JUMS

JUNIOR MANAGEMENT SCIENCE

Junior Management Science

Convincing Investors: A Study of Personal, Adapted Storytelling and Strategic Behavior in Entrepreneurial Fundraising

Pauline Tilla Rittstieg

Technische Universität München

Abstract

This inductive study explores the process, through which legitimacy is established for financial resource acquisition, by analyzing the accounts of 15 entrepreneurs on their storytelling and fundraising strategies. The findings show that consistent personal storytelling, venture story adaptation, and strategic behavior increase a venture's chances of receiving financial investments. Taking an entrepreneur-centric perspective in analyzing the practical implementation of cultural entrepreneurship theory, the findings have strong theoretical implications. They suggest extending the model of cultural entrepreneurship to include entrepreneurs' behavior throughout the resource acquisition process. They further contradict the theory that a venture's legitimacy only depends on its existing resources, suggesting that it is also based on the venture's founder's storytelling skills and behavior. As for practical implications, the findings show that investment decisions are not purely fact-based but influenced by investors' emotional involvement and the hype around the venture among investors, which are both a consequence of skillful storytelling and strategic behavior. The study illustrates several effective storytelling and fundraising strategies, providing practical examples for each.

Keywords: Venture legitimacy; Fundraising; Fundraising strategy; Storytelling; Cultural entrepreneurship.

1. Introduction

Storytelling and resource acquisition are two very prominent topics in entrepreneurship literature (Martens, Jennings, & Jennings, 2007). Being skilled in the former is today indisputably considered a precondition for mastering the latter. A key concept explaining the interrelation of storytelling and resource acquisition is cultural entrepreneurship, which refers to "the processes by which actors draw upon cultural resources (e.g., discourse, language, categories, logics, and other symbolic elements) to advance entrepreneurship or to facilitate organizational or institutional innovation" (Lounsbury & Glynn, 2019, p. 3). The concept of cultural entrepreneurship draws together theories on the effect of legitimacy, narrative, and identity in the resource acquisition process and can be used as a framework for further research within the field.

The relevant findings on the connection between resource acquisition and entrepreneurial storytelling can be summarized in a simple chain of reasoning: 1) The acquisition of external resources helps ventures to emerge, survive and become sustainable (Fisher, Kuratko, Bloodgood, & Hornsby, 2017; Lounsbury & Glynn, 2001); 2) A venture must be perceived as legitimate to attract resources (Aldrich & Fiol, 1994; Lounsbury, Gehman, & Glynn, 2019; Lounsbury & Glynn, 2001; Zimmerman & Zeitz, 2002); 3) The entrepreneur can establish perceived legitimacy for a venture with the help of audience- and context-adapted storytelling, also known as cultural entrepreneurship (Fisher et al., 2017; Lounsbury et al., 2019; Lounsbury & Glynn, 2001; Wry, Lounsbury, & Glynn, 2011).

How a story is told and perceived depends on the narratives of the entrepreneur and the audience, which act as individual frameworks for reasoning and understanding (Bartel & Garud, 2009; Garud & Giuliani, 2013; Martens et al., 2007). The information asymmetry, perceived risk, and uncertainty that is caused by differing narratives can be overcome by skillful storytelling, which conveys a comprehensible venture identity and establishes legitimacy in a process of meaning making (Garud, Schildt, & Lant, 2014; Manning & Bejarano, 2017; Martens et al., 2007). Throughout the process of storytelling, not only the venture identity becomes apparent, but also the identity of the entrepreneur is revealed. Being created and developed from the vantage point of the entrepreneur, the identities of the venture and entrepreneur

have shown to be closely connected and can often be described by the same attributes (Grimes, 2017; Lounsbury & Glynn, 2001).

Telling a story, which conveys a comprehensible venture identity and bridges narrative-based information asymmetries to establish distinctive legitimacy is a difficult, yet central task to master for entrepreneurs (Werven, Bouwmeester, & Cornelissen, 2015). In the existing literature on cultural entrepreneurship, legitimacy is predominantly presented as an outcome of storytelling, but the strategies through which the perception of legitimacy can be influenced remain mostly unexplored. Researchers who contributed to exploring this topic have mostly taken a linguistic perspective. By analyzing the wording of successful fundraising pitches, they discovered that rhetoric can be strategically used to establish legitimacy (Ruebottom, 2013) by integrating arguments (Werven et al., 2015) or analogies (Cornelissen & Clarke, 2010), which contextualize a venture within a familiar discourse (Navis & Glynn, 2011) or market category (King and Whetten, 2008).

In line with Steyaert (2007), I argue that the existing literature pulls the attention away from the entrepreneur, although he is at the center of the audience's attention. Addressing this gap, I analyze the accounts of entrepreneurs on their fundraising strategies instead of their fundraising pitches, thereby focusing on the entrepreneur's role in storytelling and resource acquisition. Moving away from the linguistic perspective, I seek to take a more structural perspective by identifying those parts of the fundraising process, in which legitimacy can be established by the entrepreneur. Overall, this study is one of the first to provide qualitative, entrepreneur-centric insights into the practical implementation of cultural entrepreneurship and the process of establishing legitimacy.

With this study, I set out to answer the following research question: *How do entrepreneurs approach the fundraising process through personal, adapted storytelling and strategic behavior to legitimize themselves and their ventures?* To do so, I conducted semi-structured interviews with 15 German entrepreneurs, who have successfully gone through the process of financial resource acquisition at least once and can offer very recent insights on the topic. The collected qualitative data was then thoroughly analyzed and propositions and a process model developed from the findings. Besides its theoretical contributions, this study intends to provide practical implications for entrepreneurs, who are looking for advice for their own fundraising process. This study focuses on the acquisition of financial resources from private or institutional investors (referred to as 'fundraising'), as this is something especially first-time founders of early-stage ventures often struggle with.

The results of this study show that entrepreneurs use three core strategies to establish legitimacy: consistent personal storytelling, venture story adaptation, and strategic behavior. The personal story makes meaning of the entrepreneur's background and commitment to the venture, establishing distinctive legitimacy and involving the investor

emotionally. The personal story remains consistent over time, while the venture story is adapted according to the audience's interests, expectations, and agenda. It changes over time, based on the venture stage or pivots, and legitimizes the venture's vision and valuation. Strategic behavior concerns the mindset the entrepreneur adopts during the fundraising process and how he approaches investor outreach and information sharing. Overall, this study demonstrates that investment decisions are not purely rational or fact-based. Instead, they are affected by investors' emotional involvement and the hype that is created around a venture in the investor community. Both of these factors are based on entrepreneurs' strategic storytelling and behavior and can significantly increase the chances for the successful acquisition of financial resources.

The results of this study advance the literature on cultural entrepreneurship in the following ways. Observing the practical implementation of the process model of cultural entrepreneurship by Lounsbury and Glynn (2019), they first confirm the effect of storytelling to establish legitimacy by bridging information symmetries between entrepreneurs and investors. Second, they offer a temporal perspective on the development of fundraising stories, differentiating between the entrepreneur's personal story and venture story. Third, they propose the extension of the process model of cultural entrepreneurship to include behavior. Although Martens et al. (2007) and Lounsbury and Glynn (2019) acknowledge that entrepreneurial behavior can have a significant impact on the investor's judgment about a venture's potential, "we know little about what specific entrepreneurs' behavior increases the propensity for this type of resource acquisition" (Pollack, Rutherford, & Nagy, 2012). This study confirms this notion, illustrating examples of strategic behavior before, during, and after the fundraising pitch, which can fundamentally affect perceived venture legitimacy. Fourth, this study challenges the idea that legitimacy is merely based on the venture's existing resources (e.g., Lounsbury & Glynn, 2001, 2019). Demonstrating the dependency of perceived legitimacy on the entrepreneur's storytelling skills and fundraising strategy, this study promotes an entrepreneur-centric perspective in cultural entrepreneurship theory.

The remainder of this paper is structured as follows: the next chapters provide an overview of the discourse on resource acquisition in the context of cultural entrepreneurship literature and the methods applied to answer the research question. Then, the final results are presented, including derived propositions and the process model. Finally, the empirical and practical implications are discussed, and conclusions are drawn.

2. Theoretical background

2.1. Cultural entrepreneurship perspective

A key concept, which was developed to explore entrepreneurial storytelling and its effect on resource acquisition is cultural entrepreneurship. Because it provides

explanations for connections between several other independent organizational and social theories, it provides a useful framework to review the relevant literature and the theoretical context for this thesis. Cultural entrepreneurship is defined as the "process of storytelling that mediates between extant stocks of entrepreneurial resources and subsequent capital acquisition and wealth creation" (Lounsbury & Glynn, 2001, p. 545). This concept was developed over the last four decades and builds on the idea that all kinds of entrepreneurial efforts are embedded in cultural processes (Lounsbury & Glynn, 2019).

To provide an overview of the sequence of events covered by the concept, Lounsbury and Glynn (2019) developed a "process model of cultural entrepreneurship" (Figure 1). It visualizes the key idea that "stories that are told by or about entrepreneurs define a new venture in ways that can lead to favorable interpretations of the wealth-creating possibilities of the venture; this enables resource flows to the new enterprise" (Lounsbury & Glynn, 2001, p. 546). The model can be applied to both, the individual and collective act of entrepreneurial storytelling (Lounsbury & Glynn, 2019) and specifically highlights the interplay between resources, stories (narratives), legitimacy, and identity, which are complex and well-researched concepts themselves. In essence, storytelling is a tool to communicate a venture's resources and distinctive organizational identity to establish legitimacy, which is the precondition to acquiring resources (Lounsbury & Glynn, 2001, 2019). How a story is perceived and reacted to is determined by cultural mechanisms.

While Lounsbury & Glynn have been primarily driving cultural entrepreneurship theory, other authors have significantly contributed to its development. In the following, the key definitions of and connections between the concepts of legitimacy, narratives, identity, and culture are reviewed more thoroughly.

2.2. Establishing legitimacy for resource acquisition

Legitimacy is a critical factor for the emergence, growth, and success of a venture (Aldrich & Fiol, 1994; Lounsbury et al., 2019; Lounsbury & Glynn, 2001; Zimmerman & Zeitz, 2002) and a central element in cultural entrepreneurship theory. As legitimacy builds up over time, the lack thereof is a challenge faced by all innovating entrepreneurs (Aldrich & Fiol, 1994). Especially the acquisition of resources, including the raising of capital, is constrained by the lack of legitimacy (Aldrich & Fiol, 1994; Garud et al., 2014; Lounsbury et al., 2019; Zimmerman & Zeitz, 2002). Being a precondition to accessing those other resources, legitimacy can be considered a resource itself, which is just as important for new ventures as capital (Zimmerman & Zeitz, 2002, p. 414).

Legitimacy is the "social judgment of acceptance, appropriateness, and desirability" as defined by Zimmerman and Zeitz (2002, p. 414), who interpret the lack of legitimacy as a form of "liability of newness". Aldrich and Fiol (1994) further distinguish between cognitive legitimacy, measured as the "spread of knowledge about a new venture" and sociopolitical legitimacy, measured as the "public acceptance

of an industry [...] or the public prestige of its leaders" (p. 648). Judging legitimacy is a highly individual and audience-dependent process (Fisher et al., 2017).

So how can ventures overcome this lack of familiarity and credibility, when they need to acquire new resources? According to the relevant literature, ventures gain legitimacy by building trust through reliability and reputation (Aldrich & Fiol, 1994) and by conforming to rules, regulations, and expectations (Zimmerman & Zeitz, 2002). Further strategies include the location of the venture in a conducive environment and manipulation through lobbying, advertising, etc. to change existing ideas and creating a new social context (Zimmerman & Zeitz, 2002). However, the effect of all of those measures is likely to show in the long term.

To establish legitimacy and acquire resources in the short term, entrepreneurs rely on employing the right symbolic language and behavior when presenting their ideas to potential investors, to increase their confidence in the entrepreneur's credibility and professional skills (Aldrich & Fiol, 1994; Zott & Huy, 2007). As the judgment of the venture's legitimacy depends on the investor, who further helps to legitimize the venture, when committing to support it, the quality of stakeholder relationship can be decisive in this process (Lounsbury et al., 2019; Zott & Huy, 2007). Hence, relying on interpersonal resources and most of all on skillful storytelling is critical to the process of legitimizing and acquiring resources for a young and unknown venture or founder (Aldrich & Fiol, 1994; Lounsbury et al., 2019; Zott & Huy, 2007).

2.3. Entrepreneurial stories and meaning making

Knowing about the importance of storytelling in the process of resource acquisition, many questions about the effect, content, and development of a good story remain. For this purpose, let's highlight the interplay of stories, narratives, and plots.

In the existing literature, stories are often set equal with narratives (Martens et al., 2007). However, while the same story can be told by different people, narratives are unique for each person, depending on their individual vantage points (Garud & Giuliani, 2013). All entrepreneurs will "have their own narratives, depending upon their recollection of past experiences and future aspirations" (Garud & Giuliani, 2013, p. 159). As narratives entail a broad temporal perspective and context, they provide a framework for individual reasoning and understanding (Bartel & Garud, 2009; Garud & Giuliani, 2013; Martens et al., 2007).

As innovative ideas develop from those unique narratives, information asymmetry is another key constraint for resource acquisition besides uncertainty and the lack of legitimacy (Martens et al., 2007). To make their ideas and narratives more comprehensible, entrepreneurs can combine information from past, present, and future in a structured plot with a beginning, middle, and end (Bartel & Garud, 2009; Garud & Giuliani, 2013; Lounsbury & Glynn, 2001; Martens et al., 2007). As the way information is connected is decisive for the actual message, the plot is considered the means through which a story acquires meaning (Garud et al., 2014).
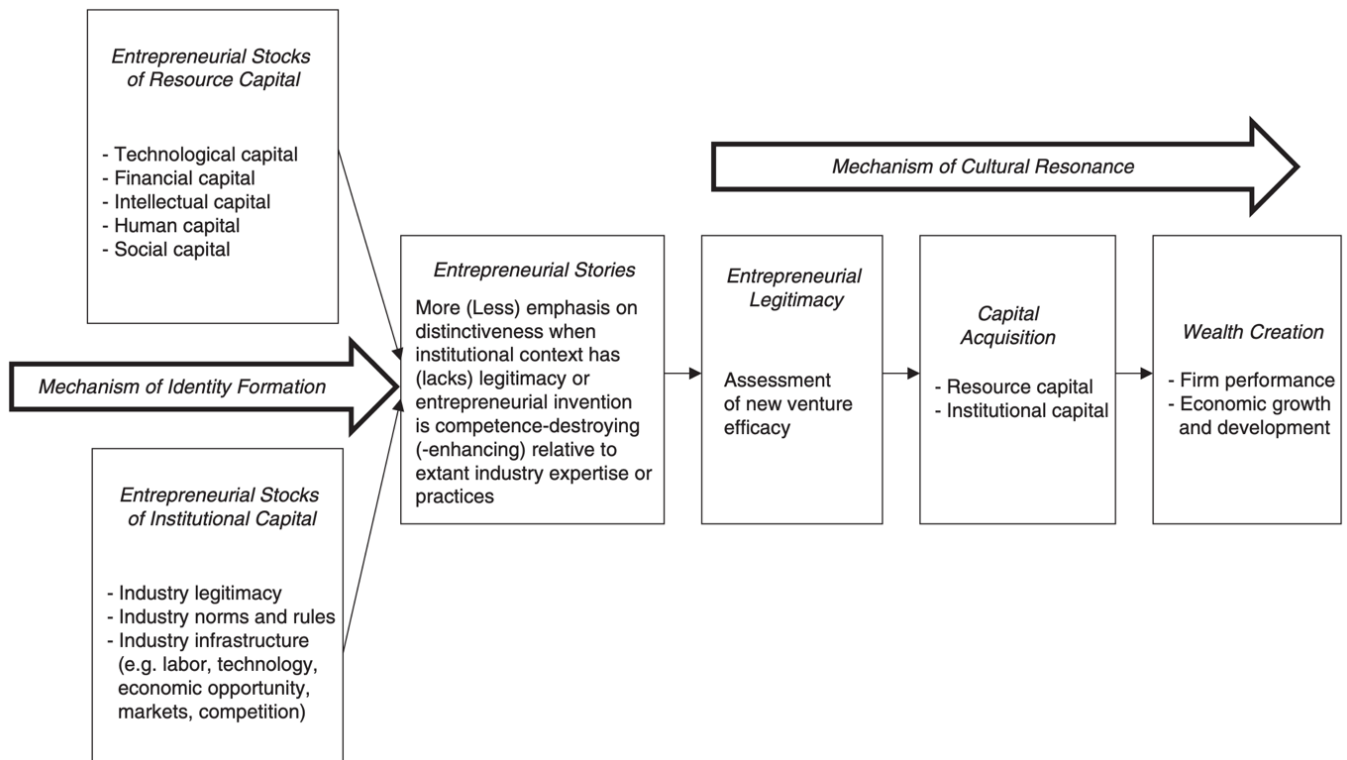
Figure 1: A process model of cultural entrepreneurship.

Source: Lounsbury and Glynn (2019, p. 11)

Meaning making is the act of defining and explaining the potential of an opportunity by relating disconnected pieces of information to each other in a new way (Garud & Giuliani, 2013; Martens et al., 2007). As the entrepreneur has relative freedom to re-present the facts in a new way (Garud & Giuliani, 2013; Martens et al., 2007), Martens et al. (2007) argue that the way the story is plotted has a bigger impact on the audience's judgment than the facts presented in the story. Therefore, the act of telling the story is a strategic moment, "where entrepreneurs can substantially influence the way projects are communicated and perceived among key audiences" (Manning & Bejarano, 2017, p. 211).

The content of a story should enable resource acquisition through the construction of a comprehensible organizational identity, invoking familiarity with the idea through contextualization and raising the interest and commitment of the audience (Martens et al., 2007). The overall goal is to decrease the perceived uncertainty and risk of supporting the venture with resources (Martens et al., 2007). To do so, entrepreneurs must develop a coherent plot, including problem definition and solution, venture resources (e.g., track record, key employees, patents), potential liquidity scenarios (i.e. exits), comparison to other growth stories, and the reasoning, why the venture requires sought resources at this point in time (Bartel & Garud, 2009; Garud et al., 2014; Lounsbury & Glynn, 2001; Manning & Bejarano, 2017; Martens et al., 2007; Wry et al., 2011).

Amongst the many findings of potential plotting strategies (which to list would go beyond the scope of this study), there has been an interesting discourse on the strategic trade-off between competitive differentiation (in terms of resources, capabilities, technology, etc.) and conformity in entrepreneurial storytelling (Lounsbury & Glynn, 2001; Zhao, Fisher, Lounsbury, & Miller, 2017). Zhao et al. (2017) argue that a story can position a venture as optimally distinct: "being different enough from peer firms to be competitive, but similar enough to peers to be recognizable" (p. 93). According to Lounsbury and Glynn (2001), competitive differentiation and conformity are the two main means of entrepreneurial storytelling, by which perceived venture legitimacy is judged. Therefore, "the content of entrepreneurial stories will focus relatively less on establishing a venture's distinctiveness when the industry context within which the entrepreneur is embedded lacks legitimacy" and the other way around (Lounsbury & Glynn, 2001, p. 559).

When establishing venture legitimacy through strategic plotting and positioning, entrepreneurs also set stakeholder expectations regarding future attributes of the venture ("cognitive expectations") and future benefits of the stakeholders ("pragmatic expectations") (Garud et al., 2014, p. 1479). When deviating from their projections, stakeholders will be disappointed, which can lead to a loss of previously established legitimacy, which Garud et al. (2014) call the "Paradox of Legitimacy".

To avoid the loss of or to regain legitimacy, stories have to be revised and replotted over time, according to the source of the deviation in markets, technologies, regulation, competitive landscape, or collective expectation (Garud et al., 2014; Manning & Bejarano, 2017). Story revision can also be triggered by the development of the narrative through increasing experience and the progression of the entrepreneurial journey (Garud & Giuliani, 2013). Trying to bridge this legitimacy gap by "ignoring the past or providing excuses" would not be perceived as credible (Garud et al., 2014, p. 1485). Instead, a new plot must be developed, which is coherent with the old version, explaining the reasons and consequences of development for each part of the story (Garud et al., 2014). Entrepreneurial storytelling can overall be seen as an ongoing process of revision: continuously adapting the strategic venture positioning "to succeed in dynamic environments" and "address the multiplicity of stakeholder expectations" (Zhao et al., 2017, p. 93).

2.4. 'Culture' in cultural entrepreneurship

The narrative is closely bound to a distinct cultural context, in which entrepreneurship and entrepreneurial storytelling always take place (Gehman & Soublière, 2017; Lounsbury & Glynn, 2001, 2019; Scott & Lane, 2000). On the one hand, this affects how an entrepreneur builds and tells his story and on the other hand, it profoundly impacts the way the story is perceived and the way that legitimacy of a venture or entrepreneur is assessed by the audience (Lounsbury & Glynn, 2019). Entrepreneurs must therefore know the importance and impact of the cultural context to increase their chances for successful resource acquisition.

Culture can be defined as an "interpretive framework through which individuals make sense of their own behavior, as well as the behavior of collectivities in their society" (Scott & Lane, 2000, p. 49) and "creates a connective thread among diverse people" (Bartel & Garud, 2009, p. 108). For an audience, culture is the basis for resource allocation decisions and to "take strategic actions based upon what the stories mean to them" (Lounsbury & Glynn, 2001, p. 545). Lounsbury and Glynn (2001) found that new ventures are perceived as more legitimate if the story matches "the expectations, interests, and agendas" of the investors, besides having narrative fidelity (p. 552).

The implication that cultural proximity is favorable for resource acquisition of new ventures relates to the concept of homophily, which is defined as "the tendency of individuals to associate with others based on shared characteristics" (Greenberg & Mollick, 2017, p. 341). Homophily has been studied with regards to class, gender, ethnicity, etc., partly also with regards to its impact on Venture Capital (VC) investment decisions (Greenberg & Mollick, 2017). Hegde and Tumlinson (2014) found that VCs are more likely to invest in ventures with founders of the same ethnicity as themselves. One reason for this may be that the investors expect "superior communication and coordination between coethnic VCs and startup executives after the investment", which may facilitate cooperation and monitoring after forming a partnership

(p. 2355). More generally speaking, Greenberg and Mollick (2017) found that "individuals are more likely to support those whom they perceive to be like them" (p. 346) or those who deal with the same structural barriers as them, because of their common social identity. From the perspective of an entrepreneur, cultural proximity may also facilitate investor selection through better access to information and opportunities within the cultural network (Greenberg & Mollick, 2017). Because the judgment process is so deeply culturally embedded, culture can be seen as a resource in itself that appears in social networks, communities, or clusters (Lounsbury et al., 2019; Lounsbury & Glynn, 2019).

Can entrepreneurs exploit the knowledge about the impact of culture on judgment processes? According to Lounsbury and Glynn (2019), entrepreneurs must understand the cultural context of each audience and adapt their communication and behavior accordingly to enable resource acquisition. For this purpose, entrepreneurs can use cultural resources, such as language, logic, or symbolic elements (Lounsbury & Glynn, 2001). With regards to narrative, the challenge is to balance distinctiveness with the contextual pressure to conform (Lounsbury et al., 2019). For this purpose, Fisher et al. (2017) introduced the concept of "emphasis framing". This entails the strategic adjustment of the salient elements within their stories "to emphasize specific legitimacy mechanisms – including identity, associative or organizational ones" that coincide with the conceptions of the audience and "improve their chances for accessing critical financial resources for venture survival and growth" (Fisher et al., 2017, p. 129).

Gehman and Soublière (2017) describe the process of legitimizing a new venture through storytelling as "deploying culture". They argue that storytelling can even create value in itself, by framing and reframing entrepreneurial aspirations over time, which they define as "cultural making" (Gehman & Soublière, 2017). According to Wry et al. (2011), cultural entrepreneurship entails the use of audience-specific vocabulary and rhetoric to deliberately shape attention and perception through storytelling and establish legitimacy. Both theories imply that the story told by entrepreneurs for the purpose of resource acquisition changes, depending on the cultural context, the entrepreneurial resources, and the stage of the venture.

2.5. Identity formation in cultural entrepreneurship: influence and outcome

This literature review has so far presented the arguments of researchers' shared belief that audience-adapted storytelling helps the establishment of legitimacy and thereby the acquisition of resources by new ventures (Fisher et al., 2017; Lounsbury & Glynn, 2001, 2019; Wry et al., 2011; Zhao et al., 2017). The means by which stories establish legitimacy is by constructing a comprehensive and optimally distinct organizational identity (Garud et al., 2014; Lounsbury & Glynn, 2019; Manning & Bejarano, 2017; Martens et al., 2007; Zhao et al., 2017). Adopting the definition of Gioia, Patvardhan, Hamilton, and Corley (2013), organizational

identity is the accumulation of those features, which are central to the character, "self-image" and distinctiveness of an organization.

What is organizational identity based on? Lounsbury and Glynn (2001) and Martens et al. (2007) argue that it is at least to some degree linked to the resources that are bound within the venture. Simultaneously, the identity is formed by the story, which is told by an entrepreneur with a unique narrative (Garud & Giuliani, 2013). Because they are so devoted to the success of their ventures, it can be argued, that venture identity is closely linked to the identity of the entrepreneurs (Grimes, 2017; Lounsbury & Glynn, 2001). Hence, not only the organizational identity is revealed through storytelling, but also the personal identity of the entrepreneur. Often, entrepreneurs would describe themselves with the same attributes, with which they would describe their venture (Lounsbury & Glynn, 2001). Martens et al. (2007) even argue that storytelling helps individuals to make sense of their own identity while shaping the perception of others. Identity can hence be seen as an influence on and an outcome of storytelling in cultural entrepreneurship.

