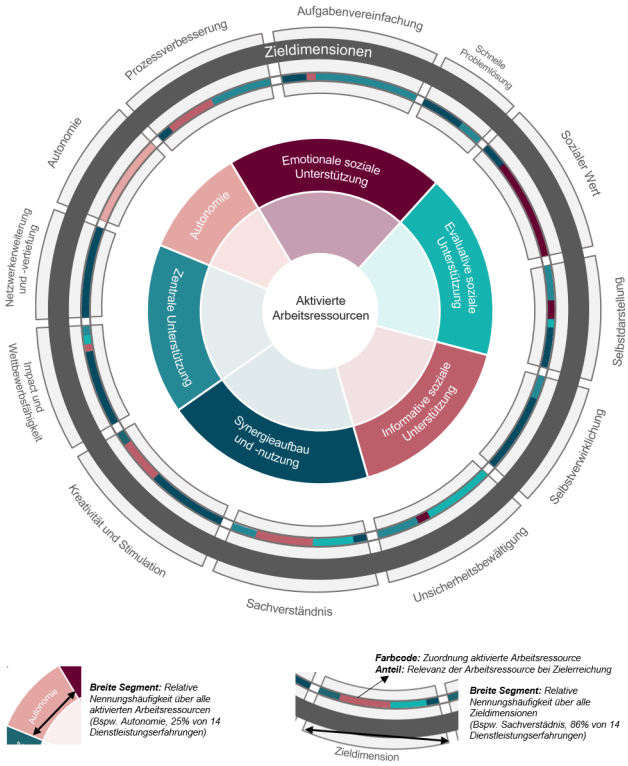

Abbildung 16: Nennungshäufigkeit und relationale Funktionalität

die Befragten auch langfristige Ziele und Arbeitsressourcen, welche über die Erfahrung der Interaktion während der Inanspruchnahme hinausgehen. Diese sind ebenfalls Teil der analysierten Ziel- und Ressourcen-Kategorien. Dennoch ist anzumerken, dass der langfristige Wert, welcher sich nach der Interaktion ergibt, kein expliziter Teil der zuvor definierten Konzeptionierung des ViU darstellt (siehe Kapitel 3.3). Der Wert kann sich jedoch auch über längere Zeiträume entwickeln (Heinonen & Strandvik, 2015). Außerdem stellt die Beziehung, welche über die Interaktionen aufgebaut wird, ebenfalls ein Wert für die begünstigten Akteure dar (Ulaga & Eggert, 2006). In diesem Zusammenhang wird von der Literatur der SD-Logik unter anderem der Aufbau von Vertrauen zwischen den Akteuren thematisiert. Dieser Aspekt des Vertrauens ist wiederum von hoher Relevanz in der Beziehung zwischen den Forschenden und anbietenden Akteuren in Forschungsverbünden (John, 2019). Die entstandene Beziehungsqualität zwischen anbietenden und begünstigten Akteuren stellt somit einen möglichen Wert von forschungsnahen Dienstleistungen dar, welcher sich über den Horizont des Dienstleistungsprozesses hinaus ergibt. Demnach wird angenommen, dass eine erweiterte Betrachtung, über den Zeitraum der Inanspruchnahme hinaus, das Verständnis des ViU von Forschenden weiter vertiefen kann.

Des Weiteren werden durch die Interviews auch Einblicke in Aspekte der Wertschöpfung gewonnen, welche den ViU negativ beeinflussen. Aufgrund der Ausrichtung der vorliegenden Arbeit auf die positive Wirkung des ViU, werden diese *behindernden Konsequenzen* aus dem Wertschöpfungsprozess (siehe Kapitel 2.2.2) in der empirischen Analyse nicht erhoben. Dennoch werden vereinzelt in den Erzählungen der Befragten Aspekte benannt, welche Einblicke in negative Facetten des Wertschöpfungsprozesses liefern, die das Erreichen der Ziele behindern. Unter anderem wird hier die subjektive Beurteilung zwischen dem wahrgenommenen Nutzen und dem persönlichen Zeitaufwand genannt.

> *„Ich finde manchmal die Treffen zu häufig und teilweise auch zu lange. Es sind schon relativ viele Treffen die [...] immer wieder angesetzt werden und [...] dann ist schon ein Teil der Arbeitszeit eines Tages weg. Also es ist recht zeitaufwändig."*

Daneben wird auch die erforderte persönliche (kognitive) Präsenz über die Dauer der Inanspruchnahme von Dienstleistungen unterschiedlich wahrgenommen. Während manche Befragten die Interaktivität und persönliche Einbindung wertschätzen, nimmt eine befragte Person den gleichen Aspekt eher negativ wahr.

> *„Das waren Formate, die ständig eine hohe Präsenz erfordert haben und das den ganzen Tag über. Da wäre kürzer oder weniger gut gewesen. Auch wenn ich die Initiative schätze. Im Prinzip war das alles gut gemacht, aber ein bisschen zu viel in der Zeitdauer."*

Im Sinne des Konzeptionierungsmodells können derartige Einsichten in die Wertbeurteilung als die Ressourcen

der Forschenden angesehen werden, welche für die Value Co-Creation investiert werden müssen. Demnach könnte das Modell um den Aspekt der (nicht-monetären) Kosten in Form von Zeit-, Zugangs-, Entscheidungs- oder Informationskosten (Geigenmüller & Lenk, 2014) erweitert werden. Sobald aus Sicht der Forschenden deren investierter Ressourcenaufwand die aktivierten Arbeitsressourcen übersteigt, kann davon ausgegangen werden, dass die Dienstleistungen nicht mehr als wertvoll oder sogar als wertvernichtend wahrgenommen werden. Solche negativen Wirkrichtungen der Wertschöpfung werden in der SD-Logik als *Value-Co-Destruction* bezeichnet (Echeverri & Skålén, 2011; Plé & Chumpitaz Cáceres, 2010) und seitens der Dienstleistungsforschung als relevantes Thema identifiziert (Ostrom et al., 2015). Neben dem ressourcenorientierten Ansatz der Implikation negativer Wirkrichtungen könnte auch der ViU in positive und negative Dimensionen (bzw. in Ziele und Vermeidungsziele) unterteilt werden. Dieser Ansatz wird unter anderem bereits von Sweeney, Plewa und Zurbruegg (2018) in Form von positiven und negativen Facetten des ViU empirisch abgeleitet.

5.5. Diskussion

Die beiden zuvor aufgeführten Aspekte des Betrachtungshorizonts und einer negativen Wertbeurteilung bieten neue Aussichten und Erweiterungsmöglichkeiten für den bestehenden Konzeptionierungsrahmen des ViU.

Im Folgenden sollen die Erkenntnisse diskutiert werden, welche sich innerhalb des konzeptionellen Rahmens durch die empirische Analyse für den ViU forschungsnaher Dienstleistungen ergeben haben. Dabei werden zunächst die Ergebnisse betrachtet (siehe Kapitel 5.5.1) und anschließend das methodische Vorgehen zu deren Erhebung und Auswertung diskutiert (siehe Kapitel 5.5.2).

5.5.1. Ergebnisse

Die gewonnen Ergebnisse basieren auf Interviews mit Forschenden, welche innerhalb ihres Arbeitsumfeldes befragt werden. Dadurch muss bei der Interpretation der Ergebnisse berücksichtigt werden, dass die Forschenden als Arbeitnehmer*innen unter Umständen ihre Bedürfnisse nicht umfänglich artikulieren, sondern sich Antworttendenzen aufgrund der *sozialen Erwünschtheit* ergeben (Richman, Kiesler, Weisband & Drasgow, 1999). Bspw. wird die theoretisch abgeleitete Zieldimension des hedonistischen Wertes aufgrund der mangelnden Relevanz in der empirischen Analyse für forschungsnahe Dienstleistungen ausgeschlossen. Dieser Ausschluss wird von der Erstprüferin der vorliegenden Arbeit in Frage gestellt. Ihrer Meinung nach stellt das Streben nach einem hedonistischen Wert eine kritische Zielsetzung innerhalb des sozialen Umfeldes von Forschenden dar. Der Einfluss dieser sozialen Komponente kann ihrer Ansicht nach ein Grund dafür sein, weshalb die Forschenden zu dieser Zieldimension wenig Bedarf bis hin zu teilweise ablehnenden Haltungen kommuniziert haben.

Dieser Aspekt stellt auch einen möglichen Einflussfaktor für die Ergebnisse der aktivierten Arbeitsressourcen dar, für welche ein induktiv entwickeltes Kategoriensystem vorgestellt wird. Dieses ergibt sich durch einen direkten Bezug zu den Zieldimensionen innerhalb der Interviews. Somit besteht, über den zuvor genannten potenziellen Einflussfaktor hinaus, allgemein die Gefahr einer Fehlerfortführung zwischen den Auswertungsebenen.

Des Weiteren deuten die Ergebnisse auch mögliche Differenzierungsschwierigkeiten zwischen den Konzeptionierungsebenen an. Die *Autonomie* wird in den Ergebnissen als angestrebtes Ziel und als aktivierte Arbeitsressource beschrieben. Auch der Aspekt *Vernetzung* findet sich ebenfalls in den Zieldimensionen wieder. Diese Überschneidungen werfen die Frage auf, ob die Zieldimensionen als solche für die Konzeptionierungszwecke richtig formuliert wurden. Deren Formulierung leitet sich aus dem Stand der Forschung ab und wird anhand der Besonderheiten aus dem Forschungsumfeld (Dienstleistungen innerhalb von Verbundprojekten) erweitert. Die beiden genannten Überschneidungen finden sich in den Zieldimensionen, welche aufgrund dieser theoretischen Vorüberlegungen in das Portfolio möglicher Zieldimensionen aufgenommen wurden. Demnach kann in Frage gestellt werden, ob diese nicht eher als Arbeitsressourcen anzusehen sind.

Neben den möglichen Differenzierungsschwierigkeiten kann auch das Ergebnis der aktivierten Arbeitsressourcen hinsichtlich des Abstraktionsgrades der formulierten Kategorien diskutiert werden. Allgemein liefern die Interviews Daten, welche theoretisch eine weitere Spezifikation der Arbeitsressourcen zulassen würden. Innerhalb der *fachlichen und methodischen Unterstützung* können spezifischere Aspekte wie bspw. *Interaktivität* oder auch *Professionalität des anbietenden Akteures* genannt werden. Eine derartige Detailebene wird jedoch selten in den Beschreibungen der Arbeitsressourcen eingenommen, weshalb sich die Arbeit auf eine höhere Abstraktion festgelegt hat. Außerdem ist anzunehmen, dass mit einer zunehmenden Spezifikation der Aspekte auch deren Abhängigkeit zu individuellen Präferenzen steigt. Demnach können die Ergebnisse zu den aktivierten Arbeitsressourcen als Rahmen betrachtet werden, wie forschungsnahe Dienstleistungen zur Zielerreichung beitragen können. Deren genaue Ausgestaltung muss situations- und personenspezifisch angepasst werden. Darüber hinaus kann der Abstraktionsgrad auch die Erkenntnisse hinsichtlich der untersuchten funktionellen Beziehungen beeinflussen. In Kapitel 5.3 wird hierzu eine Multifunktionalität für die Ressourcen *zentrale Unterstützung* und *Synergieaufbau und -nutzung* als wünschenswertes Ergebnis vorgestellt. Dieses Ergebnis könnte allerdings ebenfalls auf eine unzureichende Detaillierung bzw. Spezifikation des entwickelten Kategoriensystems hinweisen. Daneben besteht die Möglichkeit, dass die Multifunktionalität der *zentralen Unterstützung* auch durch eine mangelnde Artikulationsfähigkeit seitens der Forschenden verursacht wird. Es ist unter Umständen der Fall eingetreten, dass persönlich relevante Zieldimensionen genannt, jedoch nicht im Detail spezifiziert werden konnte, welche Arbeitsressourcen hierfür konkret notwendig sind. Im Zuge dessen werden ggf. vornehmlich Bezüge zur offen-

sichtlichen und intuitiv greifbaren Arbeitsressource *zentrale Unterstützung* hergestellt.

Abschließend ist die gesetzte Entscheidungsregel zu Identifikation von relevanten Kategorien (siehe Kapitel 4.6.1) in Frage zu stellen. Zum einen hinsichtlich der eigenen Anpassung für die Zieldimensionen, welche den Ausschluss der Dimension des hedonistischen Wertes verursacht. Übergeordnet hierzu kann zum anderen die allgemeine Zweckdienlichkeit des Einsatzes einer solchen Entscheidungsregel, unter Berücksichtigung der in Aussicht gestellten Verwendungsmöglichkeiten der Ergebnisse (siehe Kapitel 2.2.1), diskutiert werden. Die Ergebnisse sollen unter anderem Verwendung bei einer gemeinsamen Entwicklung von Dienstleistungsangeboten finden. Hierbei dienen die möglichen Zieldimensionen und aktivierten Arbeitsressourcen als Verständigungsgrundlage für die Forschenden. Diese erleichtert den Zugang zu dem komplexen inhärenten Prozess der Wertwahrnehmung und ermöglicht die Artikulation von persönlichen Präferenzen. In einem solchen gemeinsamen Gestaltungsprozess wird der Rahmen geboten, die Relevanz der Kategorien anhand der tatsächlichen Bedürfnisse der situativen Konstellation von Forschenden herauszufinden. Somit kann die Zweckdienlichkeit einer zu vorigen Reduktion der Kategorien anhand von subjektiven Erfahrungsdaten in Frage gestellt werden.

5.5.2. Methodisches Vorgehen

Neben den Ergebnissen soll folgend auch das methodische Vorgehen zu deren Gewinnung und Auswertung kritisch hinterfragt werden.

Zusätzlich zu dem zuvor dargestellten möglichen Einfluss der sozialen Erwünschtheit kann eine mögliche methodische Schwäche innerhalb der empirischen Erhebung diskutiert werden. Diese kritische Betrachtung ist insbesondere durch die Explikation der Zieldimension des hedonistischen Wertes motiviert, folgend aber auf eine allgemeine Diskussion der Ergebnisse für die Zieldimensionen zu beziehen. Der hedonistische Wert repräsentiert einen *affektiven* Wert (Holbrook, 1999; Morar, 2013). Ein solcher Wert ergibt sich aus den Dienstleistungsprozessen für die begünstigten Akteure, stellt jedoch keinen *funktionalen* Wert dar, welchr mit konkreten Erwartungen an die Dienstleistungen assoziiert wird. Während der Interviews wird mit der einleitenden Aussage „Hierbei ist mir wichtig, dass…“ der Bezug zu den Zieldimensionen aufgebaut. Durch diese Formulierung soll erreicht werden, dass die Forschenden nicht nur funktionale Ziele, sondern auch Ziele mit emotionalem und affektivem Charakter artikulieren, welche ebenfalls Teile des mehrdimensionalen ViU darstellen. Nach dem persönlichen Eindruck der Autorin und Interviewerin ist dies innerhalb der Befragungen gelungen. Dennoch sind die Ergebnisse hinsichtlich einer Gewichtung in Richtung funktionaler Ziele aufgrund des methodischen Vorgehens in Frage zu stellen. Eine alternative Analyse von Wert- bzw. Zieldimensionen stellt unter anderem das *Means-End-Chaining* und *Laddering* (Gutman, 1982; Reynolds & Gutmann, 1988) oder auch die *Repertory Grid-Methode* nach Kelly (1995) zur Ermittlung persönlicher Konstrukte dar. Diese Vorgehen erfordern jedoch einen hohen methodischen Aufwand, welcher für die vorliegende Arbeit und in Anbetracht des umfassenden Standes der Forschung nicht legitimiert werden konnte.

Über die Methodik zur Erhebung der Daten hinaus, birgt auch das Vorgehen zu deren Auswertung Diskussionspunkte. Hier kann die reliable Güte der Ergebnisse im Zusammenhang mit der qualitativen Inhaltsanalyse kritisch hinterfragt werden. Insbesondere aufgrund des zeitlichen Umfangs wird sowohl die Codierung (Extraktion der Fundstellen) als auch die Entwicklung des Kategoriensystems von der Autorin der vorliegenden Arbeit vorgenommen. Dabei werden weitere Forscherinnen in den Prozess mit einbezogen und Rücksprachen gehalten. Es findet jedoch keine parallele und unabhängige Durchführung der Auswertung einer weiteren Person statt. Dieses Vorgehen wird allerdings für qualitative Inhaltsanalysen empfohlen, um subjektive Einflüsse bei der Auswertung zu minimieren und somit die *Inter-Coder-Reliabilität* zu gewährleisten (Mayring, 2015). Außerdem haben die Abstimmungen und Validierungsgespräche bzgl. des Kategoriensystems zu einer fortlaufenden Verbesserung und Anpassung geführt. Dieser Prozess fand zum Teil nach 50 % des Materialdurchlaufs, jedoch insbesondere nach Abschluss der Materialsichtung (siehe Kapitel 4.5.1 und 4.5.2), statt. Bei den Anpassungen werden stets sämtliche Fundstellen in den iterativen Prozess mit einbezogen, eine anschließende erneute Sichtung des gesamten Interviewmaterials konnte allerdings zu dem Zeitpunkt der Arbeit nicht mehr durchgeführt werden. Nach dem Ablaufmodell der qualitativen Inhaltsanalyse ist eine solche Wiederholung des Materialdurchlaufs nicht vorgeschrieben, dennoch könnten hierbei ggf. anhand des final definierten Kategoriensystems neue Fundstellen extrahiert werden.

Darüber hinaus kann ebenfalls Kritik an der Stichprobegröße in Anbetracht der Diversität der untersuchten forschungsnahen Dienstleistungen geübt werden (Lamnek, 2010, S. 350). Dieser Problematik kann innerhalb der Arbeit teilweise begegnet werden, indem zur Interpretation der Ergebnisse Erfahrungen mit vergleichbaren Dienstleistungsangeboten gegenübergestellt werden.

6. Zusammenfassung und Ausblick

Die vorliegende Arbeit zielt darauf ab, den ViU von Forschenden zu verstehen, welcher sich aus der Inanspruchnahme von *forschungsnahen Dienstleistungen* ergibt. Um diesen wahrgenommenen Wert der Dienstleistungen abzubilden, stellt sich die Arbeit zwei Forschungsfragen, welche sich aus der Definition des ViU ableiten. Dabei sollen zum einen die Ziele von Forschenden dargestellt werden. Zum anderem sollen Einblicke hinsichtlich der Ressourcen gewonnen werden, welche durch die Dienstleistungen bereitgestellt werden und das Erreichen der Ziele erleichtern. Die beiden Forschungsfragen sind in ein Konzeptionierungsmodell für den ViU eingebettet, welches innerhalb der Arbeit aufgestellt wird. Mithilfe dieses Modells wird das folgend beschriebene

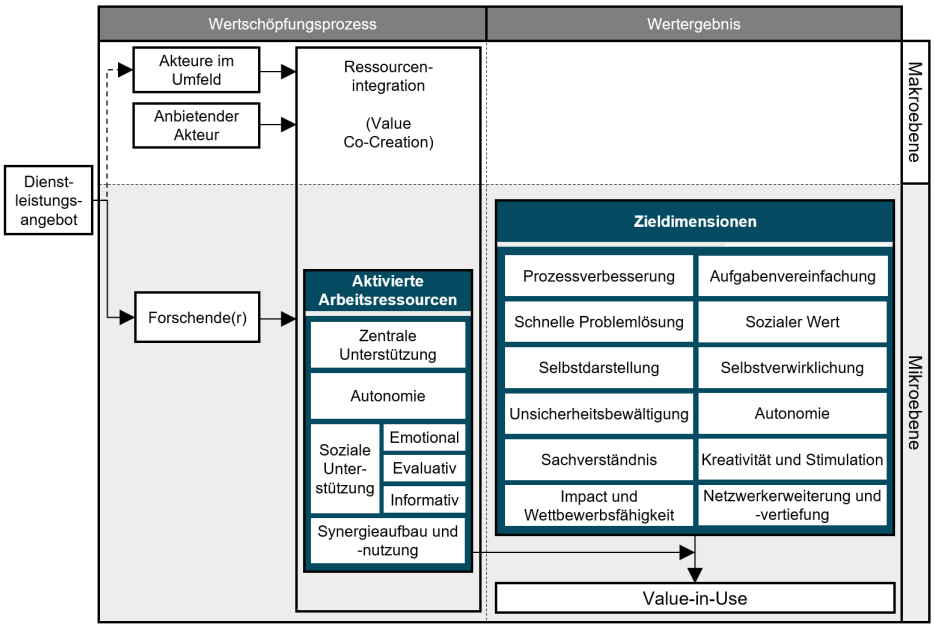

Abbildung 17: Das ViU-Modell forschungsnaher Dienstleistungen

Quelle: Eigene Darstellung

Wertverständnis für forschungsnahe Dienstleistungen eingenommen.

6.1. Beantwortung der Forschungsfragen

Im Sinne des ViU müssen die Dienstleistungen zunächst als *potenziell wertvolle* Angebote verstanden werden. Diese haben keinen Wert bis zu dem Moment ab welchem Forschende die Dienstleistung in Anspruch nehmen. Der mehrdimensionale ViU ergibt sich dabei aus dem Erreichen der Zieldimensionen. Welche Ziele die Forschenden an die Dienstleistung herantragen, bestimmt sich anhand individueller Präferenzen. Die Zielerreichung und damit der ViU wird durch die wahrgenommenen Konsequenzen aus dem Wertschöpfungsprozesses determiniert. Dieser Wertschöpfungsprozess findet entlang der Inanspruchnahme der Dienstleistungen statt. Während der Inanspruchnahme werden den Forschenden Ressourcen zur Verfügung gestellt, welche potenziell das Erreichen der Ziele ermöglichen oder erleichtern. Hierfür wird der Begriff der *aktivierten Arbeitsressourcen* definiert. Dadurch wird berücksichtigt, dass nicht alle angebotenen Ressourcen für die begünstigten Akteure wertvoll sind. Erst wenn die einzelnen Forschenden diese als funktional anerkennen und für sich persönlich integrieren können, tragen sie zur Zielerreichung bei.

Das theoretisch aufgestellte Konzeptionierungsmodell wird anschließend für den konkreten Anwendungsfall empirisch untersucht. Hierzu werden neun Forschende zu ihren persönlichen Erfahrungen mit forschungsnahen Dienstleistungen befragt und die gewonnen Daten mithilfe einer qualitativen Inhaltsanalyse ausgewertet. Teilweise beschreiben die Forschenden innerhalb der Interviews zwei Dienstleistungserfahrungen, weshalb insgesamt anhand von 14 Dienstleistungen Einsichten in die Wertbeurteilung von Forschenden gewonnen werden. Die Ergebnisse der empirischen Analyse sowie deren Einbettung in das Konzeptionierungsmodell sind zusammenfassend in Abbildung 17 dargestellt.

Die Ziele, welche Forschende an die Dienstleistungen herantragen, können anhand von zwölf Zieldimensionen charakterisiert werden, welche die erste Forschungsfrage beantworten. Die zweite Forschungsfrage wird mithilfe von sechs aktivierten Arbeitsressourcen beantwortet, deren Funktionalität sich durch insgesamt 20 Aspekte ergibt. Über diese beiden Ebenen der Konzeptionierung hinaus können außerdem erste Erkenntnisse hinsichtlich der relationalen Funktionalität der aktivierten Arbeitsressourcen mithilfe der empirischen Daten gewonnen werden, wodurch eine Verbindung der beiden Konzeptionierungsebenen geschaffen wird.

Die aufgestellten Kategoriensysteme der Zieldimensionen und aktivierten Arbeitsressourcen beziehen sich in ihrer inhaltlichen Charakterisierung auf forschungsnahe Dienstleistungen. Zusätzlich dienen repräsentative exemplarische Aussagen aus den Interviews einer besseren Zugänglichkeit zu den abstrakten Begriffen. Dadurch wird die einleitende Motivation der Konzeptionierung des ViU unterstützt, welche zum Ziel hat, eine Grundlage für die Gestaltung von forschungsnahen Dienstleistungen zu schaffen. Hierfür wird die Verwendung der Ergebnisse durch zwei Gruppen von Nutzer*innen in Aussicht gestellt.

Eine mögliche Anwendung der Ergebnisse stellt deren Berücksichtigung seitens der anbietenden Akteure bei der Gestaltung von Dienstleistungsangeboten dar. Mithilfe der Konzeptionierungsebenen werden Einblicke in die komplexe Wertbeurteilung ihrer Dienstleistungen aus der Sicht der

Forschenden geschafften. Somit kann bei künftigen Dienstleistungsentwicklungen berücksichtigt werden auf *welchen Ebenen* (Zieldimensionen) und *wie* (aktivierte Arbeitsressourcen) die Angebote für Forschende wertvoll sein können. Anhand einer zusammenführenden Betrachtung der beiden Konzeptionierungsebenen kann zusätzlich ein Verständnis dafür aufgebaut werden, welche spezifischen Ressourcen das Erreichen einzelner Ziele erleichtern und somit zur Wertschöpfung innerhalb der Zieldimensionen bereitgestellt werden sollten.

Neben den anbietenden Akteuren können die Ergebnisse auch von den Forschenden, als begünstigte Akteure der Dienstleistungen, verwendet werden. Die Ergebnisse erleichtern dabei den Forschenden die Zugänglichkeit zu deren inhärenten Prozess einer Wertbeurteilung. Anhand der Zieldimensionen und Ressourcen können die Forschenden somit besser ihre eigenen Bedürfnisse und Präferenzen reflektieren und gegenüber den anbietenden Akteuren artikulieren.

Diese Perspektive einer gemeinsamen Gestaltung von Dienstleistungsangeboten stellt eine relevante Implikation aus den theoretischen Überlegungen des ViU für die Interpretation der gewonnenen Ergebnisse dar. Der ViU wird als subjektives und phänomenologisches Wertkonstrukt beschrieben. Demnach ist das Konzeptionierungsmodell und die Ergebnisse der vorliegenden Arbeit als Orientierung für das Zustandekommen der Wertbeurteilung zu betrachten. Eine präskriptive Beschreibung der tatsächlichen persönlichen Wahrnehmung anhand der Ergebnisse ist nicht möglich, da die Beurteilung unter anderem maßgeblich durch individuelle Werte bestimmt wird (Stampacchia, Tregua & Coppola, 2020). Diese müssen durch die Forschenden innerhalb eines gemeinsamen Gestaltungsprozesses artikuliert werden, wodurch erst eine explizite Ausrichtung an den individuellen und auch situativen Bedarfen seitens der Forschenden ermöglicht wird.

6.2. Anschließende Forschungsaussichten

Die vorliegende Arbeit ermöglicht sowohl durch das aufgestellte Konzeptionierungsmodell als auch durch die Erkenntnisse der empirischen Analyse zahlreiche weitere Forschungsrichtungen.

Während die Zieldimensionen bereits innerhalb der ViU-Forschung untersucht werden, schafft die vorliegende Arbeit durch die Konzeptionierung der aktivierten Arbeitsressourcen eine neue Betrachtungsebene. Diese ergänzt die bisherige dimensionsbezogene ViU-Forschung um eine wertschöpfungsorientierte Analyse des ViU, welche ermöglicht dessen dynamischen Charakter besser abzubilden. Eine solche Analyse ist nicht auf die Dienstleistungsforschung innerhalb des Arbeitsumfeldes begrenzt, da der Begriff der Ressourcen auch unter anderem in der Humansoziologie umfassend betrachtet wird. Hier werden Ressourcen als „materielle, soziale oder persönliche Eigenschaften, die eine Person besitzt und die sie nutzen kann, um Fortschritte in Richtung ihrer persönlichen Ziele zu machen“ [übersetzt nach Diener und Fujita (1995), S. 926] definiert. Mithilfe dieses allgemein gefassten Ressourcenverständnisses lässt sich der ressourcenorientiere Konzeptionierungsansatz theoretisch auf alle Dienstleistungsbereiche übertragen.

Darüber hinaus deutet der Ressourcenbegriff durch die Weite seiner Definition auch ergänzende Forschungsansätze zu den bestehenden Ergebnissen an. So werden mit der vorliegenden empirischen Analyse Ressourcen betrachtet, welche aus dem externen Umfeld über den Dienstleistungsprozess aktiviert werden konnten. Dabei bleibt die Untersuchung der persönlichen Ressourcen zunächst außer Acht. Demnach ergibt sich ein aufbauender Forschungsansatz aus der Betrachtung von Ressourcen, welche seitens der Forschenden integriert werden müssen, um die aus dem externen Umfeld zu aktivieren. Eine Möglichkeit stellt hier die Untersuchung der notwendigen persönlichen Ressourcen in Relation zu den aktivierten Arbeitsressourcen dar. Ein solcher Ansatz wird bereits in Kapitel 5.4 anhand der Erkenntnisse aus den Interviews in Aussicht gestellt. Demnach würde eine Konzeptionierung der persönlich aufgewendeten Ressourcen eine Untersuchung des ViU ermöglichen, welche neben den externen Ressourcen auch die persönlichen (nicht-monetären) Kosten zur deren Aktivierung entlang der Wertschöpfung abbildet.

Über diesen Aspekt der wahrgenommenen Kosten hinaus, schafft das Konzeptionierungsmodell ebenfalls einen Rahmen für eine weiterführende Analyse des Wertschöpfungsprozesses. Innerhalb des Modells wird in der vorliegenden Arbeit zunächst die Mikroperspektive der einzelnen Forschenden eingenommen. Nach dem Prinzip der Value Co-Creation sind an dem Wertschöpfungsprozess weitere Akteure beteiligt. Diese sind über eine Wertschöpfungskooperation miteinander verbunden, aus welcher sich für jeden einzelnen Akteur ein individueller Wert ergibt. Für eine solche Makroperspektive bietet das Konzeptionierungsmodell bereits einen analytischen Rahmen. Innerhalb des Modells kann bspw. untersucht werden, welche persönlichen Ressourcen die Forschenden während der Inanspruchnahme von forschungsnahen Dienstleistungen integrieren, wofür Plé (2016) bereits zwölf Ressourcenkategorien definiert. Diese integrierten Ressourcen schaffen wiederum Wert für andere Akteure innerhalb des Umfeldes, weshalb sich durch deren Berücksichtigung wertvolle Implikationen für die Analyse kooperativer Wertschöpfungsprozesse ergeben können.

Darüber hinaus bietet die vorliegende Arbeit durch das aufgestellte Kategoriensystem der beiden Konzeptionsebenen die Grundlage für eine Typisierung der begünstigten Akteure. John (2019) und Wieser, Brechelmacher und Schendel (2014) stellen dazu bereits Ansätze zur charakterisierenden Einordnung der wissenschaftlichen Mitarbeitenden innerhalb von Forschungsprojekten vor. Hier werden unter anderem Kriterien wie das Streben nach Autonomie, gemeinschaftlichen Arbeiten aber auch das Streben nach Distanz eingeführt. Einer solchen Typisierung folgend könnten, anhand der gewonnenen Ergebnisse der vorliegenden Arbeit, zukünftig bedarfsgerechte Dienstleistungsangebote für die unterschiedlichen Charaktere innerhalb von Verbundprojekten angeboten werden.

Literatur

Aljets, E. & Lettkemann, E. (2012). Hochschulleitung und forscher: Von wechselseitiger nichtbeachtung zu wechselseitiger abhängigkeit. In U. Wilkesmann & C. Schmid (Hrsg.), *Hochschule als organisation* (S. 131–153). Wiesbaden: Springer VS.

Antonovsky, A. (1997). *Salutogenese: Zur entmystifizierung der gesundheit* (Bd. Bd. 36). Tübingen: Dgvt-Verl.