Research on identity in entrepreneurship can be categorized in four streams: "*distinctions* – how entrepreneurs and their corresponding identities are distinct from other groups, *variations* – variations or varieties of identity types among entrepreneurs, *constructions* – construction, creation, and evolution of entrepreneurial identities, and *intersections* – intersection of the identities of entrepreneurs and other entities" (Mmbaga, Mathias, Williams, & Cardon, 2020, p. 2). While research on identity intersections again underpins the strong correlation of entrepreneurs' and ventures' identities, the research concerning construction and variation is most relevant to explain how identities emerge and diverge among entrepreneurs.

The immense diversity among entrepreneurial identities can to some degree be explained through the role and the social group each entrepreneur associates himself with (Mmbaga et al., 2020). On the one hand, the way an entrepreneur defines his role and responsibility shapes his decisions and actions, as well as his standing and perception within his team (Stets & Burke, 2000). The entrepreneur's role identity can help predict the dynamic within and development of a venture (Mmbaga et al., 2020). On the other hand, the social identity is determined by the social groups and entrepreneur associated himself with (Stets & Burke, 2000). Although entrepreneurs are usually conceived as autonomous and independent, their social network can greatly impact their behavior and overall venture performance (de la Cruz, Verdú Jover, & Gómez Gras, 2018; Fauchart & Gruber, 2011).

Diverse entrepreneurial identities also show in the way that adversity during the resource acquisition process is perceived and dealt with. Powell and Baker (2014) found that "founders who defined the situation as an opportunity simultaneously embraced the adversity, those who defined it as a challenge sought to counter the adversity and those who defined the situation as a threat attempted to accommodate the adversity" (p. 37). Similar results have been found in entrepreneurs' reactions to feedback and their readiness to advance their idea through pivoting (Grimes, 2017). Because they identify so strongly with their idea, entrepreneurs tend to "balance demands for adaptation with the need to retain a coherent sense of self and purpose" (Grimes, 2017, p. 1693). Those findings suggest that entrepreneurial identity can sometimes act as a constraint to venture advancement (Grimes, 2017; Powell & Baker, 2014).

Just like previously discussed for narratives, social identities can develop over time, as entrepreneurs try to build or maintain legitimacy by adapting to their changing social context (Down & Reveley, 2004; Grimes, Gehman, & Cao, 2018). While the adaptation to a social context is externally triggered, identity work can also be triggered by internal needs (Marlow & McAdam, 2015). Entrepreneurs may for example adapt their identities to advance their venture more effectively (Grimes et al., 2018) or to overcome stereotypes or assumptions related to their gender (Bruni, Gherardi, & Poggio, 2004).

Deliberate identity work or construction is an important process, which helps entrepreneurs to internalize and act on the learnings from their entrepreneurial journeys (Grimes, 2017). At the same time, a carefully constructed identity, which is conveyed through storytelling, will considerably increase the perceived potential of the venture and thereby its chances for successful resource acquisition (Lee, Hiatt, & Lounsbury, 2017; Martens et al., 2007; Phillips, Tracey, & Karra, 2013). Wry et al. (2011) transfer this concept from the individual to a group of people, e.g., a founding team. When multiple members of this group tell a coherent story, indicating an aligned purpose, narrative and collective identity, this increases their distinction from other groups and considerably increases perceived venture legitimacy (Wry et al., 2011).

2.6. Implementing cultural entrepreneurship

The existing relevant literature to a large part presents legitimacy as an outcome or state and barely explains the practical strategies which can be employed by entrepreneurs to establish or influence the assessment of legitimacy (Navis & Glynn, 2011; Werven et al., 2015). Uncovering those strategies and "understanding the fundamentals of successfully pitching new ventures has become an important component of entrepreneurship research" (Pollack et al., 2012, p. 3). In analyzing the wording of pitches, many researchers have taken a linguistic approach to explore pitch strategies (e.g., Cornelissen & Clarke, 2010; Navis & Glynn, 2011; Ruebottom, 2013; Werven et al., 2015). They discovered, that using specific rhetoric can help to increase the perceived legitimacy of entrepreneurs and their ventures and thereby enable resource acquisition (Ruebottom, 2013). Such rhetoric includes arguments or analogies that help the process of meaning making (Cornelissen & Clarke, 2010; Werven et al., 2015) through the means of comparison or contextualization of the venture within familiar categories or discourse (Navis & Glynn, 2011; Werven et al., 2015). Other research has found that signaling theory can be applied in pitches,

which is the establishment of legitimacy through referral to resources, roles, partnerships, etc. of the venture, which are in turn perceived as legitimate by the audience (Higgins & Gulati, 2006).

While the linguistic approach offers valuable insights into storytelling strategies, Steyaert (2007) observes a potential danger in focusing too much on the content of the pitch, instead of the behavior of the entrepreneur within the pitch: "The problem I then encounter is that the field of entrepreneurship studies has done everything to draw the attention away from the individual entrepreneur in order to make space for understanding the complexity of the entrepreneurial process but that the cultural reception of the story is not focused on the storytelling but on the entrepreneur telling the story" (p. 734). Some research has since been conducted on the behavior of entrepreneurs in the process of resource acquisition. One example is the work by Pollack et al. (2012), which is a model showing the positive correlation between an entrepreneur's preparedness, perceived legitimacy, and the amount of funding received. Likewise, most studies within this stream of research rely on deductive quantitative data analysis or qualitative in-depth single case studies.

3. Research method

To answer the research question: *How do entrepreneurs approach the fundraising process through personal, adapted storytelling and strategic behavior to legitimize themselves and their ventures?* new constructs need to be created, which focus on the role and the insights of the entrepreneurs throughout the entire fundraising process, instead of deducting findings only from the content of the fundraising pitch. Hence, this study uses an inductive, qualitative research method in a multiple case study design, which is especially suited for seeking answers to "how"-questions in unexplored research areas (Eisenhardt, 1989; Gioia, Patvardhan, et al., 2013). Grounding my theorizing in data from the in-depth analysis of multiple individual cases allowed me to generate new generalizable insights, which contribute to the discourse on the topic of cultural entrepreneurship (Lounsbury & Glynn, 2001, 2019).

3.1. Research context

The Venture Capital (VC) market is gaining new momentum. As venture-fund returns continue to exceed public market returns, industrial and private investors enable privately-owned ventures to raise more cash than ever before, as valuations are going through the roof (Figure 2, KPMG, Moore, Smith, & Lavender, 2021). More unicorns were born in the first half of 2021 than in all of 2020, exceeding the previous annual record of 179 from 2018 and the trend is predicted to continue (KPMG et al., 2021). The availability of so much capital means that investors have to work harder to get a seat at the table of high-potential ventures, giving more

bargaining power to entrepreneurs, who seek financial resources. Those new developments create an interesting context to revisit and extend the existing literature in the field of entrepreneurial resource acquisition and explore how the dynamics of this process may have changed.

3.2. Sample selection

The cases were selected in accordance with theoretical sampling, based on the likelihood to give useful insights for theory development for storytelling in entrepreneurial fundraising (Eisenhardt, 1989; Eisenhardt & Graebner, 2007; Patton, 2014). The basic criteria for the sample selection were a German heritage, age between 20 and 40, an educational background in business studies, and strong involvement in entrepreneurial fundraising. Those criteria serve as a common ground for better analysis and comparison of the findings and enable the reader of this study to identify more easily with the sample. The interviewees were then chosen based on their exceptional skills and reputation for storytelling or entrepreneurial fundraising or because they could offer a unique angle on the topic, thereby contributing to the theory development (Yin, 1994).

The initial interviewees were approached through my personal and professional networks, who then proposed and introduced other founders, who fit the sample criteria and could therefore also offer valuable insights on the topic. This form of snowball sampling gave access to a sample population that would have otherwise been hard to identify and access (Naderifar, Goli, & Ghaljaei, 2017).

This process resulted in a sample of 15 interviewees, who differ in terms of their founding and fundraising experience, industry backgrounds, geographic locations, etc. Many of them have experience from fundraising as a founder and now also assess other founders' capabilities from an investor's standpoint themselves, which adds to their angles of reflection concerning the discussed topic. The diversity of the sample was intended to allow for data "replication, extension of theory, contrary replication, and elimination of alternative explanations" (Eisenhardt & Graebner, 2007, p. 27). This way insights from early-stage founders could be contrasted or validated with the learnings of highly experienced founders. Overall, the sample of this multiple-case study offers a very holistic view and rich data on different personal stories and how those come into play during the entrepreneurial fundraising process and therefore a strong basis for theory building (Yin, 1994). Table 1 provides an overview of the 15 cases.

3.3. Data collection

The primary source of data for the analysis was a semi-structured interview with each of the 15 cases. The interviews were conducted remotely via videotaped Zoom calls, due to the Corona pandemic and the different locations of the interviewees. All interviews were conducted and manually transcribed in German, as this was the preferred language of all interviewees. Each interview lasted 30 to 45 minutes, resulting in a total of 101 transcribed pages (Appendix 3). To

Figure 2: Global venture financing 2013 - Q2'21.
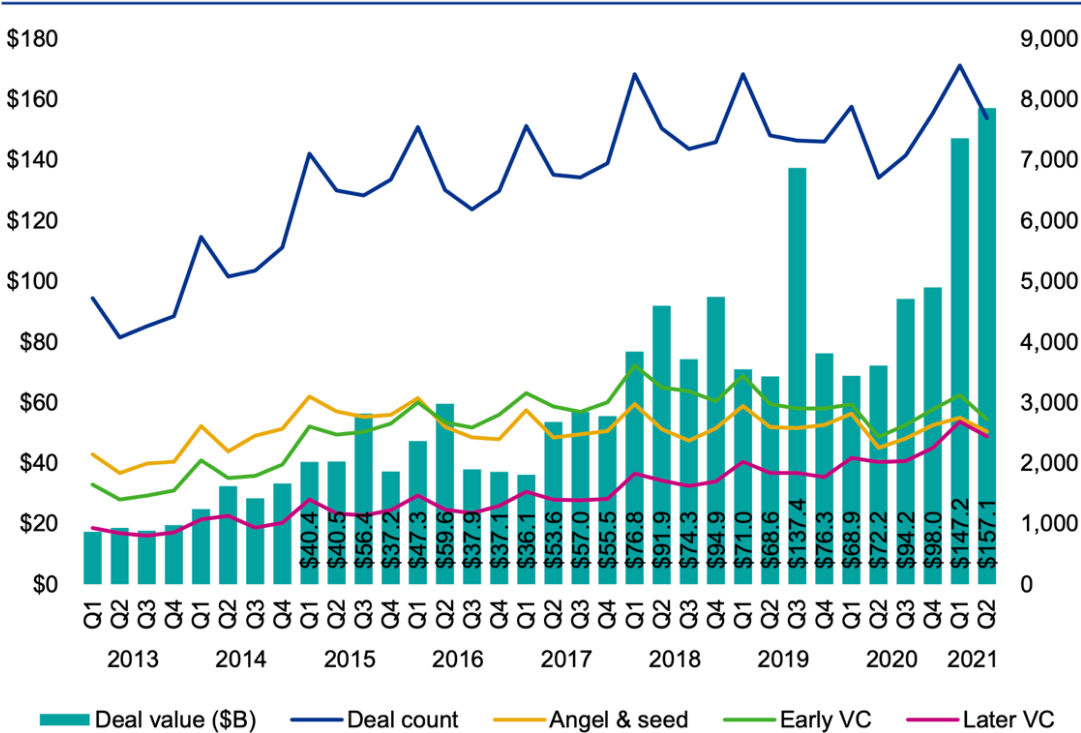
Source: Venture Pulse, Q2'21. Global Analysis of Venture Funding, KPMG et al. (2021).

Table 1: Sample overview.

Interview Partner	Synonym	Role	Location	Industry	# of ventures founded	Current venture stage
1	Ben	Founder & CEO, Angel	Munich	Logistics	4	Seed
2	Willi	Founder & MD	Munich	Mobility	2	Sold post Series B
3	Max	Founder & CEO	Munich	Nutrition	1	Series A
4	Toni	Founder & CEO	Berlin	Mental Health	1	Seed
5	Leon	Angel (Ex-Founder)	Berlin	Family Office	2	-
6	Hannah	Founder & CEO	New York City	Healthcare	1	Pre-Seed
7	Theo	Founder & CEO, Angel	San Francisco	FinTech	1	Seed
8	Jacob	Founder & CEO	Munich	Procurement	1	Pre-Seed
9	Finn	Founder & CEO	Melbourne	Nutrition	2	Sold post Series A
10	Tim	Founder & CEO	Berlin	Recruiting	2	Series A
11	Simon	CCO	Berlin	Nutrition	-	Pre-Series A
12	Julius	Founder & CTO, Angel	New York City	Cyber Security	2	Pre-Series A
13	Michael	Founder & President, Angel	London	BioTech	9	IPO
14	Nick	Partner (Ex-Founder)	Zurich	VC	1	-
15	Fred	Founder & CEO	New York City	Fashion	1	Series A

Source: own illustration.

gain new, truthful insights into this research field, this study relies on the reflection and recollection of the personal past and present experiences of each interviewee. To get open and honest answers, all interviews were conducted individually with guaranteed anonymization and confidentiality (Huber & Power, 1985). Hence, all mentioned names in this study are either fictional or initials. When speaking of 'the entrepreneur' the gender-specific pronouns "he/him/his" are used to ease the text flow, which are meant to include all genders.

To structure the line of thinking and questioning, an interview guide with a list of 30 open-ended questions was used (Appendix 1). Of those questions, only a flexible selection was asked in each interview, based on the experience of the respective interviewee and the flow of conversation. The interview guide is based on the existing literature within the research field and clustered the questions into nine potentially relevant topics (Appendix 1). The interviews generally covered three main topics: 1) the founder's background and fundraising experience, 2) the perception of the fundraising process and the attitude towards investors and 3) key learnings from their fundraising experience and development of fundraising strategies. The interview process followed the general recommendations for exploratory research (Edmondson & Mcmanus, 2007). The questions were framed carefully and the exact research question was not told until after the interview process, to avoid any biased responses (Huber & Power, 1985). At the end of each interview, the participant was asked to bring up additional points that he regards as important.

For the analysis, additional data points were included, such as the cases' LinkedIn profiles, venture websites, press releases, and articles or books, which they mentioned as an interesting source of information. During the interviews field notes were taken to recognize that the level of the interviewee's "emotional involvement with a topic or unit of analysis may either increase or decrease the accuracy of the responses" (Huber & Power, 1985, p. 175).

3.4. Data analysis and coding

The collected data was analyzed following an inductive coding strategy (Corbin & Strauss, 2014). First, all interview transcripts were analyzed on an individual level and coded by topics. To better understand the connection between collected data from different cases, I drew figures illustrating existing relationships and discussed my findings with sparring partners from my private network. I structured the collected data with the help of Condens.io, an online-based data management tool.

By recursively going through the extracted codes from the interviews, they were categorized based on underlying cross-case patterns, and those patterns were then summarized in the coding scheme and according quotation tables (Eisenhardt, 1989; Eisenhardt & Graebner, 2007; Suddaby, 2006). After constantly revising the coding scheme, the first-order codes were clustered in 10 second-order themes and

three aggregate dimensions (Figure 3). Letting the interpretation unfold bit by bit, 4 propositions were developed from analyzed qualitative data. In a final step, the findings are matched with existing theory, which either supports or is in conflict with the developed propositions (Eisenhardt, 1989).

Keeping an open mind and having no preconceptions regarding the findings, I let the data speak for themselves throughout the entire analysis (Suddaby, 2006). This process aims to ensure consistency of the empirical data from all cases with each proposition and to "provide the logical link between the constructs within a proposition" and those between the propositions and existing theory (Eisenhardt & Graebner, 2007, p. 29).

4. Results

Let me pre-empt a major finding of this study: all interviewed entrepreneurs fully concurred that an investor's decision to invest in a venture depends predominantly on the founding team. Their estimates on how much of the investment decision depends on the team varied from 80%-100%, leaving very little impact to other factors, such as the business model, market, product, etc. This finding directly supports the theory that the audience is most interested in the storyteller, not the venture (Steyaert, 2007) and presents the basic principle for all interpretation of further results of this study. The main argument to understand why this is the case was explained by Ben: "Because there are other players with a similar business model. And in the end, you just think: who do you trust most to build a category winner with this business model?" Although being the key to successful fundraising, all entrepreneurs reported not to talk much about themselves, when pitching in front of investors. Only 5%-10% of the time spent talking to investors revolves around the entrepreneur's background, expertise, and conceptions.

But how do investors assess the qualities and the potential of an entrepreneur? Or rather which strategies are entrepreneurs employing to convince investors of themselves? When diving into the analysis of the collected data to explore answers to those questions multiple patterns emerged, which could be clustered in three themes: 1) entrepreneur-based legitimacy levers, 2) investor-based story adaptation, and 3) understanding and mastering the "VC game" (Figure 3). Each of those dimensions has strong implications for the development of a fundraising strategy, which can fundamentally impact how an entrepreneur's legitimacy is perceived by investors. The identified patterns within each of the three dimensions will be presented in the following, together with a selection of strong statements, which illustrate the findings. Finally, propositions and a process model are developed, summarizing and linking the learnings from all three dimensions (Figure 4).

4.1. Entrepreneur-based legitimacy levers in the investor's judgment process

What do investors assess during a pitch? They "focus on the person, on the founding team, what they're like, and then
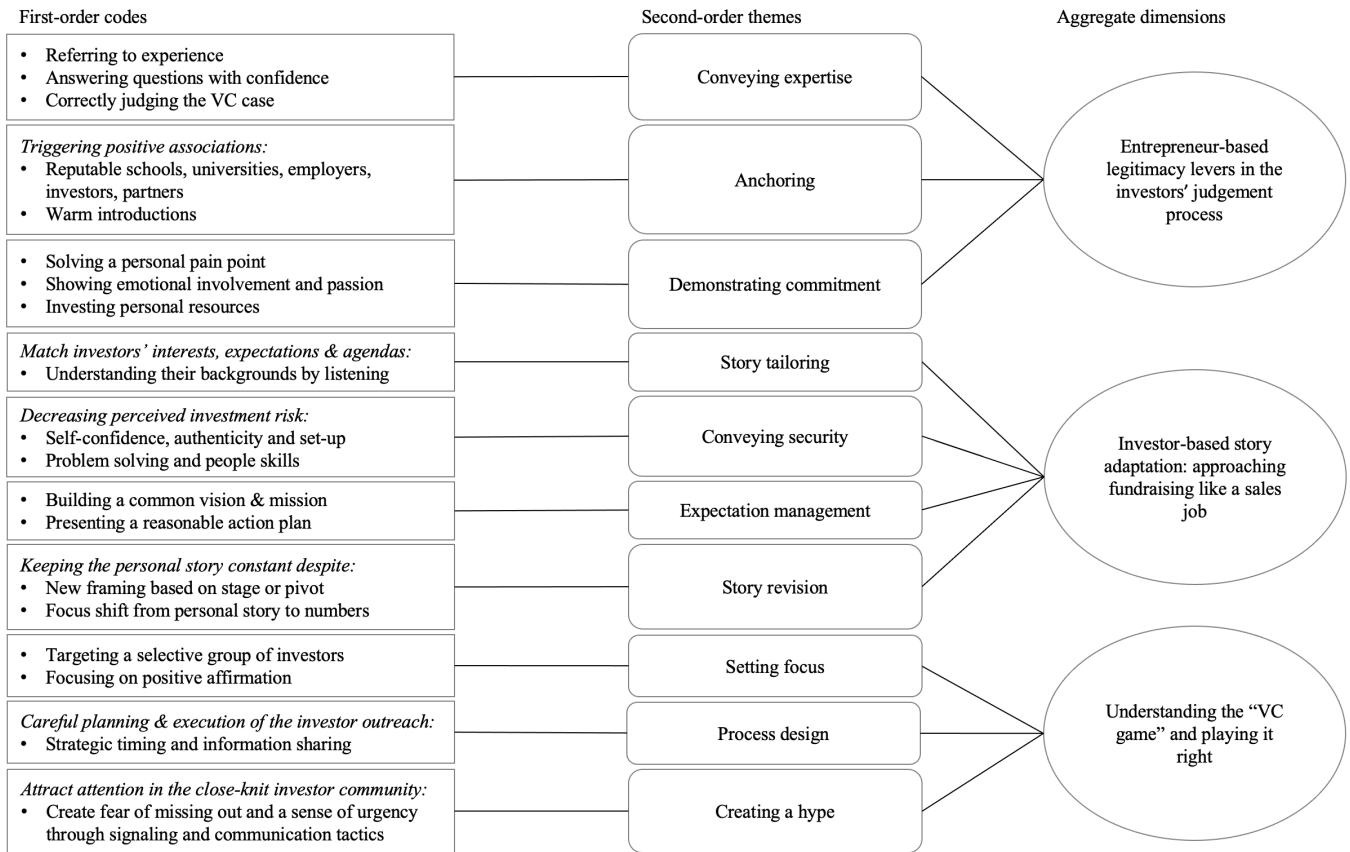
First-order codes	Second-order themes	Aggregate dimensions

- Referring to experience
- Answering questions with confidence
- Correctly judging the VC case

→ Conveying expertise

Triggering positive associations:
- Reputable schools, universities, employers, investors, partners
- Warm introductions

→ Anchoring

- Solving a personal pain point
- Showing emotional involvement and passion
- Investing personal resources

→ Demonstrating commitment

⟶ **Entrepreneur-based legitimacy levers in the investors' judgement process**

Match investors' interests, expectations & agendas:
- Understanding their backgrounds by listening

→ Story tailoring

Decreasing perceived investment risk:
- Self-confidence, authenticity and set-up
- Problem solving and people skills

→ Conveying security

- Building a common vision & mission
- Presenting a reasonable action plan

→ Expectation management

Keeping the personal story constant despite:
- New framing based on stage or pivot
- Focus shift from personal story to numbers

→ Story revision

⟶ **Investor-based story adaptation: approaching fundraising like a sales job**

- Targeting a selective group of investors
- Focusing on positive affirmation

→ Setting focus

Careful planning & execution of the investor outreach:
- Strategic timing and information sharing

→ Process design

Attract attention in the close-knit investor community:
- Create fear of missing out and a sense of urgency through signaling and communication tactics

→ Creating a hype

⟶ **Understanding the "VC game" and playing it right**

Figure 3: Data structure.

Source: own illustration.

combine that with validation of what they've done before" (Nick). "The main questions are: are they smart, do I think the team will stick together, is the market big and interesting? But above all: is the founding team good enough and are they willing to make the difficult decisions and work fast enough?" (Theo). Although only 5%-10% of the pitch directly revolves around the venture's founders, they have many options to establish legitimacy throughout the entire pitch. The three identified levers are to convey expertise, use the concept of anchoring, and demonstrate commitment.

4.1.1. Conveying expertise

"Expertise is very relevant. It doesn't have to be industry expertise, [...] but expertise in the sense of e.g., emotional intelligence, being able to build a team, at the same time being very convinced of yourself and your idea AND being humble. And that together is something that I don't think is very common, but that you need." – Willi

Entrepreneurs can convey expertise either through experience or confidence. Expertise can not only be shown during the introduction of the entrepreneur and his background but also in the way the venture pitch is prepared

and presented and the way the entrepreneur answers the investors' questions. The data reveal differences between first-time founders and serial entrepreneurs in the way that they have to present themselves and their backgrounds and answer questions to prove their expertise to investors. While a certain level of expertise from experience is taken for granted with serial entrepreneurs, first-time founders must often argue where they have acquired relevant skills, such as strategic thinking and a structured approach to problem-solving from previous consulting experience, or previous operational experience from working in another startup.

When answering questions, the perceived expertise of entrepreneurs shows to correlate with confidence. When an entrepreneur is e.g., extremely knowledgeable in his industry, he can respond confidently to investors' questions about the market potential and can give a more well-founded answer. The other way around holds also true: if an entrepreneur responds confidently to a question, he is perceived as more competent within the field. To make up for a lack of knowledge and still convey expertise, entrepreneurs often have to simulate confidence: "The important thing in fundraising is not only what story you tell, but also how you answer questions and how quickly you can answer questions, in a way that it looks professional, without you often having any idea

at all, but rather arguing something" (Toni).

Expertise does not have to be owned solely by one person within the venture. Depending on what aspects could be of interest to the investors or how responsibilities are delegated within the team, it could make sense to involve some other team members and their expertise in the fundraising process. It can be beneficial to show "on top of my expertise, I also have a few co-founders here who have already successfully implemented and built things up" (Michael) and thereby acknowledging that the entrepreneur knows his strengths and weaknesses and recruits those people who bring valuable experience. Within certain industries, involving an expert in fundraising pitches to explain the product may add credibility, as it is e.g., "important for a mental health product that the psychologist says something about it" (Toni).

Another way to convey expertise lays in the way an entrepreneur assesses his venture as an investment target. It shows that the entrepreneur understands the interests of the investor and prepares the information accordingly. "VCs invest in the team, but they also want to see a startup that can be sold for a certain multiple of their investment" (Finn). If a founder e.g., holds a minority stake of the venture before raising institutional capital, he should not argue for it to be an interesting VC case. To be perceived as knowledgeable, entrepreneurs should aim to pitch an ambitious, yet realistic growth and return potential to the investors.

This finding implies that personal fundraising stories should be carefully prepared, highlighting how previous experience will help the entrepreneur succeed with his new venture. Preparation at the same time helps to present the venture as an attractive investment target and to appear confident, increasing the perceived expertise and legitimacy of the entrepreneur (Pollack et al., 2012). Table 2 provides examples of statements about how expertise is demonstrated to investors.