Auer, C. & Herlitschka, S. (2008). Institutionelle strategieorientierung in der forschung: Strategisches forschungsmanagement an der medizinischen universität graz. *Wissenschaftsmanagement* (1), 10–18. Zugriff am 01.02.2021 auf https://www.wissenschaftsmanagement.de/dateien/dateien/management/downloaddateien/wim_2008_01_carolin_auer_sabine_herlitschka_institutionelle_strategieorientierung_in_der_forschung.pdf

Auth, G., Czarnecki, C., Bensberg, F. & Thor, A. (2019). Digitalisierung des forschungsprozesses aus sicht von forschenden – durch serviceintegration zum persönlichen forschungsinformationssystem. In T. Barton, C. Müller & C. Seel (Hrsg.), *Hochschulen in zeiten der digitalisierung* (S. 287–307). Wiesbaden: Springer Fachmedien Wiesbaden and Springer Vieweg.

Bakker, A. B. & Demerouti, E. (2007). The job demands–resources model: state of the art. *Journal of Managerial Psychology*, *22* (3), 309–328. doi: 10.1108/02683940710733115

Bakker, A. B., Demerouti, E. & Euwema, M. C. (2005). Job resources buffer the impact of job demands on burnout. *Journal of occupational health psychology*, *10* (2), 170–180. doi: 10.1037/1076-8998.10.2.170

Bakker, A. B., Hakanen, J. J., Demerouti, E. & Xanthopoulou, D. (2007). Job resources boost work engagement, particularly when job demands are high. *Journal of Educational Psychology*, *99* (2), 274–284. doi: 10.1037/0022-0663.99.2.274

Bamberger, P. (2008). From the editors beyond contextualization: Using context theories to narrow the micro-macro gap in management research. *Academy of Management Journal*, *51* (5), 839–846. doi: 10.5465/amj.2008.34789630

Baum, E. L. & Zimmermann, E. W. (1951). *World resources and industries*. New York, NY: Harper and Row.

Becker, P. (2006). *Gesundheit durch bedürfnisbefriedigung*. Göttingen: Hogrefe.

Boden, D. (2016a). Den gestellten ansprüchen entsprechen: Ganzheitliche dienstleistungsentwicklung im forschungsmanagement. *Wissenschaftsmanagement* (5), 42–47.

Boden, D. (2016b). *Systematische dienstleistungs- und organisationsentwicklung in universtären support-strukturen: Service engineering in der forschungsunterstützung*. Göttingen.

Boden, D. (2016c). Wissenschaftsmanagement an der schnittstelle zwischen wissenschaft und verwaltung: Praxisbericht forschungsmanagement. *Wissenschaftsmanagement* (2), 26–31. Zugriff am 13.10.2020 auf https://www.wissenschaftsmanagement.de/dateien/dateien/archiv/downloaddateien/wim_02_2016.pdf

Böhm, B. (2006). *Vertrauensvolle verständigung - basis interdisziplinärer projektarbeit* (Bd. Bd. 4). Stuttgart: Steiner.

Brandstäfter, S. & Sonntag, K. (2016). Interdisciplinary collaboration: How to foster the dialogue across disciplinary borders? In B. Deml, P. Stock, R. Bruder & C. M. Schlick (Hrsg.), *Advances in ergonomic design of systems, products and processes* (S. 365–409). Berlin, Heidelberg: Springer Berlin Heidelberg.

Bruhn, M. (2012). Verfahren zur messung der qualität interner dienstleistungen: Ansätze für einen methodentransfer aus dem (externen) dienstleistungsmarketing. In M. Bruhn (Hrsg.), *Internes marketing* (S. 536–575). Wiesbaden: Gabler.

Bruns, K. & Jacob, F. (2014). Value-in-use und mobile technologien: Ein allgemeiner ansatz zur messung des value-in-use und eine spezifische anwendung am beispiel der smartphone-nutzung. *Wirtschaftsinformatik*, *56* (6), 381–393. doi: 10.1007/s11576-014-0437-z

Bruns, K. & Jacob, F. (2016). Value-in-use: Antecedents, dimensions, and consequences. *Marketing ZFP*, *38* (3), 135–149. doi: 10.15358/0344-1369-2016-3-135

Capra, F. & Luisi, P. L. (2014). *The systems view of life: A unifying vision* (1st paperback ed., 8th printing Aufl.). New York: Cambridge University Press.

Carver, C. S. & Scheier, M. F. (2000). Autonomy and self-regulation. *Psychological Inquiry*, *11* (4), 284–291. Zugriff am 24.04.2021 auf https://www.jstor.org/stable/pdf/1449622.pdf?refreqid=excelsior%3A40628c4b38053fb9f52936633cfc3798

Chandler, J. D. & Vargo, S. L. (2011). Contextualization and value-in-context: How context frames exchange. *Marketing Theory*, *11* (1), 35–49. Zugriff am 27.10.2020 auf https://journals.sagepub.com/doi/pdf/10.1177/1470593110393713 doi: 10.1177/1470593110393713

Defila, R., Di Giulio, A. & Scheuermann, M. (2006). *Forschungsverbundmanagement: Handbuch zur gestaltung inter- und transdiziplinärer projekte*. Zürich: vdf, Hochsch.-Verl. an der ETH.

Defila, R., Di Giulio, A. & Scheuermann, M. (2008). *Management von forschungsverbünden: Möglichkeiten der professionalisierung und unterstützung*. Weinheim: Wiley-VCH Verl. Zugriff am 07.11.2020 auf https://www.dfg.de/download/pdf/dfg_im_profil/geschaeftsstelle/publikationen/management_forschungsverbuende.pdf

de Gregori, T. R. (1987). Resources are not; they become: An institutional theory. *Journal of Economic Issues*, *21* (3), 1241–1263. doi: 10.1080/00213624.1987.11504702

Demerouti, E., Bakker, A. B., Nachreiner, F. & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology*, *86* (3), 499–512. doi: 10.1037//0021-9010.86.3.499

Demerouti, E. & Nachreiner, F. (2019). Zum arbeitsanforderungen-arbeitsressourcen-modell von burnout und arbeitsengagement – stand der forschung. *Zeitschrift für Arbeitswissenschaft*, *73* (2), 119–130. doi: 10.1007/s41449-018-0100-4

DFG. (2018). Kennzahlen zur öffentlich finanzierten forschung in deutschland. *DFG-Förderatlas 2018*. Zugriff am 03.05.2021 auf https://www.dfg.de/sites/foerderatlas2018/epaper/files/assets/common/downloads/publication.pdf

Diener, E. & Fujita, F. (1995). Resources, personal strivings, and subjective well-being: A nomothetic and idiographic approach. *Journal of Personality and Social Psychology*, *68* (5), 926–935. doi: 10.1037/0022-3514.68.5.926

Dodge, R., Daly, A., Huyton, J. & Sanders, L. (2012). The challenge of defining wellbeing. *International Journal of Wellbeing*, *2* (3), 222–235. doi: 10.5502/ijw.v2i3.4

Durgee, J. F., O'Connor, G. C. & Veryzer, R. W. (1996). Observations: Translating values into product wants. *Journal of Advertising Research*, *36* (6), 90–99.

Echeverri, P. & Skålén, P. (2011). Co-creation and co-destruction: A practice-theory based study of interactive value formation. *Marketing Theory*, *11* (3), 351–373. doi: 10.1177/1470593111408181

Edvardsson, B., Kleinaltenkamp, M., Tronvoll, B., McHugh, P. & Windahl, C. (2014). Institutional logics matter when coordinating resource integration. *Marketing Theory*, *14* (3), 291–309. Zugriff am 22.10.2020 auf https://journals.sagepub.com/doi/pdf/10.1177/1470593114534343 doi: 10.1177/1470593114534343

Edvardsson, B., Tronvoll, B. & Gruber, T. (2011). Expanding understanding of service exchange and value co-creation: a social construction approach. *Journal of the Academy of Marketing Science*, *39* (2), 327–339. Zugriff am 26.10.2020 auf https://link.springer.com/article/10.1007/s11747-010-0200-y#citeas doi: 10.1007/s11747-010-0200-y

Eggert, A., Ulaga, W., Frow, P. & Payne, A. (2018). Conceptualizing and communicating value in business markets: From value in exchange to value in use. *Industrial Marketing Management*, *69*, 80–90. doi: 10.1016/j.indmarman.2018.01.018

Etgar, M. (2008). A descriptive model of the consumer co-production process. *Journal of the Academy of Marketing Science*, *36* (1), 97–108. doi: 10.1007/s11747-007-0061-1

Fischer, L. & Wiswede, G. (2002). *Grundlagen der sozialpsychologie* (2., überarb. und erw. Aufl. Aufl.). München and Wien: Oldenbourg.

Fleßa, S., John, M. & Mahnke, J. (2011). Zwischen vertrauen und kontrolle: Management interdisziplinärer forschungsverbünde – eine exemplarische analyse. *Wissenschaftsmanagement* (6), 38–43. Zugriff am 22.02.2021 auf https://www.wissenschaftsmanagement.de/dateien/dateien/management/downloaddateien/wim_2011_06_steffen_flessa_michael_john_juliane

_mahnke_zwischen_vertrauen_und_kontrolle.pdf

Flick, U., von Kardorff, E. & Steinke, I. (2017). Was ist qualitative forschung? einleitung und überblick. In U. Flick, E. von Kardorff & I. Steinke (Hrsg.), *Qualitative forschung* (S. 13–29). Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag.

Friedrichs, J. (1973). *Methoden empirischer sozialforschung* (54. - 60. Tsd Aufl., Bd. 28 : Sozialwiss). Reinbek bei Hamburg: Rowohlt.

Gallan, A. S. & Jefferies, J. G. (2020). Value co-creation and its meaning for customers. In E. Bridges & K. Fowler (Hrsg.), *The routledge handbook of service research insights and ideas* (S. 322–343). New York: Routledge.

Geigenmüller, A. & Lenk, K. (2014). Interner service value: Begriff, konzeptualisierung und relevanz für das interne dienstleistungsmarketing. In M. Bruhn & K. Hadwich (Hrsg.), *Service value als werttreiber* (S. 329–350). Wiesbaden: Springer Fachmedien Wiesbaden.

Gioia, D. A., Corley, K. G. & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research. *Organizational Research Methods*, *16* (1), 15–31. Zugriff am 24.04.2021 auf https://journals.sagepub.com/doi/pdf/10.1177/1094428112452151 doi: 10.1177/1094428112452151

Gläser, J. (2006). *Wissenschaftliche produktionsgemeinschaften: Die soziale ordnung der forschung* (Bd. Bd. 906). Frankfurt [am Main] and New York: Campus.

Goffin, K., Lemke, F. & Szwejczewski, M. (2006). An exploratory study of 'close' supplier-manufacturer relationships. *Journal of Operations Management*, *24* (2), 189–209. doi: 10.1016/j.jom.2005.05.003

Goffin, K., Raja, J. Z., Claes, B., Szwejczewski, M. & Martinez, V. (2012). Rigor in qualitative supply chain management research. *International Journal of Physical Distribution & Logistics Management*, *42* (8/9), 804–827. Zugriff am 09.05.2021 auf https://www.emerald.com/insight/content/doi/10.1108/09600031211269767/full/pdf?title=rigor-in-qualitative-supply-chain-management-research-lessons-from-applying-repertory-grid-technique doi: 10.1108/09600031211269767

Grönroos, C. & Voima, P. (2013). Critical service logic: making sense of value creation and co-creation. *Journal of the Academy of Marketing Science*, *41* (2), 133–150. doi: 10.1007/s11747-012-0308-3

Gummerus, J. (2013). Value creation processes and value outcomes in marketing theory. *Marketing Theory*, *13* (1), 19–46. doi: 10.1177/1470593112467267

Gutman, J. (1982). A means-end chain model based on consumer categorization processes. *Journal of Marketing*, *46* (2), 60. Zugriff am 11.10.2020 auf https://www.proquest.com/docview/1296568451/fulltextPDF/C6BF9D2F33C84397PQ/9?accountid=14431 doi: 10.2307/3203341

Hartwig, K. L. & Jacob, F. (2018). How individuals assess value-in-use: Theoretical discussion and empirical investigation. *Marketing ZFP*, *40* (3), 43–62. doi: 10.15358/0344-1369-2018-3-43

Heinonen, K. & Strandvik, T. (2015). Customer-dominant logic: foundations and implications. *Journal of Services Marketing*, *29* (6/7), 472–484. Zugriff am 06.05.2021 auf https://www.emerald.com/insight/content/doi/10.1108/JSM-02-2015-0096/full/pdf?title=customer-dominant-logic-foundations-and-implications doi: 10.1108/JSM-02-2015-0096

Helkkula, A., Kelleher, C. & Pihlström, M. (2012). Characterizing value as an experience. *Journal of Service Research*, *15* (1), 59–75. Zugriff am 26.10.2020 auf https://journals.sagepub.com/doi/pdf/10.1177/1094670511426897 doi: 10.1177/1094670511426897

Helkkula, A., Kowalkowski, C. & Tronvoll, B. (2018). Archetypes of service innovation: Implications for value cocreation. *Journal of Service Research*, *21* (3), 284–301. doi: 10.1177/1094670517746776

Hendrichs, M. (2017). Forschungsmanagement in hochschulen: Service aus einer hand: ein praxisbeispiel aus der universität zu köln. In M. Lemmens, P. Horváth & M. Seiter (Hrsg.), *Wissenschaftsmanagement* (S. 274–285). Bonn-Berlin: Lemmens Medien GmbH.

Heuer, C. (2017). *New professionals an der deutschen universität* (Dissertation, Universität Osnabrück, Osnabrück). Zugriff am 14.10.2020 auf https://repositorium.ub.uni-osnabrueck.de/bitstream/urn:nbn:de:gbv:700-2017032415660/7/thesis_heuer.pdf

Hinze, S. (2016). Forschungsförderung und ihre finanzierung. In D. Simon, A. Knie, S. Hornbostel & K. Zimmermann (Hrsg.), *Handbuch wissenschaftspolitik* (S. 413–428). Wiesbaden: Springer Fachmedien Wiesbaden. doi: 10.1007/978-3-658-05455-7{\textunderscore}15

Holbrook, M. B. (1999). *Consumer value: A framework for analysis and research*. London: Routledge. Zugriff am 13.10.2020 auf http://eprints.stiperdharmawacana.ac.id/28/1/%5BM._Holbrook%5D_Consumer_Value_A_Framework_for_Anal%28BookFi%29.pdf

House, J. S. (1981). *Work stress and social support* (Bd. 4). Reading: Addison-Wesley.

John, M. (2019). *Management interdisziplinärer forschungsverbünde: Institutionelle bedingungen nachhaltiger kooperation in der medizin* (1. Auflage 2019 Aufl.). Wiesbaden: Springer Fachmedien Wiesbaden. doi: 10.1007/978-3-658-24477-4

Johnston, R. (2008). Internal service – barriers, flows and assessment. *International Journal of Service Industry Management*, *19* (2), 210–231. doi: 10.1108/09564230810869748

Jung, H. (2014). *Persönlichkeitstypologie: Menschenkenntnis als instrument der mitarbeiterführung* (4., aktualisierte und erw. Aufl. Aufl.). München: De Gruyter Oldenbourg.

Kang, G.-D., Jame, J. & Alexandris, K. (2002). Measurement of internal service quality: application of the servqual battery to internal service quality. *Managing Service Quality: An International Journal*, *12* (5), 278–291. Zugriff am 23.10.2020 auf https://www.emerald.com/insight/content/doi/10.1108/09604520210442065/full/pdf?title=measurement-of-internal-service-quality-application-of-the-servqual-battery-to-internal-service-quality doi: 10.1108/09604520210442065

Katz, J. & Martin, B. R. (1997). What is research collaboration? *Research Policy*, *26* (1), 1–18. doi: 10.1016/S0048-7333(96)00917-1

Kelle, U. & Kluge, S. (1999). *Vom einzelfall zum typus: Fallvergleich und fallkontrastierung in der qualitativen sozialforschung* (Bd. 4). Wiesbaden: VS Verlag für Sozialwissenschaften.

Kelly, G. (1995). *The psychology of personal constructs: Volume 1: Theory and personality* (Reprinted Aufl., Bd. 1). London: Routledge.

Kleinaltenkamp, M., Brodie, R. J., Frow, P., Hughes, T., Peters, L. D. & Woratschek, H. (2012). Resource integration. *Marketing Theory*, *12* (2), 201–205. doi: 10.1177/1470593111429512

Kleinaltenkamp, M. & Dekanozishvili, D. (2019). The contextual nature of value in use. In A. Parvatiyar & R. Sisodia (Hrsg.), *Handbook of advances in marketing in an era of disruptions* (S. 223–235). Los Angeles and London and New Delhi and Singapore and Washington DC and Melbourne: SAGE Publications India.

Kleinaltenkamp, M., Löbler, H., Eßer, B. & Fennert, N. (2018). Social attribution of value – the case of electronic engagement platforms. In C. Arnold & H. Knödler (Hrsg.), *Die informatisierte service-ökonomie* (S. 93–112). Wiesbaden: Springer Fachmedien Wiesbaden. Zugriff auf https://journals.sagepub.com/doi/pdf/10.1177/1094428112452151

Kleinaltenkamp, M., Storck, F., Gumprecht, P. & Li, J. (2018). The impact of psychological ownership on value in use and relational outcomes. *Journal of Service Management Research*, *2* (2), 50–70. doi: 10.15358/2511-8676-2018-2-50

Kleinbeck, U. (1996). *Arbeitsmotivation: Entstehung, wirkung und förderung*. Weinheim: Juventa Verlag.

Klemenz, B. (2009). *Ressourcenorientierte psychologie: Ermutigende beiträge einer menschenfreundlichen wissenschaft*. Tübingen: Dgvt-Verl.

Knecht, A. & Schubert, F.-C. (2012). *Ressourcen im sozialstaat und in der sozialen arbeit: Zuteilung - förderung - aktivierung*. Stuttgart: Kohlhammer.

König, B., Diehl, K., Tscherning, K. & Helming, K. (2013). A framework for structuring interdisciplinary research management. *Research Policy*, *42* (1), 261–272. doi: 10.1016/j.respol.2012.05.006

Koskela-Huotari, K. & Vargo, S. L. (2016). Institutions as resource context. *Journal of Service Theory and Practice*, *26* (2), 163–178. Zugriff am 22.10.2020 auf https://www.emerald.com/insight/content/doi/10.1108/JSTP-09-2014-0190/full/pdf?title=institutions-as-resource-context doi: 10.1108/JSTP-09-2014-0190

Koskela-Huotari, K., Vink, J. & Edvardsson, B. (2020). The institutional

turn in service research: taking stock and moving ahead. *Journal of Services Marketing*, *34* (3), 373–387. Zugriff am 28.10.2020 auf https://www.emerald.com/insight/content/doi/10.1108/JSM-02-2019-0101/full/pdf?title=the-institutional-turn-in-service-research-taking-stock-and-moving-ahead doi: 10.1108/JSM-02-2019-0101

Kratzer, J., Leenders, R. T. A. J. & van Engelen, J. M. L. (2005). Keeping virtual r&d teams creative. *Research-Technology Management*, *48* (2), 13–16. doi: 10.1080/08956308.2005.11657300

Krohn, W. & Küppers, G. (1989). *Die selbstorganisation der wissenschaft* (1. Auflage Aufl., Bd. 776). Frankfurt am Main: Suhrkamp.

Krott, M. (1996). Interdisziplinarität im netz der disziplinen. In P. W. Balsinger, R. Defila & A. Di Giulio (Hrsg.), *Ökologie und interdisziplinarität - eine beziehung mit zukunft?* (S. 86–97). Basel: Birkhäuser Basel.

Krücken, G. (2009). Wissenschaftlerinnen und wissenschaftler als wissenschaftsmanager? In L. Guzy, A. Mihr & R. Scheepers (Hrsg.), *Wohin mit uns?* (S. 83–94). Frankfurt am Main and New York: Peter Lang.

Lamnek, S. (2010). *Qualitative sozialforschung: Lehrbuch. mit online-materialien* (5. Aufl. Aufl.). Weinheim: Beltz.

Laudel, G. (1999). *Interdisziplinäre forschungskooperation: Erfolgsbedingungen der institution "sonderforschungsbereich"*. Berlin: Edition Sigma rainer bohn verlag.

Laudel, G. (2001). Collaboration, creativity and rewards: why and how scientists collaborate. *International Journal of Technology Management*, *22* (7/8), 762–781. doi: 10.1504/IJTM.2001.002990

Lee, R. P. (2015). Corporate governance und strukturen. In A. Hanebuth, R. P. Lee, S. Meschke & M. Nicklas (Hrsg.), *Forschungskooperationen zwischen wissenschaft und praxis* (S. 37–72). Wiesbaden: Springer Fachmedien Wiesbaden.

Lemke, F., Clark, M. & Wilson, H. (2011). Customer experience quality: an exploration in business and consumer contexts using repertory grid technique. *Journal of the Academy of Marketing Science*, *39* (6), 846–869. doi: 10.1007/s11747-010-0219-0

Lippe, S. & vom Brocke, J. (2016). Situational project management for collaborative research projects. *Project Management Journal*, *47* (1), 76–96. doi: 10.1002/pmj.21561

Löbler, H. & Hahn, M. (2013). Measuring value-in-context from a service-dominant logic's perspective. *Review of Marketing Research*, *10*, 255–282. doi: 10.1108/S1548-6435(2013)0000010013

Locker-Grütjen, O. (2011). *Rahmenbedingungen für forschung an deutschen universitäten: „wissenschaft in ketten oder entfesselte hochschule?“: Ansätze eines modernen forschungsmanagements in zeiten des wandels in der deutschen hochschullandschaft* (Dissertation, Universität Duisburg-Essen). Zugriff am 23.10.2020 auf https://d-nb.info/1020790202/34

Locker-Grütjen, O., Ehmann, B. & Jongmanns, G. (2012). Definition für optimales forschungsmanagement: Wie umfassend kann forschung unterstützt werden? *Wissenschaftsmanagement* (3), 34–38. Zugriff am 13.10.2020 auf https://www.wissenschaftsmanagement.de/dateien/dateien/archiv/downloaddateien/wim_2012_03.pdf

Lusch, R. F. & Vargo, S. L. (Hrsg.). (2006). *The service-dominant logic of marketing: Dialog, debate, and directions*. London and New York: Routledge, Taylor & Francis Group.

Lusch, R. F. & Vargo, S. L. (2014). *Service-dominant logic: Premises, perspectives, possibilities*. Cambridge: Cambridge University Press.

Lusch, R. F., Vargo, S. L. & O'Brien, M. (2007). Competing through service: Insights from service-dominant logic. *Journal of Retailing*, *83* (1), 5–18. doi: 10.1016/j.jretai.2006.10.002

Lux, A. M. & Robra-Bissantz, S. (2020). Des schusters schuhe. *HMD Praxis der Wirtschaftsinformatik*, *57* (4), 697–709. doi: 10.1365/S40702-020-00629-4

Macdonald, E. K., Kleinaltenkamp, M. & Wilson, H. N. (2016). How business customers judge solutions: Solution quality and value in use. *Journal of Marketing*, *80* (3), 96–120. Zugriff auf http://dx.doi.org/10.1509/jm.15.0109 doi: 10.1509/jm.15.0109

Macdonald, E. K., Wilson, H., Martinez, V. & Toossi, A. (2011). Assessing value-in-use: A conceptual framework and exploratory study. *Industrial Marketing Management*, *40* (5), 671–682. Zugriff am 29.09.2020 auf https://www.emerald.com/insight/content/doi/10.1108/JSM-02-2015-0065/full/pdf?title=studying-customers-resource-integration-by-service-employees-in-interactional-value-co-creation doi: 10.1016/j.indmarman.2011.05.006

Madhavaram, S. & Hunt, S. D. (2008). The service-dominant logic and a hierarchy of operant resources: developing masterful operant resources and implications for marketing strategy. *Journal of the Academy of Marketing Science*, *36* (1), 67–82. doi: 10.1007/s11747-007-0063-z

May, T. Y.-m., Korezynski, M. & Frenkel, S. J. (2002). Organizational and occupational commitment: Knowledge workers in large corporations. *Journal of Management Studies*, *39* (6), 775–801. doi: 10.1111/1467-6486.00311

Mayring, P. (1983). *Qualitative inhaltsanalyse: Grundlagen und techniken* (Neuausg Aufl.). Weinheim: Beltz.

Mayring, P. (1996). *Einführung in die qualitative sozialforschung: Eine anleitung zu qualitativem denken* (3., überarbeitete Auflage Aufl.). Weinheim: Psychologie Verlags Union.

Mayring, P. (2015). *Qualitative inhaltsanalyse: Grundlagen und techniken* (12. Auflage Aufl.). Weinheim und Basel: Beltz Verlag.

Medberg, G. & Grönroos, C. (2020). Value-in-use and service quality: do customers see a difference? *Journal of Service Theory and Practice*, *30* (4/5), 507–529. Zugriff am 08.10.2020 auf https://www.emerald.com/insight/content/doi/10.1108/JSTP-09-2019-0207/full/pdf?title=value-in-use-and-service-quality-do-customers-see-a-difference doi: 10.1108/JSTP-09-2019-0207

Meschke, S. (2015). Führung und motivation von wissenschaftlern. In A. Hanebuth, R. P. Lee, S. Meschke & M. Nicklas (Hrsg.), *Forschungskooperationen zwischen wissenschaft und praxis* (S. 101–132). Wiesbaden: Springer Fachmedien Wiesbaden.

Morar, D. D. (2013). An overview of the consumer value literature – perceived value, desired value. *Marketing From Information to Decision*, *6*, 169–186.

Morelli, N. (2009). Service as value co–production: reframing the service design process. *Journal of Manufacturing Technology Management*, *20* (5), 568–590. Zugriff am 01.11.2020 auf https://www.emerald.com/insight/content/doi/10.1108/17410380910960993/full/pdf?title=service-as-value-coproduction-reframing-the-service-design-process doi: 10.1108/17410380910960993

North, D. C. (2006). *Institutions, institutional change and economic preformance* (22nd printing Aufl.). Cambridge: Cambridge University Press.

Ostrom, A. L., Bitner, M. J., Brown, S. W., Burkhard, K. A., Goul, M., Smith-Daniels, V., … Rabinovich, E. (2010). Moving forward and making a difference: Research priorities for the science of service. *Journal of Service Research*, *13* (1), 4–36. Zugriff am 06.12.2020 auf https://journals.sagepub.com/doi/pdf/10.1177/1094670509357611 doi: 10.1177/1094670509357611

Ostrom, A. L., Parasuraman, A., Bowen, D. E., Patrício, L. & Voss, C. A. (2015). Service research priorities in a rapidly changing context. *Journal of Service Research*, *18* (2), 127–159. doi: 10.1177/1094670515576315

Payne, A. F., Storbacka, K. & Frow, P. (2008). Managing the co-creation of value. *Journal of the Academy of Marketing Science*, *36* (1), 83–96. doi: 10.1007/s11747-007-0070-0

Peters, L. D., Löbler, H., Brodie, R. J., Breidbach, C. F., Hollebeek, L. D., Smith, S. D., … Varey, R. J. (2014). Theorizing about resource integration through service-dominant logic. *Marketing Theory*, *14* (3), 249–268. Zugriff am 28.10.2020 auf https://journals.sagepub.com/doi/pdf/10.1177/1470593114534341 doi: 10.1177/1470593114534341

Pfisterer, L. (2017). *Wertkreation in kundennutzungsprozessen*. Springer Gabler.

Plé, L. (2016). Studying customers' resource integration by service employees in interactional value co-creation. *Journal of Services Marketing*, *30* (2), 152–164. Zugriff am 19.10.2020 auf https://www.emerald.com/insight/content/doi/10.1108/JSM-02-2015-0065/full/pdf?title=studying-customers-resource-integration-by-service-employees-in-interactional-value-co-creation

Plé, L. & Chumpitaz Cáceres, R. (2010). Not always co–creation: introducing interactional co–destruction of value in service–dominant logic. *Journal of Services Marketing*, *24* (6), 430–437. Zugriff am 02.05.2021 auf https://www.emerald.com/insight/content/doi/10.1108/08876041011072546/full/pdf?title=not-always-cocreation-introducing-interactional-codestruction-of-value-in-servicedominant-logic doi: 10.1108/08876041011072546

Raghubir, P. & Corfman, K. (1999). When do price promotions affect pretrial brand evaluations? *Journal of Marketing Research*, *36* (2), 211–222. doi: 10.1177/002224379903600206

Ranjan, K. R. & Read, S. (2016). Value co-creation: concept and measurement. *Journal of the Academy of Marketing Science*, *44* (3), 290–315. doi: 10.1007/S11747-014-0397-2

Reichertz, J. (2014). Die konjunktur der qualitativen sozialforschung und konjunkturen innerhalb der qualitativen sozialforschung. In G. Mey & K. Mruck (Hrsg.), *Qualitative forschung* (S. 87–102). Wiesbaden: Springer Fachmedien Wiesbaden and Imprint: Springer VS.

Reynolds, T. J. & Gutmann, J. (1988). Laddering theory, method, analysis, and interpretation. *Journal of Advertising Research*, 11–31. Zugriff auf https://is.muni.cz/el/econ/jaro2013/mph_mvps/39278324/ladderingtheoy_original.pdf

Richman, W. L., Kiesler, S., Weisband, S. & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology*, *84* (5), 754–775. doi: 10.1037/0021-9010.84.5.754

Sauer, H. & Weibel, A. (2012). Formalisierung und wohlbefinden am arbeitsplatz: Neue perspektive auf eine kontroverse. In P. Conrad & J. Koch (Hrsg.), *Steuerung durch regeln* (S. 1–41). Wiesbaden: Gabler Verlag. doi: 10.1007/978-3-8349-4349-1{\textunderscore}1

Schaper, N. (2014). Neue formen der arbeit: Das beispiel telekooperation. In F. W. Nerdinger, G. Blickle & N. Schaper (Hrsg.), *Arbeits- und organisationspsychologie* (S. 541–578). Berlin, Heidelberg: Springer Berlin Heidelberg and Imprint: Springer.

Schneider, F., Buser, T., Keller, R., Tribaldos, T. & Rist, S. (2019). Research funding programmes aiming for societal transformations: Ten key stages. *Science and Public Policy*, *46* (3), 463–478. doi: 10.1093/scipol/scy074

Schneijderberg, C., Merkator, N., Teichler, U. & Kehm, B. M. (2013). *Verwaltung war gestern: Neue hochschulprofessionen und die gestaltung von studium und lehre*. Frankfurt, M. and New York, NY: Campus-Verl.

Schophaus, M., Dienel, H.-L. & von Braun, C.-F. (2004). Brücken statt einbahnstraßen: Lösungsorientiertes kooperationsmanagement für die interdisziplinäre forschung. *Wissenschaftsmanagement*, *2004* (2), 16–25. Zugriff am 19.04.2021 auf https://www.wissenschaftsmanagement.de/dateien/dateien/management/downloaddateien/wim_2004_02_malte_schophaus_hans_liudger_dienel_christop_friedrich_von_braun_bruecken_statt_einbahnstrassen.pdf

Schubert, F.-C. & Knecht, A. (2012). Ressourcen - einführung in merkmale, theorien und konzeptionen. In A. Knecht & F.-C. Schubert (Hrsg.), *Ressourcen im sozialstaat und in der sozialen arbeit* (S. 15–41). Stuttgart: Kohlhammer.

Semmer, N., Jacobshagen, N., Meier, L. & Elfering, A. (2012). Occupational stress research: The stress-as-offense-to-self perspective. In J. Houdmont (Hrsg.), *Contemporary occupational health psychology* (S. 41–58). Oxford: Wiley-Blackwell. doi: 10.7892/BORIS.26201

Siedlok, F. & Hibbert, P. (2014). The organization of interdisciplinary research: Modes, drivers and barriers. *International Journal of Management Reviews*, *16* (2), 194–210.