4.1.2. Anchoring

> "Everyone has LinkedIn, everyone can look at your CV. I don't have to stand in front of you and recite that. And I think that's also the thing that we consciously play off because each of us has a polished CV with several good names on it. [. . .] I think it's much better to undersell yourself and to be modest. You just don't have an edge to oversell." - Jacob

Exploring, how expertise can be conveyed by entrepreneurs, I found that it is often underlined by incorporating information on their academic or professional background. It is safe to assume that investors have researched an entrepreneur's background prior to the actual pitch event. Hence, it does not make a good impression to recite the entire CV but to subliminally include well-known names or logos from partners, universities, or past employers in the pitch can have a strong signaling effect. This form of anchoring triggers positive associations with the entrepreneur within the audience. This concept links back to signaling theory and the study by Higgins and Gulati (2006), who found

a positive correlation between reputable management backgrounds and investor decisions. A background in consulting may be associated with a tough interview process and disciplined work ethic, start-up experience may be associated with a hands-on attitude, and a degree from a high-ranking university may be associated with a beneficial network, which may increase the venture's potential to become successful. However, associations can vary depending on the audience's cultural background and personal experiences with each anchor point.

How much investors rely on academic or professional anchor points to assess an entrepreneur's potential varies strongly. Most interviewees stated to mention their background during a pitch and to think that this information is valuable to establish legitimacy. However, those entrepreneurs, who also started to engage in angel investments said that the academic background was barely relevant to assess a target's potential. This discrepancy indicates a bias of how entrepreneurs perceive the way their pitch is assessed versus how it is actually assessed by investors. It could also mean that entrepreneurs are differently assessed by angel and VC investors, indicating that VCs consider more information about an entrepreneur's background to make more substantiate decisions.

Besides the academic and professional background, personal references and introductions are another form of anchoring, which is judged by the interviewed entrepreneurs to have a much greater impact on perceived legitimacy. If a personally valued or reputable angel investor is already part of a venture's capitalization table, then this will trigger a positive association with an investor. The same is the case, if a person who is highly valued (personally or professionally) within a network introduces or recommends the founding team of a venture as highly skilled, then this team will automatically be perceived as more legitimate than other unknown entrepreneurs. "It also helps when other people talk about what a great team we are because that's an external signal that can be very, very, very helpful" (Theo). This finding implies that the signaling effect of a personal introduction can almost replace a formal background check and that it can therefore be highly beneficial to get people to talk about you in a positive way.

By referencing reputable universities, employers, partners, or contacts, the entrepreneur places himself within a context or category that the respective investor knows or has concrete associations with, thereby directly influencing the investor's assessment of legitimacy (Navis & Glynn, 2011; Werven et al., 2015). Similarly, personal introductions to investors through the entrepreneur's network can have a strong signaling effect. This finding supports the theory that entrepreneurs should rely on personal resources to establish legitimacy (Aldrich & Fiol, 1994; Lounsbury et al., 2019; Zott & Huy, 2007). Table 3 gives more examples of how interviewed entrepreneurs engage in anchoring.

Table 2: Conveying Expertise.

	Indicator	Statement
Max	Confidence	The more self-confident you are, the higher investors will rate your competence!
Toni	Confidence	There are a few VCs who say they don't want founders that are too big-headed. But honestly, the more confidence you have, the more they think "Fuck, this can really happen!" and all the examples you look at: WeWork, Gorillas, they all have completely megalomaniac founders!
Fred	Experience	Sure, so I've been dealing with this topic for six years now. So, they can ask me any question about shoes, and I'll have a good answer, probably.
Nick	Experience	If this person is a serial entrepreneur, then you know, this person has hired people, fired people, gone through growing pains, finds things easier than others. and you know okay, this person for example wouldn't get involved in something like this again if he or she didn't really believe in it.
Toni	Experience	At the same time, I stressed my very strategic experience. I have always dealt with strategic issues in consulting: how to approach a market, how to prevail against competitors, what to focus on, but also how to strategically set up a company, functionally and processes, etc. These are all things that make the most sense for me to oversee, because they correspond to my experience.
Michael	Experience	It helps when you have successfully founded a few times – which you can see: Valuations just go up. If I raise a seed round for a company today, the pre-money [valuation] is 20-40 million, depending on the topic. When I used to collect a seed valuation, the pre-money was 500,000, simply because the investors, especially at the beginning, think that he's already done it a few times, so at least the founder's risk is no longer so high, that the founder will somehow screw it up.
Willi	Assessing the VC case	I'll give you a stupid example: if a VC looks at a company and sees the founder still holds 30% and the 70% is held by an angel who invested the first 100k, then it can't become a VC case anymore because the founder doesn't have enough shares – basics, which I didn't know 10 years ago.
Finn	Assessing the VC case	But on the investors' side, there was this tension of my shares and the focus on my person. They were not in balance; they were out of balance. Because someone who puts himself so strongly in the center of a company is expected [. . .] to be the majority shareholder, which I was not. I think that was definitely a problem for VCs [...], because the founder doesn't have enough shares and if it doesn't work, then he leaves.
Tim	Preparation	We try to think about what data points would be important and to have everything ready, because that also makes a lot of impression in the process, when he says can you give me a cohort analysis by segments and you say yes, here it is, 15 minutes later, he just has the feeling that you have a company blatantly under control. Exactly what he's looking at, you're already looking at. That just builds a lot of trust.

Source: own illustration.

4.1.3. Demonstrating commitment

"The more personal you can tell the story of why you are the one building the startup now, the better the story." – Julius

An entrepreneur can be an expert in his field and have a perfectly matching background, still, this does not answer, why he founded the venture and is striving for it to become a success. In addition to the relatively rational assessment criteria of expertise and background, all entrepreneurs described an emotional component to be central to their pitch. "So, the story about yourself has to show that you're on a mission, that you're not just doing this because you want to earn money, but that you have some kind of connection to it, that you're highly motivated, that you're generally an ambitious guy" (Toni). Since I purposefully selected a diverse sample to discover similarities and differences in the way entrepreneurs approach the fundraising process, they all had different ways to express their motivation for founding, man-

aging, and growing their ventures. However, two aspects remained constant throughout all interviews: they deeply understand the pain point that they are solving, and they invest their personal resources to succeed.

Firstly, to be perceived as legitimate, entrepreneurs must be able to explain, where the idea for the venture originated. Often this is the personal experience or direct observance of the pain point by the entrepreneur. "I think it's important to every founder that they can somehow identify with the product and potentially be their own customer, which correlates quite well with the authenticity of the pitch" (Michael). One extremely impressive example for a personal founding story was told by Michael:

"I treated my own depression and anxiety disorder with psilocybin, or mushrooms, and that's how I got into it. And my co-founders had a son who was suicidal and suffered from obsessive-compulsive disorder and he was treated at Harvard at the time

Table 3: Anchoring.

	Association trigger	Statement
Max	Education & Work experience	So sure, you have to convince the investors that you're the [...] right person for the job. But that mostly comes through the storyline itself and not by them asking: "What's your resume?" However, [...] if you've been at McKinsey for 5 years, you'll get a lot of trust and a higher valuation when it comes to the seed round or pre-seed.
Nick	Education & Work experience	So just how does the story fit together, where you went to university, what companies you were at, what you're like, very much thinking, where has already gotten a cachet before? Is that kind of credible?
Finn	Education & Work experience	There was half a slide about me in it. It was just like... What experience do I have? And then I threw all the logos in there that were somehow connected to me.
Theo	Network referral	So, when you already have a couple of stamps and signaling points, as we do, or when you get a good intro, then they already know that the team is pretty good and don't need to hear any more about it. But if you get to know someone completely new, you focus on the team, because they want to build trust in you as a person. Therefore, the better you know someone, the less you focus on team.
Theo	Network referral	We've heard very often that in investor circles people talk about what a small but high-quality team we have, and that other people talk about what a great team we are. Of course, it also helps when other people talk about what a great team we are, because that's an external signal that can be very, very, very helpful.
Simon	Network referral	The second major dimension is networking, exactly. Because I would say that this is almost half the battle in investor acquisition. If you know the right contacts, or if they refer you properly, then you save a lot of cold contact, which is the same as when you do sales.

Source: own illustration.

and nothing worked. He then tried to kill himself three times. They then also had him treated underground, first with ketamine and then successfully with psilocybin. We then somehow came together and said we'll start C. together and figure out how to develop that for patients."

Demonstrating identification with and passion for the cause behind a venture assures the investors that a founder is highly committed and also ready to persevere times in which the venture is not performing so well. Another way to do so is by integrating anecdotes that highlight the entrepreneur's willingness to invest personal resources, such as time, energy, and money. This can include measures taken to overcome hurdles or ones, which demonstrate that an entrepreneur is ready to take tough decisions or go the extra mile. Examples from the interviews are shown in Table 4. Overall, demonstrating commitment through storytelling can quickly become very personal. It is this part of the pitch that can get investors emotionally involved and trigger connection and understanding between investors and entrepreneurs, as it did for Willi: "A very brutal, beautiful statement was made by a US investor, who said: 'One thing we trust about you, is that you're either gonna win or die trying'. And I thought that was wonderful! That really hit home, in a positive way, it made me happy, and for them, that was the decisive argument, because they totally see that

I run and run and run."

Explaining the personal reasons for founding and developing the venture creates meaning for the investors (e.g., Garud & Giuliani, 2013; Garud et al., 2014). It gives the investors a deeper understanding of the entrepreneur's personality and hence the organizational identity of the venture (e.g., Gioia, Patvardhan, et al., 2013; Lounsbury & Glynn, 2019). While multiple entrepreneurs may have the same idea or business model, their personal story and commitment are what sets them apart and determines which venture will become most successful. In other words, while expertise and a background with signaling effect are what legitimizes the entrepreneur, his personal story and commitment to the venture are what sets him apart, establishing overall legitimate distinctiveness (e.g., Lounsbury & Glynn, 2001, 2019; Werven et al., 2015) and may eventually trigger emotional involvement and conviction of investors for the venture.

4.2. Investor-specific story adaptation: approaching fundraising like a sales job

After covering what entrepreneurs should be telling investors about themselves as part of their fundraising pitch, let's explore the findings on how entrepreneurs should pitch to investors to establish legitimacy. "Looking at the big picture, I think that founders definitely make the main impact when it comes to fundraising and how they build and set up

Table 4: Demonstrating commitment.

	Indicator	Statement
Julius	Experiencing the pain point	Then we came back and said, maybe we should build this, we're not the only ones who get so fed up! [...] That was a kind of "personal pain", that is perhaps a very good description, that we ourselves experienced how annoying it was. And that was also enough motivation to build it.
Julius	Experiencing the pain point	Why are we better than our competitors? Because we experienced this problem ourselves, we were our customers, it was our problem, so we understand that.
Nick	Experiencing the pain point	I once heard pitch was about a product against erectile dysfunction and the founder then said yes, I also have this problem. It can quickly get very personal.
Michael	Experiencing the pain point	I think it's important to every founder that they can somehow identify with the product and potentially be their own customer, which correlates quite well with the authenticity of the pitch.
Theo	Overcoming hurdles	My co-founder and I shared a bed for the first nine months when we lived in San Francisco because the rents were incredibly high [...] That just shows that we're really hustling and really putting up with hurdles somehow. And I think that's one of the things that they found coolest about the team.
Max	Overcoming hurdles	They invested, because we could show that we managed to come so far with super little money compared to other participants in the market.
Jacob	Showing passion	After all, we did a couple hundred interviews and built prototypes and so on. [...] I think that the most important and often underestimated thing is just to have 100% focus on the thing and show I'm doing this, there's nothing on the side where I just waste time. And it will move forward with or without you. I think that's what underlines the seriousness.
Willi	Showing passion	And I wasn't aware of that to the extent that I'm aware of it today, that what I could tell about me and us, that's always been what a VC basically likes to hear, I think. Namely, someone is 100% committed. My wife and I, for example, quit our jobs together and said we were going to start a company. As a VC, you can't ask for more commitment than that.
Ben	Skin in the game	I mention it probably deliberately subtle, so as not to say: Hey, look, I've also invested in the company. That means I am more committed than others.

Source: own illustration.

their story. Some are good at it; some are not so good at it. Some have an awesome product, and no one invests because they can't tell their story and just shove their product in everyone's face. Versus someone who can tell a mega story and creates excitement for everyone. And maybe he has a crappy product, but people invest. That's supposed to have happened in the past (*laughs*). So that's simply human psychology, I would say, that's the absolute basis of it" (Finn). Max agreed with Finn: "Even now, you can really badmouth our business case and our numbers so that no one would invest. But you can also use [the numbers] to support your overall story, [. . .] It depends 100% on HOW you tell the story."

Speaking to the entrepreneurs about their storytelling tactics in fundraising, they often drew parallels to sales strategies. Willi argued that "fundraising is above all selling – selling yourself, selling the company, selling the idea, selling the vision, selling the team". Similarly, Ben realized that he changed his story over time because "in the beginning it always sounded much more financially driven to me, and it's basically not really anything other than doing a sales

job." Because "in the end, it's nothing other than convincing customers about you. You just have to be investor-centric instead of customer-centric" (Max). When analyzing how the entrepreneurs report to approach fundraising like a sales pitch to convince investors, four core tactics could be identified: story tailoring, conveying security, expectation management, and story revision.

4.2.1. Story tailoring

> *"I would call it active selling to adjust the behavior and story to the recipient. Doing anything else would not be very appropriate, I think.*" – Willi

When the interviewed entrepreneurs were asked about their most important learnings from their fundraising experience, most of them reported that they had stopped pitching the same exact story to all investors, because they were not successful that way. Instead, they started tailoring their story to the interests and culture of the investors, increasingly empathizing with their current situation, background, and goals. This finding supports the theory that investors perceive those stories as more legitimate, which match their

individual interests, and that the vocabulary and rhetoric of stories should therefore be tailored to the specific audience (Lounsbury & Glynn, 2001; Wry et al., 2011).

Almost half of the entrepreneurs who were interviewed have previously fundraised in both Europe and the US. This allows for interesting insights into how cultural differences required them to adapt their stories. The general notion was that "the Europeans are just rather number-driven with their Excel spreadsheets. And in America, especially in early-stage investing, it's all just story, team, and vision!" (Julius). This also shows in a different investment mentality. While they perceived European VCs as generally more risk-averse, "the good US VCs don't want to understand what can go wrong, they want to know what happens when everything goes right. [. . .] They ask themselves, what is the next 100 billion or maybe even trillion-dollar company that can be built because they have done it so often. And that is also the difference" (Michael). Consequently, entrepreneurs who fundraise in the US should adapt their story to communicate a bigger vision, which does not have to be as rigidly underpinned by data as it should be in Europe.

A second effect of the cultural difference is that US investors prefer more emotion-loaded stories than European VCs, which may be a challenge for European entrepreneurs. This is also what Hannah experienced when pitching her venture in the maternity care space in the US:

> "About my personal story: many ask me, why do you do that, right? And I really noticed in the US, the best thing is that you yourself almost died at birth, so that you have the right to work in the space, also because it is such a social issue. And I was like, I actually just want to work in the space somehow. And I think there is also a bit of a clash with Europe and the US. Europeans just say some things as they are, without making an extreme story, right? And I thought, hey, I'd almost have to invent something so that they'd believe me, that I'd stick to it, right? Well, that is, the personal story was my problem."

However, several entrepreneurs reported that there is a mindset shift taking place in Europe, making it more common to invest based on emotional involvement. This would explain, why many of them made bad experiences with leaving out a more emotional introduction, explaining the personal connection and commitment to the cause: "Well, I had two investor pitches, in which I jumped into the facts relatively quickly. Both of them dropped out. I don't know. Maybe I didn't perform in the call somehow, I don't know. But at least I didn't do the emotional part. That's one data point that I have" (Simon). Referring to the concept of "The Golden Circle" (Sinek, 2011), this finding implies that entrepreneurs should always adapt their stories to start with "the 'Why' and not start with the 'What', because the 'Why' convinces in the end" (Simon).

As culture provides an important interpretive framework (Scott & Lane, 2000) and basis for resource allocation decisions (Lounsbury & Glynn, 2001), these findings support the theory that understanding the cultural context of the pitch and adapting it accordingly can influence the perceived legitimacy of a venture (Lounsbury & Glynn, 2019). Changing the communication of the venture's vision with regards to boldness, abstraction, and emotional involvement can increase perceived cultural proximity, based on the recognition of a similar narrative of the investor and entrepreneur (Greenberg & Mollick, 2017) and thereby increase the chances for resource acquisition.

Similar to tailoring the tone of a pitch to the audience's cultural background, it is equally important to address the personal interest of each investor. "It's great when founders [. . .] report on an extremely successful business model and have crazy numbers written on the slides. But I think it's a big mistake to use that as an intro for such a call because you're completely ignoring the interests of many investors" (Simon). Those interests can be uncovered when questioning: "What do they want from you? What do they need in the end so that they look good internally so that they don't get any trouble from their limited partners? What do you have to deliver to them, i.e., what are the key points, so that the investment manager can stand up in a partner meeting and pitch [your venture] in such a way that they say 'Whoa, we have to invest!'" (Max). Entrepreneurs can e.g., make their pitch more relatable and appealing by drawing parallels to previous investments of the investor, again highlighting the significance of analogies in fundraising rhetoric (Cornelissen & Clarke, 2010; Werven et al., 2015).

Sometimes investors may have their own passions, which they like to live out by helping ventures to succeed within the field. Uncovering these and integrating them in the pitch can not only spark conviction in the investor but also turn out highly beneficial for the entrepreneur, as it did for Fred: "what was also an important learning was that every investor invests for different reasons and that you have to tailor the pitch. So, for A.B., he invested because his personal goal was to go a bit back to the roots because he loves Italy as a country and loves the production in Italy, he wanted to work with a company where he can somehow pursue his Italy production dreams a bit further." (Extended in Table 5.)

Adapting the fundraising story based on the background and interest of the investor does not mean telling a completely new story each time. "The base story always stays the same. It always depends on what you emphasize or what you highlight" (Max). While some background research may help to prepare for an investor meeting, a lot of the tailoring takes place during the pitch, in conversation with the investor. It is important to understand that to develop conviction for a venture, investors must have the feeling to understand what is important to them about the venture. To understand the entire setup, especially regarding "the team, the vision, and the market" (Toni), entrepreneurs must make time to listen and thoroughly respond to their questions. "Fundraising is kind of like speed dating. You only have half an hour and if you don't let the investor get rid of his questions, then he doesn't understand [the venture] and can't develop conviction on

your case, and then your chances of an investment decrease radically" (Nick). Similarly, entrepreneurs should make time during the pitch to ask questions themselves, which is a good way to find out about the individual investor's interests. A highly instructive example for listening and clever tailoring was given by Tim:

> *"I used to just rigidly pitch my thing. And then I learned that in good sales you actually start with open questions and try to understand what the customer wants, what the customer needs, what his problems are before you pitch anything, so to speak. So really bad sellers, they always pitch immediately, and really good sellers ask a lot of open questions, like a psychologist or a doctor or something and I do that for example now in fundraising very extremely that if they schedule a half-hour call with me, I try to ask them questions for a quarter of an hour. What does your ideal investment look like? What is critical for you guys in a successful Marketplace? And so, to understand what they pay particular attention to, and then in my pitch in the second 15 minutes, to tell them exactly these things. So, if they say: Yes, SaaS and retention metrics are totally important for us. Then I don't tell them: I'm a B2C marketplace, but rather say: yes, we have a really strong B2B business with SaaS-like retention metrics. Whereas if they say yes, we love Marketplace with local network effects, then I'm more likely to say: yes, we have the super liquid micro-market in Berlin call center agents. So, I tailor that a bit more in the meantime."*

Similar to anchoring, tailoring the pitch to the investors' interests and culture is another way of relating the unknown venture to concepts that are familiar to the investors, thereby making meaning of the venture's vision and directly influencing their assessment of legitimacy (Navis & Glynn, 2011; Werven et al., 2015). By laying out the organizational identity in a way that sounds especially promising to the investors, entrepreneurs also legitimize the venture's valuation.

4.2.2. Conveying security

> *"What's important about you personally is what contributes to the story and what gives the investor confidence, simply security. Venture capital is about being able to assess risk or thinking that you can assess risk."* – Nick

What entrepreneurs can say to convey security was already covered with what I called legitimacy levers: expertise, anchoring, and demonstrating commitment. However, the assessment of venture risk and legitimacy is not only based on the things that entrepreneurs say but also on their behavior. The kinds of risks, which investors assess to develop conviction include: "can the founder solve problems? It's about problems from finding an office to problems like managing

100 people" (Leon). Is the founder "able to recruit senior staff" and "get people behind a vision" (Michael); and "do we believe that he can fundraise successfully again in the future?" (Toni).

So how do entrepreneurs perform their pitch to decrease the perceived risk of investing in their venture? All interviewed entrepreneurs stated to rely on self-confidence for this purpose. Showing strong confidence in the venture's business model and "objectively stating why the business model makes extreme sense and why we will win with it" provides the very basis for a convincing pitch (Hannah). Proactively providing relevant data, breaking down complex topics, and presenting them in a way that signals competence, drive, and self-sufficiency is also very well received by investors. When speaking about topics, which are especially important to investors, such as the commercial development of the venture, Ben described how he goes "into detail relatively strongly, to prove prior knowledge and certain stability. On the one hand, it's not made up - but on the other hand, I do this with the goal that they have the feeling: Hey, I don't have to get involved at all."

While some entrepreneurs stated "if you go out of the pitch and don't ask yourself: Have I potentially overdone it? Then you've done something wrong" (Max), others argued that it is key to still be perceived as authentic and do not want to be perceived as arrogant or pretending to be something they are not (Table 6). However, from an investor's perspective, Nick stated: "I have met many founders who came across as almost over-confident, but I still ended up offering them a term sheet. So, arrogance is still the easiest to forgive, I would say." This implies that confidence is positively correlated with perceived security and arrogance is in fact not necessarily correlated with higher perceived risk.

To assess, whether entrepreneurs will be able to win other investors, as well as senior staff for their mission, people skills move into the focus. Being a good communicator, making tough topics comprehensible and sound easy, speaking positively about other co-founders or team members, and acting respectfully towards the investors are all strong signals. (More examples in Table 6.) That is how Hannah explained a turning point during her fundraising experience: "I was very proactive, so I rather led the conversation and also asked a lot of questions. [...] If you let the other person talk, then you bond more and it's all about this relationship. You have to like each other. If they don't like you, then no one will invest."

While the personal connection and emotional involvement certainly play a role to convey security and develop conviction in the fundraising process of early-stage ventures, different patterns emerged for later-stage ventures. As early-stage ventures cannot offer a lot of data to assess their performance, the investment decision is naturally much more intuitive and emotionally driven. The investors' decision is therefore much more dependent on the performance of the founding team, as well as the estimated market potential. When investing in a later-stage venture, more data is available based on which the potential or risk of a venture and

Table 5: Story tailoring.

	Focus	Statement
Julius	Culture	The Europeans are just rather number-driven with their Excel spreadsheets. And in America, especially in early stage investing, it's all just story, team and vision!
Michael	Culture	I think US investors know how important capital is, so you just want to fund a startup massively and in Germany everything is always very tightly sewn on the edge. You kind of make your budget, the VCs look at your budget and fund exactly your budget. In the US it's more like, yes, we have a budget, here's about how we want to spend the money and we want to raise another 30 million for some exciting opportunities and everyone thinks that's cool.
Michael	Culture	For them, it's much more important to think, okay, what's the world going to look like in 10 years. And what they are all trying to do now, the Valley VCs, is this exponential thinking. We think linearly and then question ourselves, okay, what happens in a world that doesn't develop linearly, but exponentially, and where can that lead? How do technologies grow together?
Max	Investor interest	Ah yes, if there's one thing I've learned.... I always thought it was super rational and super objective. But investors are just people, they're super emotionally driven. That means I just figure out [...] what they find exciting and what not. And then you just adapt the story to that!
Max	Investor interest	In the end, you have to think: What do they want from you? What do they need in the end so that they look good internally, so that they don't get any trouble from their LPs? What do you have to deliver to them, i.e., what are the key points, so that the investment manager can stand up in a partner meeting and pitch you in such a way that they say "Whoa, we have to invest!"
Fred	Investor interest	Every investor has a little bit of their own story and you want to respond to that. It's in our interest to pick up people where they have a value-add, where they are enthusiastic and want to get involved. Understanding that and responding to it is super important!
Ben	Meaning making	Not every VC thinks the same way, and not every partner within a VC is as analytically or emotionally inclined or as focused on people and the team as the other partner might be. I would describe that as my greatest strength, to find out what the trigger points of my counterpart are.
Ben	Investor interest	But it's also become much more about having fun, making it fun for yourself: "Hey, how do I adapt a pitch to the person sitting in front of me so that there are no questions left?
Nick	Meaning making	[When does self-confidence turn into arrogance?] That's a fine line. I think confidently answering everything, but never letting the other person's questions come across like they're stupid questions, like, that's obvious! Or something like that. But just always listen, be an active listener and try to explain to people. That's super good.

Source: own illustration.

market can be assessed. As the founding team is still very important, but no longer the only driver for the further development of the venture, C-level team members start moving into focus. This finding suggests that successful fundraising stories become less personal and more data-driven, as ventures become more sophisticated.

Whether the assessment of risk and legitimacy are based on perceived authenticity and trust in the founding team or based on the careful assessment of the venture's strategic setup may also depend on the investor. Besides experiencing the fundraising process differently depending on the venture stage, the assessment focus may also vary between VCs, angel investors, and family offices. The data suggest that angel investors and family offices rely more heavily on interpersonal connections to the founders, while VCs persist in a more holistic assessment of potential success factors.