Stampacchia, P., Tregua, M. & Coppola, M. (2020). Zooming-in value-in-use through basic individual values. *Journal of Customer Behaviour*, *19* (2), 97–125. doi: 10.1362/147539220X15929906305116

Stokols, D., Misra, S., Moser, R. P., Hall, K. L. & Taylor, B. K. (2008). The ecology of team science: understanding contextual influences on transdisciplinary collaboration. *American journal of preventive medicine*, *35* (2), 96–115. doi: 10.1016/j.amepre.2008.05.003

Strauss, B. (2010). Internes relationship management: Das beziehungsmanagement interner dienstleister. In D. Georgi & K. Hadwich (Hrsg.), *Management von kundenbeziehungen* (S. 27–51). Wiesbaden: Gabler.

Super, D. (1970). *Work values inventory.* Zugriff am 18.02.2021 auf https://www.hccs.edu/media/houston-community-college/Super_Work-Values-Inventory.pdf

Sweeney, J. C., Plewa, C. & Zurbruegg, R. (2018). Examining positive and negative value-in-use in a complex service setting. *European Journal of Marketing*, *52* (5/6), 1084–1106. doi: 10.1108/EJM-10-2016-0582

Ulaga, W. & Eggert, A. (2006). Relationship value and relationship quality. *European Journal of Marketing*, *40* (3/4), 311–327. Zugriff am 06.05.2021 auf https://www.emerald.com/insight/content/doi/10.1108/03090560610648075/full/pdf?title=relationship-value-and-relationship-quality-broadening-the-nomological-network-of-businesstobusiness-relationships doi: 10.1108/03090560610648075

Vargo, S. L., Koskela-Huotari, K. & Vink, J. (2020). Service-dominant logic: Foundations and applications. In E. Bridges & K. Fowler (Hrsg.), *The routledge handbook of service research insights and ideas* (S. 3–23). New York: Routledge.

Vargo, S. L. & Lusch, R. F. (2004). Evolving to a new dominant logic for marketing. *Journal of Marketing*, *68*, 1–17. doi: 10.1509/jmkg.68.1.1.24036

Vargo, S. L. & Lusch, R. F. (2006). Service-dominat logic: What it is, what it is not, what it might be. In R. F. Lusch & S. L. Vargo (Hrsg.), *The service-dominant logic of marketing* (S. 43–56). London and New York: Routledge, Taylor & Francis Group.

Vargo, S. L. & Lusch, R. F. (2008a). The nature and understanding of value: A service-dominant logic perspective. *European Management Journal*, *26*, 145–152. doi: 10.1108/S1548-6435(2012)0000009005

Vargo, S. L. & Lusch, R. F. (2008b). Service-dominant logic: continuing the evolution. *Journal of the Academy of Marketing Science*, *36*, 1–10. Zugriff auf http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=1&sid=5f144032-05cf-4a64-ab88-f99ac1e01c9c%40sdc-v-sessmgr01 doi: 10.1007/s11747-007-0069-6

Vargo, S. L. & Lusch, R. F. (2017). Service-dominant logic 2025. *International Journal of Research in Marketing*, *34* (1), 46–67. doi: 10.1016/j.ijresmar.2016.11.001

Vargo, S. L., Maglio, P. P. & Akaka, M. A. (2008). On value and value co-creation: A service systems and service logic perspective. *European Management Journal*, *26* (3), 145–152. doi: 10.1016/j.emj.2008.04.003

Vowe, G. & Meißner, F. (2020). Erfolg und misserfolg von kommunikationswissenschaftlichen dfg-forschungsverbünden: Bilanz, faktoren, konsequenzen. *Publizistik*, *65* (2), 151–185. doi: 10.1007/s11616-020-00575-7

Wieser, B., Brechelmacher, A. & Schendel, G. (2014). Identitäten und rollen in inter- und transdisziplinärer forschung und lehre finden. In G. Dressel (Hrsg.), *Interdisziplinär und transdisziplinär forschen* (S. 151–164). Bielefeld: Transcript.

Willutzki, U. (2003). Ressourcen: Einige bemerkungen zur begriffsklärung. In H. Schemmel (Hrsg.), *Ressourcen* (S. 91–110). Tübingen: Dgvt-Verl.

Winterhager, N. (2015). *Drittmittelwettbewerb im universitären forschungssektor*. Wiesbaden: Springer VS. Zugriff am 15.05.2021 auf https://link.springer.com/content/pdf/10.1007%2F978-3-658-07188-2.pdf

Woodruff, R. B. (1997). Customer value: The next source for competitive advantage. *Journal of the Academy of Marketing Science*, *25* (2), 139–153.

Woodruff, R. B. & Flint, D. J. (2006). Marketing's service-dominant logic and customer value. In R. F. Lusch & S. L. Vargo (Hrsg.), *The service-dominant logic of marketing* (S. 183–197). London and New York: Routledge, Taylor & Francis Group.

Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *Journal of Marketing*, *52* (3), 2–22. Zugriff am 13.10.2020 auf https://www.jstor.org/stable/1251446?origin=crossref&seq=1#metadata_info_tab_contents

Junior Management Science 7(2) (2022) 504-523

Junior Management Science

journal homepage: www.jums.academy

Inter-Organizational Relations of Social Enterprises – A Systematic Literature Review

Zur Rolle interorganisationaler Beziehungen im Social Entrepreneurship – Ein systematisches Literaturreview

Sophie Gröbner

Freie Universität Berlin

Abstract

A social enterprise is a company founded by a social entrepreneur following a social purpose. Inter-organizational relations of social enterprises are an important topic in research on social entrepreneurship. The aim of this paper is to provide an overview regarding these relations in the context of a literature review. The findings are seperated into the following three cooperation phases: initiation, process and outcomes of social enterprises' inter-organizational relations. Furthermore, insights about the functions of network management are collected. The findings show that the most articles focus on the cooperation process and strategic networks are examined most often. In addition, it could be observed that cooperative alliances with other organizations play an important part in social entrepreneurial practice. Inter-organizational relations of social enterpises represent a promising topic for future research. Research about aspects concerning the initiation phase and the outcomes of these relations is especially required. The functions of network management need to be further examined in this research context as well.

Zusammenfassung

Als Social Enterprise wird ein Unternehmen bezeichnet, das von einem Social Entrepreneur für die Umsetzung seiner sozialen Ziele gegründet wird. Interorganisationale Beziehungen dieser Unternehmen stellen eine wichtige Thematik im Forschungsbereich zu Social Entrepreneurship dar. Das Ziel dieser Arbeit ist es, im Rahmen eines Literaturreviews einen Überblick zu diesen Beziehungen zu schaffen. Die Erkenntnisse werden in die folgenden drei Kooperations-Phasen unterteilt: Anbahnung, Prozess und Folgen interorganisationaler Beziehungen von Social Enterprises. Weiterhin werden Erkenntnisse zu den Funktionen des Netzwerkmanagements gesammelt. Die Ergebnisse zeigen, dass der Kooperationsprozess interorganisationaler Beziehungen von Social Enterprises am meisten untersucht wird und dabei strategische Netzwerke häufig im Fokus stehen. Weiterhin konnte festgestellt werden, dass kooperative Bündnisse mit Organisationen für Social Enterprises in der Praxis eine herausragende Rolle spielen. Interorganisationale Beziehungen von Social Enterprises weisen viel zukünftiges Forschungspotenzial auf. Ein relativ großer Forschungsbedarf besteht bezüglich der Aspekte zur Anbahnung und zu Ergebnissen der Organisationsbeziehungen von Social Enterprises und bei den Funktionen des Netzwerkmanagements.

Keywords: Social Entrepreneurship; Social Enterprise; Interorganisationale Beziehungen; Netzwerke; Netzwerkmanagement.

1. Einleitung

Das Forschungsfeld um Social Entrepreneurship (SE) ist ein relativ junges, das in den letzten 30 Jahren zunehmend das wissenschaftliche Interesse geweckt hat (Conway Dato-on & Kalakay, 2016, S. 132) und bereits einen beachtlichen Literaturbestand vorweisen kann (Certo & Miller, 2008; Chell, 2007; Doherty, Haugh & Lyon, 2014; Perrini, Vurro & Costanzo, 2010). Zu Anfang fokussierten sich die Untersuchungen auf die Definition und Erforschung des Social Entrepreneurship-Konstrukts (Conway Dato-on & Kalakay, 2016, S. 133). Dabei wurde das Social Enter-prise als besondere Organisationsform hervorgehoben und des-

DOI: https://doi.org/10.5282/jums/v7i2pp504-523

sen charakteristische Merkmale und Entstehungsursachen näher betrachtet. Aktuelle Forschungsbeiträge beschäftigen sich mit der Entwicklung von neuen Theorien, die die Entstehung, das Management sowie ethische Aspekte von Social Enter-prises erklären sollen (Doherty et al., 2014, S. 417). Des Weiteren stellen interorganisationale Beziehungen von Social Enterprises eine wichtige und viel diskutierte Thematik in diesem Forschungsbereich dar (Di Domenico, Tracey & Haugh, 2009; Huybrechts, Nicholls & Edinger, 2017; Montgomery, Dacin & Dacin, 2012). Jedoch wurde bisher noch kein Überblick zu diesen Beziehungen in Verbindung mit Social Enterprises geschaffen.

Im Rahmen der Bachelorarbeit wird dieser Forschungslücke mit Hilfe eines Literaturreviews nachgegangen und demnach folgende Forschungsfrage beantwortet: „Welche Themen wurden bereits in der wissenschaftlichen Literatur zu interorganisationalen Beziehungen von Social Enterprises betrachtet?“ Ziel der Arbeit ist es, bisheriges Wissen herauszuarbeiten und anschließende Forschungsansätze aufzuzeigen. Die Thematik ist mit dem Begriff „Beziehungen“ weit gefasst, um auch hierarchischere und marktlichere Formen von Zusammenschlüssen in den Kontext aufzunehmen.

Zur Systematisierung der Ergebnisse werden die interorganisationalen Beziehungen in zwei Unterkategorien gegliedert: dyadische Beziehungen und interorganisationale Netzwerke. Es wird dadurch aufgezeigt, inwiefern sich Partnerschaften von komplexeren Zusammenschlüssen unterscheiden. Anschließend werden die Netzwerke nach der „Typologie interorganisationaler Netzwerke“ (Sydow, 2010, S. 382) aufgegliedert. Diese Unterteilung führt zu einer besseren Übersicht über die unterschiedlichen Eigenschaften der Bündnisse.

Bei der Untersuchung der Dyaden und Netzwerke werden horizontale, vertikale und laterale Beziehungen zwischen den Partnern unterschieden. Somit liegt der Fokus zuerst auf Organisationen, die unter Betrachtung der Wertschöpfungskette entweder auf derselben Stufe angesiedelt oder der fokalen Organisation vor- bzw. nachgelagert sind. Im Anschluss werden laterale Zusammenschlüsse von Organisationen betrachtet, die aus Sicht der Wertschöpfung weder vertikale noch horizontale Verbindungen miteinander aufweisen (Sydow & Duschek, 2011, S. 33ff.).

Die Bachelorarbeit beginnt mit der Definition wichtiger Begriffe, wie Social En-trepreneurship und Social Enterprise. Darauf folgt die Beschreibung der verschiedenen Arten und Richtungen von interorganisationalen Beziehungen sowie der Managementfunktionen im Netzwerk. Anschließend wird die Methodik des Literaturreviews erläutert und die Ergebnisse der Analyse werden präsentiert. Im Fazit werden die wichtigsten Erkenntnisse zusammengefasst und relevante Forschungsansätze benannt.

2. Social Entrepreneurship

Entrepreneurship beschreibt im Allgemeinen das Gründen eines Unternehmens (Faltin, 2011, S. 76). Schumpeter hat den Begriff in den Wirtschaftswissenschaften nachhaltig geprägt und weist der Person des Entrepreneurs die Aufgabe zu, Innovationen hervorzubringen. Weiterhin grenzt er den Entrepreneur von anderen Akteuren, wie dem Erfinder, dem Kapitalgeber und dem Manager, ab (Schumpeter, 1993).

Die Interessen des Entrepreneurs des Business Sektors unterscheiden sich von denen des Social Entrepreneurs. Für den Business Entrepreneur stellt die Erwirtschaftung von Überschüssen das Ziel seiner Unternehmung dar. Der Social Entrepreneur verfolgt dagegen das Ziel, seine soziale Aufgabe zu erfüllen. Monetäre Ergebnisse stehen nicht im Vordergrund, sondern die Ergebnisse, die er durch seine Mission erreicht. Die erwirtschafteten Überschüsse des Social En-trepreneurs betrachtet dieser nur als Mittel für sein Vorhaben. Aufgrund seiner sozialen Mission muss der Social Entrepreneur demnach spezielle Herausforderungen meistern, denen der Business Entrepreneur nicht gegenüber steht. Somit stellen Social Entrepreneure Pioniere dar, die neue Ansätze für ihre Arbeit nutzen (Faltin, 2011, S. 77f.). Faltin (2011, S. 78) wird „Social Entrepreneurship [...] daher als ‚pioneering new approaches‘ [sic] verstanden und beschreibt einen eigenen Ansatz, der eine wertvolle Ergänzung darstellt zu Vielem, das bereits existiert.“ Dies ist jedoch nur eine von vielen Definitionen, da zu dem Begriff bisher keine einheitliche formuliert worden ist (Certo & Miller, 2008; Short, Moss & Lumpkin, 2009).

2.1. Definitionsproblematik

Obwohl das Forschungsfeld des Social Entrepreneurships im letzten Jahrzehnt immer mehr an Aufmerksamkeit gewonnen hat, sind Forschung und Praxis noch weit davon entfernt, sich darauf zu einigen, was der Begriff genau bedeutet. Deshalb gibt es aktuell viele konkurrierende Definitionen (Choi & Majumdar, 2014, S. 364). Choi und Majumdar (2014, S. 364) stellen folgende Begriffserklärungen für Social Entrepreneurship aus verschiedenen Studien zusammen: (I) Verbindung des Begriffs mit Not-For-Profit Organisationen, die durch ihre Beteiligung am Markt zu neuen Finanzierungsstrategien gelangen (Lasprogata & Cotten, 2003), (II) Gründung von Unternehmen zur Unterstützung von Hilfsbedürftigen (Seelos & Mair, 2005) und (III) Nutzung von sozialen Innovationen, um soziale Probleme zu lösen und sozialen Wandel hervorzurufen, unabhängig davon, ob kommerzielle Aktivitäten dabei involviert sind (Martin & Osberg, 2007). Short et al. (2009, S. 162) fällt es durch das Fehlen der einheitlichen Definition schwer, die Legitimität des Forschungsfeldes und des Konstrukts zu vermitteln. Mehrere Forscher haben bereits die vielfältigen Bedeutungen von Social Entrepreneurship in der wissenschaftlichen Literatur zusammengetragen und untersucht. Eine Lösung der Definitionsproblematik gibt es bisher jedoch nicht (Choi & Majumdar, 2014, S. 364).

2.2. Social Enterprise

Das von einem Social Entrepreneur gegründete Unternehmen wird als Social Enterprise bezeichnet. Es stellt das Instrument dar, das der Social Entrepreneur für die Umsetzung

seiner sozialen Ziele nutzt. Hierbei herrscht in Forschung und Praxis - ähnlich wie bei dem Begriff Social Entrepreneurship – Uneinigkeit darüber, was ein Social Enterprise genau ausmacht (Littlewood & Khan, 2018, S. 391). Allerdings wurden von Littlewood und Khan (2018, S. 391) bereits einige wesentliche, in der wissenschaftlichen Literatur beschriebene Charaktereigenschaften von Social Enterprises zusammengetragen:

- Zentralität der sozialen oder ethischen Mission; Vorrang der sozialen Wertschaffung gegenüber der Schaffung von ökonomischem Wert (Defourny & Nyssens, 2006)
- Erwirtschaftung von Erträgen durch kommerzielle Aktivitäten (Langdon & Burkett, 2004)
- Beteiligung der Stakeholder an der Unternehmensführung (Defourny & Nyssens, 2006)
- Begrenzte Ausschüttung von Gewinnen (Langdon & Burkett, 2004)
- Schaffung von Innovationen zur Bewältigung sozialer Probleme (Dees, 2003)

Da Social Enterprises gegründet werden, um sowohl finanzielle als auch soziale Ziele zu erreichen, stellt diese Art der Unternehmung eine hybride Organisationsform dar (Doherty et al., 2014, S. 419). Der hybride Charakter von Social Enterprises wird besonders gut durch die Definition des Britischen Ministeriums für Handel und Industrie erfasst: „A social enterprise is a business with primarily social objectives, whose surpluses are principally reinvested for that purpose in the business or in the community, rather than being driven by the need to ma-ximise profit for shareholders and owners." (Department of Trade and Industry, 2002, S. 8). Auf dieser weit formulierten Definition beruht die Auffassung des Social Enterprise-Begriffs dieser Arbeit. Hierbei wird nämlich die soziale Mission des Unternehmens, aber auch das Auftreten der Unternehmung als Anbieter am Markt thematisiert. Dies führt zur Erwirtschaftung von Überschüssen, welche hauptsächlich für den sozialen Zweck des Social Enterprises reinvestiert werden und nicht der Gewinnausschüttung dienen.

In der wissenschaftlichen Literatur werden für den Begriff des Social Enter-prises ebenso die Bezeichnungen „Social Venture" und „Social Business" verwendet. In der Regel stellen sie Synonyme dar, die ebenfalls Unternehmen mit sozialen und finanziellen Zielen beschreiben. So definiert beispielsweise Yunus, Moingeon und Lehmann-Ortega (2010, S. 311) ein Social Business als „no-loss, no-dividend, self-sustaining company that sells goods or services and repays investments to its owners, but whose primary purpose is to serve society and improve the lot of the poor." Für das Literaturreview wurden alle drei Varianten berücksichtigt – Social Enterprise, Social Venture und Social Business.

3. Interorganisationale Beziehungen

3.1. Arten interorganisationaler Beziehungen

Eine Beziehung zwischen zwei Akteuren wird als Dyade bezeichnet (Miemczyk, Johnsen & Macquet, 2012, S. 479). Wenn hierbei ausschließlich Organisationen beteiligt sind, handelt es sich um eine dyadische interorganisationale Beziehung. Ein Zusammenschluss von mehr als zwei Organisationen zum Erreichen eines gemeinsamen Ziels stellt ein interorganisationales Netzwerk dar (Provan, Fish & Sydow, 2007, S. 482). Schließt ein solches Netzwerk Unternehmen als Teilnehmer ein, wie z.B. Social Enterprises, wird es als Unternehmungsnetzwerk bezeichnet. Sydow (1992, S. 79) stellt dies „eine auf die Realisierung von Wettbewerbsvorteilen zielende Organisationsform ökonomischer Aktivitäten dar, die sich durch komplex-reziproke, eher kooperative denn kompetitive und relativ stabile Beziehungen zwischen rechtlich selbständig, wirtschaftlich jedoch zumeist abhängigen Unternehmungen auszeichnet."

Unternehmungsnetzwerke werden im Rahmen einer Typologie nach ihrer Steuerungsform und nach Stabilität ihrer interorganisationalen Beziehungen unterschieden. Die Steuerungsform des Netzwerks kann vorwiegend hierarchische oder heterarchische Merkmale aufweisen (Sydow & Möllering, 2015, S. 232). Die Hierarchie zeichnet sich durch Vorgesetzte aus, die Untergebenen Anweisungen erteilen können. Die Machtverteilung ist somit asymmetrisch (Sydow & Möllering, 2015, S. 22). Im Gegensatz zur monozentrischen Struktur der Hierarchie ist die Struktur einer Heterarchie polyzentrisch ausgelegt, d.h. sie beinhaltet mindestens zwei Steuerungszentren und bevorzugt Selbst- statt Fremdorganisation (Sydow & Möllering, 2015, S. 232f.).

Bei der Stabilität des Netzwerks wird zwischen Formen unterschieden, die eher stabiler oder eher dynamischer sind (Snow, Miles & Coleman, 1992, S. 13f.). Die Dimensionen der Bindungsintensität und des Bindungshorizonts bestimmen, wie stabil die Beziehungen im Netzwerk sind. Dabei stellt der Bindungshorizont nicht die tatsächliche zeitliche Dauer der Beziehung dar, sondern beschreibt, ob die Beteiligten subjektiv wahrnehmen, dass die Bindung eher langfristig oder eher kurzfristig ist (Möllering, 2010, S. 782f.).

Die Betrachtung der Steuerungsform und Stabilität von Netzwerken führt zur Unterscheidung von vier charakteristischen Typen, die in der folgenden Matrix (siehe Abbildung 1) eingeordnet werden: strategische und regionale Netzwerke sowie Projektnetzwerke und virtuelle Unternehmungen (Sydow, 2010, S. 381f.).

Bei *strategischen Netzwerken* gibt es eine oder mehrere Unternehmungen, die die Netzwerkführerschaft innehaben (Dhanaraj & Parkhe, 2006, S. 659). Diese fokalen Unternehmungen geben den für das Netzwerk relevanten Markt vor sowie Strategien, Technologien und Ausgestaltung des Netzwerkes. Es sind somit hierarchische Netzwerkeigenschaften zu erkennen. Typischerweise sind hierbei Unternehmungen verschiedener Größe beteiligt, jedoch stellt tendenziell ein Großunternehmen, das nah am Endverbraucher agiert, die fokale Unternehmung dar. Formelle Regeln, vor allem in

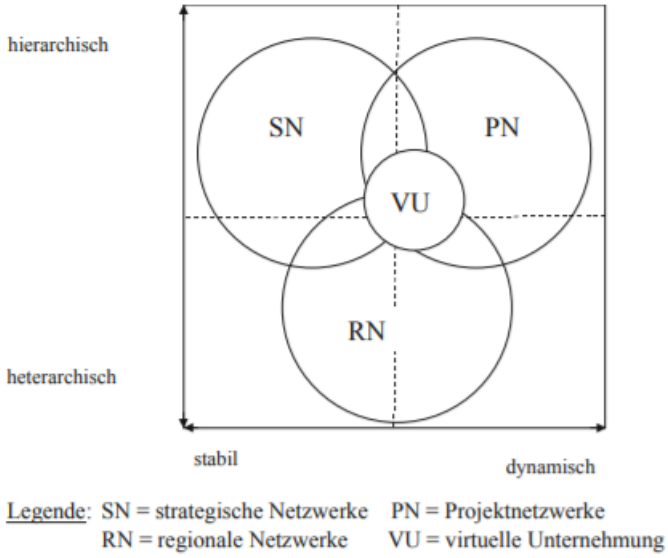

Abbildung 1: Eine Typologie interorganisationaler Netzwerke (Sydow, 2010, S. 382)

Form von Verträgen, prägen hier zwar die Koordination, allerdings spielt auch Vertrauen eine große Rolle, da Verträge unvollständig sind und so auch informelle Regeln Verwendung finden (Sydow & Möllering, 2015, S. 233f.).

Regionale Netzwerke zeichnet die Beteiligung eher kleinerer und mittlerer Unternehmen aus, die oftmals räumlich nicht weit voneinander entfernt angesiedelt sind. Hierbei liegt eine heterarchische Organisation vor, d.h. die Netzwerkführerschaft fehlt. Daraus resultiert eine emergente Strategie und informelle Regeln haben zumeist mehr Bedeutung als formelle (Sydow & Möllering, 2015, S. 234).

Projektnetzwerke unterscheiden sich durch einen festen zeitlichen Rahmen ihrer Aktivitäten von den bereits beschriebenen Typen. Die entstandenen Beziehungen bleiben nach Abschluss des Projekts latent bestehen. An sie wird bei Aufnahme eines neuen Projekts wieder angeknüpft. Eine projektübergreifende Strategie gibt es in den meisten Fällen nicht (Sydow & Möllering, 2015, S. 234).

Virtuelle Unternehmungen stellen Projektnetzwerke dar, die durch informationstechnische Verfahren unterstützt werden (Sydow & Möllering, 2015, S. 234).

3.2. Richtungen interorganisationaler Beziehungen

Es wird zwischen vertikalen, horizontalen und lateralen Interorganisationsbeziehungen unterschieden (Sydow & Duschek, 2011, S. 32).

Vertikale interorganisationale Beziehungen sind der im Fokus stehenden Organisation in der Wertschöpfungskette entweder vor- oder nachgelagert (upstream and downstream). Ein Beispiel für vorgelagerte Unternehmungen stellen Rohstofflieferanten einer fokalen Organisation dar und Handelsbetriebe für nachgelagerte Unternehmen (Sydow & Duschek, 2011, S. 33).

Bei *horizontalen* Interorganisationsbeziehungen befinden sich die betrachteten Organisationen auf derselben Stufe der Wertschöpfungskette. Dies ist bei Wettbewerbern der Fall, welche beispielsweise gemeinschaftlich ihre Produktion im Sinne einer F&E-Allianz koordinieren (Sydow & Duschek, 2011, S. 36).

Organisationen, die unter Betrachtung der Wertschöpfungskette weder eine vertikale noch eine horizontale Verbindung aufweisen, stehen in *lateraler* Beziehung zueinander. Sie verknüpfen verschiedene Stufen von mehreren Wertketten miteinander. Hierbei sind Beziehungen zwischen der fokalen Organisation mit ihren verschiedenen Stakeholder-Gruppen bedeutsam, die unterschiedliche Interessen bezüglich der Organisation verfolgen und ihre Ansprüche durchsetzen wollen. Dies sind unter dem Gesichtspunkt von interorganisationalen Beziehung z.B. Regierungsorganisationen und Gewerkschaften (Sydow & Duschek, 2011, S. 36f.).

3.3. Management interorganisationaler Beziehungen

Es ist zwischen vier zentralen Funktionen zu differenzieren, die allgemein für das Management interorganisationaler Beziehungen und insbesondere im Netzwerk angewendet werden: Selektionsfunktion, Allokationsfunktion, Regulationsfunktion und Evaluationsfunktion (Sydow, 2010, S. 394). Diese werden im Folgenden näher beschrieben:

Im Rahmen der *Selektionsfunktion* wird entschieden, welche Organisationen in das Netzwerk aufgenommen werden sollen (Sydow & Möllering, 2015, S. 195). Dabei werden die passenden Mitglieder danach ausgewählt, ob ihre Kompetenzen zum Erfüllen der Ziele des Netzwerks geeignet sind (Sydow, 2010, S. 395). Weiterhin wird die „Netzwerkdomäne“ (Sydow & Möllering, 2015, S. 396) festgelegt. Die Selektion stellt eine Aufgabe dar, die kontinuierlich durchgeführt werden muss, da es im Laufe der Zeit zum Ausschluss von Mitgliedern aus dem Netzwerk kommen kann (Sydow, 2010, S. 396). Dieser Vorgang wird als „De-Selektion“ (Sydow & Möllering, 2015, S. 195) bezeichnet. Dagegen verblei-

ben Mitglieder, die nach ihrer Aufnahme langfristig kooperativ mit den anderen Teilnehmern zusammenarbeiten, durch Re-Selektion im Netzwerk (Sydow, 2010, S. 396).

Die *Allokationsfunktion* beschäftigt sich mit der Frage, wie Ressourcen und Aufgaben sowie Zuständigkeiten auf die Netzwerkteilnehmer aufgeteilt werden sollen. Diese Verteilung basiert auf den verschiedenen Kompetenzen der Organisationen im Netzwerk und geschieht im Rahmen von Wettbewerbs- und Verhandlungsprozessen, die relativ gleichberechtigt ablaufen. Auch die Allokation ist eine kontinuierliche Aufgabe, da eine Umverteilung im Sinne der Re-Allokation jederzeit als Möglichkeit besteht (Sydow, 2010, S. 397).

Die *Regulationsfunktion* beschreibt die Aufgabe der Netzwerkmitglieder, Regeln für die Zusammenarbeit zu entwickeln und durchzusetzen. Dabei können sowohl formelle als auch informelle Verfahren festgelegt werden (Sydow, 2010, S. 397).

Bei der *Evaluationsfunktion* steht die Frage im Fokus, wie die Kosten und der Nutzen im Netzwerk ermittelt und auf die Teilnehmer verteilt werden sollen. Die Bewertung der Tätigkeiten im Netzwerk muss ebenfalls kontinuierlich, entweder formal oder implizit, vorgenommen werden (Sydow & Möllering, 2015, S. 196). Sie kann sowohl einzelne Mitglieder und Netzwerkbeziehungen als auch das ganze Netzwerk betreffen (Sydow, 2010, S. 398). Akteure innerhalb und außerhalb des Netzwerks können die Evaluation durchführen (Sydow & Möllering, 2015, S. 196).

Bei diesen vier Funktionen sind Managementpraktiken entscheidend, d.h. die kontinuierlichen, routinierten Handlungen des Netzwerkmanagements führen zur „tatsächliche[n] Ausgestaltung des Unternehmensnetzwerks“ (Sydow & Möllering, 2015, S. 197). Diese Praktiken bewirken die (Re-)Produktion der Strukturen des Netzwerks, gleichzeitig muss sich das Management dabei auch auf die bereits bestehenden beziehen (Sydow & Möllering, 2015, S. 197).

4. Methodik

Im Rahmen eines systematischen Literaturreviews wurde in der vorliegenden Arbeit der aktuelle Forschungsstand zu interorganisationalen Beziehungen von Social Enterprises erfasst. Um relevante Beiträge zu diesem Thema zu finden, wurde zuerst in den Datenbanken EBSCO Business Source Premier, Web of Science, JSTOR und Scopus (Elsevier) nach folgenden Schlüsselbegriffen gesucht: „social ent*“ AND alliance, „social ent*“ AND „network*” NOT „social network*”, „social ent*“ AND collaboration, „social ent*“ AND partnership, „social ent*“ AND cross-sector*, „social ent*“ AND interorgani?atio*, „social ent*“ AND inter-organi?atio*, „social ent*“ AND cooperation, „social ent*“ AND „joint venture”, „social ent*“ AND „inter-firm*”. Anschließend wurde „social ent*“ in den genannten Kombinationen jeweils durch die Begriffe „social venture” und „social business“ ersetzt, um den gesamten Bezeichnungskomplex von Social Enterprises im wissenschaftlichen Sprachgebrauch abzudecken, und derselbe Suchvorgang wiederholt.

Thematisch passende Beiträge mussten mindestens einen dieser Schlüsselbegriffe entweder im Titel, Abstract oder in den Keywords aufweisen. Weiterhin wurde der Abstract jeweils dahingehend geprüft, ob die Social Enterprise-Definition zum gewählten Kontext der Bachelorarbeit passt. So muss der marktliche Austausch neben der sozialen Mission des Unternehmens Beachtung finden und es darf sich demnach nicht um ein reines Non-Profit-Unternehmen handeln. Papiere, die soziale Tätigkeiten von Non-Governmental Organisationen (NGOs) betrachten, wurden ebenfalls aussortiert. Mit dem Einsatz der Begriffskombination „social ent*“ AND „network*” NOT „social network*” sollten Beiträge ausgeschlossen werden, die das soziale Netzwerk eines Social Entrepreneurs untersuchen, da hierbei nicht die interorganisationalen Beziehungen des Social Enterprises selbst im Fokus stehen. Bei der Recherche wurde nur nach Beiträgen gesucht, die bereits von Experten der Thematik geprüft (peer-reviewed journals) und in deutscher und englischer Sprache veröffentlicht wurden. Insgesamt sind bei dieser Suche 140 relevante Beiträge gefunden worden.

Im nächsten Schritt des Reviews wurden diese nach Beiträgen durchsucht, die in Predatory Journals[1] veröffentlicht wurden. Es wurden dabei jedoch keine gefunden. Danach wurden die Papiere im Rahmen eines Journal-Rankings untersucht und bewertet. Hierbei kamen der Association of Business Schools *Academic Journal Quality Guide* von 2018, veröffentlicht in der *Journal Quality List* (zusammengestellt und herausgegeben von Professor Anne-Will Harzing) von 2020 (Harzing, 2020) und der Association of Business Schools *Academic Journal Guide* von 2015 (The Association of Business Schools, 2015) zum Einsatz. Um den Beitrag zu werten, musste er in mindestens einem der genannten Rankings zu finden sein. Wenn dies nicht der Fall war, wurde das Alter des Journals ermittelt, indem das Entstehungsjahr des Journals bzw. die Anzahl der bereits veröffentlichten Volumes herausgesucht wurde. Beiträge aus relativ jungen Journals, die erst vor wenigen Jahren entstanden sind und deren Volume-Zahl dementsprechend gering ist, wurden in den weiteren Verlauf des Reviews aufgenommen. Beiträge älterer Journals wurden im Ranking mit der Ziffer 0 bewertet. Alle anderen Beiträge, die in den Rankings zu finden waren, wurden mit den Ziffern von 1 bis 4* gemäß der genutzten Quality Guides bewertet.

Alle Papiere, die eine bessere Bewertung als 2 aufwiesen, wurden in das Literaturreview aufgenommen. Dies waren insgesamt 40. Weiterhin wurden vier Buchkapitel aufgenommen, die zum Thema passten, sowie fünf Papiere, von denen zwei in einem relativ jungen Journal veröffentlicht wurden und drei die Bewertung 0 aufwiesen, aber inhaltlich zum Kontext passten. Im Anschluss wurden noch drei Konferenzpapiere aussortiert, da diese keine Journalartikel darstellen, und 12 Beiträge der Zeitschriften *Journal of Social Enterprise* (Bewertung: 0) und *Social Enterprise Journal* (Bewertung:

[1] Für die Identifikation dieser Beiträge wurde die List of Predatory Journals 2020 der folgenden Website genutzt: https://predatoryjournals.com/journals/ (Abrufdatum: 15.09.2020)

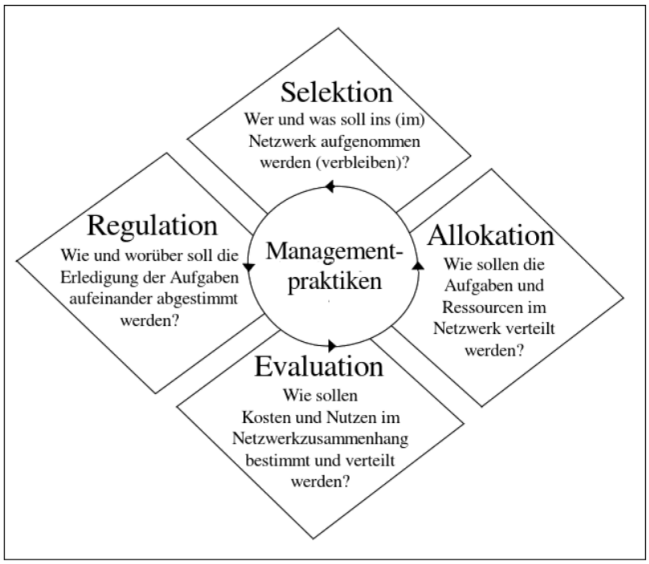

Abbildung 2: Funktionen des Netzwerkmanagements (Sydow & Windeler, 1997, S. 151)

1) in das Review aufgenommen. Dies sind themenspezifische Fachzeitschriften und passen daher besonders gut zum betrachteten Kontext. Somit wurden im Rahmen des Journal-Rankings insgesamt 61 Beiträge ausgewählt.

Unter Anwendung des Schneeballprinzips wurden im nächsten Schritt die Referenzen der 61 Papiere nach Beiträgen durchsucht, die ebenfalls für die betrachtete Social Entrepreneurship-Thematik relevant sind. Dabei wurden vier Beiträge gefunden. Außerdem wurden die Titel der Papiere jeweils in die Datenbank Web of Science eingegeben, um weitere thematisch passende Werke zu finden, die Zitationen der bereits gefundenen Beiträge des Journal-Rankings enthalten. Wenn ein Beitrag im Web of Science nicht auffindbar war, wurde derselbe Suchvorgang bei Google Scholar durchgeführt. Hierbei wurden drei Artikel gefunden. Das Schneeballprinzip hat somit insgesamt sieben zusätzliche Beiträge hervorgebracht und die Gesamtzahl der relevanten Werke auf 68 erhöht.

Diese wurden anschließend inhaltlich nach bestimmten Kriterien untersucht, um die Beiträge auszuschließen, die weniger genau zur Thematik der Abschlussarbeit passen. Zuerst wurde noch einmal geprüft, ob der in den Werken betrachtete Charakter der Social Enterprises mit der in dieser Arbeit genannten Definition übereinstimmt: das Unternehmen wird am Markt tätig und tritt aktiv als Anbieter von Produkten und/oder Dienstleistungen auf und verfolgt gleichzeitig eine soziale Mission. Weiterhin muss es in den Beiträgen um mindestens ein Social Enterprise gehen, das mit einem anderen Unternehmen bzw. einer anderen Organisation in Verbindung steht, oder es müssen Themen behandelt werden, die mit interorganisationalen Beziehungen von Social Enterprises zu tun haben. Das bedeutet, dass diese Beziehungen in den Papieren und Kapiteln möglichst im Fokus stehen sollen und keine Randthematik darin darstellen.

Außerdem wurde Literatur aussortiert, bei der Social Enterprises betrachtet wurden, die durch gemeinschaftliche Bemühungen mehrerer Individuen entstanden sind. Es kommt wirklich darauf an, dass das Social Enterprise Beziehungen unterhält und im Beitrag nicht das Ergebnis darstellt. Einige Werke haben das persönliche Netzwerk des Social Entrepreneurs bzw. des Gründers des Social Enterprises thematisiert und wurden, weil sie deshalb nicht zum betrachteten Kontext passten, ebenfalls vom weiteren Verlauf des Reviews ausgeschlossen. Abschließend wurden noch zwei Dissertationen aussortiert, da sich das Literaturreview ausschließlich auf Journalartikel und Buchkapitel konzentriert. Die Anzahl der relevanten Werke wurde durch die Prüfung der Literatur nach den genannten Kriterien auf 41 reduziert.

Im Anschluss folgte die inhaltliche Untersuchung der Papiere. Es wurden die Forschungsfrage, das Forschungsdesign (konzeptionell, empirisch-qualitativ, -quantitativ oder -mixed methods) und die geografische Herkunft jedes Werkes erfasst sowie die Art der interorganisationalen Beziehung, die jeweils thematisiert wird. Wenn es sich dabei um ein Netzwerk handelte, wurde dieses einer Kategorie der „Typologie interorganisationaler Netzwerke“ (Sydow, 2010, S. 382) zugeordnet. Weiterhin wurden vertikale, horizontale und laterale Beziehungen herausgearbeitet und thematisch unterschieden. Um diese Unterscheidung auch für Netzwerke anwenden zu können, musste eine fokale Unternehmung im Netzwerk festgelegt werden, deren Beziehungen zu anderen Organisationen genauer beleuchtet wird. Außerdem wurden alle Beiträge nach Informationen zu Funktionen des Netzwerkmanagements (Sydow & Windeler, 1997, S. 151) durchsucht und die wichtigsten Forschungsergebnisse jedes Werkes erfasst. Darauf basierend, welche Forschungsthematik im jeweiligen Papier im Vordergrund stand und welche Erkenntnisse gewonnen wurden, wurde jedes Werk in die folgende Systematisierung eingeordnet:

- Relevante Aspekte bei der Anbahnung interorganisa-

tionaler Beziehungen von Social Enterprises

- Relevante Aspekte zum Prozess der Zusammenarbeit bei interorganisationalen Beziehungen von Social Enterprises
- Ergebnisse und Folgen interorganisationaler Beziehungen von Social Enterprises

5. Ergebnisse

In Tabelle 1 ist der prozentuale Anteil der verschiedenen Forschungsdesigns dargestellt, die in den 41 analysierten Beiträgen angewendet werden.

Zu erkennen ist, dass empirisch-qualitative Methoden die Forschungsbeiträge in der formulierten Social Enterprise-Thematik dominieren. So beinhalten die meisten Werke die Untersuchung einer oder mehrerer interorganisationaler Beziehungen von Social Enterprises in der Praxis. Jedoch beschäftigt sich auch ein beachtlicher Teil der betrachteten Papiere mit theoretischen Konstrukten bezüglich dieser Beziehungsart. Die empirische Forschung, welche sowohl qualitative als auch quantitative Methoden einschließt (mixed methods), kommt hierbei nur relativ selten zur Anwendung. Kein einziges Papier beschäftigt sich mit empirisch-quantitativer Forschung.

Weiterhin wurde die geografische Herkunft der 41 analysierten Beiträge erhoben und in verschiedene Kategorien eingeteilt. Die prozentualen Ergebnisse finden sich in Tabelle 2.

Der größte Teil der betrachteten Werke stammt aus dem europäischen Raum, dominiert von Forschungsbeiträgen aus Großbritannien. Allerdings sind auch einige Papiere aus den USA und Kanada unter den 41 Papieren und bilden so eine eigene Kategorie. Die restliche Literatur besteht aus Beiträgen anderer Nationen und aus Werken, die durch interkontinentale Zusammenarbeit entstanden sind.

Bei der Analyse der Literatur konnte nicht jedes Papier einer ganz genauen Kategorie zugeordnet werden. So war es bei einigen Werken nicht möglich zu bestimmen, inwiefern die betrachteten Beziehungen im Sinne der Wertschöpfungskette miteinander verbunden sind. Diesbezüglich gab es den Fall, dass keine Informationen im jeweiligen Werk vorhanden waren (Kategorie „n/a“ in den nachfolgenden Tabellen) oder alle Verbindungstypen im Werk betrachtet wurden (Kategorie „alle“ in den nachfolgenden Tabellen). Außerdem war es bei zwei Papieren ebenfalls nicht möglich, den konkreten Netzwerktyp zu bestimmen.

5.1. Aspekte bei der Anbahnung der Kooperation

5.1.1. Dyaden

In dieser Kategorie wurden Informationen zu dyadischen interorganisationalen Beziehungen von Social Enterprises gefunden, die vertikale und laterale Verbindungen miteinander aufweisen. Ein Papier konnte bezüglich der Verbindung im Rahmen der Wertschöpfungskette nicht zugeordnet werden.

In der wissenschaftlichen Literatur wurde bezüglich dieser Beziehungsart betrachtet, welche Gründe es für Social Entrepreneure gibt, mit anderen Organisationen zusammenzuarbeiten. Dabei wird als Hauptgrund der Zugriff auf Ressourcen durch die Kollaboration genannt, wodurch das Social Enterprise Wettbewerbsvorteile erzielen kann (Schirmer & Cameron, 2012, S. 84f.). Daneben ist laut Huybrechts und Nicholls (2013, S. 142) auch organisatorische Legitimität ein wichtiger Faktor, welcher die Entstehung von sektorübergreifenden Beziehungen zwischen Organisationen im Social Entrepreneurship erklärt.

Einen weiteren Aspekt in der Literatur stellt die wichtige Rolle der Überprüfung der potenziellen Partner durch das Social Enterprise dar (Huybrechts & Nicholls, 2013, S. 139). Dieser Prozess bildet im Sinne der Selektionsfunktion die entscheidende Grundlage für die Entstehung interorganisationaler Beziehungen. Auch (Schirmer & Cameron, 2012, S. 96) weisen der Vorbereitungs- und Planungsphase von interorganisationalen Beziehungen im Social Entrepreneurship eine besondere Bedeutung zu. Hierbei sind vor allem die Definition eines strategischen Ziels sowie die Festlegung von Erwartungen und Verantwortlichkeiten für die künftige Zusammenarbeit wichtig. Bei der systematischen Suche nach Kooperationspartnern sind klar definierte Kriterien hilfreich.

Eine soziale Mission, die von beiden Partnern geteilt wird, sowie „specific organizational motivations“ (Henry, 2015, S. 153) sind wichtige frühe Einflüsse, die bestimmen, warum und wie Möglichkeiten im sozialen Non-Profit-Sektor realisiert werden (Henry, 2015, S. 153). Nicholls und Huybrechts (2016, S. 709) bringen an, dass für die Entstehung von Beziehungen zwischen Organisationen, die verschiedenen Logiken und Machtstrukturen folgen, eine vorherige „‘hybridization’ [sic]” (Nicholls & Huybrechts, 2016, S. 709) der Logiken notwendig ist. Dies ist z.B. bei der Zusammenarbeit von Social Enterprises mit herkömmlichen Unternehmen der Fall, deren hauptsächliches Ziel die Erwirtschaftung von finanziellen Überschüssen ist.

5.1.2. Netzwerke

Hier wurden Informationen zu Aspekten bei der Anbahnung von interorganisationalen Beziehungen von Social Enterprises gefunden, die Teil von strategischen Netzwerken, einem Projektnetzwerk und einem Social Franchise sind.

Bei den strategischen Netzwerken stellen Social Enterprises die fokale Unternehmung dar, welche die Netzwerkführerschaft innehat. Wie bereits bei den dyadischen Beziehungen wurde hier die Hybridisierung der unterschiedlichen Organisationslogiken durch Nicholls und Huybrechts (2016, S. 709) als notwendige Maßnahme vor der Zusammenarbeit im Netzwerk hervorgehoben. Huybrechts et al. (2017) haben diesbezüglich verschiedene Kooperationsarten von Social Enterprises mit „mainstream businesses“ (Huybrechts et al., 2017, S. 595) im Fair Trade-Sektor untersucht. Dabei wurden die Partner vor dem Eingehen der Beziehung nach bestimmten festgelegten Kriterien gemäß der jeweiligen Kooperationsart untersucht und entsprechend ausgewählt.

Tabelle 1: Forschungsdesigns der Beiträge

Anzahl der Papiere	Konzeptionell	Empirisch-quantitativ	Empirisch-qualitativ	Empirisch-mixed-methods
41	26,8%	0%	56,1%	17,1%

Tabelle 2: Herkunft der Beiträge

Anzahl der Papiere	Europa	Nordamerika	Rest der Welt	Europa & Nordamerika	Europa & Rest der Welt
41	51,2%	17,1%	22,0%	4,9%	4,9%

Bezüglich der Projektnetzwerke von Social Enterprises, die auf eine vorübergehende zeitliche Dauer angelegt sind, haben Gillett, Loader, Doherty und Scott (2019, S. 965) herausgearbeitet, dass gleiche Werte und ethische Ansichten der kooperierenden Organisationen eine wichtige Basis für die Überwindung von Hindernissen bei der Zusammenarbeit darstellen.

Zafeiropoulou und Koufopoulos (2013) haben sich mit dem Phänomen des Social Franchising beschäftigt. Dieses Papier wurde der Netzwerk-Kategorie, aber keinem speziellen Netzwerk-Typ zugeordnet. Ein Social Franchise beschreibt einen interorganisationalen Zusammenschluss, der ein Social Enter-prise als Franchisegeber beinhaltet. Dieses Unternehmen hat bereits ein standardisiertes soziales Konzept auf den Markt gebracht und tritt als Anbieter von Produkten und Dienstleistungen auf. Es steht in fortdauernder vertraglicher Verbindung mit unabhängig betriebenen Social Enterprises, den Franchisenehmern. Sie operieren unter dem Handelsnamen des Franchisegebers, um dessen soziales Konzept anzubieten und spezielle soziale Ergebnisse zu erreichen. Dies wird durch die Produktion und/oder Vermarktung von Produkten und Dienstleistungen durch die Franchisenehmer erreicht, wobei das tatsächlich genutzte Format vom Franchisegeber festgelegt wird (Zafeiropoulou & Koufopoulos, 2013, S. 76). In dem betrachteten Werk wurde dieser Prozess näher untersucht und herausgefunden, dass die Einbettung des Social Enterprises (Franchisegeber) in ein soziales Netzwerk in positiver Verbindung mit der Entstehung von Social Franchises und der Selektion angemessener Partner steht (Zafeiropoulou & Koufopoulos, 2013, S. 87).

In Tabelle 3 sind die vorgestellten Themen dieses Forschungsbereichs dargestellt und bezüglich der Zahl der Partnerorganisationen in dyadische und netzwerkförmige interorganisationale Beziehungen von Social Enterprises aufgeteilt. In der Dyaden-Kategorie steht in der Klammer hinter dem jeweiligen Thema die Verbindungsart der Beziehung im Sinne der Wertschöpfungskette. Die Beschriftung „n/a“ zeigt auf, dass diesbezüglich keine Informationen im jeweiligen Werk vorhanden waren. Die Bezeichnung „alle“ beschreibt, dass alle Verbindungstypen im Werk betrachtet wurden. In der Netzwerke-Kategorie steht in der Klammer zusätzlich der jeweilige Netzwerktyp gemäß der „Typologie interorganisationaler Netzwerke“ (Sydow, 2010, S. 382). Hierbei werden strategische Netzwerke, regionale Netzwerke und Projektnetzwerke mit den Abkürzungen „SN“, „RN“ und „PN“ bezeichnet. Dieser Systematik folgen alle Tabellen in diesem Kapitel.

5.2. Aspekte zum Prozess der Kooperation

5.2.1. Dyaden

In dieser Kategorie beschäftigen sich die gefundenen Werke mit interorganisationalen Beziehungen im Social Entrepreneurship, die im Sinne der Wertschöpfungskette alle Arten von Verbindungen aufweisen. Ein Papier beleuchtet die Zusammenarbeit von Social Enterprises mit Wettbewerbern, demnach horizontale Kooperationsbeziehungen, drei Werke betrachten vertikale Beziehungen und drei weitere Papiere laterale Beziehungen. Die überwiegende Zahl der Forschungsarbeiten in diesem Bereich war allerdings keiner bestimmten Bindungsart zuzuordnen, da entweder keine spezifischen Informationen dafür vorlagen oder alle Bindungsarten im jeweiligen Papier Beachtung fanden. Da diese Kategorie viel Literatur hergibt, werden im Folgenden die Forschungsthemen vorgestellt, die in diesem Feld besonders im Fokus standen.

Eine zentrale Thematik bezüglich des Kooperationsprozesses dyadischer Organisationsbeziehungen im Social Entrepreneurship ist der Erwerb von Ressourcen. So stellen Montgomery et al. (2012) verschiedene Strategien zum Beschaffen und zum Handeln von Ressourcen vor und betrachten dazu einige Tätigkeiten von Social Enterprises in der Praxis. Sie unterteilen ihre Ergebnisse in sektorübergreifende Vorgänge und Prozesse, die im selben Sektor ablaufen. Choi (2015) betrachtet, welche Arten von Ressourcen Social Enterprises in Korea durch verschiedene Partner zur Verfügung gestellt werden. Die Ergebnisse zeigten, dass Partner des öffentlichen Sektors Unterstützung im Bereich Finanzen und Marketing anbieten, jedoch nicht bezüglich des Managements der Social Enterprises. Partner aus dem privaten Sektor bieten dagegen nur finanzielle Unterstützung an (Choi, 2015, S. 270) und Partner des sozialen Sektors stellen finanzielle Ressourcen zur Verfügung, jedoch keine Unterstützung

Tabelle 3: Themen zu Aspekten bei der Anbahnung der Kooperation

Dyaden	Netzwerke
• Gründe für Partnerschaften (n/a) (Schirmer & Cameron, 2012) • Vorbereitung der Partnerschaft (n/a) (Schirmer & Cameron, 2012) • Legitimität als Ursache für Entstehung und Entwicklung von sektorübergreifenden Partnerschaften (vertikal) (Huybrechts & Nicholls, 2013) • geteilte soziale Mission und spezifische organisationale Motivationen ermöglichen Partnerschaft (lateral) (Henry, 2015) • Hybridisierung der unterschiedlichen Unternehmenslogiken (vertikal) (Nicholls & Huybrechts, 2016) • Screening von potenziellen Partnern sehr wichtig (vertikal) (Huybrechts & Nicholls, 2013)	• Hybridisierung der unterschiedlichen Unternehmenslogiken (SN/vertikal) (Nicholls & Huybrechts, 2016) • gleiche Werte und Ethikansicht als wichtige Basis zur Überwindung von Hindernissen bei der Zusammenarbeit (PN/lateral) (Gillett et al., 2019) • Unterschiedliche Strategien zur Auswahl der Kooperationspartner im Fair Trade-Sektor (SN/vertikal/lateral) (Huybrechts et al., 2017) • Bildung von Social Franchises, Selektion angemessener Partner durch Beteiligung von Social Enterprises in sozialen Netzwerken (Social Franchise) (Zafeiropoulou & Koufopoulos, 2013)

des Managements. Weiterhin erwarben diese die Produkte und Dienstleistungen der Social Enterprises zur Weiterveräußerung (Choi, 2015, S. 271). Weidner, Weber und Göbel (2019) fanden im Rahmen ihrer Forschung heraus, dass Social Enterprises durch ihre Beteiligung an strategischen Partnerschaften und die Entwicklung von partnerschaftlicher Legitimität sowohl Ressourcenbeschränkungen als auch Legitimitätsprobleme abmildern können (Weidner et al., 2019, S. 517).

Mehrere Werke in diesem Forschungsbereich haben sich mit interorganisationalen Beziehungen zwischen Social Enterprises und herkömmlichen For-Profit-Unternehmen befasst. Di Domenico et al. (2009) heben vier Spannungsquellen dieser Beziehungsart hervor: Ziele und Logik, Eigentümerschaft, Unternehmensführung und Unternehmenshaftung (Di Domenico et al., 2009, S. 897ff.). Weiterhin wird der Verlauf der Zusammenarbeit modellhaft dargestellt und in drei Phasen gegliedert: Thesis, Antithesis und Synthesis (Di Domenico et al., 2009, S. 901). Um zum Synthese-Zustand zu gelangen, müssen beide Partner bei ihrer Arbeit Werte verfolgen, die sich wesentlich von ihren ursprünglichen Werten außerhalb der Beziehung unterscheiden (Di Domenico et al., 2009, S. 900). Jug (2020) betrachtete Gründe von For-Profit-Unternehmen für die Zusammenarbeit mit Social Enterprises und nennt dafür u.a. das Erreichen von Zielen im Corporate Social Responsibility-Bereich. Weiterhin werden aber auch Hindernisse für die Zusammenarbeit benannt, wie z.B. der Vorrang der sozialen Mission des Social Enterprises. So kann es für das For-Profit-Unternehmen schwierig sein, das Social Enterprise als vollwertigen Business-Partner zu akzeptieren.

Huybrechts und Nicholls (2013) untersuchten den partnerschaftlichen Zusammenschluss zwischen einem Social Enterprise und einem großen Einzelhändler im Fair Trade-Sektor und stellten fest, dass der Einzelhändler die dominante Position im Bündnis einnahm. Durch seine Überlegenheit konnte er seine Bedingungen bei der Zusammenarbeit erfolgreich durchsetzen und war derjenige, der sich für das Social Enterprise als Partner entschied und nicht umgekehrt (Huybrechts & Nicholls, 2013, S. 142). Lee und Jay (2015) befassten sich mit der Reaktion von For-Profit-Unternehmen bei Markteintritt von Social Enter-prises. Davon stellt die Übernahme des Social Enterprises durch das For-Profit-Unternehmen eine mögliche Reaktion dar, wobei dies als intraorganisationale Beziehung zu bezeichnen ist. In den Bereich der interorganisationalen Beziehungen gehört das Konsultieren der Social Enterprises als mögliche Reaktion der For-Profit-Unternehmen (Lee & Jay, 2015, S. 131).

Neben dem Fokus auf der Zusammenarbeit mit For-Profit-Unternehmen werden interorganisationale Partner von Social Enterprises in der wissenschaftlichen Literatur häufig in bestimmte Kategorien unterteilt. So gliedern Kwong, Tasavori und Wun-mei Cheung (2017) die Beziehungen nach folgenden Partner-Typen: dominante, kooperierende, inaktive und komplementäre Partner (Kwong et al., 2017, S. 616). Dabei wurde untersucht, inwiefern das Bündnis mit dem jeweili-

gen Partner-Typ die Bricolage-Strategie der Social Enterprises beeinflusst. Bricolage wurde erstmals durch Lévi Strauss mit "making do with whatever is at hand" (Lévi-Strauss, 1967, S. 17) beschrieben. So stehen nur die vorhandenen Mittel als Ressourcen für das eigene Vorhaben zur Verfügung – nicht mehr, nicht weniger. Die Ergebnisse der Studie von Kwong et al. (2017) werden im Kapitel 5.3 geschildert. Dagegen unterteilen Volkmann, Tokarski und Ernst (2012, S. 86) die verschiedenen Partnerorganisationen von Social Enterprises nach Sektoren und schildern die jeweiligen Details und Besonderheiten bei Zusammenschlüssen von Social Enterprises mit Organisationen des öffentlichen, privaten und zivilen Sektors (Volkmann et al., 2012, S. 87ff.).

Ein weiterer Aspekt, der bei der Betrachtung von dyadischen interorganisationalen Beziehungen von Social Enterprises und deren Kooperationsprozess Erwähnung findet, ist das zukünftige Forschungspotenzial in diesem Bereich. de Bruin, Shaw und Lewis (2017) haben folgende Themen bezüglich der Beziehungen herausgearbeitet, die noch genauer untersucht werden müssen: Analyse der Zusammenschlüsse auf verschiedenen Ebenen (lokal, institutionell, etc.), Koordination, Leitung, Lernen und Wissensvermittlung, Kontext (geografisch, historisch, etc.), Identität und konkurrierende Logiken innerhalb des Bündnisses (de Bruin et al., 2017, S. 578ff.). Auch Phillips, Lee, Ghobadian, O'Regan und James (2015) haben sich mit dem aktuellen Wissensstand im Bereich Social Entrepreneurship beschäftigt und heben dabei sektorübergreifende Partnerschaften im Social Entrepreneurship als ein wichtiges Forschungsfeld hervor (Phillips et al., 2015, S. 444ff.).

Im Rahmen dieses Literaturreviews wurden noch einige weitere interessante Themen zum Kooperationsprozess der dyadischen interorganisationalen Beziehungen von Social Enterprises herausgearbeitet, die allerdings jeweils nur von einem einzelnen Werk betrachtet wurden. Diese sind in Tabelle 4 zu finden.

5.2.2. Netzwerke

In Bezug auf den Kooperationsprozess interorganisationaler Beziehungen von Social Enterprises im Netzwerk wurden Papiere gefunden, die sich mit strategischen, regionalen und Projektnetzwerken auseinandersetzen. Ein Werk konnte keinem konkreten Netzwerktyp zugeordnet werden und eine weitere Forschungsarbeit in diesem Bereich beschäftigt sich mit Aspekten von Social Franchising, einer besonderen Form der Zusammenarbeit im Netzwerk, die auch nicht in die genannte Typologie eingeordnet werden kann. Dieses Phänomen wurde bereits in Kapitel 5.1 näher beschrieben. Die Netzwerkkategorie gibt bezüglich des Kooperationsprozesses der interorganisationalen Beziehungen von Social Enterprises relativ viel Literatur her. Deshalb werden, wie schon bei der Dyaden-Kategorie in diesem Kapitel, im Folgenden die am häufigsten betrachteten Themen in diesem Forschungsfeld dargestellt.

Wie bei den dyadischen Beziehungen werden auch bei den netzwerkförmigen interorganisationalen Beziehungen von Social Enterprises verschiedene Kollaborations- und Partnertypen in der wissenschaftlichen Literatur herausgearbeitet. So findet bei strategischen Netzwerken von Social Enterprises ebenfalls die Aufteilung der Partnerschaften nach verschiedenen Sektoren gemäß Volkmann et al. (2012) Anwendung. Hierbei werden Partner von Social Enterprises aus dem öffentlichen, privaten und zivilen Sektor voneinander unterschieden (Volkmann et al., 2012, S. 87ff.). Lyon (2012) betrachtet strategische Netzwerke im Social Entrepreneurship dagegen im Sinne der Wertschöpfungsverbindungen. Er bezeichnet Beziehungen von Social Enterprises mit Käufern, Geldgebern und Subunternehmen als vertikal und grenzt diese von horizontalen Beziehungen ab, die Social Enterprises mit anderen Dienstleistern eingehen können (Lyon, 2012, S. 14).

Eine weitere zentrale Thematik bei der Kooperation von Social Enterprises im strategischen Netzwerk ist der Einfluss von und der Umgang mit Stakeholdern. Dies sind die verschiedenen Interessengruppen des Unternehmens. Im Rahmen dieser Arbeit stehen dabei vor allem Organisationen als Stakeholder im Fokus, wie z.B. Gewerkschaften und Regierungsorganisationen. Phillips, Alexander und Lee (2019) haben in diesem Bereich eine „Social Innovation-Stakeholder Relation-ship Matrix" (Phillips et al., 2019, S. 328) entwickelt, welche vier Typen von Social Enterprises bezüglich der Ziele ihrer Stakeholder-Beziehungen unterscheidet. So sind diese Typen danach eingeteilt, inwiefern das Social Enterprise mithilfe der Beziehungen zu seinen Stakeholdern neue Möglichkeiten für soziale Innovation sucht und diese implementiert (Phillips et al., 2019, S. 328).

Des Weiteren betrachten Ramus, La Cara, Vaccaro und Brusoni (2018), ob externe Stakeholder Einfluss auf die Innovationsstrategie von Social Enterprises ausüben. Sie fanden heraus, dass diese vor allem in unruhigen Situationen die genannte Strategie des Social Enterprises signifikant beeinflussen (Ramus et al., 2018, S. 465). Außerdem zeigten ihre Ergebnisse, dass Social Enterprises eine Innovationsstrategie entwickeln, die mehr zur sozialen als zur kommerziellen Unternehmensperformance beiträgt, wenn soziale externe Stakeholder mehr Druck ausüben als kommerzielle externe Stakeholder (Ramus et al., 2018, S. 479). Soziale externe Stakeholder sind beispielsweise Non-Profit-Partnerorganisationen und Sozialeinrichtungen (Ramus et al., 2018, S. 470) und kommerzielle externe Stakeholder For-Profit-Partnerunternehmen und investierende Organisationen (Ramus et al., 2018, S. 471).

Die Eigenschaften guter Netzwerkbeziehungen sind ein weiteres Thema, das in der wissenschaftlichen Literatur des Social Entrepreneurships im Fokus steht. Im Strategischen Netzwerk haben Nicholls und Huybrechts (2016) u.a. folgende Eigenschaften herausgearbeitet, die Voraussetzung für langfristige interorganisationale Beziehungen von Social Enterprises sind: Konfliktttoleranz, Akzeptanz der Machtasymmetrie und Beteiligung aller Netzwerkteilnehmer bei der Entwicklung von Regeln und Praktiken (Nicholls & Huybrechts, 2016, S. 709). In Bezug auf eine erfolgreiche Zusammenarbeit im Projektnetzwerk, das Social Enterprises als Teilnehmer einschließt, stellen Gillett et al. (2019) folgende Eigen-

schaften als Voraussetzungen fest: zentrale Steuerung des Netzwerks und Entstehung von Synergien durch das Teilen von Ressourcen (Gillett et al., 2019, S. 965).