4.2.3. Expectation management

"The question is simply, how to build the product, what does the timeline look like, how to do operations, how to hire the right people? That you communicate everything in advance and if the investor doesn't agree, he's probably not the right person."
– Leon

For the collaboration of entrepreneurs and investors to become a success, it is important "how the founders handle the entire investor relationship management and expectation management" (Leon) right from the beginning. This is why all entrepreneurs stated to paint a clear picture of their vision and include some kind of action plan on how to reach this vision in their fundraising story (Table 7). Understanding "what is the vision of the person? How does he or she envision the next six or twelve months?" (Ben) legitimizes the presented vision, makes meaning of the process of reaching this vision, and gives investors a deeper understanding of the entrepreneur's mode of work and personal and organiza-

Table 6: Conveying security.

	Assurance	Statement
Julius	Authenticity	It is important to be authentic. If I now also go by the example of L. Every interaction I had with her was super sympathetic [. . .] and came across as very, very authentic and honest and likeable.
Simon	Authenticity	I think on a business angel level, it's probably a lot more personal touch and let's go for a drink together and the business plan, we'll put that to the side. It will develop if we believe in you.
Jacob	Authenticity	For us, it's more about professional competence and sobriety and a willingness to learn, and I think it's more like Hey, we're up for it. We don't know it all, nor do we pretend to do so.
Hannah	Confidence	Well, then of course there are those who somehow tell you they believe in the market and so on. But at the end of the day, I think the reason why this worked out was confidence. The more security you radiate, the more the investor also thinks, if I put my money in there, then it will multiply.
Hannah	Confidence	On the other hand, I was very proactive, so I rather led the conversation and also asked a lot of questions. Which I think was very key to making it more successful.
Nick	Confidence	I have met many founders who came across as almost over-confident, but I still ended up offering them a term sheet. So, arrogance is still the easiest to forgive, I would say.
Willi	Confidence	From what I have learned and experienced, an investor always wants a strong founder personality who leads the way, with all the advantages and disadvantages that this entails.
Tim	Confidence	So I love to invest in very strong founders. I think it's also good if they are in a very large or interesting market. I think where you should pay less attention is: what exactly is their idea?
Michael	People skills	That means the risk they don't have is that I'm not able to get people behind a vision. And that's already a problem that startups fail at, if they're not able to recruit senior staff.
Michael	People skills	Exactly the same with the fundraising risk, especially with the early-stage VCs who are then not able to do the follow-on rounds, there is always the question, are you able to go out and collect the capital at the end from the funds that are doing the late-stage rounds?
Julius	People skills	That's why it's still an advantage when I speak in the fundraising rounds, because I'm relatively extroverted and I can break down and explain very complex architectures or very technical things to investors relatively well. That definitely helps a lot!
Theo	Resilience	I am very good at breaking down complex problems and, above all, I have a very strong resilience. It doesn't bother me so much to run into walls all the time. I just see if I have to jump over them or dig under them.

Source: own illustration.

tional identity (e.g., Garud et al., 2014; Lounsbury & Glynn, 2019; Martens et al., 2007).

When presenting an action plan, the challenge for the entrepreneurs is to strike a balance of showing drive yet being realistic, as described by Ben:

> *"I have always seen a very great danger in overpromising and have therefore been very defensive in everything we say about deadlines or sales indications, and fortunately have not yet really been in a position to say: 'We are significantly below.' Instead, we tend to always be on the verge of overperforming. But that's a matter of framing. You simply have to set yourself a goal that you know is achievable for you. Conversely, you must not fall into the trap of someone saying, 'your goals are not ambitious at all.' I think that's a bit of a tightrope walk that you have to do."*

By closely managing expectations like that, investors also get information about the kind of collaboration that the entrepreneur is looking for. They can for example assess whether the entrepreneur is more driven by numbers or by emotions, how far he plans ahead, how much he wants investors to get involved, and how he communicates his vision. As for the investors in front of which Tim pitched his ventures, he believes "that many investors really appreciate working with a founder who really has his business and his numbers under control. I often get feedback that our board meetings, for example, are very, very good. That they are very well prepared, that they are very goal-oriented and very deep discussions." If an investor does not sympathize with the founder's vision or way of thinking, it is not likely that he will invest in the venture. This can be seen as a process of natural selection that is beneficial for both the investors and the entrepreneur, as it would be difficult to align their vision for the venture (see section 4.3.1.). However, if their vision and expectations are well aligned, this sets a great basis for a conducive collaboration and takes away the pressure

that entrepreneurs may feel that they could disappoint their investors.

This finding implies that investors, who have a similar narrative to the entrepreneur perceive the presented action plan and vision as more realistic and legitimate, supporting the theory that cultural proximity between the investor and entrepreneur correlates with a higher probability for investment (Greenberg & Mollick, 2017; Lounsbury & Glynn, 2001, 2019). It also suggests that the transparent and honest management of investors' cognitive and pragmatic expectations helps to contain the risk of losing previously established legitimacy (see "Paradox of Legitimacy", (Garud et al., 2014) and may be key to an expedient fundraising process, independent of venture stage or type of investor. The finding overall confirms that careful stakeholder management is the basis for a successful fundraising process (Lounsbury et al., 2019; Zott & Huy, 2007).

4.2.4. Story revision

"You don't tell the same story that you told a year ago. So sure, parts, components remain the same. But overall, it develops over time. You understand the company better, you understand the market better, you have different numbers to show. So yes, the story changes." – Theo

Interviewing many entrepreneurs who covered the fundraising process for their ventures over a longer time allowed me to assess their reflection on the development of their fundraising stories. Those insights do not refer to investor-based story adaptation, which was covered in section 4.2.1., but to the revision of the content and focus of the fundraising story. First, story revision in terms of wording can be based on investor feedback and learning-by-doing: "If you pitch all the time and have your 10 investor meetings per day, then you immediately learn which sentence is good and realize from the reaction which sentence is not so good. And then the story just changed that way" (Hannah). Second, "fundraising changes a lot with the stage, so pre-seed is different from seed, is different from series A, is different from collecting 100 million in debt. In other words, fundraising must always be very individually tailored to the respective round" (Theo). The findings show that what changes are the venture story, as well as the story focus, which shifts from founding team to numbers. The venture story changes over time because markets change, the team grows and learns, and the business model changes and manifests with traction. The founders can revise their story to say "look, we are able to generate learnings from the first traction and we are able to derive clear to-dos from these generated learnings, which we also implement and realize" (Ben).

Generated learnings can cause a venture to pivot, which of course triggers the need for story revision. While inexperienced entrepreneurs may hold on to their original idea in order not to disappoint investors, the findings show that pivots are mostly anticipated by the investors. As Leon learned: "If a person sends me a pitch deck in an early-stage, the idea will never end up being the idea or business model that the founders will still be interested in 2-3 years from now. The founder has to realize quickly enough when to pivot." To legitimize the revised vision and new valuation of a venture in each fundraising round, new meaning has to be created, explaining the development of the venture and the reasons for the pivot (Garud et al., 2014; Martens et al., 2007). The fact that pivoting is very common offers another explanation for why the founding team is the absolute fundament of investment decisions in early-stage ventures.

Interestingly, while the venture story changes with each stage, all interviewed entrepreneurs agreed that their personal stories remain the same to maintain their established personal legitimacy. They may only convey more expertise in later rounds, as the entrepreneurs become more confident and knowledgeable with growing experience. They also described that over time the focus of the fundraising story shifts away from the personal story of the founding team. While in early-stage rounds the team is the biggest concern of the investors, "at some point, the team has already been checked, people know that it is legit, and especially if other externals have talked about it a few times in a positive way, then the team is not the concern of the investors anymore" (Theo). This finding is also valid for serial entrepreneurs founding a new startup, regardless of the fundraising stage, as they have been legitimized several times before.

Aside from the legitimization of the founding team, a more established venture depends on more success factors than just the founders, such as "team, traction, execution" (Ben), which therefore move into focus. Simultaneously, numbers regarding the venture's traction become more important as its vision is slowly becoming more tangible (extended in Table 8). As fundraising decisions rely less on the emotional involvement and personal connection of the investor to the entrepreneur in later-stage funding rounds, the founder may have to use storytelling to ensure investors of still being the right person to manage and lead a bigger company, as highlighted by Toni:

"What I see with people who are later [stage], like [Series] C, you no longer have to explain why you are motivated and so on. [. . .] It's more about being able to show that you've got what it takes to be the right man at the top of the company, even in the later phase. Because it's a completely different job to lead a seed company as CEO and founder than it is to lead a Series C company."

All those findings imply that fundraising stories are dynamic. They can be revised, to mirror learnings, market developments, and subsequent shifts in the product, business model, and vision of the venture. The fundament of the story to be perceived as legitimate remains the personal story of the founding team, which stays the same with regards to motivation, background, and personality, but must reflect personal growth with regards to management skills and expertise.

Table 7: Expectation management.

	Approach	Statement
Max	Common vision	And I have to say quite honestly that I don't really care very much whether I disappoint them or not. To be honest, we have a common goal, and in our stage, there are not really any discrepancies yet.
Tim	Common vision	I'm honestly not that guy who scores so much with personality, I think. I try to be very clear about the business so that the investors have the feeling: 'oh, this is the next big thing! And the founding team or management has a blatant handle on that. And that's why I want to join in!' I think there are other founders who are much more about themselves. I try to get them to say okay, that's a good guy, because he's got his numbers under control, rather than saying that's a good guy because he's selling himself so well.
Willi	Common vision	It is important to communicate the goal and to describe your goal in a picture. To say I want to build a company that makes 100 million in sales is relatively banal, I think. That's a relatively simple story that anyone can tell. But to paint a picture and say the future of mobility looks like this: XYZ. And to get there I have already hired 10 strong people [...] and the picture where we want to go is: we want to become a listed company, because this market has a huge potential, and we believe that we can win 2% of this market by doing ABCDE. You must get very specific to paint that kind of picture.
Max	Realistic action plan	It's not that we say that we want to somehow talk ourselves up as a team. It comes across automatically when you ultimately draw up the storyline of the equity story and so on, because the investors already see that they have a strategy behind it. They have a vision. They have a clear plan of how they want to get there and can tell me exactly what they need to do it and what they can't do.
Tim	Realistic action plan	What I wouldn't pitch is: we stay 150 people and next year, by magic, so to speak, sales triple or something, because then I would feel like that falls back on me, as bullshit.
Ben	Realistic action plan	Being the lubricant, setting structures, defining focus, pulling out more and more from day-to-day business and actually trying to somehow continuously think about: Okay, where do we want to be in two quarters and what has to happen today to get there.
Nick	Realistic action plan	The biggest and strongest driver, where I say these are top top top founders and I'm excited, these are the people who bring across that they know exactly what they're doing. You just know, they know the space so much better than you do, and they just say we need the money now and then we're going to do this, this, this and this to make this happen. If you ask questions there, they can answer that so well that you just develop complete trust. Trust building is very important there.

Source: own illustration.

4.3. Understanding the 'VC game' and playing it right

After covering what entrepreneurs are pitching to investors about themselves and how they tell the fundraising story, I will next explore which mindset and process structure entrepreneurs adopt during fundraising. The entrepreneur's mindset with which he approaches fundraising determines his behavior throughout the process and fundamentally impacts how successful he will be in acquiring financial resources. At the same time, the way he structures the fundraising process determines, how his story is perceived by investors.

Understanding the dynamics between different investors, who are all on the lookout for the next big, promising deal, has strong implications for the way entrepreneurs interact with them and the signals that they want to send. "The VC scene is a very close-knit community, and everyone talks to one another" (Nick). A well-structured fundraising process can thereby increase the dynamic and trigger "fear of missing out" (hereinafter FOMO) among investors, which can be highly beneficial to the success and speed of the fundraising process. "It's a bit of a psychology game to be honest" (Julius). To win this game, the interviewed entrepreneurs have developed strategies, which will be presented in the following, by looking at the way they focus themselves before and throughout fundraising, structure the process and create hype among investors.

4.3.1. Adopting the right mindset for the fundraising process
"The good thing about fundraising is that, as I keep telling myself, you don't have to convince everyone. In the end, you only have to convince one or two, and I think it's almost better if there's a bit of nat-

Table 8: Story revision.

	Development	Statement
Michael	Constant	We are still working with my personal story and have now raised almost half a billion in capital and the story is still working quite well.
Theo	Constant	The core components remain extremely constant. I think I know very well why I am where I am, or what is important to me. [...] So, everything up to the startup I tell almost the same.
Ben	Constant	The story behind the founding idea has remained the same. Of course, when I talk a bit more about day-to-day business or the frequently asked question "What keeps you up at night?", that has changed.
Theo	Focus shift	In the beginning we still had [our personal story] very, very much as an integral part of what we do because it's also a little bit related to why we started the company.
Fred	Focus shift	In the beginning, people just looked at your resumé [...] and numbers were relatively unimportant. And now, if the sales and the gross margin are not right, then even the best story in our case doesn't help.
Fred	Focus shift	I think if you look at it a little bit analytically, then the success factors were quite few at the beginning and two of them were C. and me. And now the success factors have multiplied, and we are still only two of them. So, in terms of share, we actually make up less of the entire company than before.
Willi	Focus shift	But if you've built two startups, raised capital and sold them, then people assume you're not a total talker. So that [personal] part can be smaller, and it can be more about the business you're building.
Willi	Focus shift	It depends on the stage of the company. If you're looking for angel or seed funding, or maybe Series A, then there's a very very very strong focus on the founders, on the team. Logically, because otherwise there isn't much there yet. What revenues are you looking at in the seed phase? Maybe there are already sales, but they are not very meaningful yet. [...] And the further you get, i.e., in a Series B / Series C or in the exit process, it's very little about you as a person anymore.
Leon	Pivot	If a person sends me a pitch at the early stage, the idea will never end up being the idea or business model that the founders will still be interested in 2-3 years from now. The founder has to realize quickly enough when to pivot.
Tim	Pivot	We also completely pivoted H., the product and the target group and so on. I think experienced founders simply do that, while some of the younger founders, who are a bit naïve, stick very rigidly to their ideas. And that's not always healthy, so to speak.

Source: own illustration.

ural selection." – Tim

All interviewed entrepreneurs described fundraising as a very intense and consuming process. Especially first-time founders who raise capital for their early-stage ventures face a lot of rejection and have to try not to get discouraged. According to Theo, "you're just used to catching all the No's at some point. It still really hurts. It hurts your ego, and it also hurts you personally. Maybe you've heard this before, the startup is just such an extension of [the founder] so to say, so you're very intertwined with it, which means that every no you get is almost a personal attack, or at least it feels that way." Keeping in mind that only very few investors must be convinced of the venture to secure funding may help the entrepreneurs to save time and emotions. "Don't waste time on the No's!" (Tim) is one learning that many of the entrepreneurs share. Focusing to convince those investors who were already excited about the founders and their venture, instead of trying to turn around those who were not, helped

them to stay focused and optimistic.

One way to avoid facing too much rejection is by strategically targeting investors, who could generally be interested in the venture: "My key learning is that you have to talk to the right investors and that you can't expect to convince someone who has no idea about your product. Getting into the room with the right people, that's the most important thing" (Fred). Product-related knowledge is one way to select target investors, another is the focus on a bold vision or the size of the funding round, which requires speaking to investors with a certain mindset or cultural background (see section 4.2.1.). For this reason, Michael for example mainly spoke to investors in the US and targeted angel investors, who are "always involved in exciting topics that somehow seem to be on the edge of feasibility."

Again, targeting investors with a similar educational, professional or cultural background to the entrepreneur can help to foster a personal connection and thereby perceived legiti-

macy of the venture (Greenberg & Mollick, 2017). Knowing that the decision to invest in early-stage ventures is predominantly based on personal connection can put a lot of pressure on the entrepreneur throughout and after the fundraising process. Analyzing the interviews showed that how well entrepreneurs cope with this pressure is again related to their mindset (see Table 9). Some see the responsibility to live up to the investors' expectations to generate good returns for their investment very much on themselves and feel like "kind of disappointing them if I don't make it now" (Hannah). This mindset links to the Paradox of Legitimacy, which describes that during fundraising expectations are created, which may be hard to meet and could lead to a subsequent loss of legitimacy (Garud et al., 2014).

Other entrepreneurs try to manage the pressure by rationally reminding themselves that those who decide to invest act in their own interest and also have to carry the consequences themselves: "at the end of the day, they want to make more money out of a lot of money. And that's a relatively cold investment business" (Toni). Again, others remind themselves that both parties follow the same interest: "In the very beginning it was really like 'Crazy there's somebody with a lot of money and me with my little stupid idea and I want something from that person.' And that actually changed completely into 'Hey, they actually want something from you'" (Ben).

One last targeting strategy, which can be highly beneficial for the development of venture is focusing on investors with a good network, e.g., angels who are also LPs at other investment funds and can help to secure follow-on rounds as described by Michael:

> "I think it's super important [. . .] to consider which are the angels who have a real network. In the meantime, I myself, but in the past my friend C.A., who sometimes got into companies at brutal discounts, but then made billion-dollar companies out of them with very high reliability, because he was able to get the right investors in. In that case, you can easily afford to give up a few percentage points for that."

Of course, this kind of selective investor targeting is easiest for serial entrepreneurs, in follow-on rounds, or for those founders who are generally in a very strong fundraising position. However, this finding implies that strategically targeting investors may increase the efficiency of the fundraising process for all entrepreneurs, while a more focused mindset also helps entrepreneurs not to get discouraged.

4.3.2. Process design

> "To structure the process well is so, so, so important. A structured process makes everything easier because then you can just focus completely on the fundraise." – Theo

One reason why fundraising was described as such an intense and consuming process throughout all interviews is that most conversations and negotiations with investors take place in the course of a very few weeks. If entrepreneurs have full control over shaping the process, why would they make it so stressful for themselves? This is because the process design impacts the perception of the fundraising story and hence the decision of investors. As investors speak to each other a lot about current investment opportunities, being contacted by an entrepreneur months after first hearing about the venture from other investors implies that the entrepreneur was not able to secure funding fast enough and makes the deal appear less attractive.

"Doing a process", as it was described by many interviewees means putting full focus on fundraising and contacting all target investors at the same time. "A structured process helps brutally because then you have everyone on the same timeline and can create urgency" (Tim). "You send all the emails in one day with a scheduling link. Then you have 2-3 weeks of pure meetings. You can't really do anything at the same time anyway; after every pitch, you're so excited, it's like when you just come off stage" (Hannah). The same urgency should be maintained throughout the negotiations: "From term sheet to closing, somehow try to get it through in four weeks. Push the lawyers every day and ask how far the draft is now or what is causing the hold-up to speed up the process" (Tim, extended in Table 10). That way investors do not have the time to question or back out of the deal.

Michael, who has lived through the fundraising process many times summarizes the work that has to be done in this short amount of time and highlights the importance of warm introductions and in-person meetings:

> "In the end, it's just hard work. Building Excel lists and scrubbing through them, finding out via LinkedIn who knows whom to get warm intros - always trying to get warm intros! Everything that comes in via the info@ address is kind of dead. And just don't give up, keep calling people and ideally meeting them! That's also something that you think about, okay which are the relevant cities? For a European, it's kind of Berlin, Zurich, London, Stockholm and say hey, I'm renting an Airbnb for two weeks, I'm on the ground I'm meeting the VCs live, they want to see you, they want to shake your hand and they want to see that you're a proactive, happy person. That's it, it's all not witchcraft."

Two aspects of the process were repeatedly mentioned to require careful planning: strategic timing and distribution of information. To determine the right timing to raise capital, entrepreneurs can first look at the burn rate of their venture. To build a power position in negotiations with investors, they should fundraise rather earlier than later. "If you are in urgent need of money, then you are in such a bad position that every investor can take advantage of you" (Jacob). Second, entrepreneurs can look at business performance. When are the numbers best to paint a convincing picture of the venture's traction and potential? "If it's a seasonal business, then

Table 9: Setting focus.

	Approach / Mindset	Statement
Fred	Choosing target investors	You have to talk to the right people. In our case, we just needed people who were into brand, who were into direct-to-consumer and who understood a little bit about this space.
Fred	Choosing target investors	We're better at picking and choosing our conversations. So now we only talk to people where we know okay, they're probably interested, [. . .] and go into this meeting with more authority.
Leon	Choosing target investors	If you have certain investors on your CapTable, most of whom are also LPs at certain funds, it gets easier.
Finn	Choosing target investors	It may very well be that I go fundraising again. But then much more selectively and with an all or nothing approach. So I only take those I want to work with and not everyone who gets in front of me.
Max	Dealing with rejection	You need a thick skin. No matter what you do, investors think they know everything better anyway. [. . .] But no one knows the company as well as you do as a founder. If someone has a different opinion, then it doesn't fit. I think we could have saved ourselves a lot of time and emotions in the seed round.
Tim	Dealing with rejection	I think many inexperienced founders try to somehow turn someone around [...] but I'm not even doing that.... So if the VC is not in the first call already half in love with what I'm doing, then there is no point for me to invest any more time on him, but I would then invest all the time with the 4, 5, 6, who have somehow fallen in love with it and try to convince them.
Ben	Handling pressure	Funnily enough, just at the beginning, I was told: "The reason why I invest is you". Which of course builds up a bit of pressure.
Toni	Handling pressure	I'm a very, very, very rational person and think to myself, if I screw up, then they've lost a bit of money. But they haven't invested as much blood, sweat and tears in it as I have. For myself it is a bigger fail. With Angels I would be sorrier, because maybe they are not all multi-billionaires.

Source: own illustration.

you should raise just before the peak, so that positive numbers pour in during the process" (Nick). Third, entrepreneurs must keep in mind that "VC is very trend-driven. When something is hyped, it's relatively easy to raise money" (Michael), so depending on the product or business model of the venture, the trend development plays a role in timing. Fourth and last, the availability of cash in VC can play a role in good timing. The current situation (Section 3.1.) can present "an opportunity to do the next fundraise even earlier" (Toni) at potentially higher valuations.

Regarding the distribution of information to investors, the interviewed entrepreneurs follow different strategies. "There are two different philosophies on the market among the founders. There are those, who do five coffee chats with investors every week. [...] Or there's me, who says be raising or not be raising. I don't talk to any investor for two years and then go out and try to talk to 50 in a week" (Tim, extended in Table 10). The advantage of Tim's strategy may be that the process is easier to structure and all investors have the same information. However, reaching out to or keeping in touch with investors without actually fundraising can create momentum and accelerate the upcoming funding round, while sharing less information. Whichever strategy

they are following, all entrepreneurs stressed the importance of being careful who to share sensitive information about the venture with. One reason for this is that information travels very fast among investors and the other is that investors may have other portfolio companies this information could be interesting for.

Those findings suggest that the way an entrepreneur designs the fundraising process directly influences the perceived legitimacy of the venture. Sharing the right information at the right time increases the perceived potential of the venture, legitimizing its vision and valuation. Simultaneously, a strategic fundraising process draws a lot of attention to the venture and creates a sense of urgency among the investor community, which increases the dynamic and hype around the deal.

4.3.3. Creating hype

> "It's always a bit of an issue that you want something from someone else. It means you're in a weaker position, to begin with. [. . .] What you have to do as a founder is turning the tables and say, hey, here is this opportunity and actually we don't need any money, but if you say please, we

Table 10: Process structuring.

	Strategy	Statement
Tim	Outreach	There are two different philosophies on the market among the founders. There are those, who do five coffee chats with investors every week. [...] Or there's me, who says be raising or not be raising. I don't talk to any investor for two years and then go out and try to talk to 50 in a week and do a hell of a structured process. Hopefully 10 will find it exciting, 4 do a term sheet and I'll take one. And then - I don't want to say lock myself back in for 2 years, but then I'm working on the company again.
Nick	Outreach	My biggest tip for founders is, start talking to VCs before you go fundraising. Then you already have the contacts, had a touchpoint and you can be very, very selective and strategic about what data you want to give out. And then when you really go into fundraising, you just call XYZ and say hey, here we go. We've developed so and so well, let's do it. That way you can squeeze many calls and momentum into a very, very small period of time and that already creates a momentum.
Simon	Information sharing	Think very strategically about what information you want to share at what point in the journey. We also had a VC, who took all our documentation from the data room and we never heard from him again. [. . .] We have now seen, oops, they invested in a venture 3-4 months ago, they do something similar to what we do. They have now sucked our information and passed it on to their portfolio company. You have to be prepared for something like that, we were a bit too gullible sometimes.
Michael	Information sharing	There are a few tricks you can do. If you're a first-time founder, or you're just starting out with your seed company, you can put together a nice deck where you don't give away too much information, distribute it through your network and say hey, I'm not raising any money right now, I'm self-funded, but wanted to get some feedback. That way you don't get a No from VCs, but some possibly approach you proactively and say, hey I really want to get into the deal, how about we do a pre-round.
Nick	Timing	As a VC you often think, oh, why does he have to go fundraising now? In 2 months, we would have certainty on the topics XY, and could make a decision with a much better conscience.
Tim	Timing	We would have raised money last year if it hadn't been for Corona, which cost us a bit of growth. We've gotten through it really well, but it wasn't the best time to do a huge round for a recruiting company. Let's rather wait until Covid is over and growth is really explosive again and then raise.
Tim	Timing	From term sheet to closing, somehow try to get it through in four weeks. Push the lawyers every day and ask how far the draft is now or what is causing the hold-up to speed up the process. Sometimes it's good if it's not done by the founder himself, but by his right-hand man. [. . .] I've seen with many junior founders that they somehow leave it open for too long and then it's all lingering. [. . .] I'm someone who asks, do you want to invest now or not? Both are fine for me. But let's get to a decision.