Die Beschaffung von Ressourcen steht, wie schon bei den dyadischen interorganisationalen Beziehungen, auch beim Kooperationsprozess der Netzwerkbeziehungen von Social Enterprises im Vordergrund. Hier haben sich Meyskens, Carsrud und Cardozo (2010) mit der Erlangung von Ressourcen im strategischen Netzwerk beschäftigt, in dem Social Enterprises mit herkömmlichen Unternehmen, Regierungsorganisationen und anderen Social Enterprises zusammenarbeiten. Durch die Kooperation beschaffen die Social Enterprises laut Meyskens et al. (2010, S. 446) u.a. finanzielle Ressourcen, wie z.B. Fördergelder, und humane Ressourcen im Rahmen des Wissensaustausches. Weiterhin heben Jayawarna, Jones und Macpherson (2020) hervor, dass das Social Enterprise als fokale Unternehmung im strategischen Netzwerk die Beziehungen effektiv managen und den Ressourcenaustausch fördern muss, damit es Bootstrap-Ressourcen erlangt (Jayawarna et al., 2020, S. 63). Hiermit sind Ressourcen gemeint, die Social Enterprises nutzen, um schwierige Situationen zu überstehen und das Überleben der Unternehmung zu sichern.

Eine besondere Aufmerksamkeit im Bereich des Social Entrepreneurships erfährt der Fair Trade-Sektor (Huybrechts & Nicholls, 2013; Nicholls & Huybrechts, 2016). Im Rahmen ihrer Forschungsarbeit haben Huybrechts et al. (2017) verschiedene Kooperationsstrategien von Social Enterprises mit herkömmlichen Unternehmen im strategischen Netzwerk im Fair Trade-Sektor untersucht. Sie haben dabei drei verschiedene Strategien herausgearbeitet, welche bereits in Kapitel 5.1 erwähnt wurden. Im Sinne der ersten Strategie, bezeichnet als „Sector solidarity“ (Huybrechts et al., 2017, S. 595), sollte der Fair Trade-Sektor als Nischenmarkt erhalten bleiben und so arbeiteten nur spezialisierte Importeure und Einzelhändler des Fair Trades zusammen (Huybrechts et al., 2017, S. 595f.). Bei der zweiten Strategie, „Selective engagement“ (Huybrechts et al., 2017, S. 597), wird das Vorgehen der ersten Strategie beibehalten, gleichzeitig aber auch der Massenmarkt durch ausgewählte Kollaborationen mit For-Profit-Unternehmen bedient (Huybrechts et al., 2017, S. 597). Die dritte Strategie, „Active appropriation“ (Huybrechts et al., 2017, S. 599), beschreibt das Vorgehen von Organisationen, die das Fair-Trade-Zertifikat bzw. das Logo auf dem Massenmarkt vertreiben, ohne die herkömmlichen Unternehmen gezielt auszuwählen. So sollte das Verkaufsvolumen im Fair Trade-Sektor erhöht werden (Huybrechts et al., 2017, S. 599). Informationen zu den Ergebnissen der einzelnen Strategien werden in Kapitel 5.3 in der Netzwerke-Kategorie erläutert.

Risiken und hindernde Faktoren der interorganisationalen Kooperation von Social Enterprises im Netzwerk stellen die letzte Thematik dar, die durch ihre häufige Erwähnung in der Literatur hier hervorgehoben werden soll. Schirmer und Cameron (2012) stellen fest, dass die Zusammenarbeit im strategischen Netzwerk erschwert wird, wenn die Partner ihre Erwartungen an die Kooperation nicht ausdrücklich kommunizieren. Weiterhin stellt es für die beteiligten Organisationen eine Herausforderung dar, vor dem Eingehen der Beziehung festzustellen, ob die Partner tatsächlich genug Gemeinsamkeiten für eine erfolgreiche Kooperation aufweisen. Als Risiko der Zusammenarbeit wird die Schädigung des Rufs der Organisationen genannt. Diese kann auftreten, wenn sich einer der Beteiligten im Netzwerk falsch verhält oder die Kooperation gänzlich scheitert. Bei der Beteiligung von Social Enterprises und Organisationen des privaten Sektors im Netzwerk ist dieses Risiko erhöht (Schirmer & Cameron, 2012, S. 94). Maase und Bossink (2010) geben im Rahmen ihrer Forschungsarbeit an, dass Interessenkonflikte und das unterschiedliche Risikoverhalten der beteiligten Organisationen im strategischen Netzwerk die Zusammenarbeit behindern. Das Social Enterprise handelt gelegenheitsorientiert und sucht im Rahmen der Kooperation nach neuen Möglichkeiten, sein Geschäft zu erweitern. Andere Organisationen sind jedoch meist nicht so risikoaffin, was zu Auseinandersetzungen im Netzwerk führt (Maase & Bossink, 2010, S. 68).

Bezüglich des Kooperationsprozesses interorganisationaler Beziehungen von Social Enterprises im Netzwerk wurden im Rahmen dieses Literaturreviews noch weitere Themen herausgearbeitet. Jedoch werden diese jeweils nur von einem einzelnen Werk betrachtet und stellen demnach keine zentrale Thematik in diesem Bereich dar. Sie werden hier nicht näher erläutert, sind allerdings in der Tabelle 4 aufgeführt. Der überwiegende Teil der hier vorgestellten Aspekte zum Kooperationsprozess betrifft strategische Netzwerke, da meist das Social Enterprise im jeweiligen Papier die fokale Unternehmung darstellt. Die relevanten Aspekte zu den anderen Netzwerktypen sind in Tabelle 4 nachzulesen. Sie stellt eine Übersicht zu den beschriebenen Themen in diesem Forschungsbereich dar.

5.3. Ergebnisse und Folgen der Kooperation

5.3.1. Dyaden

Hier wurden Papiere gefunden, welche interorganisationale Beziehungen von Social Enterprises mit vertikalen und lateralen Verbindungen untersuchen. Zwei Werke konnten bezüglich der Verbindung im Rahmen der Wertschöpfungskette nicht zugeordnet werden.

Kwong et al. (2017) haben die Auswirkungen der Beteiligung verschiedener Partner-Typen auf die Bricolage-Strategie von Social Enterprises untersucht. Sie fanden heraus, dass die Beteiligung inaktiver und komplementärer Partner dazu führt, dass Social Enterprises ihre Unabhängigkeit von der Partnerschaft beibehalten können und ihrer benannten Zielgruppe treu bleiben. Die Beteiligung dominanter und kollaborativer Partner führt hingegen dazu, dass Social Enterprises ihre Bricolage-Prozesse nicht mehr selbstständig steuern können und letztendlich von ihrer sozialen Mission abweichen (Kwong et al., 2017, S. 631).

Choi (2015) hat ebenfalls den Einfluss verschiedener Partner-Typen auf Social Enterprises untersucht, jedoch im Besonderen deren Einfluss auf die soziale Unternehmensperformance. Partner des öffentlichen und sozialen Sektors stärkten die soziale Performance von Social Enterprises,

Tabelle 4: Themen zu Aspekten des Kooperationsprozesses

Dyaden	Netzwerke
• **Ressourcen:** ⇒ Strategien im Social Entrepreneurship, um Ressourcen zu bekommen und zu handeln (vertikal) (Montgomery et al., 2012) ⇒ Bereitstellung verschiedener Ressourcen durch verschiedene Partner-Typen (n/a) (Choi, 2015) ⇒ Abmilderung von Ressourcen- und Legitimitätsbeschränkungen von Social Enterprises durch die Schaffung von strategischen Partnerschaften und Legitimität zwischen Partnern (alle) (Weidner et al., 2019) • **Beziehungen zwischen Social Enter- prises und For-Profit-Unternehmen:** ⇒ Anreiz, Spannungen, Modell des Verlaufs (lateral) (Di Domenico et al., 2009) ⇒ Gründe für For-Profit-Unternehmen, mit Social Enterprises zusammenzuarbeiten (n/a) (Jug, 2020) ⇒ Dominante Position des For-Profit-Unternehmens in der Beziehung (vertikal) (Huybrechts & Nicholls, 2013) ⇒ Reaktionen von For-Profit-Unternehmen auf Markteintritt von Social Enterprises => Zusammenarbeit als eine Möglichkeit (horizontal) (Lee & Jay, 2015) • **Partner-Kategorien von Social Enter-prises:** ⇒ verschiedene Arten von Social Enterprise-Partnern (lateral) (Kwong et al., 2017) ⇒ verschiedene Partner-Typen (n/a) (Schirmer & Cameron, 2012) • **Zukünftige Forschung:** ⇒ Literaturanalyse (n/a) (Phillips et al., 2015) ⇒ Zukünftiges Forschungspotenzial für Zusammenarbeit von Social Enterprises mit anderen Organisationen (n/a) (de Bruin et al., 2017)	• **Ressourcen:** ⇒ Beschaffung von Ressourcen durch interorganisationale Beziehungen des Social Enterprises (SN/n/a) (Meyskens et al., 2010) ⇒ Bootstrapping von Ressourcen durch interorganisationale Beziehungen des Social Enterprises (SN/alle) (Jayawarna et al., 2020) • **Partner-Kategorien von Social Enterprises:** ⇒ Unterschiedliche Typen interorganisationaler Beziehungen von Social Enterprises (SN/alle) (Lyon, 2012) ⇒ Verschiedene Partner-Typen von Social Enterprises (SN/n/a) (Schirmer & Cameron, 2012) • **Interorganisationale Stakeholder-Beziehungen:** ⇒ Verschiedene Social Enterprise-Arten bezüglich der Ziele ihrer Stakeholder-Beziehungen (SN/n/a) (Phillips et al., 2019) ⇒ Beeinflussung der Innovationsstrategie von Social Enterprises durch externe Stakeholder (SN/alle) (Ramus et al., 2018) • **Eigenschaften fortdauernder interorganisationaler Netzwerk-Beziehungen:** ⇒ Eigenschaften von guten, andauernden Beziehungen eines Social Enterprises im Netzwerk (SN/vertikal) (Nicholls & Huybrechts, 2016) ⇒ Eigenschaften einer effektiven Zusammenarbeit im Netzwerk (PN/lateral) (Gillett et al., 2019) • **Fair Trade:** ⇒ Kooperationen von Social Enter-prises untereinander und Social Enterprises und For-Profit-Unternehmen sind im Fair Trade entstanden (SN/lateral) (Huybrechts et al., 2017) ⇒ Business-Netzwerke von Importeuren und Einzelhändlern arbeiten zusammen (SN/vertikal) (Huybrechts et al., 2017) • **Hindernisse bei der Zusammenarbeit:** ⇒ Risiken/Herausforderungen (SN/n/a) (Schirmer & Cameron, 2012) ⇒ Hindernde Faktoren für Social Enterprise-Partnerschaften (SN/n/a) (Maase & Bossink, 2010)

Tabelle 4—continued

- **Weitere Themen:**
  - ⇒ Risiko/Herausforderungen der Zusammenarbeit (n/a) (Schirmer & Cameron, 2012)
  - ⇒ Entwicklungs- und Erneuerungsphasen der Zusammenarbeit (n/a) (Schirmer & Cameron, 2012)
  - ⇒ Vorteile und Notwendigkeit der Social Enterprise-Partnerschaft (lateral) (Henry, 2015)
  - ⇒ Relationelle Identifikation von Organisationen im Verhältnis zu Beständigkeit in Strategischen Allianzen von Social Enterprises (n/a) (Smith, Meyskens & Wilson, 2014)
  - ⇒ Vergrößerung von Social Enterprises u.a. durch die Schaffung von Allianzen möglich (n/a) (Bloom & Chatterji, 2009)
  - ⇒ Sektorübergreifende Zusammenarbeit führt zu mehr Social Entrepreneurship (n/a) (Sinthupundaja, Kohda & Chiadamrong, 2020)
  - ⇒ Interorganisationale Beziehungen gehören zu Ergebnis beeinflussenden Faktoren von Social Enterprises (alle) (Moura, Comini & Teodósio, 2015)
  - ⇒ Eigenschaften einer guten, andauernden interorganisationalen Beziehung eines Social Enterprises (vertikal) (Nicholls & Huybrechts, 2016)

- **Weitere Themen:**
  - ⇒ Entwicklungs- und Erneuerungsphase der Zusammenarbeit (SN/n/a) (Schirmer & Cameron, 2012)
  - ⇒ Managementtechniken (SN/alle) (Clark, Woods & Adams, 2017; Lyon, 2012)
  - ⇒ Vorteile der interorganisationalen Zusammenarbeit im Netzwerk (SN/n/a) (Meyskens et al., 2010)
  - ⇒ Interorganisationale Beziehungen führen u.a. zu Schaffung einer Organisationskultur, die für Social Enterprises günstig ist; einige Gründe für Partnerschaften (SN/n/a) (Yaari, Blit-Cohen & Savaya, 2019)
  - ⇒ Interorganisationale Beziehungen im Netzwerk gehören zu Ergebnis beeinflussenden Faktoren von Social Enterprises (SN/alle) (Moura et al., 2015)
  - ⇒ Netzwerkbeziehungen eines sozialen Restaurants (SN/vertikal/lateral) (Sigala, 2019)
  - ⇒ Netzwerk-Bricolage von Social Enterprises (SN/n/a) (Tasavori, Kwong & Pruthi, 2018)
  - ⇒ Netzwerkerrichtung als wichtige Eigenschaft des Social Entrepreneurs (SN/n/a) (Sigala, 2016)
  - ⇒ Unterschiedlicher Netzwerkaufbau von kommerziellen Unternehmen und Social Enterprises (SN/alle) (Folmer, Nederveen & Schutjens, 2018)
  - ⇒ Erklärung von Social Franchising, positive Eigenschaften des Social Franchisor-Systems (Social Franchise) (Zafeiropoulou & Koufopoulos, 2013)
  - ⇒ Kooperation eines Social Business Orchestrators mit Social Enterprises (SN/lateral) (Gold, Chowdhury, Huq & Heinemann, 2020)
  - ⇒ Was ist ein Social Impact Bond (SIB)? Gute Bedingungen und Wirkungen der Zusammenarbeit im Netzwerk, Antreiber der Kollaboration (RN/lateral) (Smeets, 2017)
  - ⇒ Vorteile vom Tripartite Collaborative Model im Social Entrepreneurship (RN/lateral) (Lam, 2016)
  - ⇒ Beziehung zwischen Diversität der Partnerschaften und Erfolg von nascent green-tech ventures (SN/alle) (Meyskens & Carsrud, 2013)
  - ⇒ Vertrauen in Beziehungen mit dem öffentlichen Sektor (RN/lateral) (Seanor, 2018)
  - ⇒ Teamselektion und Team-/Job-zufriedenheit in Cross-sector social partnership (CSSP) im Verhältnis zueinander (RN/n/a) (Intindola, Pittz, Rogers & Weisinger, 2019)
  - ⇒ Nischen-Marketing und Coopetiton bei Schaffung von sozialem Wert (SN/horizontal/vertikal) (Herbst, 2019)
  - ⇒ Literaturanalyse (n/a) (Littlewood & Khan, 2018)

Partner des privaten Sektors schwächten diese. Alle Partner schwächten die soziale Performance der Social Enterprises, wenn sie finanzielle Unterstützung anboten, und trugen nicht zur sozialen Performance bei, wenn sie Unterstützung im Marketing bereitstellten (Choi, 2015, S. 271).

Bei der langfristigen Untersuchung einer Kooperation zwischen einem Social Enterprise und einem Einzelhändler im Fair Trade-Sektor kamen Huybrechts und Nicholls (2013, S. 142) zu dem Ergebnis, dass das Social Enterprise nicht in der Lage war, die Kollaboration gemäß seiner Normen zu gestalten und seine Macht langfristig zu erhalten. Dagegen zeigt das Papier von Calò, Teasdale, Donaldson, Roy und Baglioni (2018) positive Ergebnisse von Kooperationen im Social Entrepreneurship auf. So führt die Arbeit von Social Enterprises in einem kollaborativen Umfeld zu verbesserten Ergebnissen, wie Verbundenheit, Wohlbefinden und Selbstbewusstsein (Calò et al., 2018, S. 1791).

5.3.2. Netzwerke

In dieser Kategorie wurden Informationen zu strategischen Netzwerken und Projektnetzwerken gefunden, die mindestens ein Social Enterprise als Teilnehmer beinhalten.

Bezüglich der strategischen Netzwerke im Social Entrepreneurship hat sich Herbst (2019) mit der Kooperation von Social Enterprises mit Wettbewerbern befasst, was als Coopetition bezeichnet wird. Ihre Forschung zeigt, dass durch das Zusammenspiel von Kooperation und Wettbewerb Synergien entstanden sind, welche zu positiven Ergebnissen führten Herbst (2019).

Weitere positive Ergebnisse der Kollaboration von Social Enterprises im strategischen Netzwerk wurden von Maase und Bossink (2010) herausgearbeitet. Demnach können Social Enterprises (hier im Besonderen: Start-Up Social Enterprises), welche erfolgreich interorganisationale Beziehungen knüpfen, Folgendes kontrollieren: die Eigenschaften des Social Entrepreneurs und die kooperativen Prozesse zwischen dem Social Entrepreneur und den beteiligten Organisationen. Social Enterprises, die keine Partnerschaften aufbauen konnten, sind dazu nicht in der Lage (Maase & Bossink, 2010, S. 82).

Die Arbeit von Huybrechts et al. (2017), deren Inhalt bereits in Kapitel 5.1 und 5.2 Erwähnung fand und sich mit verschiedenen Partner-Strategien von Social Enterprises im Fair Trade-Sektor beschäftigt, beleuchtet auch die jeweiligen Ergebnisse. So führte die Strategie der „Sector solidarity“ (Huybrechts et al., 2017, S. 595) nur zu niedrigen Verkaufszahlen (Huybrechts et al., 2017, S. 596). Die Strategie des „Selective engagement[s]“ (Huybrechts et al., 2017, S. 597) führte dagegen zu viel höheren Verkaufszahlen und Umsätzen (Huybrechts et al., 2017, S. 598). Die dritte Strategie, genannt „Active appropriation“ (Huybrechts et al., 2017, S. 599), war aus kommerzieller Sicht die erfolgreichste (Huybrechts et al., 2017, S. 600).

Gillett et al. (2019) betrachteten ein Projektnetzwerk, dessen Teilnehmer drei Social Enterprises und ein Gemeinderat waren. Ein spezielles Ergebnis dieser vorübergehenden Zusammenarbeit stellte die öffentliche Aufmerksamkeit dar, die das Projekt für alle Beteiligten hervorbrachte (Gillett et al., 2019, S. 965).

Tabelle 5 stellt eine thematische Übersicht der erläuterten Aspekte dieses Forschungsbereichs dar.

5.4. Funktionen des Netzwerkmanagements

Das Management der interorganisationalen Beziehungen von Social Enterprises wird ebenfalls in der wissenschaftlichen Literatur dieses Forschungsbereiches betrachtet. In diesem Kapitel werden die wichtigsten Erkenntnisse im Sinne der vier beschriebenen Funktionen des Netzwerkmanagements gegliedert. Eine Zusammenfassung aller relevanten Aspekte ist in Tabelle 6 zu finden. Auch dyadische Beziehungen werden gemäß den Funktionen des Netzwerkmanagements gesteuert und kontrolliert. Deshalb wurden im Rahmen dieses Reviews sowohl die Papiere, welche sich mit netzwerkförmigen Organisationsbeziehungen von Social Enterprises beschäftigen, als auch die Werke, welche Dyaden in diesem Kontext betrachten, nach Informationen bezüglich der vier Funktionen untersucht.

5.4.1. Selektionsfunktion

Die Selektion geeigneter Partner durch das Social Enterprise stellt in der wissenschaftlichen Literatur bezüglich Organisationsbeziehungen im Social Entrepreneurship eine wichtige und ausschlaggebende Aufgabe dar. So beschreiben Di Domenico et al. (2009) die Partnerselektion bei dyadischen Organisationsbeziehungen von Social Enterprises als „critical“ (Di Domenico et al., 2009, S. 895). Auch Intindola et al. (2019) heben die wichtige Rolle der Selektionsfunktion, allerdings im regionalen Netzwerk, hervor und beschreiben diese als „one specific practice of possible importance“ (Intindola et al., 2019, S. 16). Weiterhin zeigen Zafeiropoulou und Koufopoulos (2013) in ihrer Forschungsarbeit auf, dass die Partnerwahl ebenfalls ein fester Bestandteil beim Formen des Netzwerks im Social Franchising ist (Zafeiropoulou & Koufopoulos, 2013, S. 90).

Im Rahmen dieses Literaturreviews wurden auch einige Selektionskriterien herausgearbeitet. Di Domenico et al. (2009, S. 895) empfehlen, dass Social Enterprises keine dyadischen Beziehungen mit Organisationen eingehen sollen, deren Ziele und Werte nicht mit der sozialen Mission des Unternehmens einhergehen, unabhängig von den potenziellen ökonomischen Erträgen. Dagegen arbeiten herkömmliche Unternehmen nur mit Social Enterprises zusammen, wenn sie dadurch wirtschaftliche Vorteile, wie die Vermittlung neuer Kunden oder die Verbesserung des eigenen Firmenansehens, erlangen.

In Bezug auf strategische Netzwerke von Social Enterprises stellen Huybrechts et al. (2017) verschiedene Selektionsprozesse von Social Enterprises im Fair Trade-Sektor vor. Hierbei werden geeignete Partner gemäß der jeweiligen Markteintrittsstrategie des Social Enterprises ausgewählt. Während die erste Strategie nur die Zusammenarbeit mit anderen Social Enterprises erlaubt, wird bei den anderen beiden auch mit herkömmlichen Unternehmen im For-Profit-Bereich kooperiert (Huybrechts et al., 2017, S. 595ff.). Clark

Tabelle 5: Themen zu Ergebnissen und Folgen der Kooperation

Dyaden	Netzwerke
• Einfluss verschiedener Partner auf die soziale Performance von Social Enterprises (n/a) (Choi, 2015) • Auswirkung verschiedener Partner-Typen auf Bricolage-Ergebnisse von Social Enterprises (lateral) (Kwong et al., 2017) • Ergebnisse der Beziehung zwischen einem Social Enterprise und einem For-Profit-Unternehmen im Fair Trade-Sektor (vertikal) (Huybrechts & Nicholls, 2013) • Verbesserung der Ergebnisse von Social Enterprises durch Arbeit in einem kollaborativen Umfeld (n/a) (Calò et al., 2018)	• Eigenschaft von Social Enterprises nach der Schaffung einer erfolgreichen Partnerschaft, bestimmte Dinge zu kontrollieren (SN/n/a) (Maase & Bossink, 2010) • Coopetition (Kooperation von Wettbewerbern) führt zur Erreichung von Synergien, die positive/machtvolle Ergebnisse hervorbringen (SN/horizontal/vertikal) (Herbst, 2019) • Ergebnisse von Strategien zur Partnerwahl im Fair Trade-Sektor (SN/vertikal/lateral) (Huybrechts et al., 2017) • Bekanntheit/öffentliche Aufmerksamkeit für alle Teilnehmer am Projektnetzwerk (PN/lateral) (Gillett et al., 2019)

et al. (2017) haben in ihrer Studie ebenfalls die Managementprozesse in strategischen Netzwerken von Social Enterprises untersucht und stellen bezüglich der Selektion die bevorzugten Partner-Typen der jeweiligen sozialen Unternehmen in ihrer Arbeit dar. Dazu gehören in den meisten Fällen profitorientierte Unternehmen und Organisationen des Gemeinschaftssektors (Clark et al., 2017, S. 44f.).

5.4.2. Allokationsfunktion

Bezüglich der Allokationsfunktion wurden in der wissenschaftlichen Literatur drei Werke gefunden. Montgomery et al. (2012) beschäftigen sich mit dem Bündeln und Handeln von Ressourcen in dyadischen Organisationsbeziehungen von Social Enterprises. Dabei stellen sie jeweils sowohl sektorübergreifende Prozesse als auch Vorgehensweisen im gleichen Sektor vor (Montgomery et al., 2012, S. 378ff.).

In Bezug auf strategische Netzwerke von interorganisationalen Beziehungen der Social Enterprises stellt Sigala (2019) die Kooperationsprozesse eines sozialen Restaurants dar. Die Arbeit im Netzwerk macht es für das Restaurant möglich, Ressourcen zu erlangen und mit den anderen Beteiligten auszutauschen. Soziale Vorteile, die durch die gemeinsam ausgeführten Tätigkeiten erreicht werden, werden gleichmäßig auf alle Teilnehmer aufgeteilt (Sigala, 2019, S. 37).

Mit einem speziellen Fall der Zusammenarbeit von Social Enterprises mit Organisationen im strategischen Netzwerk beschäftigt sich das Papier von Gold et al. (2020). Hier kooperiert ein Social Business Orchestrator mit mehreren Social Enterprises, um diese beim Erreichen ihrer sozialen Mission zu unterstützen. Die Ressourcen im Netzwerk, wie z.B. finanzielle Mittel und technisches Fachwissen, werden von allen Teilnehmern im Bündnis gesammelt (Harzing, 2020, S. 272) und somit jeder Organisation zur Verfügung gestellt.

5.4.3. Regulationsfunktion

Im Sinne dieser Funktion entscheiden sich die beteiligten Partner für Verfahren zur Regulation ihrer Kooperationsprozesse. Diese können sowohl formeller als auch informeller Natur sein (Sydow, 2010, S. 397). Dies bestätigt Lyon (2012) in Bezug auf Organisationsbeziehungen von Social Enterprises im strategischen Netzwerk: „These types of relationships can be both formal (based on contract) and informal (based more on word-of-mouth and a common understanding)" (Lyon, 2012, S. 14).

Gillett et al. (2019) untersuchten ein Projektnetzwerk, an dem drei Social Enterprises und ein Gemeinderat teilnahmen. Sie fanden u.a. heraus, dass sich hierbei eine Gruppe von Organisationsvertretern wöchentlich traf, um Aspekte zur Projektsteuerung zu besprechen. Entscheidungen wurden einstimmig getroffen und bindende Vereinbarungen auf informellem Weg ohne Vertrag geschlossen (Gillett et al., 2019, S. 959). Auch Huybrechts und Nicholls (2013) stellten bei der Untersuchung einer dyadischen Beziehung zwischen einem Social Enterprise und einem Einzelhändler im Fair-Trade-Sektor einen Mangel an formalen Verfahren bei Vereinbarungen fest (Huybrechts & Nicholls, 2013, S. 140).

Neben den genannten Aspekten zur Allokationsfunktion hat Sigala (2019) auch Informationen zum Regulationsprozess bei der interorganisationalen Kooperation eines sozialen Restaurants im strategischen Netzwerk herausgearbeitet. So werden in diesem Netzwerk u.a. Regeln, Normen und soziale Werte für die gemeinsame Gestaltung von sozialem Wert in zeitlichen Abständen immer wieder neu definiert. Weiterhin entwickeln die Beteiligten gemeinsam eine geläufige Sprache für die Zusammenarbeit im Netzwerk (Sigala, 2019, S. 37).

Auch Clark et al. (2017) haben sich im Rahmen ihrer Forschungsarbeit mit der Regulationsfunktion in strategischen interorganisationalen Netzwerken von Social Enterprises beschäftigt. Bei der Untersuchung von sechs spezifischen sozialen Unternehmen wurden sowohl informelle als auch formelle Verfahren zur Regulation der Beziehungen im Netzwerk dargestellt. So schloss eines der Social Enterprises beispielsweise formale Produktionsverträge ab, nutzte jedoch überwiegend informelle Prozesse für das Management der einzelnen Netzwerkbeziehungen (Clark et al., 2017, S. 44f.).

5.4.4. Evaluationsfunktion

Die Evaluation der Tätigkeiten innerhalb der interorganisationalen Beziehungen von Social Enterprises wird ebenfalls in der wissenschaftlichen Literatur betrachtet. Im Rahmen dieses Literaturreviews wurden zwei Werke herausgearbeitet, die sich mit dieser Thematik beschäftigen.

Gold et al. (2020), die sich mit der Kooperation eines Social Business Orches-trators mit Social Enterprises in einem strategischen Netzwerk beschäftigt haben, stellen fest, dass der Orchestrator u.a. die Verkaufszahlen und die Höhe des Umsatzes der Social Enterprises überwacht und auswertet (Gold et al., 2020, S. 273).

Clark et al. (2017), die die Managementprozesse von sechs Social Enterprises bezüglich ihrer Organisationsbeziehungen im strategischen Netzwerk untersucht haben, haben bei der Hälfte der betrachteten Unternehmen formale Evaluationsverfahren festgestellt (Clark et al., 2017, S. 49f.). Bei einem der sechs Social Enterprises wurde speziell herausgearbeitet, dass der Fortschritt und das Ergebnis der Allianzen beobachtet und ausgewertet wurden (Clark et al., 2017, S. 50). In Tabelle 6 sind alle Themen bezüglich der Funktionen des Netzwerkmanagements des Literaturreviews dargestellt.

6. Diskussion und Fazit

Das Literaturreview hat gezeigt, dass interorganisationale Beziehungen im Social Entrepreneurship eine bereits viel erforschte Thematik darstellen und die Verwendung aktueller Beiträge (Gold et al., 2020; Herbst, 2019; Jug, 2020) macht deutlich, dass derzeit ein hohes Forschungsinteresse besteht. Die Forschungsfrage dieser Arbeit, die im Anschluss zusammenfassend beantwortet wird, lautet: „Welche Themen wurden bereits in der wissenschaftlichen Literatur zu interorganisationalen Beziehungen von Social Enterprises betrachtet?“

Im Rahmen dieser Arbeit wurde betrachtet, welche Aspekte bezüglich der Anbahnung, des Prozesses und der Ergebnisse interorganisationaler Beziehungen von Social Enterprises bereits in der wissenschaftlichen Literatur Beachtung fanden. Durch die Aufteilung des Forschungsbereiches in diese drei genannten Phasen konnte Folgendes festgestellt werden: Die Forschung bezüglich des Kooperationsprozesses interorganisationaler Beziehungen von Social Enter-prises ist am weitesten fortgeschritten. Zwar wurden Aspekte zur Anbahnung und zu Ergebnissen dieser Beziehungen ebenfalls untersucht, jedoch fällt die Anzahl der Beiträge, die sich damit beschäftigt haben, im Vergleich zur zweiten Phase relativ gering aus.

In Bezug auf die untersuchten Netzwerktypen werden in der wissenschaftlichen Literatur strategische Netzwerke am häufigsten betrachtet. Hier stellt ein Social Enterprise die fokale Unternehmung dar, welche die Netzwerkführerschaft übernimmt. Die Analyse von regionalen Netzwerken und Projektnetzwerken kommt bisher nur selten vor und eine virtuelle Unternehmung wurde im Verlauf des Literaturreviews in keinem Papier entdeckt.

Weiterhin konnte durch das Literaturreview festgestellt werden, dass kooperative Bündnisse mit Organisationen für Social Enterprises in der Praxis eine herausragende Rolle spielen. Dies zeigt die überwiegende Menge empirisch-qualitativer Forschungsbeiträge, die konkrete interorganisationale Zusammenschlüsse von Social Enterprises in der Praxis untersucht haben. Gründe von Social Enterprises für die interorganisationale Zusammenarbeit sind u.a. das Erlangen von Ressourcen (Schirmer & Cameron, 2012, S. 84f.) und Legitimität (Weidner et al., 2019, S. 517). Dagegen stellt für For-Profit-Unternehmen z.B. die Stärkung ihrer Corporate Social Responsibility einen Grund für die Kooperation mit Social Enterprises dar (Jug, 2020). Neben den herkömmlichen Unternehmen kooperieren u.a. auch Non-Profit-Organisationen und Regierungsorganisationen mit Social Enterprises. Im Rahmen des Literaturreviews wurde kein bestimmter Partnertyp in interorganisationalen Beziehungen von Social Enterprises ausgemacht, der in der wissenschaftlichen Literatur überwiegt.