Source: own illustration.

will take you with us. That's something you have to practice over time, to somehow build up this fear of missing out." – Michael

According to all interviewed entrepreneurs, investment decisions are often made under the pressure to be part of a deal that is highly popular among investors. In this case, investors start competing for high potential deals and develop FOMO, which can lead to a real power shift, as entrepreneurs suddenly get to pick and choose between investors. "Until you get the first yes, [fundraising is] a real pain, but as soon as you do, you're already kind of overwhelmed with offers. It's really just about getting that first good term sheet" (Theo). The creation of hype relies on the fact that investors are part of a very close-knit community, in which information travels fast. Hence, when hearing about a high-potential ven-

ture, investors may proactively reach out to the entrepreneur to communicate their interest, which further increases the hype. All investors who are then trying to get in on the deal, start competing and outbidding each other with regards to their offered terms, transaction speed, valuation, etc. This implies that hype legitimizes and correlates with higher venture valuations.

This kind of hype can be created through a lead investor with a strong signaling effect (see section 4.1.2.): "there are certain characters, if you have them as angels, then there are VCs who get FOMO" (Toni). Similarly, reputable investors who speak positively about a venture or its founding team can create FOMO. However, the dynamic among investors can also turn negative. "The problem is that if one of the good funds with a signaling effect drops out, word gets around

very, very, very quickly and then a deal loses momentum" (Nick). Entrepreneurs can partly avoid losing dynamic by designing a structured process (see section 4.3.2.): "Doing a proper process is extremely important because these investors all talk to each other and create so much FOMO. And that's the best thing that can happen to you as a founder because then people start going, 'Hey, by the way, you don't know us yet, but we wanted to talk to you too!' And there you have it. It's so powerful and you don't really have to do much, except schedule meetings correctly" (Hannah).

Hype can also be artificially created through playing investors off against each other through communication techniques: "We've gone from the point of 'we want money now and we're happy if we can get some' to saying 'this is what we want, and these are the terms. Are you up for it? If so, you get a data room. Period.' [...] Many are impressed and think 'Fuck, if they act like this, there must be others who already have the data room'" (Max). Table 11 contains more examples of how entrepreneurs gamble with investors' attention to create hype around their ventures.

Overall, those findings imply that the hype around a venture within the investor community increases its perceived legitimacy. As hype is closely connected to the design of the fundraising process, the strategic behavior of the entrepreneurs foreseeing the dynamic within the investor community and creating a sense of urgency around the deal directly influences the investor's decision-making process.

4.4. Summary: Propositions and process model development

The results were presented in three dimensions: 1) entrepreneur-based legitimacy levers, 2) investor-based story adaptation, and 3) understanding and mastering the 'VC game'. In essence, they give insights into the interviewees' accumulated learnings on 1) how they effectively pitch themselves, 2) how they effectively pitch their ventures, and 3) which mindset and process structure they adopt to establish legitimacy and increase their chances for positive investment decisions. From the presented results, four propositions can be derived.

Entrepreneurs' personal story, although only covered in a small part of the pitch, plays a highly important role in establishing legitimacy. Especially first-time founders of early-stage ventures can increase their chances to get funding by constructing a meaningful, consistent and emotional story around their background and commitment, conveying expertise and emotional conviction for driving the development of their venture. It is also their chance to distinguish themselves from other entrepreneurs, who may be working on a similar project. Even when the venture story is adapted or revised, the personal story must remain predominantly constant to maintain previously established legitimacy. The findings on personal storytelling induct that

P1: *Telling investors a consistent personal story to make meaning of the entrepreneur's background and commitment establishes perceived legitimate distinctiveness.*

For investors to perceive the vision and valuation of a venture as legitimate, the venture story must be adapted within each pitch. Approaching fundraising like a sales job, finding out about the interests of investors, and tailoring the story to match those interests can decrease the perceived risk of the investment. When progressing to the next venture stage or pivoting the venture's business model or product, the venture story must be revised accordingly to closely manage the investor's expectations and aligning the entrepreneur's and investor's vision of further venture development. The findings on adapted storytelling induct that

P2: *Adapting the venture story to match the investor's interests and the venture stage is necessary to establish or maintain the perceived legitimacy of the venture's vision and valuation.*

Besides the pitch, the entrepreneur's behavior throughout the fundraising process impacts the venture's perceived legitimacy. This includes the mindset and endurance the entrepreneur approaches the process with and specifically shows in the way the entrepreneur structures the investor outreach and information sharing. By strategically using the exchange between investors and keeping them on the same timeline, the dynamic and perceived urgency of the process and thereby perceived legitimacy of the venture can be influenced. The findings on strategic behavior induct that

P3: *The entrepreneur can impact the dynamic of the fundraising process and thereby the venture's perceived legitimacy through strategic investor outreach and information sharing.*

Overall, the findings of this study show that investment decisions are not purely rational or fact-based. Instead, they can be influenced by the entrepreneur through strategic storytelling and behavior. Two factors can be highlighted as having a considerable impact on perceived legitimacy and hence the investment decision and can be considered "irrational". The first one is emotional involvement by the investor, triggered by compelling storytelling, perceived cultural proximity, and a personal connection to the entrepreneur. The second factor is perceived urgency that is based on the dynamic of the venture's fundraising process in the investor community, which is caused by the strategic behavior of the entrepreneur. Those findings induct the final proposition

P4: *The investor's decision-making process is influenced by his emotional involvement in the fundraising story and the perceived urgency of the investment decision.*

The four developed propositions can be combined in the process model of strategic storytelling and behavior in entrepreneurial fundraising, shown in Figure 4. The model does not include a temporal perspective, as the entrepreneur's strategic behavior would then cover the entire fundraising process, while the personal and venture story only concerns the pitch event. For simplicity, the model focuses on describing the impact of storytelling (including the personal and venture story) and strategic behavior on the venture's perceived legitimacy, which then builds the basis for the investment decision. It implies, that emotional involvement and perceived urgency not only impact perceived legitimacy but can have a direct effect on the final investment

Table 11: Creating hype.

	Trigger	Statement
Jacob	FOMO	You have to consistently try to create an artificial or non-artificial dynamic in fundraising and this dynamic is best when you don't need money in that moment.
Jacob	FOMO	They have a herd mentality. I mean, they all get FOMO when you've already gotten a commitment, and then the other one gets scared that they don't get in on the deal.
Toni	FOMO	Then there are real investors, who want to invest in your venture just because you have him as an angel. [...] If he just stands in the cap table and says nothing, it's different than if he somehow also says in reference calls, yes, the guys are the raddest and the hottest. Then they get FOMO again, just because they want to be in the cap table with H.K.
Michael	FOMO	This then very quickly takes on the character of a self-fulfilling prophecy, because you have P.T. with you, he's always had a golden touch, here's my chance to make a deal with P.T., which then suddenly gets you a completely different interest for your venture.
Julius	Gamble	You can game so much, it's all about psychology. You build a list with all investors who were interested at some point, then send them emails, then don't answer them, then the first ones become interested because you don't answer. And all the investors talk. They are like chatty high school girls. It's always quite funny when you say something to one of them, to see whether it gets through to the others.
Julius	Gamble	This week, we have scheduled again when we send which emails to which investors. Our strategy right now is, we have an investor who is the Rabbit who will give the first term sheet with a high probability. And once you have the first term sheet, it becomes much easier with all the other investors because then you can always say, I already have a term sheet, I like the investor, you then bring them to the decision more or less. It's fun!
Toni	Gamble	To be really tough in negotiations, to be blatantly ready to take the risk to play them off against each other, to create FOMO, to tell in doubt: we already have 300 term sheets, even though you don't - that can make the difference between a valuation of 10 or 20 million for the same company.
Nick	Gamble	Play the Game right. The VC scene is a very close-knit community, and everyone talks to one another. The problem is that if one of the good funds with a signal effect drops out, word gets around very, very, very quickly and then a deal loses momentum.

Source: own illustration.

decision. The theoretical and practical implications of the propositions and model are discussed in the following.

5. Discussion

This inductive study of 15 entrepreneurs and investors and the analysis of their fundraising experiences and strategies have opened up new perspectives on the establishment of legitimacy. The findings have theoretical implications for further research in cultural entrepreneurship and important practical implications for entrepreneurs who seek to acquire financial resources.

5.1. Theoretical implications

This study offers one of the first assessments of personal accounts of experiences and strategies in entrepreneurial fundraising. Although the establishment of distinctive legitimacy is a difficult and important task for entrepreneurs, cultural entrepreneurship literature mainly depicts the establishment of legitimacy as an outcome and not as a strategic process (Werven et al., 2015). Those researchers who so far contributed to unraveling the underlying process, have mainly taken a linguistic perspective, analyzing the rhetoric of fundraising pitches (Cornelissen & Clarke, 2010; Ruebottom, 2013; Werven et al., 2015). In this study I have taken a more structural perspective, identifying three ways in which legitimacy can be established throughout the fundraising process: personal storytelling, venture storytelling, and strategic behavior. Interviewing entrepreneurs and investors who can offer different perspectives on the topic or have gone through the fundraising process multiple times allows assessing their learnings and the development of their approach to fundraising over a longer time. This study, therefore, enables the identification of underlying fundraising strategies which could not be discovered using the linguistic perspective and therefore offers a more holistic assessment of the practical establishment of distinctive legitimacy (Gioia, Corley, & Hamilton, 2013; Lounsbury & Glynn, 2001, 2019).

By assessing the process of establishing legitimacy through the lenses of entrepreneurs and investors, this study also tests whether the theoretical process model of cultural entrepreneurship (Figure 1, Lounsbury & Glynn, 2019) can be
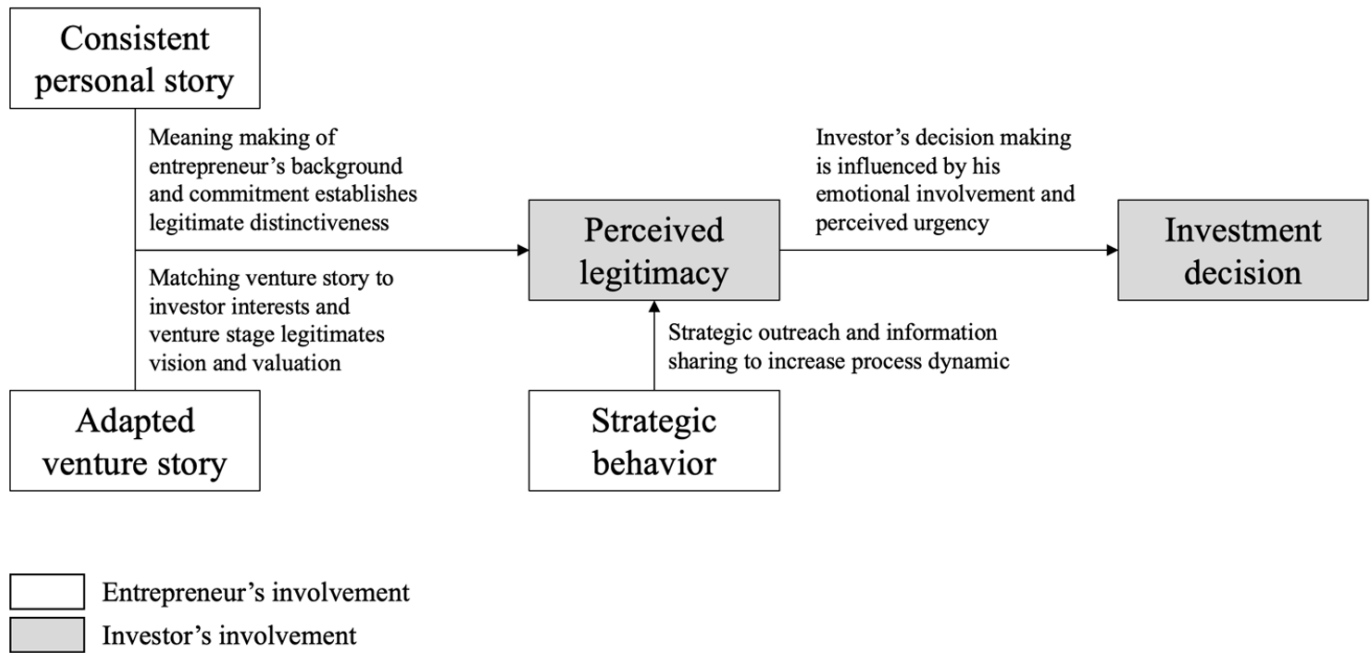
Figure 4: A process model of strategic storytelling and behavior in entrepreneurial fundraising.

Source: own illustration.

transferred into practice. The core concept of cultural entrepreneurship, which is for entrepreneurs to use storytelling to establish legitimacy in the process of resource acquisition holds true (Lounsbury & Glynn, 2001, 2019). The findings of this study confirm that entrepreneurs adapt the venture story based on the audience to bridge information asymmetries stemming from differing cultural backgrounds and narratives of the entrepreneur and the investor (Garud et al., 2014; Manning & Bejarano, 2017; Martens et al., 2007). Using audience-specific rhetoric (Wry et al., 2011) and referring to familiar concepts, categories, or discourses to make meaning of the venture's organizational identity and vision (Garud et al., 2014; Lounsbury & Glynn, 2019; Martens et al., 2007) is common practice among the interviewed entrepreneurs.

This study also confirms that fundraising stories emphasize distinctiveness more, the more established the venture is (Lounsbury & Glynn, 2019). The findings show a shift of the story focus away from explaining the venture idea and background, towards numbers and performance metrics that distinguish the venture as it progresses to later stages. However, this study shows that story development also depends on the cultural context (Zhao et al., 2017), as e.g., US VCs may assess distinctive legitimacy differently than European investors. Another interesting finding concerns the Paradox of Legitimacy (Garud et al., 2014), showing that the loss of legitimacy is not caused by disappointing investors' expectations (Zimmerman & Zeitz, 2002) by pivoting the venture's product, business model, or vision. Instead, pivots are mostly anticipated by investors but require their constant expectation management and the revision of the venture story to

explain the changes and consequences. While revising the venture story, entrepreneurs must keep their personal story constant to maintain previously established legitimacy after a pivot.

Two key findings of this study contradict the current design of the process model of cultural entrepreneurship (Figure 1, Lounsbury & Glynn, 2019). First, this study suggests that besides storytelling, entrepreneurs' strategic behavior plays an important role in establishing legitimacy (Figure 4). Although Martens et al. (2007) and Lounsbury and Glynn (2019) acknowledge that entrepreneurial behavior can have a significant impact on the investor's judgment about a venture's potential, "we know little about what specific entrepreneurs' behavior increases the propensity for this type of resource acquisition" (Pollack et al., 2012). This study contributes to filling this gap in the existing literature by showing how entrepreneurs can directly influence investors' perceived legitimacy and decision-making processes concerning the venture through strategic investor outreach and information sharing. Hence, this study recommends the extension of the cultural entrepreneurship process model by Lounsbury and Glynn (2019) to include the dimension of behavior.

Finally, this study contradicts the theory that the legitimacy of a venture is based on its' existing resources (Lounsbury & Glynn, 2001, 2019), which is deeply anchored in the process model of cultural entrepreneurship and overall in the existing literature on cultural entrepreneurship. The findings demonstrate, how the establishment of legitimacy directly depends on the entrepreneur's personal story, his ability to adapt the venture story according to the audience, and his

strategic behavior. Hence, it can be argued that especially in early-stage ventures, the entrepreneur is at the core of the fundraising process, instead of the venture's resources, as not many resources exist at that point of the venture's development. Although the theory that ventures' legitimacy is established through the narrative of the entrepreneur and that the venture's organizational identity builds on the entrepreneur's identity (which was also confirmed by this study's findings) already suggests that fundraising success does not depend on the venture's resources, but on the entrepreneur, this notion is still underrepresented in the existing literature on cultural entrepreneurship. In line with Steyaert (2007), this study argues that the existing literature pulls the attention away from the entrepreneur, although he is at the center of the audience's attention and hence the decisive factor for successful resource acquisition.

5.2. Practical implications

The main practical implication of this study is that investment decisions are not purely rational. Although investors may study hard facts, such as the entrepreneur's educational background or the venture's traction metrics to assess the potential or risk of an investment, the story that is built around those facts strongly influences the investor's decision. Entrepreneurs who seek to acquire financial resources for their venture need to be aware that telling a consistent personal story, adapting the venture story based on the audience and venture stage, and behaving strategically fundamentally affect the perceived legitimacy of the venture and can significantly increase chances for successful fundraising.

Consistent personal storytelling helps investors to understand the entrepreneur's distinctive background and commitment to the success of the venture. Involving the investor emotionally by sharing the personal story behind the venture helps to build trust and foster a personal connection between investors and entrepreneurs. Tailoring the venture story according to the interests, expectations, and agenda of the audience and revising it based on the venture stage helps to convince investors by legitimizing the venture's vision and valuation. Lastly, strategic behavior in terms of strategic investor outreach and information sharing helps to draw attention to the venture and create competition and a sense of urgency among investors to become part of the deal. Understanding the way, the investor community exchanges information about high-potential deals helps entrepreneurs to strategically plan and implement their own participation in the 'VC game'.

Three factors need to be kept in mind to impact the fundraising story and perceived venture legitimacy: the venture stage, the founder's entrepreneurial experience, and the targeted type of investor. First, as ventures progress to later stages and the founding team's legitimacy has already been established in earlier fundraising rounds, the focus of the story shifts from emotional personal storytelling, towards presenting numbers and facts, which show the venture's progression and realization of its goals and vision. Still the entrepreneur's personal story concerning his background,

commitment, and personal connection to the investors needs to be kept consistent, to maintain previously established legitimacy. Second, first-time founder fundraising stories may differ from those of serial entrepreneurs. A serial entrepreneur, who has successfully founded and exited ventures in the past, benefits from previously established legitimacy when raising early-stage funding for a new venture. Accordingly, his personal story may still have to explain his commitment to the new venture but does not need to focus on proving his expertise, as his background is already legitimate and distinctive. Third, establishing legitimacy with an angel investor, family office or VC may work differently. Each investor has an individual way of assessing a venture's potential and values legitimacy levers differently. The results of this study suggest that especially VCs often conduct a more thorough and holistic assessment, while angel investors' and family offices' investment decisions rely more heavily on emotional involvement and personal connection to the entrepreneur. This may be connected to the respective venture stage, as e.g., angels who invest in early-stage ventures have fewer data to assess and base their decision on than VCs who invest in later-stage ventures with a more extensive track record. Although the impact of the venture stage, founding experience, and investor type on the fundraising story and process are very case-dependent, all three factors need to be considered by entrepreneurs when preparing a new fundraising round.

5.3. Limitations and suggestions for future research

Due to the nature of case studies, the developed propositions are based on the accounts of a relatively small sample of interviewees and are limited in their generalizability (Eisenhardt, 1989). Additionally, the sample is relatively heterogeneous, including first-time founders and serial entrepreneurs in different venture stages, industries, and locations, as well as investors, thereby giving a high-level overview of different fundraising strategies. Future research should validate the proposed effect of consistent personal storytelling, adapted venture stories, and strategic behavior on perceived legitimacy on a larger sample. Each of the three categories should be researched individually, exploring the respective fundraising strategies in more depth, and jointly for a more homogenous sample e.g., first-time founders targeting angel investors. As this study gives a predominantly entrepreneurial perspective on the topic, future research may also test the propositions from an investor perspective e.g., interviewing only VCs.

As this study relies on the subjective accounts of interviewees, another limitation regards the potential misjudgment of effects of certain fundraising strategies and biased recollection of past fundraising experiences. Future research should therefore use other methods to test and triangulate the developed propositions, e.g., conducting longitudinal case studies to explore the development of fundraising stories over time (Eisenhardt, 1989; Yin, 1994).

This study focuses only on the acquisition of financial resources, while the process model of cultural entrepreneur-

ship concerns all forms of entrepreneurial resource acquisition (Lounsbury & Glynn, 2019). Although this limitation does not imply that the findings of this study cannot be transferred to other forms of resource acquisition, future research needs to further explore the role of the entrepreneur's personal story, adaptation skills, and behavior in the acquisition of all types of resources (Martens et al., 2007). That way, future research can contribute to introduce and manifest an entrepreneur-based narrative within the cultural entrepreneurship literature (Steyaert, 2007).

5.4. Conclusion

This inductive study suggests a strong positive impact of consistent personal storytelling, venture story adaptation, and strategic behavior on perceived venture legitimacy in the process of entrepreneurial fundraising. The findings show that besides the factual assessment of venture potential and risk, emotional involvement and the dynamic around a venture within the investor community influence investment decisions. Entrepreneurs can therefore significantly increase their chances for financial resource acquisition by strategically approaching the fundraising process. The findings of the learnings and strategies of experienced entrepreneurs offer strong practical insights for entrepreneurs, who seek to improve or validate their fundraising preparation. By exploring the establishment of legitimacy as a process, not as an outcome, this study proposes an entrepreneur-centric perspective on the process of resource acquisition through cultural entrepreneurship.

References

Aldrich, H. E., & Fiol, C. M. (1994). Fools Rush in? The Institutional Context of Industry Creation. *Academy of Management Review*, *19*(4), 645–670.

Bartel, C. A., & Garud, R. (2009). The Role of Narratives in Sustaining Organizational Innovation. *Organization Science*, 12.

Bruni, A., Gherardi, S., & Poggio, B. (2004). Entrepreneur-mentality, gender and the study of women entrepreneurs. *Journal of Organizational Change Management*, *17*(3), 256–268.

Corbin, J., & Strauss, A. (2014). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. SAGE Publications.

Cornelissen, J. P, & Clarke, J. S. (2010). Imagining and Rationalizing Opportunities: Inductive Reasoning and the Creation and Justification of New Ventures. *The Academy of Management Review*, *35*(4), 539–557.

de la Cruz, M. E., Verdú Jover, A. J., & Gómez Gras, J. M. (2018). Influence of the entrepreneur's social identity on business performance through effectuation. *European Research on Management and Business Economics*, *24*(2), 90–96.

Down, S., & Reveley, J. (2004). Generational Encounters and the Social Formation of Entrepreneurial Identity: 'Young Guns' and 'Old Farts'. *Organization*, *11*(2), 233–250.

Edmondson, A. C., & Mcmanus, S. E. (2007). Methodological fit in management field research. *Academy of Management Review*, *32*(4), 1246–1264.

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.

Fauchart, E., & Gruber, M. (2011). Darwinians, Communitarians, and Missionaries: The Role of Founder Identity in Entrepreneurship. *Academy of Management Journal*, *54*(5), 935–957.

Fisher, G., Kuratko, D. F., Bloodgood, J. M., & Hornsby, J. S. (2017). Legitimate to whom? The challenge of audience diversity and new venture legitimacy. *Journal of Business Venturing*, *32*(1), 52–71.

Garud, R., & Giuliani, A. P. (2013). A Narrative Perspective on Entrepreneurial Opportunities. *Academy of Management Review*, *38*(1), 157–160.

Garud, R., Schildt, H. A., & Lant, T. K. (2014). Entrepreneurial Storytelling, Future Expectations, and the Paradox of Legitimacy. *Organization Science*, *25*(5), 1479–1492.

Gehman, J., & Soublière, J.-F. (2017). Cultural entrepreneurship: from making culture to cultural making. *Innovation*, *19*(1), 61–73.

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

Gioia, D. A., Patvardhan, S. D., Hamilton, A. L., & Corley, K. G. (2013). Organizational Identity Formation and Change. *Academy of Management Annals*, *7*(1), 123–193.

Greenberg, J., & Mollick, E. (2017). Activist Choice Homophily and the Crowdfunding of Female Founders. *Administrative Science Quarterly*, *62*(2), 341–374.

Grimes, M. G. (2017). The Pivot: How Founders Respond to Feedback through Idea and Identity Work. *Academy of Management Journal*, *61*(5), 1692–1717.

Grimes, M. G., Gehman, J., & Cao, K. (2018). Positively deviant: Identity work through B Corporation certification. *Journal of Business Venturing*, *33*(2), 130–148.

Hegde, D., & Tumlinson, J. (2014). Does Social Proximity Enhance Business Partnerships? Theory and Evidence from Ethnicity's Role in U.S. Venture Capital. *Management Science*, *60*(9), 2355–2380.

Higgins, M. C., & Gulati, R. (2006). Stacking the Deck: The Effects of Top Management Backgrounds on Investor Decisions. *Strategic Management Journal*, *27*(1), 1–25.

Huber, G. P., & Power, D. J. (1985). Retrospective reports of strategic-level managers: Guidelines for increasing their accuracy. *Strategic Management Journal*, *6*(2), 171–180.

KPMG, Moore, C., Smith, K., & Lavender, J. (2021). *Q2'21 Venture Pulse Report – Global trends - KPMG Global.* https://home.kpmg/xx/en/home/campaigns/2021/07/q2-venture-pulse-report-global.html.

Lee, B., Hiatt, S., & Lounsbury, M. (2017, 5). Market Mediators and the Trade-offs of Legitimacy-Seeking Behaviors in a Nascent Category. *Organization Science*, *28*.

Lounsbury, M., Gehman, J., & Glynn, M. A. (2019). Beyond Homo Entrepreneurus: Judgment and the Theory of Cultural Entrepreneurship. *Journal of Management Studies*.

Lounsbury, M., & Glynn, M. A. (2001, 6). Cultural Entrepreneurship: Stories, Legitimacy and the Acquisition of Resources. *Strategic Management Journal*, *22*, 545–564.

Lounsbury, M., & Glynn, M. A. (2019). *Cultural Entrepreneurship: A New Agenda for the Study of Entrepreneurial Processes and Possibilities*. Cambridge University Press.