Die Untersuchung der Ergebnisse interorganisationaler Beziehungen im Social Entrepreneurship zeigt, dass die Partnerorganisationen einen großen Einfluss auf den Arbeitsprozess und die Wirkung von Social Enterprises haben. Die Kooperation mit anderen Organisationen kann für Social Enterprises sowohl positive Ergebnisse, wie eine größere öffentliche Bekanntheit (Gillett et al., 2019, S. 965), als auch negative Folgen, wie Kontrollverlust im Rahmen der Zusammenarbeit (Kwong et al., 2017, S. 631), bewirken.

Bezüglich der Funktionen des Netzwerkmanagements bei interorganisationalen Beziehungen von Social Enterprises wurde bisher vor allem die Selektionsfunktion in der wissenschaftlichen Literatur betrachtet. Dabei wurde im Besonderen die Wichtigkeit dieser Funktion hervorgehoben und Kriterien für die Auswahl geeigneter Partnerorganisationen durch das Social Enterprises in der Praxis untersucht. Dagegen wiesen relativ wenige Forschungsbeiträge Informationen zu den Funktionen der Allokation und der Regulation auf und mit der Evaluationsfunktion haben sich nur zwei der betrachteten Papiere beschäftigt. Im Allgemeinen war festzustellen, dass die Funktionen des Netzwerkmanagements bisher wenig untersucht worden sind und in dem Forschungsbereich der interorganisationalen Beziehungen im Social Entrepreneurship keine zentrale Thematik darstellen.

Interorganisationale Beziehungen von Social Enterprises weisen viel zukünftiges Forschungspotenzial auf. Wie bereits beschrieben, ist der Kooperationsprozess bereits relativ häu-

Tabelle 6: Themen zu Funktionen des Netzwerkmanagements

Selektion	Allokation	Regulation	Evaluation
• **Wichtige Rolle dieser Funktion:** ⇒ Notwendigkeit der Partner-Selektion (Dyade/lateral) (Di Domenico et al., 2009) ⇒ Selektion ist unverzichtbar (RN/n/a) (Intindola et al., 2019) ⇒ Selektion ist fester Bestandteil im Netzwerk und leichter, wenn traditionelle Businesses beteiligt sind (Social Franchise) (Zafeiropoulou & Koufopoulos, 2013) • **Selektionskriterien:** ⇒ Spezielle Selektionskriterien und Ausschlussgründe (Dyade/lateral) (Di Domenico et al., 2009) ⇒ Selektionskriterien im Fair Trade-Sektor (SN/vertikal/lateral) (Huybrechts et al., 2017) ⇒ Papier mit detaillierten Selektionsprozessen (SN/alle) (Clark et al., 2017) • **Weitere Themen:** ⇒ Screening der potenziellen Partner vor Auswahl (Dyade/vertikal) (Huybrechts & Nicholls, 2013) ⇒ Entwicklungsphase und Diversität der Partnerauswahl (SN/alle) (Folmer et al., 2018) ⇒ Einfluss der Teamselektion auf Zufriedenheit des Teams (RN/n/a) (Intindola et al., 2019)	• Strategien zum Beschaffen und Handeln von Ressourcen im Social Entrepreneurship (Dyade/vertikal) (Montgomery et al., 2012) • Vorgänge in einem sozialen Restaurant bezüglich Allokationsfunktion (SN/vertikal) (Sigala, 2019) • Kooperation von Social Business Orchestrator mit Social Enterprises (SN/lateral) (Gold et al., 2020)	• Formelle und informelle Vorgänge (SN/n/a) (Lyon, 2012) • Informelle Vereinbarungen ohne Vertrag (PN/lateral) (Gillett et al., 2019) • Regeln bei Beziehung zwischen Social Enterprise und einem For-Profit-Unternehmen (Dyade/vertikal) (Huybrechts & Nicholls, 2013) • Vorgänge in einem sozialen Restaurant (SN/vertikal) (Sigala, 2019) • Papier mit Details zu verschiedenen Vorgängen bezüglich Regulationsfunktion (SN/alle) (Clark et al., 2017)	• Kooperationsvorgang von Social Business Orchestrator mit Social Enterprises bezüglich Evaluationsfunktion (SN/lateral) (Gold et al., 2020) • Papier mit Details zu verschiedenen Vorgängen bezüglich Evaluationsfunktion (SN/alle) (Clark et al., 2017)

fig in den betrachteten Werken untersucht worden. Deshalb besteht bezüglich der Aspekte zur Anbahnung und zu Ergebnissen der Organisationsbeziehungen von Social Enterprises aktuell ein vergleichsweise größerer Forschungsbedarf. Dieser besteht auch bei den Funktionen des Netzwerkmanagements, da im Rahmen des Literaturreviews ein Mangel an relevanten Beiträgen diesbezüglich festgestellt wurde.

In Bezug auf das Forschungsdesign der betrachteten Werke ist ein Mangel an empirisch-quantitativer Forschung festgestellt worden. Dies kann daran liegen, dass der Forschungsbereich zum Social Entrepreneurship noch relativ jung ist und es bisher nicht möglich war, statistische Analysen größerer Datenmengen durchzuführen. In Zukunft sollte es als Ziel formuliert werden, diese Forschungslücke zu verringern.

Bei der Betrachtung der geografischen Herkunft der jeweiligen Papiere ist aufgefallen, dass die meiste Forschung zu interorganisationalen Beziehungen von Social Enterprises bisher zwar in Europa betrieben wurde, dabei allerdings nur eine geringe Zahl der Werke aus Deutschland stammt. Deshalb sollten sich deutsche Wirtschaftswissenschaftler zukünftig mehr mit diesem Forschungsbereich auseinandersetzen.

Literatur

Bloom, P. N. & Chatterji, A. K. (2009). Scaling Social Entrepreneurial Impact. *California Management Review*, *51* (3), 114–133.

Calò, F., Teasdale, S., Donaldson, C., Roy, M. J. & Baglioni, S. (2018). Collaborator or competitor: assessing the evidence supporting the role of social enterprise in health and social care. *Public Management Review*, *20* (12), 1790–1814.

Certo, S. T. & Miller, T. (2008). Social entrepreneurship: Key issues and concepts. *Business Horizons*, *51* (4), 267–271.

Chell, E. (2007). Social Enterprise and Entrepreneurship: Towards a Convergent Theory of the Entrepreneurial Process. *International Small Business Journal*, *25* (1), 5–26.

Choi, Y. (2015). How Partnerships Affect the Social Performance of Korean Social Enterprises. *Journal of Social Entrepreneurship*, *6* (3), 257–277.

Choi, Y. & Majumdar, S. (2014). Social entrepreneurship as an essentially contested concept: Opening a new avenue for systematic future research. *Journal of Business Venturing*, *29* (3), 363–376.

Clark, S., Woods, M. & Adams, D. (2017). Managing alliance portfolios: Lessons From Social Enterprise. In T. K. Das (Hrsg.), *Managing Alliance Portfolios and Networks* (S. 33–69). Information Age Publishing.

Conway Dato-on, M. & Kalakay, J. (2016). The winding road of social entrepreneurship definitions: a systematic literature review. *Social Enterprise Journal*, *12* (2), 131–160.

de Bruin, A., Shaw, E. & Lewis, K. V. (2017). The collaborative dynamic in social entrepreneurship. *Entrepreneurship and Regional Development*, *29* (7-8), 575–585.

Dees, G. (2003). *Social Entrepreneurship is About Innovation and Impact, Not Income.* Zugriff am 2020-09-15 auf https://centers.fuqua.duke.edu/case/wp-content/uploads/sites/7/2015/02/Article_Dees_SEisAboutInnovationandImpactNotIncome_2003.pdf

Defourny, J. & Nyssens, M. (2006). Defining social enterprise. In M. Nyssens (Hrsg.), *Social Enterprise. At the crossroads of market, public policies and civil society* (S. 3–26). Routledge.

Department of Trade and Industry. (2002). *Social Enterprise: A Strategy for Success*. DTI.

Dhanaraj, C. & Parkhe, A. (2006). Orchestrating Innovation Networks. *Academy of Management Review*, *31* (3), 659–669.

Di Domenico, M., Tracey, P. & Haugh, H. (2009). The Dialectic of Social Exchange: Theorizing Corporate-Social Enterprise Collaboration. *Organization Studies*, *30* (8), 887–907.

Doherty, B., Haugh, H. & Lyon, F. (2014). Social Enterprises as Hybrid Organizations: A Review and Research Agenda. *International Journal of Management Reviews*, *16* (4), 417–436.

Faltin, G. (2011). Social Entrepreneurship – Zwischen Entrepreneurship und Ethik. In P. Jähnke, G. B. Christmann & K. Balgar (Hrsg.), *Social Entrepre-neurship. Perspektiven für die Raumentwicklung* (S. 75–85). VS Verlag für Sozialwissenschaften.

Folmer, E., Nederveen, C. & Schutjens, V. (2018). Network importance and use: commercial versus social enterprises. *Social Enterprise Journal*, *14* (4), 470–490.

Gillett, A., Loader, K., Doherty, B. & Scott, J. M. (2019). An Examination of Tensions in a Hybrid Collaboration: A Longitudinal Study of an Empty Homes Project. *Journal of Business Ethics*, *157* (4), 949–967.

Gold, S., Chowdhury, I. N., Huq, F. A. & Heinemann, K. (2020). Social business collaboration at the bottom of the pyramid: The case of orchestration. *Business Strategy and the Environment*, *29* (1), 262–275.

Harzing, A. (2020). *Journal Quality List 2020 (67th Edition).* Zugriff am 2020-09-15 auf https://harzing.com/download/jql67_title.pdf

Henry, C. (2015). Doing Well by Doing Good: Opportunity Recognition and the Social Enterprise Partnership. *Journal of Social Entrepreneurship*, *6* (2), 137–160.

Herbst, J. M. (2019). Harnessing sustainable development from niche marketing and coopetition in social enterprises. *Business Strategy and Development*, *2* (3), 152–165.

Huybrechts, B. & Nicholls, A. (2013). The role of legitimacy in social enterprise-corporate collaboration. *Social Enterprise Journal*, *9* (2), 130–146.

Huybrechts, B., Nicholls, A. & Edinger, K. (2017). Sacred alliance or pact with the devil? How and why social enterprises collaborate with mainstream businesses in the fair trade sector. *Entrepreneurship and Regional Development*, *29* (7-8), 586–608.

Intindola, M. L., Pittz, T. G., Rogers, S. E. & Weisinger, J. Y. (2019). Partner Selection in Social Entrepreneurship Collectives: How Team Selection Control Can Enhance Satisfaction in Cross-Sector Social Partnerships. *Journal of Social Entrepreneurship*, 1–26.

Jayawarna, D., Jones, O. & Macpherson, A. (2020). Resourcing Social Enterprises: The Role of Socially Oriented Bootstrapping. *British Journal of Management*, *31* (1), 56–79.

Jug, A. (2020). Why Should Corporations Partner With Social Enterprises and Entrepreneurs? *IEEE Engineering Management Review*, *48* (2), 202–205.

Kwong, C., Tasavori, M. & Wun-mei Cheung, C. (2017). Bricolage, collaboration and mission drift in social enterprises. *Entrepreneurship and Regional Development*, *29* (7-8), 609–638.

Lam, I. K. K. (2016). Tripartite Collaborative Model Value Creation Experience of iEnterprise with Corporate and Nongovernmental Organization. In T. Dudycz, G. Osbert-Pociecha & B. Brycz (Hrsg.), *The Essence and Measurement of Organizational Efficiency* (S. 137–173). Springer International Publishing.

Langdon, D. & Burkett, I. (2004). *Defining social enterprise. Enterprising ways to address long-term unemployment*. PI Productions.

Lasprogata, G. A. & Cotten, M. N. (2003). Contemplating enterprise: the business and legal challenges of social entrepreneurship. *American Business Law Journal*, *41* (1), 67–114.

Lee, M. & Jay, J. (2015). Strategic Responses to Hybrid Social Ventures. *California Management Review*, *57* (3), 126–147.

Lévi-Strauss, C. (1967). *The Savage Mind*. University of Chicago Press.

Littlewood, D. & Khan, Z. (2018). Insights from a systematic review of literature on social enterprise and networks. *Social Enterprise Journal*, *14* (4), 390–409.

Lyon, F. (2012). Social innovation, co-operation and competition: inter-organizational relations for social enterprises in the delivery of public services. In A. Nicholls & A. Murdock (Hrsg.), *Social Innovation. Blurring Boundaries to Reconfigure Markets* (S. 139–161). Palgrave Macmillan.

Maase, S. J. F. M. & Bossink, B. A. G. (2010). Factors that inhibit partnering for social start-up enterprises. *Journal of Enterprising Communities: People and Places in the Global Economy*, *4* (1), 68–84.

Martin, R. L. & Osberg, S. (2007). Social entrepreneurship: the case for definition. *Stanford Social Innovation Review*, 29–39.

Meyskens, M. & Carsrud, A. L. (2013). Nascent green-technology ventures: a study assessing the role of partnership diversity in firm success. *Small Business Economics*, *40* (3), 739–759.

Meyskens, M., Carsrud, A. L. & Cardozo, R. N. (2010). The symbiosis of entities in the social engagement network: The role of social ventures. *Entrepreneurship and Regional Development*, *22* (5), 425–455.

Miemczyk, J., Johnsen, T. E. & Macquet, M. (2012). Sustainable purchasing and supply management: a structured literature review of definitions and measures at the dyad, chain and network levels. *Supply Chain Management: An International Journal*, *17* (5), 478–496.

Möllering, G. (2010). Kartelle, Konsortien, Kooperationen und die Entstehung neuer Märkte. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, *62* (7), 770–796.

Montgomery, A. W., Dacin, P. A. & Dacin, M. T. (2012). Collective Social Entrepreneurship: Collaboratively Shaping Social Good. *Journal of Business Ethics*, *111* (3), 375–388.

Moura, A. M., Comini, G. & Teodósio, A. D. S. (2015). The International Growth of a Social Business: A Case Study. *Revista de Administração de Empresas*, *55* (4), 444–460.

Nicholls, A. & Huybrechts, B. (2016). Sustaining Inter-organizational Relationships Across Institutional Logics and Power Asymmetries: The Case of Fair Trade. *Journal of Business Ethics*, *135* (4), 699–714.

Perrini, F., Vurro, C. & Costanzo, L. A. (2010). A process-based view of social entrepreneurship: From opportunity identification to scaling-up social change in the case of San Patrignano. *Entrepreneurship and Regional Development*, *22* (6), 515–534.

Phillips, W., Alexander, E. A. & Lee, H. (2019). Going It Alone Won't Work! The Relational Imperative for Social Innovation in Social Enterprises. *Journal of Business Ethics*, *156* (2), 315–331.

Phillips, W., Lee, H., Ghobadian, A., O'Regan, N. & James, P. (2015). Social

Innovation and Social Entrepreneurship. *Group and Organization Management*, *40* (3), 428–461.

Provan, K. G., Fish, A. & Sydow, J. (2007). Interorganizational Networks at the Network Level: A Review of the Empirical Literature on Whole Networks. *Journal of Management*, *33* (3), 479–516.

Ramus, T., La Cara, B., Vaccaro, A. & Brusoni, S. (2018). Social or Commercial? Innovation Strategies in Social Enterprises at Times of Turbulence. *Business Ethics Quarterly*, *28* (4), 463–492.

Schirmer, H. & Cameron, H. (2012). Collaborations and Partnerships. In C. K. Volkmann, K. O. Tokarski & K. Ernst (Hrsg.), *Social entrepreneurship and social business. An introduction and discussion with case studies* (S. 83–101). Gabler Verlag.

Schumpeter, J. A. (1993). *Theorie der wirtschaftlichen Entwicklung. Eine Untersuchung über Unternehmergewinn, Kapital, Kredit, Zins und den Konjunkturzyklus*. Duncker & Humblot.

Seanor, P. (2018). Of course, trust is not the whole story: narratives of dancing with a critical friend in social enterprise–public sector collaborations. In P. Dey & C. Steyaert (Hrsg.), *Social Entrepreneurship* (S. 159–181). Edward Elgar Publishing.

Seelos, C. & Mair, J. (2005).): Social entrepreneurship: Creating new business models to serve the poor. *Business Horizons*, *48* (3), 241–246.

Short, J. C., Moss, T. W. & Lumpkin, G. T. (2009). Research in social entrepreneurship: past contributions and future opportunities. *Strategic Entrepreneurship Journal*, *3* (2), 161–194.

Sigala, M. (2016). Learning with the market. *International Journal of Contemporary Hospitality Management*, *28* (6), 1245–1286.

Sigala, M. (2019). A market approach to social value co-creation: Findings and implications from "Mageires" the social restaurant. *Marketing Theory*, *19* (1), 27–45.

Sinthupundaja, J., Kohda, Y. & Chiadamrong, N. (2020). Examining Capabilities of Social Entrepreneurship for Shared Value Creation. *Journal of Social Entrepreneurship*, *11* (1), 1–22.

Smeets, D. J. A. (2017). Collaborative Learning Processes in Social Impact Bonds: A Case Study from the Netherlands. *Journal of Social Entrepreneurship*, *8* (1), 67–87.

Smith, B., Meyskens, M. & Wilson, F. (2014). Should We Stay or Should We Go? 'Organizational' Relational Identity and Identification in Social Venture Strategic Alliances. *Journal of Social Entrepreneurship*, *5* (3), 295–317.

Snow, C. C., Miles, R. E. & Coleman, H. J. (1992). Managing 21st century network organizations. *Organizational Dynamics*, *20* (3), 5–20.

Sydow, J. (1992). *Strategische Netzwerke*. Gabler Verlag.

Sydow, J. (2010). Management von Netzwerkorganisationen – Zum Stand der Forschung. In J. Sydow (Hrsg.), *Management von Netzwerkorganisationen. Beiträge aus der "Managementforschung"* (S. 373–470). Gabler Verlag.

Sydow, J. & Duschek, S. (2011). *Management interorganisationaler Beziehungen. Netzwerke, Cluster, Allianzen.* Kohlhammer.

Sydow, J. & Möllering, G. (2015). *Produktion in Netzwerken. Make, buy & co-operate*. Verlag Franz Vahlen.

Sydow, J. & Windeler, A. (1997). Komplexität und Reflexivität in Unternehmensnetzwerken. In H. W. Ahlemeyer & R. Königswieser (Hrsg.), *Komplexität managen: Strategien, Konzepte und Fallbeispiele* (S. 147–162). Gabler Verlag.

Tasavori, M., Kwong, C. & Pruthi, S. (2018). Resource bricolage and growth of product and market scope in social enterprises. *Entrepreneurship and Regional Development*, *30* (3-4), 336–361.

The Association of Business Schools. (2015). *Academic Journal Guide 2015.* Zugriff am 2020-09-15 auf https://gsom.spbu.ru/files/abs-list-2015.pdf

Volkmann, C. K., Tokarski, K. O. & Ernst, K. (2012). *Social entrepreneurship and social business. An introduction and discussion with case studies*. Gabler Verlag.

Weidner, K., Weber, C. & Göbel, M. (2019). You Scratch My Back and I Scratch Yours: Investigating Inter-Partner Legitimacy in Relationships Between Social Enterprises and Their Key Partners. *Business and Society*, *58* (3), 493–532.

Yaari, M., Blit-Cohen, E. & Savaya, R. (2019). Hybrid Organizational Culture: The Case of Social Enterprises. *Journal of Social Entrepreneurship*, 1–23.

Yunus, M., Moingeon, B. & Lehmann-Ortega, L. (2010). Building Social Business Models: Lessons from the Grameen Experience. *Long Range Planning*, *43* (2-3), 308–325.

Zafeiropoulou, F. A. & Koufopoulos, D. N. (2013). The Influence of Relational Embeddedness on the Formation and Performance of Social Franchising. *Journal of Marketing Channels*, *20* (1-2), 73–98.

Junior Management Science 7(2) (2022) 524-542

Junior Management Science

journal homepage: www.jums.academy

Investigating Market Behavior Correlations between Classified Tokens using the International Token Classification Framework

Felix van Walbeek

Technische Universität München

Abstract

This paper explores the novel International Token Classification framework, creates a large sample set of tokens classified according to the framework, clusters the tokens into groups, and performs statistical analysis regarding the selected token's correlation. It investigates how the current token landscape looks by classifying 200 tokens. These tokens are clustered in three different groups, payment token, DeFi ecosystem token, and network utility tokens. We first investigate whether tokens tend to move in the same direction with the tokens from their group, and secondly, we use a created average portfolio return to compare the single token return with the different groups.

According to the results, we mainly found utility and payment tokens from the IT and Finance industries. Out of the three groups, tokens clustered in the payment token group showed the highest correlations within the group and with their own group portfolio average. Overall, we conclude that the classification indeed has an impact on the relationship of token pairs. However, the results show that many more factors influence the market behavior of tokens, which should also be considered.

Keywords: Blockchain; token; correlation; classification; Bitcoin.

1. Introduction

More than ten years ago, a developer named Laszlo Hanyecz ordered two large pizzas from Papa John's and made history. Hanyecz used a novel digital payment method called Bitcoin, which is now known as the first transaction using cryptocurrencies to pay for a tangible product. At this time, Hanyecz paid 10.000 BTC for two pizzas, where each Bitcoin was worth way less than a penny (McCall, 2020; Moore, 2020). However, at the time of writing and the current exchange price for Bitcoin with over 45 thousand US Dollars (USD), he would be able to buy a tenth of the whole pizza chain Papa John's for 450 million USD (CoinMarketCap, 2021b; yahoo! finance, 2021). Since 2010 the market for cryptocurrencies and mainly the growth of the world's first cryptocurrency, Bitcoin, have gained more attention (Chen, 2018). However, the technology behind Bitcoin has many more use cases than only serving as a digital currency. For example, blockchain is already of great relevance for healthcare, the supply chain sector, and many more industries (Knight, 2017). Adding assets to a blockchain is referred to as tokenization, and the blockchain representation of that asset is called token (Roth, Schär, & Schöpfer, 2019; Schär, 2020). The rising market leads to a growing need for standardization in classifying a token (Sandner & Ketz, 2019).

This paper will give insights into the basic concepts of the token economy, the classification of a token, and how this can be done using the International Classification Framework (ITC) published by the International Token Standardization Association (ITSA) (Sandner, Ketz, Tumasjan, & Lentge, 2019). Finally, we will provide insights on how classified tokens correlate in terms of price regarding their classification. Over the last year's first research papers have investigated the cross-correlation between price changes of different cryptocurrencies (Stošić, Stosic, & Ludermir, 2018). As the subject of tokenization is still at an early stage, it should be

First, I want to thank Prof. Dr. Philipp Sandner for giving me the opportunity to write my bachelor thesis, always having great suggestions, providing new and great ideas to improve my thesis and generally supporting me throughout the time of writing. In addition, I want to thank the chair of Strategy and Organization and Prof. Dr. Isabell M. Welpe, who was my supervisor from Technical University Munich. Lastly, I am grateful for Constantin Ketz, who always supported my work and encouraged me and my ideas.

DOI: https://doi.org/10.5282/jums/v7i2pp524-542

mentioned that a lot of research, and sources are still distributed among tweets, forums and blog posts. Only a small part of the material is of academic origin. As the world has not yet decided on a standardized token classification framework, it is also new to the academic world. This paper adds value to the topic of a unique and standardized classification framework by providing a first sample set of tokens classified according to the ITC. Afterward, we will provide a strategy to build individual clusters using the classification and group tokens to perform statistical analysis regarding the token's price behavior. The results of this paper will create one of the first use cases for the ITC. All different stakeholders taking part in the token market, reaching from private users, investment managers, and regulators to academics, can use the classification of the tokens to identify specific patterns in the token market. According to the tokens classification, private users and investors can use the correlation analysis results for diversifying their token portfolio regarding the price movement of tokens.

2. Theoretical Background and Research Gap

Prior to the main work of the paper, including a deep dive into the classification work and the following quantitative empirical analysis of token market behavior, it is essential to give a brief introduction into blockchain technology, the token market, classification frameworks, as well as the use of correlation coefficients for time series data in the financial market.

2.1. Distributed Ledger Technology, Blockchain, and Bitcoin

In 2008, someone used the pseudonym Satoshi Nakamoto to publish the first conception of the digital currency Bitcoin. With that, he introduced a digital payment infrastructure with cryptographic proof instead of trusting centralized authorities (Nakamoto, 2008). The concept of distributed ledger technology (DLT) describes a technique used to document data or transactions. While in the classical approach of a ledger, a centralized authority manages the ledger. A distributed ledger has any number of identical copies of the ledger maintained in a decentralized manner by the different participating stakeholders. Although often used as synonyms, the DLT and blockchain are not the same. Blockchain, the technology known behind Bitcoin and Ethereum, is a specific implementation of the DLT. Blockchain got its name from the process of grouping transactions or other data into blocks and attaching those blocks to a chain of already verified blocks. A cryptographic signature called hash is used to connect the blocks. Appropriate measures called consensus mechanisms are used to ensure that newly added transactions are adopted in all copies of the ledger and that there is consensus on the ledger's current status at any given time (Drescher, 2017). After the used consensus algorithm validates the transaction, the transactions are irreversible, verifiable, permanent, and secure (Chen, 2018). The first ideas to use the concept of a blockchain go even back to 1991, where Haber and Stornetta (1991) introduced a computationally practical procedure to guarantee the unalterable digital time-stamping of documents (Haber & Stornetta, 1991). However, blockchain technology's real breakthrough came with the Bitcoin whitepaper in 2008 (Cong & He, 2019).

2.2. The Concept of Smart Contracts and Possible Use-Cases of Blockchain

Bitcoin and other digital cryptocurrencies like Litecoin serve as a virtual payment method and could potentially disrupt the financial industry (Böhme, Christin, Edelman, & Moore, 2015). Beyond the use-case of the digital payment functionalities of many blockchain applications called Blockchain 1.0, it is essential to introduce two more generations of blockchain. Blockchain 2.0, which covers the concept of smart contracts, mainly focuses on decentralized finance (DeFi) and finally Blockchain 3.0, which describes in the broadest sense all those use cases of the blockchain technology beyond the finance sector.

For this paper, it is essential to introduce a key concept called smart contracts, as they are among other things used to bring assets onto the blockchain. Smart contracts should firstly fulfill the same use case as a physical contract, but in the context of the blockchain, smart contracts are programmed computer scripts stored on a blockchain. However, unlike the classic physical contract that only defines specific rules and penalties related to an individual agreement between two parties, the smart contract also automatically executes once the requirements are fulfilled to minimize the external participation and, therefore, the risk of fraud by a third party. With the smart contract, it is now possible that everything noted in the contract, even though both participants do not trust or know each other, will automatically be executed (Christidis & Devetsikiotis, 2016).

With the continuous further development of the technology over the last years, many more industries than the financial sector showed interest in the new technology. For example, entrepreneurship, digital rights management, the public sector, the energy sector, the healthcare market, and the supply chain management are highly interested in the blockchain market (Casino, Dasaklis, & Patsakis, 2019; Hughes, Park, Kietzmann, & Archer-Brown, 2019; Knight, 2017). In the healthcare sector, a common problem is that data is stored inefficiently and centralized in local systems. Different stakeholders including the doctor, healthcare provider, the patient and insurance companies each store the data locally. Blockchain technology now offers the possibility of storing data anonymously on a blockchain network (Hughes et al., 2019). A specific use case was proposed by Ekblaw, Azaria, Halamka, and Lippman (2016), who introduced a decentralized electronic health records management system to overcome the problem of centralized stored patient data. Furthermore, one of the most affected industries will be the supply chain market, where blockchain will disrupt the market by providing detailed information about cost inefficiency. Perboli, Musso, and Rosano (2018) provided a specific use case

of how blockchain can overcome exactly these hidden cost-inefficient structures and optimize logistic operations.

2.3. The Growing Concept of Tokens

After introducing blockchain technology and the smart contracts' basic concepts, this paper gives an overview of the concept of tokenization. In the beginning, the concept of tokens was reserved to the crypto asset, which was native to their respective blockchain, e.g., Ether for the Ethereum Blockchain or Bitcoin for the Bitcoin blockchain. But as the interest in the concept was rising, the idea was to store an asset onto the ledger, which is not native to the blockchain. The digital asset representation on the blockchain is called token, and the process of bringing the asset to the blockchain is called tokenization (Roth et al., 2019; Schär, 2020).

The concept of tokenization can now, for example, be used to digitize a famous drawing. Imagine a famous Picasso drawing worth one million USD. The number of people who are able to buy the painting is minimal. The painting's value is defined through ownership of the physical version that cannot be divided easily. By creating a token, each worth one USD, the drawing's value and ownership can now be distributed between an infinite number of people, each owning a small piece of the expensive drawing. Adding to the advantage of transferring fractional ownership as described, more advantages come along with the concept. Global access for investments from everywhere using a smartphone into different markets is allowed, and also, it has become much simpler to trade digital valuables thanks to the high level of liquidity. As soon as certain conditions are fulfilled, the smart contract automatically executes itself almost immediately, allowing real-time transactions. Because blockchain technology allows direct trades between buyers and sellers, intermediatory participants are reduced. Two more advantages that come with tokenization are the given transparency and the immutability the blockchain technology is providing. Not only is every transaction visible, but also impossible to change or delete a transaction once it is validated and added to the blockchain, which adds a very high-security standard to the transparency factor (Sazandrishvili, 2020).

Regarding the technical implementation of those tokens, often specific standards are set using smart contract templates. A famous example is the widely used standard called ERC-20 for the Ethereum blockchain introduced by Vogelsteller and Buterin (2015), which defines specific rules and functions that every token using the ERC-20 standard should implement. These technical standards can now be used to implement protocols, for example, on top of the Ethereum blockchain, which will later be referred to as application layer protocol. Protocol is

2.4. Token Classification Frameworks

While not only the token market is having its second renaissance in terms of total market capitalization, the variety of tokens offered is also increasing. This leads to a growing need for standardization in token classification as the market is currently still lacking a standardized way to distinguish between the nature of tokens (Sandner & Ketz, 2019).

However, before heading into specific different approaches of token classification, it is essential to distinguish between the term "token" and "coin". CoinMarketCap, as one of the most common market trackers, follows the rule that the term "coin" refers to all those assets which are native to a blockchain, e.g., Bitcoin, Ethereum, XRP. In contrast, the term "token" refers to those built on top of a blockchain and governed by smart contracts (CoinMarketCap, 2021a). Previous scientific research tends to agree on this distinction between "token" and "coin" or "cryptocurrency" using technical-based differentiation aspects (Chen, 2018; Massey, Dalal, & Dakshinamoorthy, 2017).

A second approach to differentiate between different crypto assets is the purpose of creation. The German Federal Financial Supervisory Authority called BaFin divides token into three subclasses payment-token, equity- and other investment token and utility token. Payment-token usually have the exclusive function of serving as a payment method, for example, Bitcoin. In the case of equity- and other investment token, the token holder gets provided with claims under debt law with monetary content and membership rights, similar to shares and securities. The subclass utility token covers tokens that are not designed to serve with payment functionalities across different ecosystems but to be used to purchase the token provider's real economic good (Fußwinkel & Kreiterling, 2018). The private consulting company Deloitte is certifying this division into payment, investment token, and utility token and is following a similar approach (Deloitte, 2019). The Swiss Financial Market Supervisory Authority (FINMA) also tends to agree on this diversification but labels the beforehand called investment token as asset token (FINMA, 2018).