Manning, S., & Bejarano, T. A. (2017). Convincing the crowd: Entrepreneurial storytelling in crowdfunding campaigns. *Strategic Organization*, *15*(2), 194–219.

Marlow, S., & McAdam, M. (2015). Incubation or Induction? Gendered Identity Work in the Context of Technology Business Incubation. *Entrepreneurship Theory and Practice*, *39*(4), 791–816.

Martens, M. L., Jennings, J. E., & Jennings, P. D. (2007). Do the Stories They tell get them the Money They Need? The Role of Entrepreneurial Narratives in Resource Acquisition. *Academy of Management Journal*, *50*(5), 1107–1132.

Mmbaga, N. A., Mathias, B. D., Williams, D. W., & Cardon, M. S. (2020). A review of and future agenda for research on identity in entrepreneurship. *Journal of Business Venturing*, *35*(6), 106049.

Naderifar, M., Goli, H., & Ghaljaei, F. (2017). Snowball Sampling: A Purposeful Method of Sampling in Qualitative Research. *Strides in Development of Medical Education*.

Navis, C., & Glynn, M. A. (2011). Legitimate Distinctiveness and The Entrepreneurial Identity: Influence on Investor Judgments of New Venture Plausibility. *Academy of Management Review*, *36*(3), 479–499.

Patton, M. Q. (2014). *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. SAGE Publications.

Phillips, N., Tracey, P., & Karra, N. (2013). Building entrepreneurial tie portfolios through strategic homophily: The role of narrative identity work in venture creation and early growth. *Journal of Business Venturing*, *28*(1), 134–150.

Pollack, J. M., Rutherford, M. W., & Nagy, B. G. (2012). Preparedness and Cognitive Legitimacy as Antecedents of New Venture Funding in Televised Business Pitches. *Entrepreneurship Theory and Practice*, *36*(5), 915–939.

Powell, E. E., & Baker, T. (2014, 1 23). It's What You Make of It: Founder Identity and Enacting Strategic Responses to Adversity. *Academy of Management Journal*, *57*(5), 1406–1433. (publisher: Academy of Management) doi: 10.5465/amj.2012.0454

Ruebottom, T. (2013). The microstructures of rhetorical strategy in social entrepreneurship: Building legitimacy through heroes and villains. *Journal of Business Venturing*, *28*(1), 98–116.

Scott, S. G., & Lane, V. R. (2000). A Stakeholder Approach to Organizational Identity. *Academy of Management Review*, *25*(1), 43–62.

Sinek, S. (2011). *Start With Why: How Great Leaders Inspire Everyone To Take Action*. Penguin UK.

Stets, J. E., & Burke, P. J. (2000). Identity Theory and Social Identity Theory. *Social Psychology Quarterly*, *63*(3), 224–237.

Steyaert, C. (2007). Of course that is not the whole (toy) story: Entrepreneurship and the cat's cradle. *Journal of Business Venturing*, *22*(5), 733–751.

Suddaby, R. (2006). From the Editors: What Grounded Theory is Not. *Academy of Management Journal*, *49*(4), 633–642.

Werven, v. R., Bouwmeester, O., & Cornelissen, J. P. (2015). The Power of Arguments: How Entrepreneurs Convince Stakeholders of the Legitimate Distinctiveness of their Ventures. *Journal of Business Venturing*, *30*(4), 616–631.

Wry, T., Lounsbury, M., & Glynn, M. A. (2011). Legitimating Nascent Collective Identities: Coordinating Cultural Entrepreneurship. *Organization Science*, *22*(2), 449–463.

Yin, R. K. (1994). *Case Study Research: Design and Methods*. SAGE Publications.

Zhao, E. Y., Fisher, G., Lounsbury, M., & Miller, D. (2017). Optimal Distinc-

tiveness: Broadening the Interface between Institutional Theory and Strategic Management. *Strategic Management Journal*, *38*, 93–113.

Zimmerman, M. A., & Zeitz, G. J. (2002). Beyond Survival: Achieving New Venture Growth by Building Legitimacy. *Academy of Management Review*, *27*(3), 414–431.

Zott, C., & Huy, Q. (2007, 3). How Entrepreneurs Use Symbolic Management to Acquire Resources. *Administrative Science Quarterly - ADMIN SCI QUART*, *52*, 70–105.

Big Data Analytics Capabilities: A Systematic Literature Review on Necessary Skills to Succeed in Big Data Analytics

Marc A. Richly

Ludwig-Maximilians-Universität München

Abstract

While the amount of data keeps growing, managers ask themselves whether they already retrieve full value from their data. To maximize the value of big data, literature offers first insights in building BDA capabilities (Gupta & George, 2016; Mikalef, Framnes, Danielsen, & Krogstie, 2018). Nevertheless, BDA remains a new field to researchers and companies. BDA frameworks, still offered scarcely, discuss roughly the same dimensions (incorporating some technical, human, and cultural aspects), but are only superficially discussed. This thesis builds a framework of the different approaches offered in literature. Furthermore, it is important to distinguish whether a new development as BDA can be seen as a trend topic or rather a long-lasting game changer for businesses. Here, this thesis discusses differences among digital capabilities, IT capabilities, that research stared addressing by 1990, and BDA capabilities. A major finding is that building IT capabilities is considered as an isolated responsibility of IT departments by, i.e., offering IT infrastructure to the whole company. BDA capabilities, on the contrary, cannot be planned and rolled out from one specific department – those need to be developed in every organizational unit; therefore, a data-driven culture is a key element in building BDA capabilities.

Keywords: Big data analytics; Big data; Data analytics; Dynamic capabilities; Resource-based view.

1. Introduction

Several studies have shown that firms using Big Data Analytics (BDA) in their company are more successful. For instance, authors of a MIT Sloan Management article could show that "top-performing organizations use analytics five times more than low performers" (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011, p. 22).

Moreover, research has shown in many studies that companies that run BDA activities reach stronger results (Günther, Huysman, & Feldberg, 2017; Gupta & George, 2016). Hao, Zhang, and Song (2019, p. 9-12) have found that big data improved sales growth and gross margins, whilst Wamba, Gunasekaran, Akter, Ren, and Dubey (2017, p. 9, 16) could show that BDA capabilities have direct, positive effects on firm performance. In general, many researchers see high potential of BDA enhancing innovation, competition, and productivity (Manyika et al., 2011).

While the amount of data keeps growing, managers ask themselves whether they already retrieve full value from their data. To maximize the value of big data, literature offers first insights in building BDA capabilities (Gupta & George, 2016; Mikalef et al., 2018).

Nevertheless, BDA remains a new field to researchers and companies. Thus, current results change quickly, and researchers still discuss capabilities a company needs for BDA. BDA frameworks, still offered scarcely, discuss roughly the same dimensions (incorporating some technical, human, and cultural aspects), but are only superficially discussed. Most BDA papers do simply orientate at IT capabilities and only offer marginal adoptions to the analytical component. Only recently, first papers start to quantify the proposed effects of BDA capabilities (and not analytics on firm performance in general) to enrich literature (Yasmina, Tatoglua, Kilicb, Zaimc, & Delen, 2020).

This leads to the first research question.

> RQ1: "Which capabilities are proposed in literature to conduct Big Data Analytics in firms and how do they influence success of Big Data Analytics?"

Furthermore, BDA was not situated in the context of other big technological developments as e.g., IT capabilities in the 1990s, yet. For managers and researchers, it is nevertheless important to be able to weigh different trends or to even iden-

tify whether a new development as BDA can be seen as a trend topic or rather a long-lasting game changer for businesses.

Consequently, the second research question reads as follows:

> RQ2: "In what respect do Big Data Analytics Capabilities resemble and differ from IT and Digital capabilities?"

2. Conceptual Foundation

2.1. Definitions of Key Terminology

For the beginning, key terminology will be defined (Big Data, BDA, BDA capabilities), clearly delineated from each other, and situated in current research.

2.1.1. Big Data

Initially, Big Data was coined to reflect the "bigness" or *volume* of data "generated as a result of using new forms of technology (e.g., social media, radio-frequency identification (RFID) tags, smart phones, and sensors)" (Gupta & George, 2016, p. 1050). It is measured not only by a large set of observations itself, but also many variables.

With volume, also velocity and variety were introduced as the "three Vs" characteristics to define the term of Big Data (McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012, p. 4-5).

Velocity reverberates the speed at which data is collected (near to or at real time) from sales transactions, sensors, social media posts, and sentiment data from breaking news and social trends (Gupta & George, 2016).

The term *variety* describes the plurality of the data (as texts, graphics, videos, networks among others) as it can emerge in structured, partly structured, and unstructured sources (Gupta & George, 2016).

Today, most definitions of big data have shifted from including the analytical tools and visualization of results to only integrate the V-characteristics of Big Data, as Davis 2014 defined: "Big data consists of expansive collections of data (large volumes) that are updated quickly and frequently (high velocity) and that exhibit a huge range of different formats and content (wide variety)."

Some researchers have extended the existing body of definitions by further characteristics as *veracity* (unreliability inherent in some sources of data), *variability* (variation in the data flow rate), and *value* (showing the potential in transforming low value raw data into high value data via BDA) (Gandomi & Haider, 2015, p. 139). The term *value* is nonetheless highly firm-dependent as it is connected to the strategic goals of a company (Günther et al., 2017, p. 191).

2.1.2. Big Data Analytics

BDA definitions – in comparison to big data – focus on multiple analytics methods "that address the diversity of big data to provide actionable descriptive, predictive, and prescriptive results" (Lamba & Dubey, 2015, p. 5). The data

can be analyzed with technologies (e.g., data mining tools or database) and different techniques (e.g., analytical methods) to generate insights of big data (Kwon, Lee, & Shin, 2014, p. 387).

Analytics and Data Style

As one of Big Data's characteristics is variety (video, text, audio, social media data etc.), also BDA must adopt with suitable techniques to the style of data (Choi, Wallace, & Wang, 2018; Mikalef et al., 2018).

Textual data

Firstly, textual data as texts, obtained from emails, social media posts, blogs, surveys, news etc., can be analyzed with analytical tools to extract information. So far, three approaches exist to work with textual data: Information Extraction (IE), Text summarization, and Question Answering (QA) (Gandomi & Haider, 2015, p. 140).

IE can, for instance, obtain specific data as drug name and dosage from a medical prescription. Text summarization are (or will be) useful in news articles and emails to provide a short overview with the important details. QA most important examples include Siri from Apple, or Alexa from Amazon. Siri can provide answers on orally asked questions (Gandomi & Haider, 2015, p. 140).

Video data

This form of data is not only more complex to analyze, also in terms of volume it is more difficult to process video analytics. Nevertheless, analytical methods for video data are progressing well (Gandomi & Haider, 2015, p. 141). Use cases for video analytics are broader, from closed-circuit television (CCTV) cameras to video sharing websites (as YouTube that automatically checks videos for e.g. copyright violations (Agrawal & Sureka, 2013)). Furthermore, video analytics can help for marketing purposes to - for instance - identify the demographics, gender and similar characteristics of people that go shopping (Hu, Xie, Li, Zeng, & Maybank, 2011).

Audio Data

For audio analytics (in most cases: speech analytics, when analysis is about spoken words), two main systems are in place: LVCSR systems (transcript -based approach) and Phonetic-based systems (Gandomi & Haider, 2015, p. 141).

LVCSR systems uses automatic speech recognition (ASR) to transcribe spoken word into text with the help of a dictionary. A popular application is the dictation function on smartphones. With the text output, standard text-based analytics can be conducted (Gandomi & Haider, 2015, p. 141).

Phonetic-based systems work with "perceptually distinct units of sound" (Gandomi & Haider, 2015, p. 141) that distinguish one word from another and thus create text. As with LVCSR systems, this text can then be analyzed with text-based analytics.

Main use cases for audio analytics are in call centers and hospital (for automatically documenting what a doctor has

said) (Wang, Kung, & Byrd, 2018).

Social Media Data

Social media data can consist of varying online platforms, as social networks (e.g., Instagram, LinkedIn), blogs, social news (e.g., Digg and Reddit), media sharing (e.g., YouTube), wikis (e.g., Wikipedia), review sites (e.g., Trivago, Yelp), and questions-and-answer sites (e.g., Ask.com) (Gandomi & Haider, 2015, p. 142; Mikalef et al., 2018).

Social media analytics has emerged in the early 2000s and can be classified into two groups: *Content-based analytics* (focuses on data posted by users) and *Structure-based analytics* (synthesising structural attributes of a social network) (Gandomi & Haider, 2015, p. 142).

Technologies and Tools

Analytics itself can be conducted on different data types. But also different technologies are used to analyse data sets (here: descriptive, predictive, and prescriptive results (Lamba & Dubey, 2015, p. 5)) that will be discussed in the following paragraph.

Descriptive Analytics – What has happened? Raw data, as sales, customers, and operations data, from the past can be analyzed on insights. Organizations use this form of analytics to gasp a deeper understanding of their business and identify relationships that have happened previously (Soltanpoor & Sellis, 2016, p. 247). With that, insights can be obtained as: Which products sell better or worse compared to other products or how has the demand developed over a year.

To generate reports on the historical data or to extract information from raw data, techniques as statistical analytics, data integration, data augmentation, and data reduction can be used (Soltanpoor & Sellis, 2016, p. 247). In general, these techniques are rather simple to comply (creating graphs) and normally include basic descriptive statistics as "measures of central tendency (mean, median, mode), measures of dispersion (standard deviation), charts, graphs, sorting methods, frequency distributions, probability distributions, and sampling methods" (Ajah & Nweke, 2019, p. 7).

Statistics, nevertheless, were identified as one major technique in BDA by Choi et al. (2018). With this already well-established field, relationships as correlations or statistical regression are often used to generate insights of data sets.

Predictive Analytics – What will happen?

Prediction can vary a lot in big data use cases – "the failure of jet engines based on the stream of data from several thousand sensors, to predicting customers' next moves based on what they buy, when they buy, and even what they say on social media" (Gandomi & Haider, 2015, p. 143).

The aim here is to foresee opportunities and risks in the future. Two core techniques can be used to generate these insights: regression techniques (e.g., multinomial logit models) and machine learning techniques (Gandomi & Haider, 2015, p. 143).

Machine learning was identified as another major BDA

technique by Choi et al. (2018). It compromises not only Artificial Intelligence, but also neural networks, support vector machines, and statistical machine learning. Even though machine learning takes time until it is trained and can return results, it provides helpful algorithms to capture complex behaviors.

With machine learning, the concept of "data lakes" also became relevant as a new technique. Data lakes were firstly defined about ten years ago by Dixon (2010) and compromise of four main characteristics: they are (1) massively scalable in terms of volume, and the data is (2) stored as raw data ("as is") and thus in an unstructured format (in contrast to data warehouses, where structured data is stored). Thereby, the lakes also use (3) dynamic analytics applications as machine learning (not like data warehouses: pre-build static). Lastly, data becomes (4) accessible, as soon as the lake is created (again, different to data warehouses which are designed for only slowly changing data) (Miloslavskaya & Tolstoy, 2016, p. 302). As the real time analysis has been identified as a major trend in Business Intelligence (BI) (Russom, 2011, p. 26), data lakes can be expected to play a big role in BDA.

Prescriptive Analytics – What should be done?

Prescriptive analytics "provides enterprises with adaptive, automated, time-dependent, and optimal decisions" (Soltanpoor & Sellis, 2016, p. 247).

The aim is to generate recommended business decisions, optimal courses of action or similar. These outputs are normally generated through techniques as optimization, simulation, operations research, or management science (Soltanpoor & Sellis, 2016, p. 247).

Optimization was also highlighted as another major BDA technique by Choi et al. (2018). This technique can be used in quantitative decision-making problems to find the (near) optimal solutions.

2.1.3. Big Data Analytics Capabilities

A business capability is defined as "the organization's capacity to successfully perform a unique business activity" (Keller, 2009, p. 2).

Looking at the field of digital transformation, Soto Setzke, Opderbeck, Böhm, and Krcmar (2020) identified several configurations and strategies that are needed to build capabilities. Their findings reach from the importance of centralized decision making to the rather negative impact of competitive pressure. Also, they suggest that smaller or medium-sized companies shall focus on strong partnerships that can (partly) replace missing capabilities (Soto Setzke et al., 2020, p. 14).

In IS literature, researchers have proposed to take a bird's eye view to better understand how IS investments enhance business value (Bharadwaj, 2000; Wamba et al., 2017). Hence, IT capabilities were introduced to investigate these connections. They are defined as the "firm's ability to mobilize and deploy IT-based resources in combination or co-
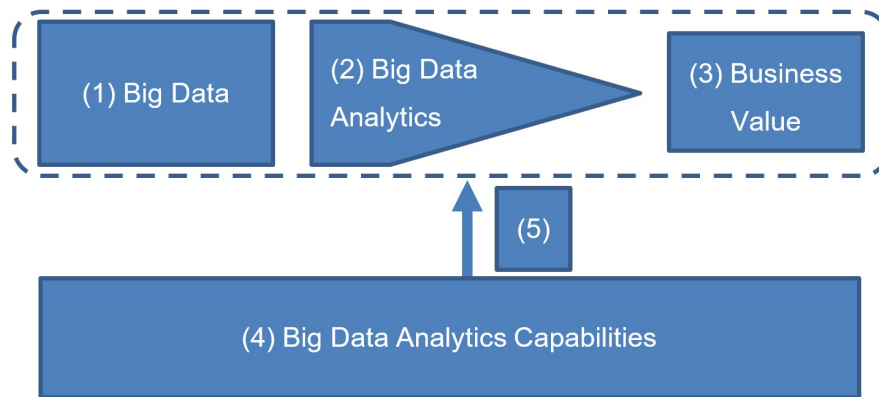
Figure 1: Delimitation of Terminology

present with other resources and capabilities" (Bharadwaj, 2000).

Also for BDA, capabilities were introduced. Akhtar, Frynas, Mellahi, and Ullah (2019, p. 252) defined BDA capabilities as a balanced combination of advanced technologies supported by large datasets to generate analytical reports and actionable insights utilized, produced, and processed by mathematical, statistical techniques, and machine learning tools for enhanced performance, but also essential human resources with big-data skills. In short – a balance must be found between the human part and the technical part of a business.

In regards to a successful implementation of BDA, Wang et al. (2018, p. 8-9) suggest five strategies in their analysis on the healthcare setting: (1) the implementation of big data governance, (2) the creation of an information-sharing culture, (3) the training of key personnel to use BDA, (4) the integration of cloud computing into the organization's BDA, and (5) the generation of new business ideas from BDA.

2.1.4. Delimitation of Terminology

While (1) big data is about data sets with exact characteristics (s. 2.1.1), (2) BDA encompasses the tools and technologies used to analyze big data (Mikalef et al., 2018, p. 555-556). Some researchers even include (3) business value in terms of presentation and performance impact of BDA in their definitions.

This paper will focus on (4) Big Data Analytics capabilities, to understand what firms need to work with Big Data and how firms can build BDA capabilities. Addressing the research question, firstly this paper will show current capabilities proposed by literature to conduct BDA (4). Secondly, this paper will show how these capabilities influence the success of BDA (5).

2.2. Definition and Types of Capabilities

2.2.1. Resource Based Theory

To describe why and how BDA is important and can enhance firms' performances, researchers use different strategic backgrounds. During the 1980s most strategic theories

focused externally (e.g. Porter's five forces model, firstly published in 1979), in the 1990s researchers' frameworks have focused more internally as the resource-based view emerged (Barney, Ketchen Jr., & Wright, 2011, p. 1300). Also for BDA, internal frameworks are applicable, and thus, the resource-based view that later emerged in the resource-based theory (RBT) is commonly used by researchers to describe how BDA capabilities can add value (Gupta & George, 2016, p. 1050).

The RBT understands a company as a sum of different firm resources that are composed of "assets, capabilities, organizational processes, firm attributes, information knowledge etc." (Barney, 1991, p. 101). Initially three types of assets were proposed – physical (e.g., specialized equipment), human (e.g., expertise in technology), and organizational (e.g., superior sales force) – that can be used to implement value creating strategies (Eisenhardt & Martin, 2000, p. 1106-1107). Gupta and George (2016, p. 1054) argue that data can also be an important asset to an organization – if the firm has implemented a data-driven culture.

Firm resources must fulfil the characteristics of value, rareness, imperfect imitability, and substitutability (s. Tab. 1). With those a company can implement a value-generating strategy by creating a competitive advantage over competitors (Barney, 1991, p. 105-112).

Then, the firm resources can be used to archive a competitive advantage. A competitive advantage exists, when a firm has implanted a value creating strategy not used by any other firm (Barney, 1991, p. 102).

Later, the understanding of firm assets was split into two parts, resource-picking and capability-building to further explain the mechanism how assets can be turned into a competitive advantage (Makadok, 2001; Mikalef et al., 2018).

On the one hand, *resource-picking* describes the process of identifying (often also buying) resources that promise to be VRIN (valuable, rare, inimitable, and non-substitutable). Nevertheless, these resources are still rather easy replicable (Santhanam & Hartono, 2003).

On the other hand, *capability-building* facets are concerned with the orchestration of resources. Capabilities, in contrast to resources, are highly specific to an organization

Table 1: Conditions for sustainable advantage, (Barney, 1991, p. 105-106)

Conditions for sustainable advantage	Definition of condition
Value	Asset exploits opportunities / neutralizes threats in a firm's environment
Rareness	Asset is rare among current and potential competitors
Imperfect Imitability	Asset must be difficult (or impossible) to copy
Substitutability	No other asset that is rare and imperfect imitable can replace this asset
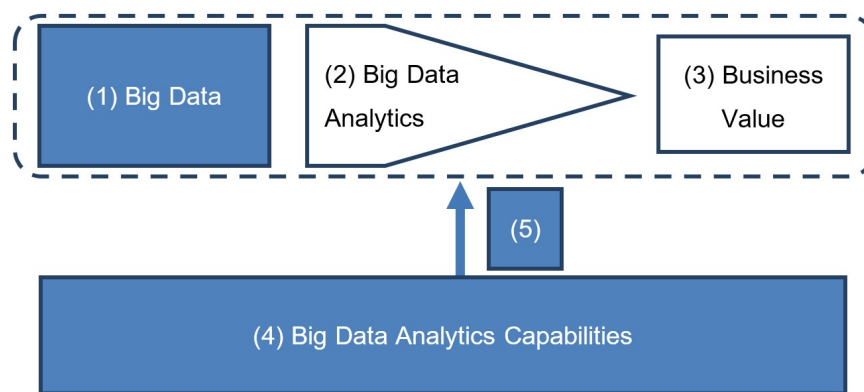
Figure 2: Resource-picking (1) and capability-building (4) in context of the BDA process

and thus is hard to replicate. With a distinctive set of capabilities, a sustained competitive advantage can be reached (Santhanam & Hartono, 2003).

In the context of BDA, big data can be thus seen as a *resource-picking activity* (s. (1) in Fig. 2) that can be bought from external suppliers or developed from internal data. BDA, in contrast, is the result of *capability-building activities* (s. (4) in Fig. 2). These capabilities are firm specific and cannot be bought or sold as they are in some way connected to the organization.

Research also suggests that firms do need a resource before they start building capabilities for it (Mikalef et al., 2018). This is important for firms as it shows that they first need some level of access to big data, before they should start building capabilities to analyse it.

2.2.2. Dynamic Capabilities View of an Organization

With the resource-based theory explaining how firms can build a competitive advantage against competitors, the RBV cannot explain how companies can defend its competitive advantage against competitors in rapidly changing environments. The Dynamic capabilities theory (DC) tries to fill this gap as an extension to the RBV (Mikalef et al., 2018, p. 560).

DC is defined in a seminal article as "the ability of a firm to integrate, build and reconfigure its resources and capabilities

to address changing environments" (Teece, Pisano, & Shuen, 1997, p. 516). Researchers furthermore make a distinction between *ordinary capabilities* (also: "zero order" capabilities) that determine how "a firm makes a living at the moment" (Tiguint & Hossari, 2020, p. 2), in contrast to *dynamic capabilities* that enable an organization to change.

The impact of such dynamic capabilities depends on the environment. In moderately dynamic markets, dynamic capabilities are difficult to identify or to separate from ordinary capabilities. In high-velocity markets, routines (here: dynamic capabilities are simple, experiential, and iterative) are adaptive to changing circumstances – but at the cost of unstable processes and uncertain outcomes (Eisenhardt & Martin, 2000). This results in uncomfortable but still success bringing situations, as one manager put it:

"We do everything on the fly ... I've done some things at IBM and other companies where there is a very structured environment—these companies are failing and we're leading the way. I'm not comfortable with the lack of structure, but I hesitate to mess with what is working." (Brown & Eisenhardt, 1997, p. 28)

To conclude, how can firms eventually ensure their long-term competitive advantage (especially in high-velocity markets)? They must use dynamic capabilities "sooner, more astutely, or more fortuitously than the competition to create re-

source configurations that have that advantage" (Eisenhardt & Martin, 2000, p. 1117).

2.3. Challenges in Building BDA Capabilities

First, research on BDA implications and capabilities is still scare, which also makes it difficult for companies to start with the implementation of BDA as they cannot linchpin on a broad base of literatures results (Gupta & George, 2016).

In terms of BDA, the proposed capabilities itself can also entail a challenge for companies. Big data – in its terms of volume, velocity, and variety - is still challenging to current IT architecture, networks, servers, and software (Ajah & Nweke, 2019, p. 22). Some research suggests that firms are in particular worried about the unstructured nature of data (the variety part) rather than the volume of data (Gupta & George, 2016, p. 1051). For that, new tools and technologies are required (e.g., Hadoop) that will be further discussed in the literature review (s. 4.1).

Moreover, before starting with the implementation, a company should make sure to have fully covered needed capabilities. If, for instance, measures to ensure privacy are not implemented in the BDA capability-building stage, it will hardly be considered in the BDA process (in this paper's BDA capabilities framework privacy and security are recognised under *Governance* in 4.3.2). This literature review will enrich current literature by discussing different proposed capability frameworks and showing, what impact the different characteristics have on the success of BDA (as described in RQ1).

2.4. Challenges in Building IT Capabilities

Research in IT Capabilities is much more mature than BDA capabilities – offering a better overview on the challenges that have occurred over time.

Nevertheless, also research on IT capabilities has fought some major problems. In the early 1990s, for instance, several studies have examined the so-called *IT productivity paradox* that referred to the failure to show a positive relationship between IT investments and firm productivity (Gupta & George, 2016, p. 1049). Also, renown economists as Robert Solow (holds a Nobel prize in economics) stated that "we see computers everywhere except in the productivity statistics" (Brynjolfsson, 1993).

In his paper, Brynjolfsson (1993) summarized four possible reasons for the occurrence of this paradox:

1. Mismeasurement of outputs and inputs,
2. Lags due to learning and adjustments,
3. Redistribution and dissipation of profits, and
4. Mismanagement of Information Technology.

Later, in some industries a positive correlation could be clearly shown, and the paradox was eventually resolved. Consequentially, it was suggested to take several resources besides pure investments into account, as managerial and technical IT skills, firm's intellectual capital, and IT infrastructure (Gupta & George, 2016, p. 1049). Also, so far no *BDA productivity paradox* has been identified.

Today, some researchers argue that IT capabilities no longer offer any competitive advantage (in contrast to BDA) but have turned to an organizational commodity (Chae, Koh, & Prybutok, 2014, p. 307).

2.5. Challenges in Building Digital capabilities

Digital capabilities are broadly defined and cover most of digital extensions that exceed IT capabilities (s. 5.2 - comparison of BDA and digital capabilities).

Thus, it is also difficult to define concrete challenges of digital capabilities. Mostly, organizations will work on concrete areas as BDA or Artificial Intelligence (AI) that have their respective own challenges.

3. Methodology

For this thesis, a structured literature review on existing papers in IS literature dealing with BDA capabilities will be conducted.

The literature review is divided into two parts, as we must expect that no broad use of BDA capabilities is proposed in current literature, yet. Thus, in the first stage, a structured literature review on existing papers dealing with BDA capabilities is conducted. The literature review is based on Webster and Watson (2002).

In the second stage, the basic capabilities framework will be discussed and for a deeper analysis further punctual research on specific parts of the proposed framework will be conducted.

### 3.1. Literature Review on Big Data Analytics Capabilities

The literature review in the first stage is conducted in four phases.

In stage one (1), articles of the AIS *Basket of Eight* IS journals were collected. As literature on BDA capabilities is still scarce, the search radius will also be extended to highly cited BDA publications from other journals (via Google Scholar and Web of Science). Additionally, three leading IS conferences (International Conference on Information Systems (ICIS), European Conference on information Systems (ECIS), and Americas Conference on Information Systems (AMCIS)) were considered to cover the latest research on BDA capabilities, too. For the research on journals, only papers published in the past eleven years (2010-2021) were considered, as this field is still relatively new and thus the emphasize lies on the latest developments in BDA research. For conferences, only the years 2018 to 2021 were considered, assuming good conference articles would have reached a journal outlet by today.

The following key words were used to find relevant studies (all words were also searched with their respective abbreviations): Big Data Analytics, Big Data Analytics Capabilities.

After stage one, 17 papers were found in the Basket of Eight, 15 highly cited journals from other journals, and eleven conferences could be identified.

In stage two (2), the titles, keywords, and abstracts were evaluated. Papers were excluded that focused too much on
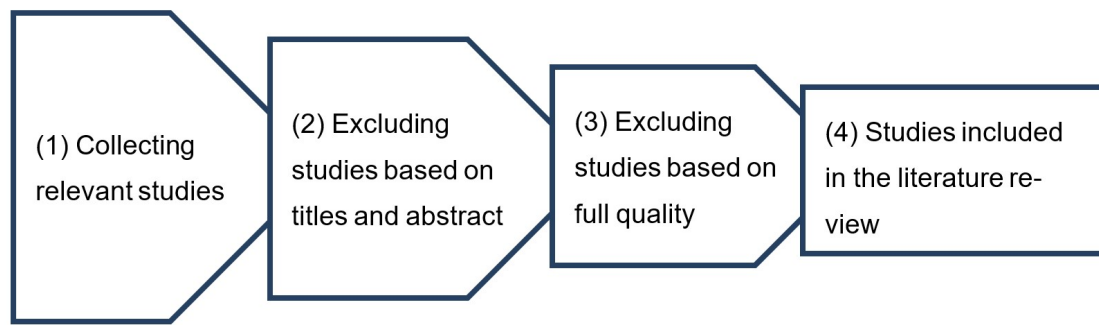
Figure 3: Process of literature review, based on (Webster & Watson, 2002)

the technical side of BDA or privacy. Papers with a focus on specific industry or case studies were not excluded, as still major concepts or the underlying theory could be useful for an evaluation in the literature review.

After stage two, 24 papers and seven conferences were considered for the next stage.

At stage three (3), the full paper was accessed. Here, papers were excluded that did not discuss actual capabilities needed to perform BDA projects in firms and where BDA only played a minor role.

Ultimately, twelve papers and two conferences were accessed and included in the literature review of this thesis.

3.2. Discussion on Big Data Analytics Capabilities versus IT and Digital Capabilities

After the literature review, an extension in form of an argumentative evaluation of the different capabilities a company must have to perform BDA compared to IT and Digital capabilities will be provided. For this, a small research via Google Scholar and Science of Web was conducted to retrieve most influential papers dealing with IT and digital capabilities.

Conclusively, five papers on IT capabilities and nine papers on digital capabilities were evaluated.

4. Literature-Review on Big Data Analytics Capabilities

To build a BDA capability a firm needs different resources. One of the first frameworks "to put analytics to work in [...] a business" (Davenport, Harris, & Morison, 2010, p. 19) was offered by Davenport et al. (2010) that was later further discussed by Seddon, Constantinidis, Tamm, and Dod (2017). To implement analytics, they suggested a framework consisting of five pieces, called DELTA (Davenport et al., 2010, p. 19):

- *D*ata: accessible and high quality

- *E*nterprise orientation

- *L*eadership in analytics

- *T*argets: Strategic goal

- *A*nalysts: the human input needed

Nevertheless, this framework was not further discussed or developed in the specific BDA literature. Moreover, Davenport et al. (2010) focus on analytics in general, but not discuss the more specific BDA. Thus, this framework will not be investigated further.

In a highly influential paper, Gupta and George (2016) have built a framework of needed resources, dividing them by tangible, human and intangible resources.

Initially, Gupta and George (2016) derived this separation from research on IT capabilities, using then the framework to expand it on BDA capabilities (Bharadwaj, Sambamurthy, & Zmud, 1999; Chae et al., 2014; Gupta & George, 2016; Santhanam & Hartono, 2003).

Gupta and George's three pillars framework was taken as a foundation by Mikalef et al. (2018), who have adjusted the framework by some characteristics. As this literature review by Mikalef et al. (2018) can be seen as an extension of the Gupta and George (2016) paper, the adopted version of Mikalef et. al. will be taken as a basis for further analysis.

The other big framework used by researchers for BDA capability distinguishes them among (sometimes called slightly different but entailing the same ideas) *BDA technology / infrastructure / data capabilities*, *BDA management capabilities*, and *BDA talent capabilities* (Akter, Wamba, Gunasekaran, Dubey, & Childe, 2016; Ransbotham, Kiron, & Prentice, 2015; Wamba et al., 2017).

Also Yasmina et al. (2020) defined three similar pillars (similar differentiation: Infrastructure capabilities, Human Resource Management Capabilities, and Management Capabilities) in one of the latest researches on BDA.

Both, Mikalef and Yasminas' frameworks, cover same areas, as technical requirements for BDA, human knowledge, managerial roles, culture, and alignment with a firm's business strategy (s. Fig. 4).

Nevertheless, both frameworks set different emphasis. While Mikalef et al. (2018) summarize analytical skills of employees and managerial skills in one pillar "Human Resources", Yasmina et al. (2020) highlight employees and managerial skills by dividing both on two of the three pillars. Looking closer at differences between the third cluster of
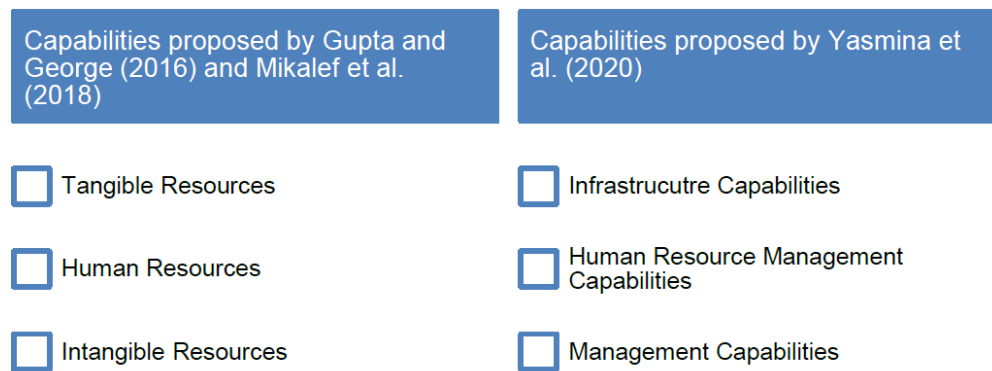
Figure 4: Proposed BDA capability framework by Gupta and George (2016) and Mikalef et al. (2018) (left side) and Yasmina et al. (2020) (right side)

Mikalef et al. (2018) – Intangible resources – and the third cluster of Yasmina et al. (2020) – Management Capabilities – reveals that both discuss cultural and strategy alignment aspects, which straightens out the expected major difference of both researchers.

A detailed comparison can be seen in the following table 2.

Nonetheless, for the following analysis Yasmina et al. (2020) will not be taken as a foundation (even though it is the most recent paper published on BDA capabilities), mainly because the paper does not clearly distinguish IS from BDA capabilities (e.g., in the detailed description of the pillars they refer to IS, but not to BDA characteristics). Furthermore, the paper does not provide deep theory backgrounds, but focuses on testing its hypothesis. Also other papers' frameworks (Akter et al., 2016; Ransbotham et al., 2015; Wamba et al., 2017) with similar separation of capabilities did often not clearly separate BDA from IS capabilities. Thus, in this thesis the provided framework from Mikalef et al. (2018) that clearly focuses on BDA and its respective theories will be used.

4.1. Tangible Resources

The first pillar in the BDA capabilities framework is about *tangible resources* (Gupta & George, 2016; Mikalef et al., 2018). Initially, Gupta and George (2016) have defined the following characteristics: *Data* (internal, external, merging of both), *Technology* (Hadoop, NoSQL), and *Basic Resources* (time, investment).

The characteristic *data* was established by Mikalef et al. (2018), too. Nevertheless, they have exchanged the two other terms with *Software and IS*, and *Infrastructure*.

4.1.1. Data

The most obvious ingredient to conduct BDA is data itself. As a continuously trend, the cost for storage of data is still (even exponentially) declining and consequently data storage is not a main challenge for companies for conducting

BDA anymore (Rosenthal et al., 2012). Nevertheless, to recap, big data is not only about its volume, but comes along with other characteristics as velocity, variety, value, and veracity (s. 2.1.1).

Whereas it is easy for companies to measure the size of their data, it already gets more complicated to evaluate the data in terms of quality. Research shows that improved quality helps to turn data into a competitive advantage (Ransbotham et al., 2015). Consequently, leaders in organizations ask themselves "whether they are getting full value from the massive amounts of information they already have within their company" (LaValle et al., 2011, p. 21).

Moreover, organizations have moved on from focusing on enterprise-specific structured data to an increasing amount of unstructured data (variety characteristic of big data) – that does increasingly include external data from outside the organization - to make business decisions (Gupta & George, 2016, p. 1051; Manyika et al., 2011). Thus, not only the amount of data increases (volume), but data also gets more diverse (variety).

Therefore, it is important to evaluate available data in terms of quality, for which research offers some standards to evaluate data sets. Cai and Zhu (2015, p. 4) propose a data quality framework that includes five dimensions (s. Tab 3).

4.1.2. Software and Information System

Next to the increasing amount of data (volume) and the growth in unstructured data (variety), also the major trend in analyzing data in real-time (velocity) have led to the development of new software and Information Systems (IS) (McAfee et al., 2012).

Today, most prominent example of a BDA software is Hadoop. Introduced as an open source project in 2007, it has evolved to support the whole BDA workflow, "including data collection, storage, processing, and much more" (Landset, Khoshgoftaar, Richter, & Hasanin, 2015, p. 5). Besides, research states that the real value of Hadoop does not only lie in its ability to handle large volumes of data, but to manage

Table 2: Detailed comparison of BDA capabilities framework by Mikalef et al. (2018) and Yasmina et al. (2020)

Framework by (Mikalef et al., 2018)	Similarities	Differences	Framework by (Yasmina et al., 2020)
Tangible Resources	Describes technical needs	Yasmina focuses solely on IS infrastructure, not on special infrastructure needed for analytics	**Infrastructure Capabilities**
Human Resources	Technical knowledge needed from employees	Only Mikalef describes managerial skills (Yasmina does so in third pillar)	**Human Resource Management Capabilities**
Intangible Resources	Focus on cultural alignment	Yasmina connects cultural alignment and management capabilities	**Management Capabilities**
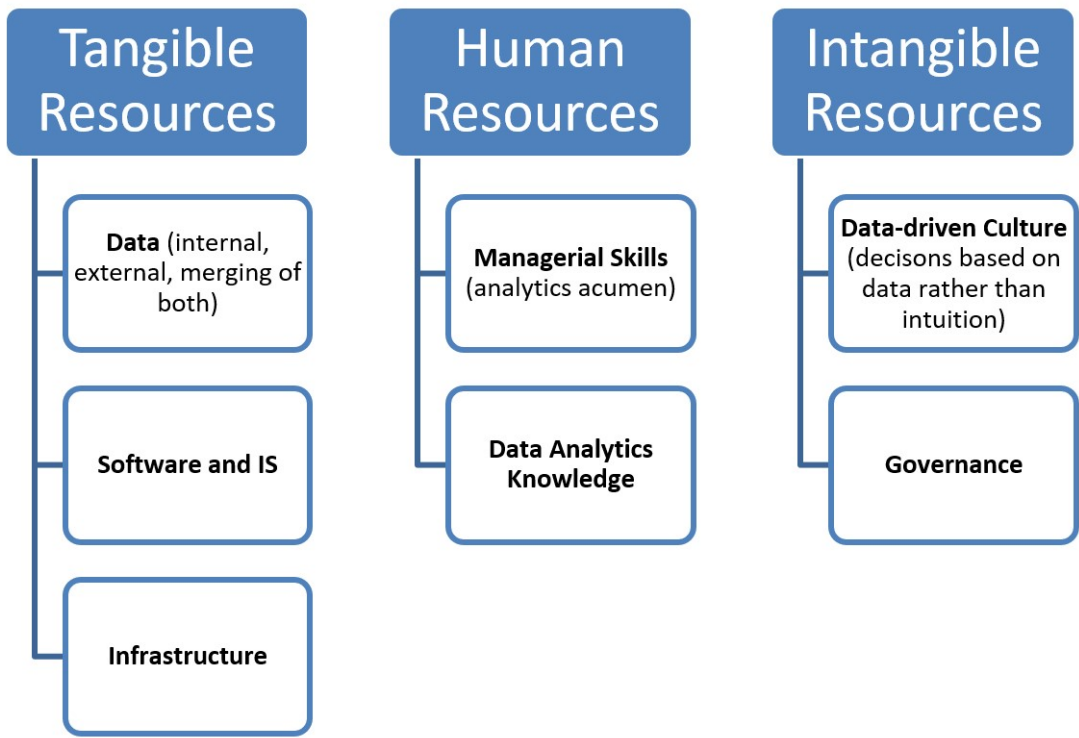
Figure 5: BDA resources to build BDA Capabilities (obtained from Mikalef et al. (2018))

data's variety - a broad range of data types (Russom, 2011, p. 7).

Hadoop combines commodity hardware with its open-source software. Large data streams can then be distributed (also in real-time) onto cheap disks for the analysis part (McAfee et al., 2012, p. 8). The distribution is done by the Hadoop distributed file system (HDFS), while the MapReduce data processing engine is the heart of the analytics part.

MapReduce consists of two primary parts, a mapper and a reducer (Ajah & Nweke, 2019, p. 13). The mapper filters and transforms data by cleaning out dirty data (also duplicates). After this, an intermediate output file is generated that is then distributed to reducers. Here, the reducer phase begins, sorting each file by key and aggregating them into one large file again (Landset et al., 2015; White, 2012). This process allows to run through large and heterogenous data sets in a very effective way.

Other software solutions of nonrelational databases to parallel process large unstructured datasets and to visual-

Table 3: Data quality characteristics (Cai & Zhu, 2015, p. 5-6)

Term	Meaning
Availability	Having access to data in a reasonable amount of time
Usability	Data can satisfy the user's needs
Reliability	Data is complete, from a trusted source, consistent, and meets integrity criteria
Relevance	"Fitness of data" – data can possibly answer what an organization wants to know
Presentation Quality	Format of data is clear / understandable

ize the results are Apache Cassandra, MongoDB, Monet, and Hazelcat (Mikalef et al., 2018, p. 563).

4.1.3. Infrastructure

Wamba et al. (2017, p. 17) suggest three elements that BDA infrastructure is composed of: *connectivity*, *compatibility*, *modularity*. Here, *connectivity* grasps that all offices and departments have access to / can provide their data for BDA. *Compatibility* is a further step of connectivity, which means that not only the data flow is ensured among departments/offices, but also that analysis can be conducted on different platforms and consequently results are accessible for all. Lastly, *modularity* describes an efficiency component: can software modules be re-used or adjusted for different tasks?

Sometimes, also software solutions as Apache Hadoop and cloud-based computing are called in regards to infrastructure (Akter et al., 2016, p. 13). In this thesis, this falls in the category of Software and IS (s. 4.1.2).

$$\phi$$

According to LaValle et al. (2011, p. 23) the most difficult adoption that organizations face are not rooted in tangible assets as data and technology, but in managerial and cultural dimensions. This will be further investigated in the following two pillars.

4.2. Human Resources

The second pillar in the BDA capabilities framework is about *human resources* (Gupta & George, 2016; Mikalef et al., 2018). Between both papers only a minor difference occurred, as Gupta and George (2016) named the knowledge characteristic toward employees *Technical Skills* that was specified by Mikalef et al. (2018), calling it *Data Analytics knowledge*.

4.2.1. Managerial Skills

Companies do not just perform better because of more or better data, but "because they have leadership teams that set clear goals, define what success looks like, and ask the right questions" (McAfee et al., 2012, p. 8). Gupta and George (2016, p. 1053) explained further that it is in the manager's responsibility to foresee potential of generating insights from data and they then must have a good understanding of how

and where to apply the insights retrieved from their technical teams.

Also, Mikalef et al. (2018) stated that it is more important for managers to learn how to access companies' big data sets in the best possible way and how to perform BDA that is aligned with an organization's competitive strategy than to "simply perform raw data analysis on large data sets without a clear direction" (Mikalef et al., 2018, p. 572) or unclear contribution to overall firm strategy.

Moreover, also in contrast to employees' technical skills that can be developed by hiring new staff, managerial skills are highly firm specific and thus must be internally developed over time (Gupta & George, 2016, p. 1053).

To enhance this management capability, research suggests improving the quality of planning, investment, coordination, and control (Akter et al., 2016, p. 41).

4.2.2. Data Analytics Knowledge

A challenging question for managers is how to structure technical and human resources in a firm. The implementation of *competency centres* as a centralized approach to bundle analytics competence in one place have be considered as an effective concept, even though the idea of *competency centres* is also controversial, because of the difficulty to reach an alignment between technical and business units (Günther et al., 2017, p. 197).

In terms of *Data Analytics Knowledge* different roles with respective skills are proposed in literature. For that, Ajah and Nweke (2019) describe three distinct teams. They suggest an (1) *Analytics Team*, compromised of *Data Scientist* and *Business Analysts / Data Analysts*. While the *Data Scientist* is required to have more the technical background with "advanced skills in mathematics and statistical modelling" (Ajah & Nweke, 2019, p. 20), a *Business Analysts* should be aligned with business goals and strategic directions. They also have technical knowledge and are skilled in varying data modelling tools, nevertheless, they primarily keep a bird's-eye view to oversee the alignment of analytics with business strategy / performance outcome.

Additionally, Ajah and Nweke (2019, p. 21) describe the (2) *Big Data Architect Team*, comprimising of *Global Architects / Platform Engineers* and *Data architect / Data Wranglers*. Global architects are experts in supercomputing platforms, focusing their work on performance tuning and root

cause analysis.

Data architects are important because of their industry knowledge and further for their strong mathematics/statistics skills and thus are mostly working on special use cases (Ajah & Nweke, 2019, p. 21).

Lastly, Ajah and Nweke (2019, p. 21) describe the (3) *Big Data Hadoop operators team* consisting of *Hadoop Engineers*, responsible for cluster performance, debugging etc., while *Hadoop Operators* take care of the quality of results. To emphasize Hadoop teams that much might be surprising first, but it is nevertheless the most prominent BDA tool and needs knowledge and skills that exceed the abilities found in most IT departments (McAfee et al., 2012, p. 8).

Overall, McAfee et al. (2012, p. 8) suggest that next to statistical knowledge for the core analytics part, also skills as cleaning and organizing large data sets will become increasingly demanded as well as tools and techniques for visualising results.

Mikalef et al. (2018, p. 565-566) have proposed an additional dimension to the otherwise homogenous dimensions of skills and knowledge that are offered in literature (*Technical Knowledge*, *Business Knowledge*, and Business Analytics Knowledge):

Since teams with an analytics background must work well together with teams that have a business background, "Relational Knowledge" as the "communication and collaboration skills between employees of different backgrounds" (Mikalef et al., 2018, p. 565-566) do play an important role in ensuring strong results from the different teams.

4.3. Intangible Resources

The third and last pillar in the BDA capabilities framework is about *intangible resources* (Gupta & George, 2016; Mikalef et al., 2018). While Gupta and George (2016) have also included the pillar *Intensity of Organizational Learning* this was not considered by Mikalef et al. (2018). Reason for this could be that already existing aspects on learning or knowledge of employees are already covered in the second *Human resources* pillar.

4.3.1. Data-driven Culture

For the implementation of a successful BDA process, researchers suggest a change in the organization's culture to reach a data-driven culture. The baseline is to build a culture where "decisions [are taken] based on data rather than intuitions" (Gupta & George, 2016, p. 1051).

Here it is important to distinguish between the core BDA teams that possess extensive analytics knowledge and are responsible for developing insights from big data and must identify which parts of an organization have the highest potential when analytics is brought in. Whereas other business units' employees from HR, marketing etc. must adopt to a certain degree to a data-centric thinking (a) to increase their knowledge on analytical models and technologies that are or will be used in their department, and (b) to question their daily work routine: Where can analytics be helpful for my work or department? When do we base our decisions on intuition that could be better evaluated with data?

To implement a data-driven culture, it is not enough to support data-to-knowledge initiatives, an organization must ensure leaders behave as a role model and insist that also others take decisions based on data rather than intuition (Davenport, Harris, De Long, & Jacobson, 2001, p. 136).

Regarding culture, a further extension grasps the alignment of IT and business strategy. Akter et al. (2016) have investigated a mediating effect of business strategy alignment on firm performance via BDA capabilities for which they have found a positive significant relationship. Besides, Akhtar et al. (2019) could also show a positive significant relationship among big-data-savvy teams, big-data-driven actions, and firm performance. Moreover, Court (2015) found that organizations can increase operating margins by more than 60% when they can ensure a good alignment between analytics efficiency and business strategy.

4.3.2. Governance

In newer research, experts underscore the importance of data governance (Llave, Hustad, & Olsen, 2018). In literature, two distinctive specifications are offered regarding governance in BDA.

For the first specification, governance grasps compliance with privacy regulations (legal boundaries of how much data may be collected / what degree of personalisation is allowed) and ensuring security of data (data must be secured against hacker attacks). Ajah and Nweke (2019, p. 23) state that organizations often (if so) only have immature tools and practices in place to ensure privacy laws and data security of BDA. For Hadoop for instance, extensions as Kerberos, an open-source project initiated at the MIT, are needed to store data encrypted on Hadoop clusters.

This specification of governance only came later to a scientific discussion. Also both papers of Gupta and George (2016); Mikalef et al. (2018) have not mentioned this characteristic, but solely focused on the second specification in governance:

The second specification focuses on internal rules to ensure an effective use of BDA. Specifically for BDA, Tallon, Ramirez, and Short (2013) propose a governance framework consisting of three types: (a) structural (defining responsibilities, directing, and planning), (b) procedural (cost control, resource allocation, and shaping user behaviours trough value analysis), and (c) relational (alignment / partnerships between IT and business units, idea exchange and conflict resolution).

To ensure that governance is taken to action, literature suggest that an organization must follow a top-down approach, which also implies that top management must commit to data-driven decisions (Vidgen, Shaw, & Grant, 2017) (this is also important for a successful IT/business alignment, s. 4.3).

Concluding, many researchers stress the importance of having governance schemes implemented when working

with BDA. Some researchers even see a missing or badly implanted governance model as a main reason why companies cannot extract full value of big data (Mikalef et al., 2018, p. 565).