A different approach, called The Token Classification Framework (TCF), has been developed by Euler (2018), who divided the classification into five major dimensions. The first dimension covers aspects of the token's primary purpose. This is in line with a previously made distinction regarding the term's token, cryptocurrency and investment token. A token can be a cryptocurrency to serve as a digital payment method, a network token to enable a specific network and speed up its growth, or an investment token to provide the opportunity to invest in an entity or an asset. The second dimension portraits the utility of the token. Here they divide into usage tokens, which should give access to the network or service feature and work tokens, allowing the token holder actively to participate in the system. A third dimension covers the legal status, whether the token should be treated as a cryptocurrency, a utility token, or a security token. The token's underlying value is covered in the fourth dimension of the framework, where the question is covered where the token derives its value. Here is a distinction made between asset-backed tokens, where the value comes from the asset the token is backed by, share-like tokens that would most likely be regarded as securities and network value tokens, where the token is tied to the value of a network. The

last dimension covers the technological implementation of the token, whether it is implemented as the blockchain's native token token, on top of a different blockchain and therefore non-native-protocol token, or on the application level as (d)App token (Euler, 2018).

2.5. The International Token Classification Framework

After reviewing some of the different existing frameworks spread across the crypto asset market for this paper, the ITC will be used. The following chapter's information is obtained from the ITC questionnaire (Sandner et al., 2019) and the documentation (ITSA, 2020). The ITC published first in 2019 by ITSA is a new approach to classify crypto tokens in a standardized way. It was created to provide a flexible tool to classify a token that different market stakeholders can use. One of the framework's explicit goals is to provide clear and transparent characteristics for tokens in multiple different dimensions. Public institutions like central banks or governments can use the ITC to gain more in-depth knowledge about the token landscape. Private investors or investment funds active in the crypto space can use the framework to run detailed market analyses or diversify their investment portfolios by knowing about the token's classification.

The framework is designed to adapt continuously and is always open for further development. For the sake of retaining an overview, the framework is vertically split into levels. Level 1 is pointing to the highest level. In version 1.0, published in October 2020, the framework covers four so-called dimension groups (level 1). Each dimension group can have an infinite number of dimensions (level 2). At the framework's current, two out of four dimension groups cover more than one dimension. Figure 1 provides an overview of the existing dimensions.

Every dimension contains then different categories (level 3), which are then further divided into subcategories (level 4), classes (level 5), subclasses (level 6), groups (level 7), and finally subgroups (level 8). The ITC uses unique ITC codes for each level. A code consists of each level's individual level segment code's composition and therefore represents a hierarchical classification path. This paper will strictly stick to the code labels used and described in the ITC and the complementary documentation. To avoid confusion every time we talk about the classification label, we will provide the ITC Code of the token and write the label in capital letters, e.g., Utility Token (EEP22).

Table 1 provides a hierarchical overview of the vertical levels and how an ITC code is assembled, using the subclass USD-Pegged Payment Token (EEP21PP01USD). The framework's design was created to leave enough space for adding further subdivisions on every hierarchical level.

As the six different dimensions on the second-highest level are the fundamental concept of the ITC, we want to provide further information for each of the dimensions. A more detailed description of all levels below level two, the dimensions can be found in the ITC framework documentation (ITSA, 2020).

Economic Purpose (EEP): The first dimension of the ITC, called Economic Purpose (EEP), uses the same distinction into three different categories as the BaFin, Deloitte, and FINMA, which has been described in the previous chapter 2.4. Comparing it to the Token Classification Framework by Euler (2018), this dimension is in line and covers aspects of the TCF's three dimensions, the purpose, the underlying value, and the utility aspects of a token. The currently used categories are Payment Tokens (EEP21), Utility Tokens (EEP22), and Investment Tokens (EEP23). Regarding the classification Payment Token (EEP21), it is important to state that the token should serve in the same way as a real-world currency in different environments. Payment Tokens (EEP21) should most likely be compared to US Dollar, Euro, the Chinese Yuan, or any other currency. A special case of payment token which is covered by the dimension are the so-called Pegged Payment Tokens (EEP21PP), often in the literature referred to as stablecoins, which try to follow a particular stable peg (e.g., USD) (Klages-Mundt & Minca, 2020). The second category of the dimensions is Utility Token (EEP22). Utility Tokens (EEP22) are designed to be used within the given environment created by the issuer. Its utility functionality can reach from serving as an access voucher to the ecosystem to a specific economic good or functionality of the issuer or distributing rewards for ecosystem participants.

Furthermore, the category Utility Token (EEP22) covers those tokens which provide the token holder with a governance functionality or a particular ownership right. The last category within the economic dimension is Investment Token (EEP23) and can be compared to equity- and other investment tokens earlier introduced by the BaFin (Fußwinkel & Kreiterling, 2018). Investment Tokens (EEP23) include those tokens "that are designed to provide institutional and/or retail investors with an instrument for investment (incl. trading, speculation, and/or hedging)" (ITSA, 2020, p. 17). As previously stated, this category is often labeled in the market as "security token". The ITC avoids using the term "security token" as the term illuminates not the token's economic purpose but on its regulatory status (Sandner & Ketz, 2019).

Issuer Industry (EIN): The second dimension, which is also part of the Economic Dimension Group (E) called Issuer Industry (EIN), covers different industries the issuer of the token is active. It is important to emphasize that the industry of the issuer of the token is the crucial point to look at and not the industry where the token finds its primary use case.

Technological Setup (TTS): The dimension Technological Setup (TTS) covers the discussion previously held about whether to call a digital asset "token" or "cryptocurrency". The ITC disagrees with the previous research presented to use the term "coin" or "cryptocurrency" as it lacks a clear definition. Regarding the technological setup, the ITC splits into two categories. On the one hand, Ledger-Native Token (TTS41), "which captures every Token that is implemented by means of a Distributed Ledger Protocol and thus forms an integral part of such software protocol (incl. the consensus mechanism defined for the Distributed Ledger)" (ITSA, 2020, p. 21). And on the other hand the category Applica-

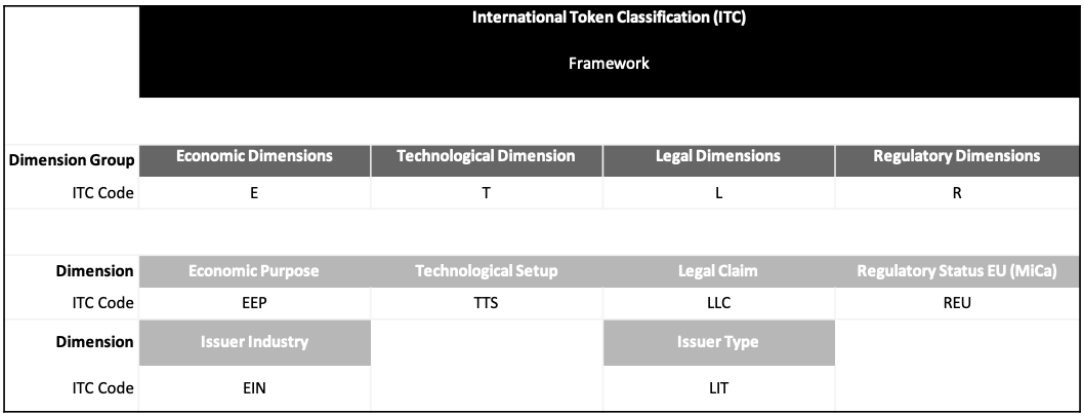

Figure 1: Dimension Overview of the ITC v1.0

Source: Own illustration based on (ITSA, 2020)

Table 1: ITC Code Composition Example for the Subclass "USD-Pegged Payment Token

Level	Level Label	Level Segment	Level Segment Code	ITC Code
1	Dimension Group	Economic Dimensions	E	E
2	Dimension	Economic Purpose	EP	EEP
3	Category	Payment Token	21	EEP21
4	Subcategory	Pegged Payment Token	PP	EEP21PP
5	Class	Fiat-Pegged Payment Token	01	EEP21PP01
6	Subclass	USD-Pegged Payment Token	USD	EEP21PP01USD
7	Group	[n/a]	[n/a]	[n/a]
8	Subgroup	[n/a]	[n/a]	[n/a]

Source: Own illustration based on (ITSA, 2020, p. 5)

tion Layer Token (TTS42), that "captures every Token that is implemented by means of an Application Layer Protocol on top of a Distributed Ledger" (ITSA, 2020, p. 22).

Legal Claim (LLC): The dimension Legal Claim (LLC) captures information on whether the token provides any legal rights to the token holder. At the current status of the ITC framework, this dimension needs to be treated with special caution. Many tokens manage their ecosystem in a decentralized nature, and often, no real third party exists where the legal claim could be raised against (Sandner & Ketz, 2019).

Issuer Type (LIT): The second dimension, Issuer Type (LIT) within the Legal Claim Dimension Group (L), covers information about the background of the token issuer. The dimension splits itself into two categories, one covering those where a legal entity behind the token can be found called Legal Entity (LIT61), and those types of issuers without a legal entity called Entity without Legal Personality (LIT62). A Legal Entity (LIT61) can be either be a Private Sector Legal Entity (LIT61PV), which covers companies or foundations issuing a token, or a Public Sector Legal Entity (LIT61PC) which refers to governments, central banks, or ministries. Token issuers, which are not clearly defined or stated in the official token documents, would fall under the category Entity without Legal Personality (LIT62). A further distinction here is whether the token is either issued by a distributed ledger directly, subcategory Distributed Ledger Protocols (LIT62DL), or is based on the distributed ledger of another protocol, which would then result in a classification into the subcategory of Application Layer Protocols (LIT62AL).

Regulatory Status EU (REU): The last dimension is the first dimension covering aspects of a token's regulatory status. This category is entirely based on the European Commission's proposal for a Regulation on Markets in Crypto Assets (MiCA) (European Commission, Directorate-General for Financial Stability, & Financial Services and Capital Markets Union, 2020). The MiCa regulation includes a description of the largest digital assets regulation as of to date and tries to provide detailed regulation rules for the entire crypto asset market (Sandner & Blassl, 2020). As the MiCa framework is still only a proposal and has not yet entered into force, it is essential to state that the dimension does not represent any official classification made by the European Commission. The ITC aims to test the potential applicability of the MiCa proposal even before it enters into force. With the category Crypto Asset in Scope of MiCA (REU51) tokens, which are either defined as Payment Token (EEP21) or Utility Token (EEP22) in the first dimension Economic Purpose (EEP) are covered. Investment Token (EEP23) are for now out of the

scope of the MiCa and therefore falls under the second category Crypto Asset out of Scope of MiCa (REU52) of the dimension Regulatory Status EU (REU).

2.6. Price Correlation in the Token Economy and Research Gap

Correlation plays a crucial role in the traditional finance market and has been studied broadly in academic literature. Analysts and investment fund managers have used correlation analysis to diversify and allocate their assets across different sectors and industries. Using cross-correlation matrices allows dividing stocks into different subsets that are similar to previously identified business sectors. Identifying sectors can be useful to find an investment that can earn a return without increasing the risk (Gopikrishnan, Rosenow, Plerou, & Stanley, 2000). Due to the early stage, the market for crypto assets is in academic research regarding token price correlation is only very limited. Stošić et al. (2018) showed several collective behaviors in the crypto asset market, which can help construct a crypto asset portfolio. The literature is mainly focused on the correlation between Bitcoin and other assets such as Gold. Klein, Pham Thu, and Walther (2018) revealed that Bitcoin in shock-like moments does not negatively correlate with the market, and the price declines whenever markets are declining. Apart from the academic literature, Binance Research (2019c) suspects that the type of consensus algorithm impacts the token's price behavior. It seems that Proof-of-Work (PoW) tokens exhibit a higher correlation with each other than with non-PoW tokens indicating the impact of the technical setup of the token on the price. Another point mentioned in two different reports from Binance Research (2019b, 2020) is that programmable blockchains such as EOS, NEO, and Ethereum often have a higher correlation with each other than with non-programmable assets. As shown in a third report by Binance Research (2019a), payment token with a particular focus on privacy such as Monero or Dash shows higher than average correlations with each other compared to other tokens.

This paper will provide a sample set of tokens classified according to the new ITC v1.0 to help establish the ITC as a standardized way on how tokens should be classified. Afterward, we will use the unique dataset to build certain groups of tokens to investigate the token's price behavior. First, we will look at the correlations between tokens that are clustered in the same group. Secondly, we will investigate whether tokens correlate more with their groups than the groups they are not part of.

3. Methodology

The following chapter describes a three-step process used for this paper. A novel dataset has been created in the first step, where tokens are classified using the ITC framework introduced in chapter 2.5. Afterward, this classification data is used to create and cluster different groups of tokens according to their classification. Finally, by adding historical data of the grouped tokens, interesting correlation aspects are investigated. Each subchapter will include how the data is obtained and processed.

3.1. Classification of Tokens

In the first step in co-creation with the Project Working Group 2 – ITC (PWG2) of ITSA, a novel dataset is created. Approximately 200 tokens are analyzed and classified according to the before described ITC v1.0. The dataset, which includes the top 200 tokens according to market capitalization, is extracted from CoinMarketCap on November 29, 2020, at 1:00 pm and is sorted by the tokens market capitalization. To come up with a specific classification for the token, every token is first classified by three different participants of the PWG2 independently. Afterward, a consensus is found during a discussion with all members of the PWG2. To identify the classification for each token only official materials are used. This includes the tokens website, the whitepaper, and medium articles published by the token entity.

To follow a clear and unified line for the classification of a token according to the ITC, this paper follows the provided questionnaire and the classification guidelines provided by ITSA, which can be found in appendix A. We found additional guidelines during our work of classifying tokens, which we want to provide in this paper while remaining in line with the questionnaire and the previously defined guidelines by ITSA (ITSA, 2020; Sandner et al., 2019).

Suppose a Transactional Utility Token (EEP22TU) only provides access to the decentralized network or application that the token is implemented on and not to any additional service, product, or functionality within that network or application. In that case it will not classify as a Settlement and Access Token (EEP22TU02). However, it shall be classified as Settlement Token (EEP22TU01) as the provision of access to the decentralized network or application that the token is implemented on is considered to be a necessary prerequisite for its transactional functionality and not a functionality of its own.

Suppose a Utility Token (EEP22) is implemented as Blockchain-Native Token (TTS41BC), and the token is designed as an integral part of a Proof-of-Stake (PoS) consensus mechanism. In that case, the token will classify as Settlement and Governance Token (EEP22TU03) since the token provides governance functionality as part of the PoS consensus mechanism, which governs the distributed ledger. If a Utility Token (EEP22) is implemented as Application Layer Token (TTS42) and the application layer protocol assigns certain governance functionality to the token (e.g., voting rights), it will classify as Settlement and Governance Token (EEP22TU03). However, the Application Layer Token's classification does not depend on the type of consensus mechanism of the underlying distributed ledger.

Regarding the regulatory status of a token according to MiCA (Regulatory Status EU (MiCA)), the following rules apply:

1. If a token is classified as Utility Token (EEP22) with respect to its Economic Purpose (EEP), it will also classify

as Utility Token (REU51UT) with respect to its regulatory status in the EU (REU).
2. If a token is classified as Unpegged Payment Token (EEP21UP) according to its Economic Purpose (EEP), it will classify as Other Crypto Asset in Scope of MiCA (REU51ZZ), as the current version of MiCA does not feature a dedication definition for this category of tokens.
3. If a token is classified as Pegged Payment Token (EEP21PP) with respect to its Economic Purpose (EEP), it will either classify as E-Money Token (REU51EM) or as Asset-Referenced Token (REU51AR). Further information on the differentiation of both categories can be found in the Classification Questionnaires and/or the ITC Code Descriptions.
4. Whenever a token is classified as Investment Token (EEP23) according to its Economic Purpose (EEP), it will most likely classify as Crypto Asset out of Scope of MiCa (REU52).

Suppose a token provides access or governance functionality and is classified accordingly in the Economic Purpose (EEP) dimension. In that case, it does not necessarily imply that the token provides such functionality (e.g., access or voting rights) in the form of real relative rights against a counterparty, and hence it does not directly imply a classification as Relative Rights Token (LLC32). Each case has to be analyzed individually, and no generalizations can be made.

Regarding the classification of a "wrapped" version of a token, this paper follows a clear rule of thumb. As a wrapped version of a token helps to overcome the problem of interoperability between blockchains, one of the primary purposes is to represent the token on other blockchains to serve with payment functionalities, e.g., Wrapped Bitcoin. As its price is pegged to the original token, the token should always be classified as Asset Pegged Payment Token (EEP21PP02).

3.2. Token Groups according to the Token Classification

After classifying the top 200 tokens according to their market capitalization, it is important to build specific clusters to investigate correlation behavior. As the market for tokens is still in its early stage and projects are rising with very high speed into top positions regarding their market capitalization, this paper focuses only on tokens, which are classified and have a market capitalization of over 100 million USD. Out of the 99 remaining tokens, token groups are built. In the following section, we want to describe each group's requirements and if and why this group is included in the further analysis. The tokens grouping mainly focuses on the classification of the tokens regarding the Economic Dimension Groups (E) and the Technological Setup Dimension (TTS).

Group 1: Payment Token: The first group we are looking at is the so-called group payment token. The first condition for a token to be considered as a payment token is that the token is not only classified on a category level as a Payment Token (EEP21) but as an Unpegged Payment Token (EEP21UP). This is done to exclude the before introduced stablecoins. As their value is pegged to a particular value, mainly USD, stablecoins are in this paper not of relevance regarding the correlation of the tokens price. Also, we exclude all the wrapped versions of tokens. For example, Wrapped Bitcoin, the ERC-20 version of Bitcoin, is almost 100% correlating with Bitcoin due to its nature. The second condition for the group payment token is the Issuer Industry (EIN) of a token. A crucial factor is that the token issuer is active in the industry of Payment Services and Infrastructure (EIN06PS) or within the industry Cyber Security, Data Privacy and Digital Identity (EIN05DA04). This is done to include payment tokens focusing on privacy, as their focus is the privacy functionality, and are therefore not classified in the industry Payment Services and Infrastructure (EIN06PS) but still exist with the primary purpose to serve as a digital currency.

Group 2: DeFi Ecosystem Token: For our second group, we focus on all tokens active in the DeFi space. The group will include all classified tokens that fall under the subcategory Decentralized Finance (EIN06DF), including tokens from all different DeFi classes. We decided to look at the DeFi tokens as an aggregated group and not as further divided single groups to get a sufficiently large enough sample set.

Group 3: Network Utility Token: With the third group network utility token, we want to introduce a group of tokens serving as the native asset for their blockchain to power their own ecosystem. Regarding the classification in the Issuer Industry (EIN) dimension, we have to make sure that the token is classified as Cloud Computing, Distributed Systems, and Decentralized Applications (EIN05DA03). This class covers all tokens which are created to power a decentralized ecosystem such as Ethereum or Polkadot. All of these tokens are created to power their ecosystem created around their blockchain. The token has to be classified as some kind of Utility Token (EEP22) as the token should provide a specific type of utility within the defined environment. As the third criteria, the token has to be classified as Blockchain-Native Token (TTS41BC) regarding the Technological Setup (TTS) to make sure we only include tokens native to their own blockchain. An overview of the requirements and examples for the groups can be found in figure 2.

3.3. Correlation between Classified Tokens

The underlying dataset contains the token's historical price and is obtained from CoinGecko (2021), which is "a leading source of information on cryptocurrencies" (S. Wang & Vergne, 2017, p. 5) on January 27, 2021. For each token assigned to one of the before defined groups, we downloaded the token's price data for the last 90 days and calculated the token's daily arithmetic return. We used a short-term period of 90 days to ensure that we have the same data for each token. Some of the projects included in this analysis are very new to the market, e.g., Uniswap had its token launch only in September 2020. It is hard to compare with tokens such as Bitcoin that provide multiple years of price data. Furthermore, we follow the approach of Coinbase (2021), who are also using a time period of 90 days to calculate their correlation coefficients.

Group 1: Payment Token	*Group 2: DeFi Ecosystem Token*	*Group 3: Network Utility Token*
Requirements:	***Requirements:***	***Requirements:***
Economic Purpose: Unpegged Payment Token (EEP21UP)	Economic Purpose: -	Economic Purpose (EEP): Utility Token (EEP22)
Issuer Industry: Payment Services and Infrastructure (EIN06PS) / Cyber Security, Data Privacy and Digital Identity (EIN05DA04)	Issuer Industry: Decentralized Finance (EIN06DF)	Issuer Industry (EIN): Cloud Computing, Distributed Systems, and Decentralized Applications (EIN05DA03)
Technological Setup: -	Technological Setup: -	Technological Setup (TTS): Blockchain - Native Token (TTS41BC)
Examples:	***Examples:***	***Examples:***
Bitcoin, Litecoin, Bitcoin Cash	Link, Uniswap, Aave	Ethereum, Cardano, Polkadot

Figure 2: Requirements and Examples for the Token Groups

Source: Own illustration

As prices heavily differ in their total number, ranging from more than thousands of dollars to cents per token, it is vital to use returns instead of prices (Birch, Pantelous, & Soramäki, 2016; Meucci, 2010). We start by defining a return: r_i at the time *i,* where p_i is the price at the time i and $j = (i - 1)$:

$$r_i = \frac{p_i - p_j}{p_j}, \tag{1}$$

Using returns instead of raw price data comes with the benefit of normalization and is ubiquitous in finance. Due to the high differences in the absolute price of a token, it is unavoidable to normalize them to measure the price variable in a comparable metric. Besides, it is widespread in finance to use logarithmic returns instead of arithmetic returns, this paper sticks to arithmetic returns. This is possible as the considered timeframe is very short, and therefore, the logarithmic return does not significantly differ from the arithmetic return (Meucci, 2010).

3.3.1. Price Correlation within a Token Group

In a first step, we look at each previously defined group individually and compare each token within its own group. To do so, we use the pair-wise Pearson correlation coefficient between all pairs of daily returns. The Pearson correlation coefficient C_{ij} between the token i and j is defined:

$$C_{ij} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sqrt{\langle r_i{}^2 - \langle r_i \rangle^2 \rangle \langle r_j{}^2 - \langle r_j \rangle^2 \rangle}}, \tag{2}$$

where r_i and r_j are the token return vectors for token i and j respectively, and $\langle . \rangle$ is an average over the period investigated. For n tokens, the Pearson correlation matrix C is

$$C = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}, \tag{3}$$

with all entries ranging from [-1,1], where -1 indicates a total negative linear correlation, 0 no linear correlation at all, and +1 total positive linear correlation. Correlation coefficients between 0 and 0.3 (0 and -0.3), can be interpreted as a weak positive (negative) linear relationship. If a coefficient has its value between 0.3 and 0.7 (-0.3 and -0.7) we talk about a moderate positive (negative) linear relationship and about a strong positive (negative) linear correlation for values between 0.7 and 1.0 (-0.7 and -1.0) (Ratner, 2009). As a correlation is symmetric, we will only show the lower triangular of the correlation matrix C (Birch et al., 2016; Orac, 2017; G. J. Wang, Xie, & Stanley, 2018).

3.3.2. Price Correlation between Token Groups

In a second step, this paper aims to compare correlation coefficients between the different groups. Therefore, we need to provide a way on how we can compare correlation coefficients. For each of our groups, we created a naive portfolio according to DeMiguel, Garlappi, and Uppal (2009), where each of the assets is weighted the same using a weight of *1/n.* We follow this simple approach of asset-allocation for the following reasons. Due to the ease of its implementations, we do not have to consider any optimization of the weighting process. Secondly, it is still widespread across investors to use simple allocation rules to diversify their portfolios (Benartzi & Thaler, 2001; DeMiguel et al., 2009). Compared to an approach where we would have used market capitalization as a weighting factor, we avoid that single assets like Bitcoin or Ethereum with an enormous market capitalization are dominating a whole group portfolio. As of January

11, 2021, Bitcoin and Ethereum together makeup 81.15% of the whole crypto market (CoinMarketCap, 2021a). The average return gr of a group portfolio k including n tokens with their return r_i is now defined as:

$$gr_k = \frac{1}{n}\sum_{1}^{n} r_i, \tag{4}$$

For this part of the analysis, we modify Equation (2), where the Pearson correlation coefficient C_{ik} now indicates the correlation between token i and group k.

$$C_{ik} = \frac{\langle r_i gr_k\rangle - \langle r_i\rangle\langle gr_k\rangle}{\sqrt{\langle {r_i}^2 - \langle r_i\rangle^2\rangle\langle {gr_k}^2 - \langle gr_k\rangle^2\rangle}}, \tag{5}$$

where r_i is the token return vector and gr_k is the average group portfolio return vector for token i and group portfolio k, respectively, and $\langle . \rangle$ is an average over the period investigated.

As each token we compare is included in one of the three groups, the token would automatically correlate with the group more than with other groups, only because the token itself is included in the group. Therefore, we always do not include the return of the considered token to calculate the average group of the tokens group. For example, looking at Bitcoin, we calculate the Payment Token portfolio's average group return while not including Bitcoin in the calculations. The group portfolio returns for Network Utility Tokens and DeFi Ecosystem Tokens remain untouched. The analysis will give us a matrix with three correlation coefficients C_{ik} per token for each of the defined groups. For all calculations, we used the python packages pandas and SciPy (The Pandas Development Team, 2020; Virtanen et al., 2020). The python script can be found in appendix B.

4. Results and Discussion

The following chapter will describe the obtained results in the same structure as the previous chapter regarding the paper's methodology. We will start with descriptive statistics on the top 200 tokens classified, followed by presenting the token groups and the statistical investigations' results of the statistical investigations regarding the token's price behavior.

4.1. Classification of the Top 200 Tokens

The full dataset including, the classification data of each token included in the top 200 list that was created in joint work with the project working group 2 of ITSA, can be found in appendix C. We only need the two economic dimensions and the technological setup dimension for the further grouping of tokens. Therefore, we only provide descriptive statistics about the Economic Purpose (EEP), the Issuer Industry (EIN), and the Technological Setup (TTS) in the appendix. Highlights are stated in the following sections. The results are sorted according to the ITC framework and only include the lowest level applicable if at least one token has been classified.

Economic Purpose: Looking only into the level of categories out of 200 classified tokens, 170 are Utility Tokens (EEP22), which makes a utility token dominance of 85%. 28 and correspondingly 14,0% are Payment Token (EEP21), and only 2 out of 200 tokens are classified as Investment Token (EEP23). While looking closer into the observed Payment Token (EEP21), we found mainly Unpegged Payment Tokens (EEP21UP) such as Bitcoin or Litecoin and USD-Pegged Payment Token (EEP21PP01USD) such as Tether or USD Coin. Payment tokens pegged to a different fiat currency have only rarely been found, with one token classified as EUR-Pegged Payment Token (EEP21PP01EUR).

Those tokens that fall under the category Utility Token (EEP22) do most likely serve with a mean of transaction settlement within the defined environment by the token's issuer and are therefore classified as Transactional Utility Token (EEP22TU) on a category level. The subcategory is divided into three classes. Those tokens that serve within the defined environment only with the means of transaction settlement (Settlement Token (EEP22TU01), 35 token). Or the transaction settlement functionality combined with access to a particular service, good or functionality (Settlement and Access Token (EEP22TU02), 37 token). The third option, Settlement and Governance Token (EEP22TU03), 90 tokens, combines not only the transaction settlement and access functionality but adds also a governance functionality. To cover the Non-Transactional Utility Token (EEP22NT), we find three tokens classified as Access Token (EEP22NT01) and nine tokens classified as Governance Token (EEP22NT02). A detailed breakdown of the classification can be found in appendix D.

Issuer Industry: Regarding the Issuer Industry (EIN), the dataset is dominated by two Categories. 46% of the tokens can be classified in the category Information, Communication and IT (EIN05) and 46% in the category Finance and Insurance (EIN06). Considering the subcategory level, mainly three subcategories are containing a significant amount of all tokens, Software, Data Storage and Processing (EIN05DA) with 41.5% of all 200 tokens, Payment Services and Infrastructure (EIN06PS), with 17.5% and the Decentralized Finance (EIN06DF) with 18.5%. A detailed breakdown of the classification numbers can be found in appendix E.

Technological Setup: While analyzing the technological setup of the top 200 tokens, the first thing we notice is that the total number of tokens is 209 instead of 200. This is due to the technological implementation of the following seven tokens, Tether, UNUS SED LEO, Binance USD, TrueUSD, Chiliz, Serum, and ShareToken, that exist on more than one blockchain. For example, the token Tether is implemented on four different blockchains. Tether has parallel running versions of the token on the four different blockchains Ethereum, Omni, EOS, and Tron. Regarding the distribution of the tokens, mainly two classifications are dominant. 46,41% of the tokens under consideration are Blockchain-Native Tokens (TTS41BC), and 42,58% are Ethereum ERC-20 Standard

Tokens. A detailed breakdown of the classification numbers can be found in appendix F.

Regarding the three additional dimensions, we want to provide some insights without going into further detail. In the Legal Claim Dimension (LLC), we only find two different classifications. One hundred ninety-three tokens do not provide any legal claim, while only seven tokens provide some relative rights and are, therefore, classified as Relative Rights Token (LLC32). Regarding the Issuer Type (LIT) Dimension out of the top 200 tokens, 156 are classified as Private Sector Legal Entity (LIT61PC), 29 as Distributed Ledger Protocol (LIT62DL), and 15 as Application Layer Protocol (LIT62AL). In the last dimension, Regulatory Status EU (MiCa) of the ITC, tokens are classified according to the MiCa. Eight tokens are classified as Non-Authorized Significant E-Money Token (REU51EM12), 169 as Utility Token (REU51UT), 22 as Other In-Scope Crypto Assets (REU51ZZ), and one as Crypto Asset out of Scope of MiCa (REU52)

4.2. Correlation between Classified Tokens

For the following chapter, we look at each of the beforehand defined groups individually. We first present the groups, followed by the correlation coefficients between each token and those clustered in the same group. In the last step, we compare each token with the beforehand created group portfolio returns. We repeat the process for each of the groups. A whole list of the three groups with the tokens they contain and their corresponding tickers, which we will use, can be found in the in table 2.

Group 1: payment token: The first group covers all privacy focuses, and non-privacy focused tokens, which are created to serve as a digital payment method. The group covers a total number of 13 tokens, including the most popular token, Bitcoin, Litecoin, or Monero. As previously mentioned in this paper's methodology part, so-called stablecoins are not included in the category as they are not relevant for this correlation analysis. Regarding the correlation within this group, we find that Bitcoin only shows a strong positive linear relationship with Litecoin. The highest observed correlation coefficient is between Bitcoin Cash and Bitcoin SV with 0.851. Bitcoin Cash also is the only token showing four strong positive linear relationships, followed by Zcash and Bitcoin Diamond with three. Especially low correlation coefficients are found for Dogecoin with eight weak, four moderate, and no strong positive linear relationships within the token group. A detailed list of all correlation coefficients between each of the token pairs can be found in table 3.

We observe that five of the 13 tokens show a strong positive linear relationship with their own group portfolio. Bitcoin Cash shows the highest correlation with 0.839 and Dogecoin the lowest with 0.345, which supports our findings from before. The eight remaining tokens all show a moderate positive linear relationship. For twelve of the 13 payment tokens, we find a moderate positive linear relationship with the DeFi ecosystem token portfolio. Only Dogecoin with a coefficient of 0.2380 shows a weak positive linear relationship. Looking at the correlation between the single tokens and the third group, network utility tokens, we observe two tokens, Litecoin and Zcash showing a strong positive linear relationship with the group portfolio. The remaining tokens are indicating a moderate positive linear relationship. Furthermore, we can see that eleven out of 13 correlate the most with the payment token portfolio. Only Litecoin and Decred correlate more with group 3: network utility tokens while showing a minimal difference in the coefficient. Table 4 shows the correlation coefficients between each of the payment tokens and the three portfolios.

It is also notable that we only see positive linear relationships between all the single token pairs and the correlations between the payment tokens and all of the portfolios.

Group 2: DeFi Ecosystem Token: Our second group includes all tokens active in the DeFi space. A total number of 17 out of the 99 tokens which we considered are from the DeFi ecosystem. This group includes six tokens classified as Decentralized Exchanges, Markets and Market Making (EIN06DF01) such as Uniswap, five tokens in the space of Decentralized Lending, Saving and Asset Management (EIN06DF02), e.g., AAVE. Two tokens are classified as Decentralized Derivatives, Synthetic Assets and Insurance (EIN06DF03), three tokens as Decentralized Data, Oracles and Infrastructure (EIN06DF04), and one token as Other Decentralized Finance (DeFi) (EIN06DF05). Regarding the Technological Setup (TTS), this group mainly consists of tokens implemented on top of the Ethereum blockchain. Looking into the correlation coefficients between the single token pairs within this group, we only find five strong positive linear relationships between the following pairs: Link and 0x, Link and Band Protocol, Aave and Uniswap, SNX and Aave and 0x and Ravencoin. All other relationships between the single token pairs are mainly moderate positive or sometimes weak positive. The token CEL, which powers Celsius's asset management platform, shows mainly weak positive linear relationships with the other tokens. CEL has eight correlation coefficients within the group are below 0.300, and the highest being 0.513. A detailed listing of the correlation coefficients is provided in table 5.

Regarding the correlation between the groups, eight out of 17 DeFi ecosystem tokens show the highest correlation with their own group compared to the other portfolio groups. In contrast, eight show the highest correlation with the portfolio of group 3 network utility tokens. Only Maker shows the highest correlation with group 1 payment tokens portfolio. While six DeFi ecosystem tokens show a strong, eleven show a moderate positive linear relationship with the own portfolio. We find mostly moderate positive linear relationships between the single DeFi ecosystem tokens and the payment token portfolio. Only the token of Ravencoin shows a strong positive relationship, while the tokens of SushiSwap and AAVE show only weak positive linear relationships with the payment token portfolio. We find six strong and eleven moderate positive linear relationships looking into the correlations between the DeFi ecosystem tokens and the network utility token portfolio. Again, we only see positive linear relationships between all single token pairs and the correlations

Table 2: Token Groups according to the Classification

Token Group	Name	Ticker
Group 1: Payment Token	Bitcoin	BTC
	Bitcoin Cash	BCH
	Litecoin	LTC
	Bitcoin SV	BSV
	Monero	XMR
	Dash	DASH
	Zcash	ZEC
	Dogecoin	DOGE
	Decred	DCR
	Bitcoin Gold	BTG
	Nano	NANO
	Verge	XVG
	Bitcoin Diamond	BCD
Group 2: DeFi Ecosystem Token	Chainlink	LINK
	Uniswap	UNI
	Aave	AAVE
	yearn.finance	YFI
	Celsius	CEL
	Maker	MKR
	Synthetix Network Token	SNX
	UMA	UMA
	Compound	COMP
	0x	ZRX
	Loopring	LRC
	SushiSwap	SUSHI
	Kyber Network	KNC
	Augur	REP
	THORChain	RUNE
	Band Protocol	BAND
	Ravencoin	RVN
Group 3: Network Utility Tokens	Ethereum	ETH
	Cardano	ADA
	Polkadot	DOT
	EOS	EOS
	TRON	TRX
	Tezos	XTZ
	NEM	XEM
	NEO	NEO
	Cosmos	ATOM
	Ethereum Classic	ETC
	Waves	WAVES
	Kusama	KSM
	Algorand	ALGO
	DigiByte	DGB
	Zilliqa	ZIL
	Ren	REN
	Qtum	QTUM
	ICON	ICX
	Hedera Hashgraph	HBAR
	NEAR Protocol	NEAR
	Lisk	LSK
	Blockstack	STX
	Horizen	ZEN

Source: Own illustration.

Table 3: Return Correlation between Payment Tokens

Platform	BTC	BCH	LTC	BSV	XMR	DASH	ZEC	DOGE	DCR	BTG	NANO	XVG	BCD
BTC	1.000												
BCH	0.535	1.000											
LTC	0.745	0.681	1.000										
BSV	0.373	0.851	0.475	1.000									
XMR	0.474	0.510	0.462	0.323	1.000								
DASH	0.414	0.647	0.493	0.458	0.761	1.000							
ZEC	0.453	0.707	0.575	0.553	0.713	0.842	1.000						
DOGE	0.428	0.352	0.371	0.251	0.220	0.205	0.249	1.000					
DCR	0.569	0.327	0.491	0.222	0.247	0.215	0.313	0.326	1.000				
BTG	0.539	0.743	0.539	0.679	0.505	0.650	0.671	0.222	0.323	1.000			
NANO	0.393	0.378	0.407	0.178	0.170	0.289	0.413	0.095	0.304	0.389	1.000		
XVG	0.486	0.540	0.463	0.380	0.310	0.524	0.527	0.261	0.380	0.441	0.710	1.000	
BCD	0.247	0.705	0.303	0.757	0.483	0.605	0.562	0.196	0.114	0.721	0.071	0.280	1.000

Source: Own illustration.

Table 4: Correlation between single Payment Token Returns and Group Portfolio Returns

Token	Payment Token Portfolio	DeFi Ecosystem Token Portfolio	Network Utility Token Portfolio
Bitcoin	0.6725***	0.5790***	0.5054***
Bitcoin Cash	0.8391***	0.7578***	0.6515***
Litecoin	0.7025***	0.7027***	0.6674***
Bitcoin SV	0.6335***	0.5822***	0.5274***
Monero	0.5763***	0.5238***	0.4214***
Dash	0,7042***	0.5822***	0.4294***
Zcash	0.7786***	0.7272***	0.6110***
Dogecoin	0.3449***	0.3072**	0.2380*
Decred	0.4461***	0.4826***	0.4258***
Bitcoin Gold	0.7596***	0.6334***	0.5453***
Nano	0.4517***	0.4437***	0.3946***
Verge	0.6697***	0.5985***	0.4658***
Bitcoin Diamond	0.5561***	0,4565***	0.4209***

Source: Own illustration. Note: Pearson correlation coefficients with significance levels: * $0.01 < p \leq 0.05$; ** $0.001 < p \leq 0.01$; *** $p \leq 0.001$.

between the DeFi ecosystem tokens and the portfolios. The detailed correlation coefficients can be found in table 6.

Group 3: Network Utility Tokens: Our last group covers the largest set of tokens. A total number of 23 tokens have been assigned to the network utility token group, including tokens like Ethereum, Cardano, and Polkadot. Throughout this group, we mostly find correlation coefficients that indicate a moderate positive linear relationship. However, two tokens show almost only weak positive linear relationships throughout the whole group, Blockstack, and Hedera Hashgraph. Also, it is notable that the correlation coefficient between Blockstack and Hedera Hashgraph is the only one that is negative and shows a weak negative linear relationship between the two tokens. The tokens with the most correlation coefficients indicating a strong positive linear relationship are Lisk and Tezos, with nine correlation coefficients over 0.700. In table 7, we can find the detailed correlation coefficients.

Looking into the correlation between the single network utility tokens and the different portfolio returns in table 8, we observe that 16 out of the 23 tokens correlate the most with their own portfolio. Thirteen of the tokens show a strong while 10 show a moderate positive linear relationship. The correlations between the single tokens with the payment and the DeFi ecosystem token portfolio show mainly moderate positive linear relationships. For the native asset of the Ethereum blockchain ETH, we discover that it shows the

Table 5: Return Correlation between DeFi Ecosystem Token

Platform	LINK	UNI	AAVE	YFI	CEL	MKR	SNX	UMA	COMP	ZRX	LRC	SUSHI	KNC	REP	RUNE	BAND	RVN
LINK	1.000																
UNI	0.527	1.000															
AAVE	0.484	0.732	1.000														
YFI	0.377	0.536	0.655	1.000													
CEL	0.361	0.390	0.243	0.288	1.000												
MKR	0.600	0.417	0.308	0.297	0.275	1.000											
SNX	0.465	0.578	0.751	0.458	0.352	0.436	1.000										
UMA	0.408	0.382	0.259	0.352	0.297	0.262	0.270	1.000									
COMP	0.507	0.613	0.556	0.339	0.375	0.502	0.506	0.350	1.000								
ZRX	0.802	0.584	0.521	0.433	0.344	0.638	0.551	0.445	0.581	1.000							
LRC	0.447	0.418	0.495	0.364	0.224	0.316	0.531	0.324	0.398	0.574	1.000						
SUSHI	0.435	0.637	0.616	0.515	0.237	0.266	0.480	0.280	0.399	0.472	0.275	1.000					
KNC	0.664	0.478	0.387	0.328	0.221	0.524	0.379	0.419	0.502	0.625	0.306	0.365	1.000				
REP	0.564	0.392	0.345	0.441	0.379	0.521	0.422	0.484	0.432	0.645	0.338	0.315	0.517	1.000			
RUNE	0.614	0.617	0.669	0.483	0.513	0.428	0.609	0.432	0.571	0.670	0.531	0.582	0.450	0.482	1.000		
BAND	0.717	0.494	0.498	0.479	0.280	0.379	0.444	0.404	0.507	0.626	0.413	0.445	0.685	0.476	0.534	1.000	
RVN	0.667	0.475	0.361	0.390	0.406	0.569	0.452	0.439	0.568	0.790	0.466	0.365	0.655	0.668	0.568	0.612	1.000

Source: Own illustration.

Table 6: Correlation between single DeFi Ecosystem Token Returns and Group Portfolio Returns

Token	Payment Token Portfolio	DeFi Ecosystem Token Portfolio	Network Utility Token Portfolio
Chainlink	0.6364***	0.7544***	0.7795***
Uniswap	0.4283***	0.7506***	0.6057***
Aave	0.2843**	0.7337***	0.5049***
yearn.finance	0.4514***	0.6055***	0.4947***
Celsius	0.4269***	0.4443***	0.4580***
Maker	0.6842***	0.5688***	0.5773***
Synthetix Network Token	0.3945***	0.6986***	0.5324***
UMA	0.4695***	0.4964***	0.5299***
Compound	0.5443***	0.6781***	0.6077***
0x	0.6737***	0.8195***	0.8305***
Loopring	0.3627***	0.5625***	0.5057***
SushiSwap	0.2621*	0.6023***	0.4418***
Kyber Network	0.5820***	0.6431***	0.7513***
Augur	0.6782***	0.6319***	0.7413***
THORChain	0.4641***	0.7875***	0.6476***
Band Protocol	0.5393***	0.6975***	0.7386***
Ravencoin	0.7324***	0.7300***	0.7929***

Source: Own illustration. Note: Pearson correlation coefficients with significance levels: * $0.01 < p \leq 0.05$; ** $0.001 < p \leq 0.01$; *** $p \leq 0.001$

highest correlation with our DeFi ecosystem group portfolio.

Looking at the results of the different groups, we also find more general results. First, it is worth mentioning that we only see positive linear relationships between the single token pairs besides one exception. Another finding was that network utility tokens that are indicating a strong positive linear relationship between the token return and the portfolio return of its own group often also show a high correlation coefficient with the other group portfolio returns, e.g., Ethereum, EOS, and Lisk. This also holds true for the other two groups, DeFi ecosystem token group, e.g., Link, and the payment token group, e.g., Ravencoin, Bitcoin Cash.

4.3. Discussion

During the classification efforts of the top 200 tokens according to market capitalization, we have mainly seen tokens classified as Utility Token (EEP22) and Payment Token (EEP21) regarding the Economic Purpose (EEP) dimension. An explanation for this could be that tokens that fall under the category Payment Token (EEP21) were the first tokens issued, followed by the tremendous rise of utility token initial coin offerings in 2017 and early 2018 (Howell & Niessner, 2020). Regarding the Issuer Industry (EIN) dimension, we found mainly token issuers in the industries of Information, Communication and IT (EIN05) and Finance and Insurance (EIN06). As the concept of tokenization is still very new and we looked at the largest projects in terms of market capitalization, we expect to find more projects from different industries in the future while looking at projects with lower market capitalization. As a lot of the tokens in our list are the native assets to run their blockchain, we found many Blockchain-Native Token (TTS41BC). The second dominating group was the widespread Ethereum ERC-20 Standard Token (TTS42ET01), which is not only the most spread token standard of the most prominent blockchain but also most of the tokens from our large DeFi ecosystem token group are implemented as an Ethereum ERC-20 Standard Token (TTS42ET01). During the screening of the token market, we also made an interesting finding regarding the tokens that are active in the DeFi space. The project behind the token describes itself often decentralized many of the tokens are classified as Private Sector Legal Entity (LIT61PV) regarding the Issuer Type (LIT). For example, the token Link, Uniswap, or SushiSwap are issued by a registered company and are therefore not entirely decentralized.

The group of payment tokens showed the highest correlations between the single tokens and between the tokens and the group portfolio returns. As previously described, we have found almost only positive linear relationships. A potential reason for that could be the rise of the total market during the time period we used. Another interesting finding was that tokens like Dogecoin, Hedera Hashgraph, or Blockstack that show mainly weak positive linear relationships within the group also show weak positive linear relationships with the three portfolio groups. We also found that the density of positive linear relationships was higher within the payment token group than in the other two groups.

Table 7: Return Correlation between Network Utility Tokens

Platform	ETH	ADA	DOT	EOS	TRX	XTZ	XEM	NEO	ATOM	ETC	WAVES	KSM	ALGO	DGB	ZIL	REN	QTUM	ICX	HBAR	NEAR	LSK	STX	ZEN
ETH	1.000																						
ADA	0.720	1.000																					
DOT	0.505	0.524	1.000																				
EOS	0.660	0.721	0.424	1.000																			
TRX	0.679	0.631	0.384	0.880	1.000																		
XTZ	0.682	0.703	0.492	0.776	0.782	1.000																	
XEM	0.404	0.479	0.223	0.556	0.532	0.515	1.000																
NEO	0.644	0.604	0.322	0.701	0.751	0.802	0.537	1.000															
ATOM	0.507	0.604	0.599	0.617	0.597	0.765	0.424	0.616	1.000														
ETC	0.720	0.615	0.420	0.741	0.818	0.775	0.480	0.787	0.563	1.000													
WAVES	0.455	0.525	0.296	0.445	0.424	0.513	0.271	0.493	0.417	0.462	1.000												
KSM	0.442	0.496	0.728	0.418	0.439	0.482	0.297	0.427	0.570	0.430	0.328	1.000											
ALGO	0.558	0.647	0.433	0.681	0.623	0.764	0.543	0.605	0.586	0.597	0.459	0.498	1.000										
DGB	0.578	0.445	0.325	0.584	0.603	0.516	0.401	0.574	0.390	0.596	0.270	0.328	0.513	1.000									
ZIL	0.327	0.406	0.355	0.329	0.340	0.386	0.374	0.308	0.422	0.327	0.227	0.332	0.434	0.522	1.000								
REN	0.654	0.649	0.561	0.612	0.577	0.675	0.435	0.527	0.608	0.556	0.474	0.476	0.682	0.401	0.397	1.000							
QTUM	0.570	0.620	0.445	0.760	0.741	0.810	0.491	0.766	0.595	0.736	0.460	0.472	0.695	0.513	0.342	0.544	1.000						
ICX	0.604	0.721	0.442	0.706	0.673	0.641	0.507	0.634	0.613	0.600	0.365	0.459	0.669	0.474	0.482	0.578	0.631	1.000					
HBAR	0.224	0.294	0.258	0.218	0.173	0.222	0.143	0.164	0.134	0.188	0.134	0.256	0.297	0.184	0.132	0.385	0.239	0.447	1.000				
NEAR	0.431	0.455	0.457	0.554	0.509	0.554	0.297	0.437	0.542	0.448	0.300	0.488	0.571	0.394	0.281	0.551	0.487	0.461	0.422	1.000			
LSK	0.661	0.711	0.479	0.793	0.821	0.764	0.614	0.742	0.612	0.780	0.509	0.542	0.702	0.638	0.423	0.653	0.752	0.719	0.337	0.540	1.000		
STX	0.256	0.125	0.067	0.223	0.189	0.221	0.168	0.218	0.135	0.141	0.105	0.194	0.275	0.232	0.024	0.248	0.127	0.181	-0.083	0.116	0.171	1.000	
ZEN	0.321	0.415	0.148	0.432	0.500	0.511	0.405	0.585	0.308	0.507	0.267	0.151	0.387	0.353	0.233	0.276	0.429	0.493	0.152	0.351	0.457	0.160	1.000

Source: Own illustration.

Table 8: Correlation between single Network Utility Token Returns and Group Portfolio Returns

Token	Payment Token Portfolio	DeFi Ecosystem Token Portfolio	Network Utility Token Portfolio
Ethereum	0.7326***	0.7555***	0.7444***
Cardano	0.6569***	0.6985***	0.7846***
Polkadot	0.4210***	0.5058***	0.5753***
EOS	0.7899***	0.7165***	0.8228***
TRON	0.8418***	0.6786***	0.8105***
Tezos	0.7143***	0.7933***	0.8652***
NEM	0.4914***	0.4729***	0.5768***
NEO	0.8031***	0.6897***	0.7849***
Cosmos	0.4961***	0.6703***	0.7268***
Ethereum Classic	0.8278***	0.6906***	0.7849***
Waves	0.4591***	0.5697***	0.5188***
Kusama	0.4507***	0.5569***	0.6030***
Algorand	0.5471***	0.7920***	0.7962***
DigiByte	0.6854***	0.5240***	0.6269***
Zilliqa	0.3480***	0.4109***	0.4724***
Ren	0.5478***	0.7412***	0.7509***
Qtum	0.7072***	0.6720***	0.7828***
ICON	0.6018***	0.6641***	0.7916***
Hedera	0.1852**	0.2750**	0.3175
NEAR Protocol	0.4748***	0.6221***	0.6298***
Lisk	0.8271***	0.7585***	0.8704***
Blockstack	0.2102*	0.2748**	0.2127*
Horizen	0.4612***	0.3593***	0.4894***

Source: Own illustration. Note: Pearson correlation coefficients with significance levels: * $0.01 < p \leq 0.05$; ** $0.001 < p \leq 0.01$; *** $p \leq 0.001$

5. Conclusion

The last chapter of the thesis summarizes the classification work, the results on the investigated tokens price correlations, highlights the theoretical and practical implications, points out the limitations, and discusses future research opportunities for this topic.

5.1. Summary

This paper intends to provide a large sample set of classified tokens according to the ITC and how dependencies in the market can be identified. We created a dataset containing the top 200 tokens according to market capitalization using the introduced additional classification guidelines. Looking into the top 200 tokens according to market capitalization, we mainly find Utility Tokens (EEP22). Secondly, we find that most of the token issuers are from the industries of Information, Communication and IT (EIN05) and Finance and Insurance (EIN06). We created token groups only according to the obtained classification data and used a simple approach to create our own group portfolio returns. Using the classification dataset and the obtained groups, we found exciting dependencies between tokens while taking the classification of the token into account. We computed the correlation coefficients between the single token pairs within those groups and between the groups using the group portfolio returns. To some extent, we found that the token's classification indeed can explain some of the correlation between tokens, as the tokens show very high correlation within the groups and with the introduced average of their own portfolio.

5.2. Theoretical and Practical Implications

This paper expands the current state of research regarding the classification of a token and introduces a practical use case for the classification framework. We have reviewed different classification approaches and highlighted that the ITC covers most of the aspects that the other classification approaches are describing. The introduced additional guidelines can help further research to classify tokens according to the ITC and should be considered. The created dataset can be used by further research to investigate the token market more in detail. Also, literature can use the sample set as an indicator of how tokens should be classified. Regulators or governments can use the approach of ITSA and the dataset to differentiate between the tokens while introducing laws for the token market. Investors can use the dataset and the correlation computations results to construct new analysis and investment strategies while knowing which assets tend to move in the same direction. Due to the wide range of information the dataset provides, people interested in a specific field such as DeFi space can now use the information to

gather information, whether the project behind the token is entirely decentralized or if there is a hidden company behind the project. As the token market is very new to most traditional investors, they can use the dataset to overview the current token landscape and find their potential investment target. Every person interested in the token economy can now search for a specific kind of token and compare it to a completely different one. For example, using the dataset, it is now possible to compare an Utility Token (EEP22), issued by a Private Sector Legal (LIT61PV) active in Decentralized Finance (EIN06DF) and implemented as an Ethereum ERC-20 Standard Token (TTS42ET01) to the native asset of the Ethereum blockchain ETH.

5.3. Limitations and Further Research

Due to the limited length of the paper, some aspects are not covered in more detail, such as the results of the three more applied dimensions or the correlation between single tokens and market-dominating tokens like Bitcoin or Ethereum. As this paper is one of the first scientific papers using the ITC, we covered fundamental statistical analysis instead of using more advanced methods. Furthermore, it is essential to note that we only have used 90 days of price data to calculate the correlation coefficients. As the market is currently volatile, it is not predictable what the correlation will be in a year, a month, or even a week while looking only at a short-term data set. Even during this paper's working time, some tokens dropped out of the top 200 tokens according to market capitalization, while many new tokens entered. The positive correlation we have observed in this paper may also have occurred by chance but should be analyzed again after the market provides consistent data for a longer time frame. A second limitation that should be mentioned is the equally weighted portfolio we have applied to compare single assets with a whole group. This was done to avoid a portfolio constructed, for example, by market capitalization, which would be entirely dominated by single assets such as Bitcoin for our first Group payment token or Ethereum for the second group Network Utility Token. However, a more advanced strategy to construct a real index for a group should be considered and applied. The third limitation we want to mention is the current bull market in which we currently are. The analysis should be repeated in a somehow different market, such as a bear or a stable market. Therefore, more research is necessary to evaluate whether the classification impacts the price behavior of a token. As we have seen in this paper's results, ETH shows the strongest linear relationship with the DeFi ecosystem token group portfolio. At the same time, we have discovered that the DeFi ecosystem token group portfolio mainly consists of Ethereum based tokens. Therefore, an exciting research topic could be to see whether tokens implemented on top of a blockchain tend to move together with the native asset.

References

Benartzi, S., & Thaler, R. H. (2001). Naive Diversification Strategies in Defined Contribution Saving Plans. *American Economic Review*, *91*(1), 79–98.

Binance Research. (2019a). *2019 Q2 Crypto-Correlations Review.* Retrieved from https://research.binance.com/en/analysis/correlations-q2-2019

Binance Research. (2019b). *2019 Q3 Crypto-Correlations Review.* Retrieved from https://research.binance.com/en/analysis/correlations-q3-2019

Binance Research. (2019c). *Are Cryptoassets Highly Correlated?* Retrieved from https://research.binance.com/en/analysis/correlations-q1-2019

Binance Research. (2020). *2019—Annual Crypto-Correlations Review.* Retrieved from https://research.binance.com/en/analysis/annual-crypto-correlations-2019

Birch, J., Pantelous, A. A., & Soramäki, K. (2016). Analysis of Correlation Based Networks Representing DAX 30 Stock Price Returns. *Computational Economics*, *47*(4), 501–525.

Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: Economics, Technology, and Governance. *Journal of Economic Perspectives*, *29*(2), 213–238.

Casino, F., Dasaklis, T. K., & Patsakis, C. (2019). A systematic literature review of blockchain-based applications: Current status, classification and open issues. *Telematics and Informatics*, *36*, 55–81.

Chen, Y. (2018). Blockchain tokens and the potential democratization of entrepreneurship and innovation. *Business Horizons*, *61*(4), 567–575.

Christidis, K., & Devetsikiotis, M. (2016). Blockchains and Smart Contracts for the Internet of Things. *IEEE Access*, *4*, 2292–2303.

Coinbase. (2021). *Bitcoin Preisdiagramm.* Retrieved from https://www.coinbase.com/price/bitcoin

CoinGecko. (2021). *Preise und Marktkapitalisierung in Krypto-Währung.* Retrieved from https://www.coingecko.com/de

CoinMarketCap. (2021a). *Bitcoin price today, BTC marketcap, chart, and info.* Retrieved from https://coinmarketcap.com/currencies/bitcoin/

CoinMarketCap. (2021b). *Cryptocurrency Prices, Charts And Market Capitalizations.* Retrieved from https://coinmarketcap.com/

Cong, L. W., & He, Z. (2019). Blockchain Disruption and Smart Contracts. *The Review of Financial Studies*, *32*(5), 1754–1797.

Deloitte. (2019). *Are token assets the securities of tomorrow?* Retrieved from https://www2.deloitte.com/content/dam/Deloitte/lu/Documents/technology/lu-token-assets-securities-tomorrow.pdf

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies*, *22*(5), 1915–1953.

Drescher, D. (2017). *Blockchain Basics—A Non-Technical Introduction in 25 Steps.* Retrieved from https://www.springer.com/de/book/9781484226032

Ekblaw, A., Azaria, A., Halamka, J., & Lippman, A. (2016). *A Case Study for Blockchain in Healthcare: "MedRec" prototype for electronic health records and medical research data.* Retrieved from http://www.truevaluemetrics.org/DBpdfs/Technology/Blockchain/5-onc_blockchainchallenge_mitwhitepaper_copyrightupdated.pdf

Euler, T. (2018). *The Token Classification Framework: A multi-dimensional tool for understanding and classifying crypto tokens.* Retrieved from http://www.untitled-inc.com/the-token-classification-framework-a-multi-dimensional-tool-for-understanding-and-classifying-crypto-tokens/

European Commission, Directorate-General for Financial Stability, & Financial Services and Capital Markets Union. (2020). *Proposal for a Regulation of the European Parliament and of the Council on Markets in Crypto-assets, and amending Directive (EU) 2019/1937.* Retrieved from https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52020PC0593

FINMA. (2018). *Guidelines for enquiries regarding the regulatory framework for initial coin offerings (ICOs).* Retrieved from https://www.finma.ch/en/news/2018/02/20180216-mm-ico-wegleitung

Fußwinkel, O., & Kreiterling, C. (2018). Blockchain-Technologie—Gedanken zur Regulierung. *BaFin Perspektiven*(1), 48–67.

Gopikrishnan, P., Rosenow, B., Plerou, V., & Stanley, H. E. (2000). *Identifying Business Sectors from Stock Price Fluctuations.* Retrieved from http://arxiv.org/abs/cond-mat/0011145

Haber, S., & Stornetta, W. S. (1991). How to Time-Stamp a Digital Document. *Journal of Cryptology*, *3*(2), 99–111.

Howell, S. T., & Niessner, D., M.and Yermack. (2020). Initial Coin Offerings: Financing Growth with Cryptocurrency Token Sales. *The Review of Financial Studies*, *33*(9), 3925–3974.

Hughes, A., Park, A., Kietzmann, J., & Archer-Brown, C. (2019). Beyond Bitcoin: What blockchain and distributed ledger technologies mean for firms. *Business Horizons*, *62*(3), 273–281.

ITSA. (2020). *International Token Classification (ITC)—Framework Documentation (1.0).*

Klages-Mundt, A., & Minca, A. (2020). *(In)Stability for the Blockchain: Deleveraging Spirals and Stablecoin Attacks.* Retrieved from http://arxiv.org/abs/1906.02152

Klein, T., Pham Thu, H., & Walther, T. (2018). Bitcoin is not the New Gold – A comparison of volatility, correlation, and portfolio performance. *International Review of Financial Analysis*(59), 105–116.

Knight, W. (2017). The Technology Behind Bitcoin Is Shaking Up Much More Than Money. *MIT Technology Review*.

Massey, R., Dalal, D., & Dakshinamoorthy, A. (2017). Initial coin offering: A new paradigm. *Deloitte US*. Retrieved from https://www2.deloitte.com/us/en/pages/consulting/articles/initial-coin-offering-a-new-paradigm.html

McCall, M. (2020). A $200 Million Pizza! Here's How Bitcoin Made That Possible ... *Nasdaq*. Retrieved from https://www.nasdaq.com/articles/a-$200-million-pizza-heres-how-bitcoin-made-that-possible-...-2020-12-02

Meucci, A. (2010). Quant Nugget 2: Linear vs. Compounded Returns – Common Pitfalls in Portfolio Management. *SSRN Scholarly Paper ID 1586656*. Retrieved from https://papers.ssrn.com/abstract=1586656

Moore, G. (2020). 10 Years After Laszlo Hanyecz Bought Pizza With 10K Bitcoin, He Has No Regre. *Yahoo! Finance*. Retrieved from https://finance.yahoo.com/news/10-years-laszlo-hanyecz-bought-143000184.html

Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System.* Retrieved from https://bitcoin.org/bitcoin.pdf

Orac, R. (2017). Cryptocurrency Analysis with Python—Buy and Hold. *Roman Orac Blog*. Retrieved from https://romanorac.github.io/cryptocurrency/analysis/2017/12/25/cryptocurrency-analysis-with-python-part2.html

Perboli, G., Musso, S., & Rosano, M. (2018). Blockchain in Logistics and Supply Chain: A Lean Approach for Designing Real-World Use Cases. *IEEE Access*, *6*, 62018–62028.

Ratner, B. (2009). The correlation coefficient: Its values range between +1/-1, or do they? *Journal of Targeting, Measurement and Analysis for Marketing*, *17*(2), 139–142.

Roth, J., Schär, F., & Schöpfer, A. (2019).). The Tokenization of Assets: Using Blockchains for Equity Crowdfunding. *SSRN Scholarly Paper ID 3443382*.

Sandner, P., & Blassl, J. (2020). Crypto-Europe: Comprehensive European Regulation For Crypto Assets Has Been Presented. *Forbes*. Retrieved from https://www.forbes.com/sites/philippsandner/2020/09/21/crypto-europe-comprehensive-european-regulation-for-crypto-assets-has-been-presented/

Sandner, P., & Ketz, C. (2019). Market standards for the global token economy. *Medium*. Retrieved from https://medium.com/@itsa_global/market-standards-for-the-global-token-economy-7d9f3d0cde37

Sandner, P., Ketz, C., Tumasjan, A., & Lentge, K. (2019). *International Token Classification (ITC) - A Taxonomy for Classifying Cryptographic Tokens.* Retrieved from https://itsa.global/what-we-do/#ITC

Sazandrishvili, G. (2020). Asset tokenization in plain English. *Journal of Corporate Accounting & Finance*, *31*(2), 68–73.

Schär, F. (2020). Decentralized Finance: On Blockchain- and Smart Contract-based Financial Markets. *SSRN Scholarly Paper ID 3571335*.

Stošić, D., Stosic, D., & Ludermir, T. (2018). Collective behavior of cryptocurrency price changes. *Physica A: Statistical Mechanics and Its Applications*, *507*, 499–509.

The Pandas Development Team. (2020). pandas-dev/pandas: Pandas 1.0.3. *Zenodo*.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., … van Mulbregt, P. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272.

Vogelsteller, F., & Buterin, V. (2015). ERC-20 Token Standard. *GitHub*. Retrieved from https://github.com/ethereum/EIPs

Wang, G. J., Xie, C., & Stanley, H. E. (2018). Correlation Structure and Evolution of World Stock Markets: Evidence from Pearson and Partial Correlation-Based Networks. *Computational Economics*, *51*(3), 607–635.

Wang, S., & Vergne, J.-P. (2017). Buzz Factor or Innovation Potential: What Explains Cryptocurrencies' Returns? . *PLOS ONE*, *12*(1), e0169556.

yahoo! finance. (2021). *Papa John's International, Inc. (PZZA) Stock Price, News, Quote & History—Yahoo Finance.* Retrieved from https://finance.yahoo.com/quote/PZZA/