5. BDA Capabilities Compared to IT and Digital Capabilities

5.1. Overview IT Capabilities

In the 1990s first papers were published in which IT capabilities (Sabherwal & Kirs, 1994) were defined – in contrast to BDA capabilities in the early 2010s (LaValle et al., 2011). On the first sight, IT capabilities are not only older than BDA capabilities, but also facilitate "day-to-day running of the firm" (Bharadwaj, 2000, p. 175). Moreover, BDA specialists need significantly different skills to IT specialists in a firm (Gupta & George, 2016, p. 1061).

Nevertheless, researchers of BDA and IT capabilities both use the resource-based view to describe the mechanism how these capabilities can enhance value creation in a company (e.g., in terms of overall performance, innovation performance, (Hao et al., 2019; Wamba et al., 2017)). This is not too surprising as both – IT and BDA – are a firm resource that, if it is valuable, rare, inimitable, and non-substitutable, and if this resource is managed well (s. 2.2.1), can lead to a competitive advantage over competitors.

While the theoretical mechanism relies on the same theory, it is not trivial to distinguish the frameworks. Even though Gupta and George (2016, p. 1050) have adopted their framework (tangible-human-intangible resources) from the IS literature, content-wise they do differ in some aspects.

For this thesis, the following IT capabilities papers are considered for a comparison to the BDA capabilities (s. Tab. 4).

5.1.1. Tangible Resources

Structure-wise, in both capabilities frameworks data and technology are defined as important tangible resources. Nevertheless, they differ content-wise.

BDA capabilities are far more specified than IT capabilities. For BDA, exact technologies are defined as machine learning or optimization (s. 2.1.2). Moreover, precise software solutions as Hadoop (s. 4.1.2) are described. Still, the biggest difference between both capabilities lies in data itself. In BDA, data is of major importance, while it is not important per se in a company's IT.

In addition, "big data IT fundamentally differ from traditional IT" (Miloslavskaya & Tolstoy, 2016, p. 301). While for traditional IT a processing device (computer, cloud) is put at the centre of a data processing process, big data IT is seen as a "continuous flowing substance" (Miloslavskaya & Tolstoy, 2016, p. 301). The key lies in the constant float rates of data (which is not a priority for traditional IT), otherwise a theoretical infinite overload of data could occur.

IT capabilities, in contrast, are much broader regarding its tangible elements. Here, the digital support of initial analogous processes is of central importance to an IT department (e.g., CRM, ERP systems).

5.1.2. Human Resources

In all papers the human factor was identified to play an important role in building capabilities. On the one hand, it was described that employees with IT knowledge are needed (for IT capabilities, e.g.: "firm -relevant IT knowledge and competence" (Ross et al., 1996, p. 33)). On the other hand, also support from top management was described to be important.

IS literature, nevertheless, limits the need for IT competence to IT staff and suggests that through the mix of IT and business units the IT staff can gain a "business understanding" (Ross et al., 1996, p. 33). Especially for the development of IT accompaniment of business processes (e.g., digitalization of an ordering process, implementation of an ERP system, or similar), it is understandable why a good skilled IT staff is recommended to understand the needs of business units to then offer suitable IT solutions to the business units.

This stands in wide contrast to BDA capabilities. Certainly, researchers suggest specific knowledge of some employees to perform analytics on big data. But literature also states the need of a comprehensive change in culture – from intuitive decision making to a data-driven culture (s. 4.3.1).

5.1.3. Intangible Resources

In BDA research the focus of intangible resources lies on a data-driven culture and the intensity of organizational learning (s. 4.3). For IT capabilities, intangible resources are focused differently on knowledge assets, customer orientation and synergies (Bharadwaj, 2000, p. 174).

Similarly, BDA and IT capabilities encompass characteristics of a learning organization to enhance adoption to novel BDA or IT solutions (BDA: as a capability: organizational learning, IT, as a resource: knowledge assets).

In fact, IT capabilities stress two other components: *customer orientation* and *synergies*.

In general, *customer orientation* has been found to positively impact firm performance (Jaworski & Kohli, 1993). To satisfy a customer orientation in a firm, often IT is used (e.g., for a classic CRM system). But also, to better understand and predict customer needs, IT relates to a firm's strategy at its foundation.

Synergies describe the idea that IT systems can be re-used in a company (e.g., in different departments, countries, similar cases).

To sum up, BDA intangible resources point on organizational aspects, while IT intangible resources pin on efficiency and quality aspects (being customer-centric with IT applications, re-using technology in a company).

5.2. Overview Digital Capabilities

In 2016, Junior, Maçada, Brinkhues, and Montesdioca (2016, p. 2) conducted a literature review on *Digital Ca-*

Table 4: Overview of proposed IT capabilities

Author(s)	Capabilities proposed
Bharadwaj et al. (1999)	IT business partnership, External IT linkages, Business IT strategic thinking, IT business process integration, IT management, IT infrastructure
Bharadwaj (2000)	IT infrastructure, Human IT Resources, IT enabled Intangibles
Chae et al. (2014)	Rebottle of Bharadwaj (2000)
Ross, Beath, and Goodhue (1996)	Technology Asset, Human Asset, Relationship Asset
Santhanam and Hartono (2003)	Rebottle of Bharadwaj (2000)

pabilities, searching the Basket of Eight as well as Ebscohost and Google Scholar. They could identity 26 relevant papers of which only six papers offered a clear definition of digital capabilities.

Besides the small number of definitions, they are also largely heterogenous. Westerman, Bonnet, and McAfee (2012), for instance, define digital capabilities broadly as skills that go beyond pure IT, and even move in the direction of BDA capabilities, speaking of analytics to retrieve value from big data. Also Lyytinen, Yoo, and Boland Jr. (2016) define digital capabilities as an ability with the aim of creating "new outputs, structures, and behaviours [...] without deliberate planning from the originator of the system" (Lyytinen et al., 2016, p. 53).

Even though there is no homogenous understanding of digital capabilities in research yet, most definitions point in a similar direction: Researchers expect a firm's functioning IT landscape to be a basic prerequisite and aggregate possible extensions under the umbrella of digital capabilities.

5.3. Comparison of the Different Types of Capabilities

IT capabilities have been found to explain how organizations can transform analogous work and processes into digital processes. Doing so, different systems as CRM and ERP systems can be implemented that offer not only more efficient solution for companies, but also provide useful extensions compared to analogous old processes (as connecting an ERP system with suppliers to automatically re-order etc.).

Digital Capabilities are different, as they do not primarily focus on the digital transformation in firms but focus on finding extensions that technology offers to their business. Thus, digital capabilities can be seen as a broad extension of IT capabilities, while BDA capabilities – as a very specific extension of digital capabilities – form a part of digital capabilities.

To also understand how the three technical capabilities – IT, digital, and BDA capabilities – relate to each other, a good indicator is to look at the chronological development (s. Fig. 6).

For most companies, IT became a topic in the beginning of the 1970s when first computers moved in corporations. Researchers and practitioners have deeply discussed the potential and possible pitfalls of bringing IT in an organization.

One major discussion was the emergence of the *IT productivity paradox* (s. 2.4). Researchers could not show how higher IT spending would result in higher performance, but eventually this paradox could be resolved (Gupta & George, 2016).

Furthermore, researchers were already discussing IT Capabilities in the 1990s before the term big data with its popular 3 V's characteristic has been defined in 2001 (Laney, 2001). And another ten years later – after the successful implementation of today's widely used software framework of Hadoop in 2005 (Apache, 2021) – first papers about BDA capabilities were published (LaValle et al., 2011).

With the accelerating supply of cloud computing in the beginning of the 2010s, the processing of big data became easier and - most important - cheaper for organizations, which has most probably also given a boost to BDA usage (West, Battleson, Kim, & Ramesh, 2014; Winegar, 2021). This aspect was out of scope of this paper and thus not further discussed. Nevertheless, it is still relevant to name cloud computing when looking at the chronological development of BDA.

6. Discussion

6.1. Pitfalls for organizations after implementing Big Data Analytics

To answer the first research question, this thesis could clearly show and discuss proposed BDA capabilities frameworks. Nevertheless, a further discussion in conducted to also answer the second part of the first research question: "How do BDA capabilities influence the success of BDA?"

6.1.1. Tangible Resources

Researchers commonly describe distributed and real-time analytics as the next trends in BDA's infrastructure (Landset et al., 2015).

In terms of data itself, on possible pitfall for organizations is to concentrate on building big data resources and improving the BDA process, but not implementing quality checks on the data to "screen out redundant, inaccurate, and duplicated data" (Boldosova, 2019, p. 1142).

Apart from quality, quantity is a major game changer in the world of big data. Google's director of research, Peter Norvig, once stated: "We don't have better algorithms. We
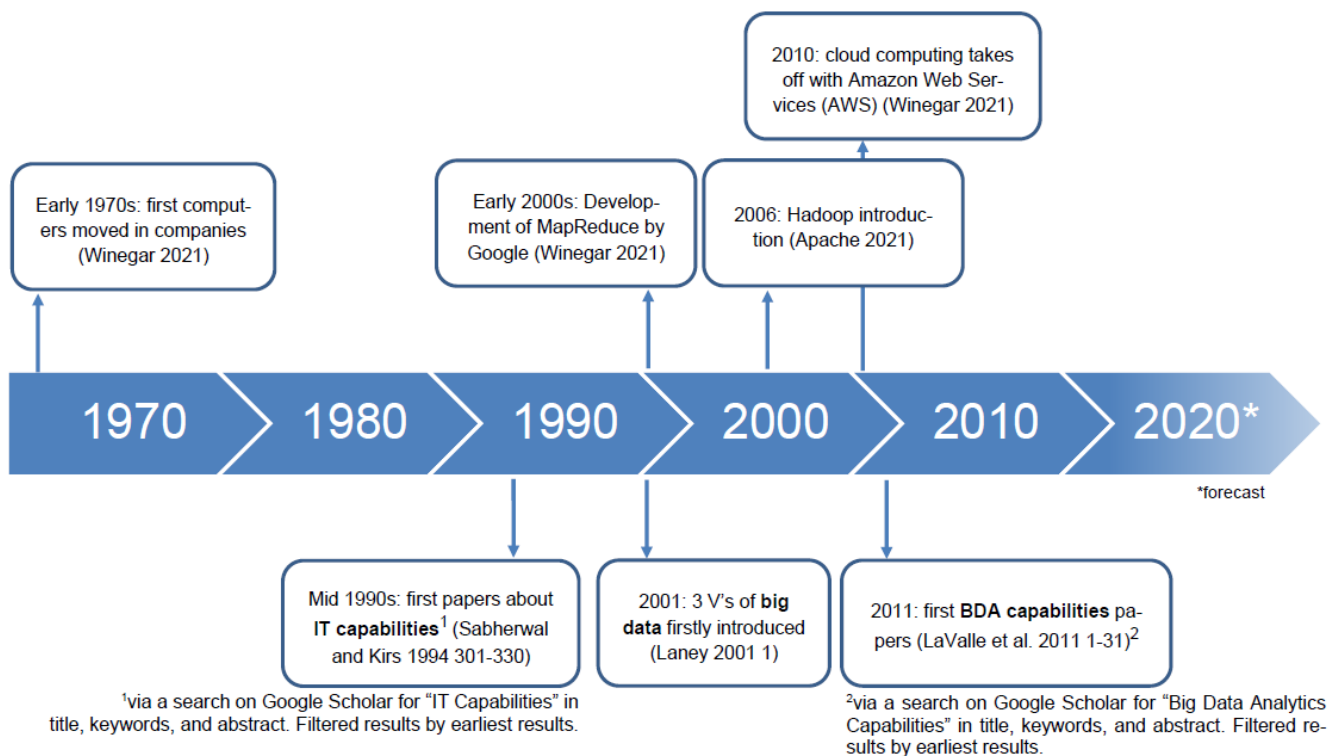
Figure 6: Timeline of technological developments

just have more data." (McAfee et al., 2012, p. 5). In the same direction pins a report from the Executive Office of the President of the United States suggesting that: "technical capabilities of big data have reached a level of sophistication" (Podesta, Pritzker, Moniz, Holdren, & Zients, 2014, p. 5).

It seems that in terms of BDA, data itself can be the differentiator for companies. Ultimately, reaching a competitive advantage in BDA could come down to the simple size of data available to an organization.

6.1.2. Human Resources

A key factor in terms of BDA human resources is the availability of skilled staff. In 2011, McKinsey has predicted a shortage in the US of 140,000 to 190,000 big data skilled individuals by 2018. However, a study by Bain from 2018 also said that they expect that the "global supply of advanced analytics talent will double" (Brahm, Sheth, Sinha, & Dai, 2019, p. 1) from half a million to one million within the next two years.

Another reason why finding skilled employees might not be the key difficulty within the execution of BDA capabilities is offered by a further study from McKinsey in 2020 on AI. In this article it is questioned whether an organization still needs to hire that many data scientist to build machine learning models (Hürtgen, Kerkhoff, Lubatschowski, & Möller, 2020). In this case (based on AI, not BDA), the authors "don't foresee demand for substantial, functional data-science expertise going away anytime soon" (Hürtgen et al.,

2020, p. 4). Nevertheless, one can expect that also the field of BDA will become more accessible as tools and techniques can get easier to handle also for not specialised employees.

6.1.3. Intangible Resources

After implementing the first BDA pipelines (big data via analytics to results), it is easy to miss closing the cycle to full value creation. One possible scenario are managers that now have access to novel data insights, but still keep to a decision-making process based on intuition (Mikalef et al., 2018, p. 571). Referring to the BDA capabilities framework discussed in the literature review, this occurrence is mostly connected to a bad execution of an intuition-based culture to a data-driven culture. Especially managers are visible as role models and thus must take the results of BDA to decisions and value creation (Davenport et al., 2001, p. 136). Also, McAfee et al. (2012) could show that data-driven firms are six percent more profitable and five percent more productive than their rivals.

Another pitfall regarding the organizational aspect is to implement BDA only in specific departments as marketing (Mikalef et al., 2018, p. 570). This also hinders an organization wide change to a data-driven culture.

Furthermore, one aspect that only came up recently and was not considered in the papers from Gupta and George (2016) and Mikalef et al. (2018) is *governance*. When working with data, firms might see themselves confronted with a dilemma of collecting as much data as possible on the one

side and obeying governmental privacy rules (and ultimately limiting the value of BDA) on the other side (Llave et al., 2018). Moreover, data regulations are not settled yet, and changes in the next few years can be expected.

6.2. BDA: Temporary trend versus mega game changer

Concerning the second research question, BDA capabilities were compared to IT and digital capabilities in the second part of this thesis. Here, papers were investigated whether BDA can be seen as a new phenomenon or BDA rather is "old wine in new bottles".

Looking back, the aspiration of making sense of data is nothing new – the statistical field dates to the 18th century. Still, a clear enhancement in analytics has occurred over the past years. Now, data not only moves online at a rapid pace, but can also be analysed in (near to) real-time (Agarwal & Dhar, 2014, p. 444).

This development has highly increased possibilities that can be reached with BDA. Today, the old – more theoretical - idea of making sense of data can eventually lead to big impact in two ways. First, as described in the introduction of this thesis, several papers have already shown that BDA has a positive influence on firm performance (s. 1). Secondly, the new tools and techniques offer new (business) possibilities and insights that could have not been detected before (Wiener, Saunders, & Marabelli, 2020). In this context, literature stresses the importance for incumbent organizations to "rethink their existing business models and how these may be affected by big data" (Günther et al., 2017, p. 197).

A difficult aspect to judge is its chance of substitution – simply put: Will BDA still be important in the next years, while new technical possibilities as AI are emerging, too? To evaluate this question, it is useful to consider the industry an organization is working in.

Regarding adaptability of BDA in an organization, some managers may shrink when thinking about the implementation of BDA and be unsure about the final payoffs for moving from an institutional to an data-driven organization (Gupta & George, 2016). This thesis could show that BDA capabilities are indeed different to IT capabilities and with that the implementation is not straight forward as one might expect. Managers cannot simply rely on their strong IT power of their firm to also conduct analytics. Certainly, literature does suggest similar fields a company has to pay attention to when working with IT or BDA (e.g., both do incorporate human skills), but the needed abilities and adoptions for BDA are far more specific compared to an IT departments' abilities. Here, Barton and Court (2012, p. 81) also suggest that especially managing unstructured data remains beyond IT capabilities. Not to forget, for BDA (in difference to IT) a cultural change among all department is needed. One common misadaptation of BDA is to only use it in specific operational departments as marketing or finance (Mikalef et al., 2018, p. 570), which hinders BDA to spread its full potential.

Concluding, the conceptual idea of analytics has been around for a long time, but for only some years it is feasible to reach high impact results (with new tools and techniques that can handle large and diverse amounts of data quickly). A good IT infrastructure is helpful, but to conduct BDA a firm must invest in the more specific BDA capabilities to reach satisfying results (Barton & Court, 2012). Thus, BDA from its structural architecture is similar to IT capabilities but looking at the specificity of BDA capabilities suggest that they are more than simply "old wine in new bottles".

7. Implications, Limitations, and Outlook

7.1. Implications

This thesis enriches literature on BDA capabilities by presenting researchers' most recent views on the topic of BDA capabilities and delimiting its boards to similar areas of IT and digital capabilities.

First, a clear distinction of important terms regarding big data was proposed, which has not been separated by literature that clearly, yet. With it, also theory frameworks are better accessible, as this paper could provide a clear differentiation between the resource-picking activities that entail the creation of accessible big data, and its capability-building activities that rotate around implementing BDA (s. 2.2).

In the structured literature review different frameworks were shown and discussed and an overview of current's literature status quo was provided. Moreover, a deep dive beyond the core framework of proposed capabilities was conducted, describing in detail what characteristics each capability entails (Gupta & George, 2016; Mikalef et al., 2018).

In the second part of this thesis, BDA was compared to IT and digital capabilities. This differentiation is a novel approach and enriches current literature to offer a clear distinction between the three technical capabilities. Moreover, this thesis could show how BDA has emerged in the beginning of the 2000s and developed since then.

7.2. Limitations

Also, this thesis has limitations. The main limitation is due to the methods used. First, the research was narrowed on papers that matched specific search words ("big data analytics", "big data analytics capabilities") in their title, keywords or abstract. Doing so, studies were filtered that have not used the exact wording but could have added valuable insights. Another problem of literature reviews in general is the chance to miss complementary insights from other fields. To partly prevent this, the research in the second part of this paper was helpful, where near-field capabilities (IT and digital capabilities) were considered and compared to BDA capabilities.

7.3. Outlook

This thesis could also show how BDA has emerged in the early 2000s (s. 5.3). Now, we already see formations of new applications in various directions. A considerable example form Gupta and George (2016) and Mikalef et al. (2018) with their highly cited paper on BDA capabilities. This year,

Mikalef (who has written the extension on Gupta's BDA paper) and Gupta have published a paper on Artificial Intelligence (AI) capabilities together (Mikalef & Gupta, 2021). Gupta has described this paper himself to be an extension to his initial BDA paper.

In this thesis, capabilities from other technological areas were identified and discussed, still, both (IT and digital capabilities) have emerged before BDA. Thus, it is interesting to investigate which new technologies have emerged after BDA. Moreover, no clear distinction is made, yet that weighs parameters to help organizations to decide which technologies should be implemented next or could be potentially substitute by a later technology (or, e.g., could a company skip the implementation of BDA and directly implement AI extensions?).

8. Conclusion

With all the focus of this thesis on needed capabilities for firms to extract the maximum value for firms, one should not forget that big data is not only helpful for organizations. Big data is also helpful for people.

In healthcare, for instance, big data tools are used to support the early detection of diseases or during the COVID-19 crisis to offer an overview about a country's status in the pandemic (Ajah & Nweke, 2019, p. 23).

Also in our daily life, we get in touch with BDA even though we might not notice it at first. When watching Netflix, we value the highly ranked recommendation system that suggests new movies and series to watch – this is possible because of BDA (Lycett, 2013). When driving with our car, we value to have real-time information on traffic – this is possible because of BDA (Ajah & Nweke, 2019, p. 25). When shopping online, we value that stores can predict what we want – before we even know it (Ajah & Nweke, 2019, p. 25).

This thesis has focused on implications for companies when implementing BDA in their work. But one should not forget – big data is not just some company advancement project; it is impacting peoples' lives.

References

Agarwal, R., & Dhar, V. (2014). Big data, data science, and analytics: The opportunity and challenge for IS research. *INFORMS*.

Agrawal, S., & Sureka, A. (2013). *Detecting copyright infringement on YouTube videos using YouTube metadata* (Unpublished doctoral dissertation).

Ajah, I. A., & Nweke, H. F. (2019). Big data and business analytics: Trends, platforms, success factors and applications. *Big Data and Cognitive Computing*, 3(2), 32.

Akhtar, P., Frynas, J. G., Mellahi, K., & Ullah, S. (2019). Big data-savvy teams' skills, big data-driven actions and business performance. *British Journal of Management*, 30(2), 252–271.

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, 182, 113–131.

Apache, H. (2021). *Hadoop Releases - an Apache Webpage*. https://archive.apache.org/dist/hadoop/common/. (Accessed: 2021-06-25)

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

Barney, J. B., Ketchen Jr., D. J., & Wright, M. (2011). The Future of Resource-Based Theory: Revitalization or Decline? *Journal of Management*, 37(5), 1299–1315.

Barton, D., & Court, D. (2012). Making advanced analytics work for you. *Harvard Business Review*, 90(10), 78–83.

Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: an empirical investigation. *MIS quarterly*, 169–196.

Bharadwaj, A. S., Sambamurthy, V., & Zmud, R. W. (1999). IT Capabilites: Theoretical Perspectives And Empirical Operationalization. *ICIS*, 378–385.

Boldosova, V. (2019). Deliberate storytelling in big data analytics adoption. *Information Systems Journal*, 29(6), 1126–1152.

Brahm, C., Sheth, A., Sinha, V., & Dai, J. (2019). *Solving the New Equation for Advanced Analytics Talent* (Tech. Rep.).

Brown, S. L., & Eisenhardt, K. M. (1997). The art of continuous change: Linking complexity theory and time-paced evolution in relentlessly shifting organizations. *Administrative Science Quarterly*, 1–34.

Brynjolfsson, E. (1993). The productivity paradox of information technology. *Communications of the ACM*, 36(12), 66–77.

Cai, L., & Zhu, Y. (2015). The challenges of data quality and data quality assessment in the big data era. *Data Science Journal*, 14.

Chae, H.-C., Koh, C. E., & Prybutok, V. R. (2014). Information technology capability and firm performance: contradictory findings and their possible causes. *MIS quarterly*, 38(1), 305–326.

Choi, T.-M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. *Production and Operations Management*, 27(10), 1868–1883.

Court, D. (2015). Getting big impact from big data. *McKinsey Quarterly*, 1–8.

Davenport, T. H., Harris, J. G., De Long, D. W., & Jacobson, A. L. (2001). Data to knowledge to results: building an analytic capability. *California Management Review*, 43(2), 117–138.

Davenport, T. H., Harris, J. G., & Morison, R. (2010). *Analytics at work: Smarter decisions, better results*. Harvard Business Press.

Dixon, J. (2010). *Pentaho, Hadoop, and Data Lakes*. https://jamesdixon.wordpress.com/2010/10/14/pentaho-hadoop-and-data-lakes/. (Accessed: 2021-06-16)

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they? *Strategic Management Journal*, 21(10-11), 1105–1121.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144.

Günther, M., Wendy Mehrizi, Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *Journal of Strategic Information Systems*, 26, 191–209.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Elsevier: Information & Management*, 53, 1049–1064.

Hao, S., Zhang, H., & Song, M. (2019). Big Data, Big Data Analytics Capability, and Sustainable Innovation Performance. *Sustainability*, 11, 1–15.

Hu, W., Xie, N., Li, L., Zeng, X., & Maybank, S. (2011). A Survey on Visual Content-Based Video Indexing and Retrieval. *IEEE Transactions on Systems*, 41(6), 797.

Hürtgen, H., Kerkhoff, S., Lubatschowski, J., & Möller, M. (2020). *Rethinking AI talent strategy as automated machine learning comes of age* (Tech. Rep.).

Jaworski, B. J., & Kohli, A. K. (1993). Market orientation: antecedents and consequences. *Journal of Marketing*, 57(3), 53–70.

Junior, J. C. d. S. F., Maçada, A. C. G., Brinkhues, R., & Montesdioca, G. P. Z. (2016). *Digital Capabilities as Driver to Digital Business Performance*. AMCIS.

Keller, W. (2009). Using Capabilities in Enterprise Architecture Management [White paper].

Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, 34(3), 387–394.

Lamba, H. S., & Dubey, S. K. (2015). Analysis of requirements for big data adoption to maximize IT business value. *IEEE - 4th International Conference on Reliability, Infocom Technologies and Optimization*, 1–6.

Landset, S., Khoshgoftaar, T. M., Richter, A. N., & Hasanin, T. (2015). A survey of open source tools for machine learning with big data in the Hadoop ecosystem. *Journal of Big Data*, 2(1), 1–36.

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note*, 6(70), 1.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, S. (2011). Big Data, Analytics and the Path From Insights to Value. *MITSloan Management Review*, 1–31.

Llave, M. R., Hustad, E., & Olsen, D. H. (2018). *Creating value from business intelligence and analytics in SMEs: Insights from experts*. AMCIS.

Lycett, M. (2013). "Datafication': making sense of (big) data in a complex world. *European Journal of Information Systems*, 22, 381–386.

Lyytinen, K., Yoo, Y., & Boland Jr., R. J. (2016). Digital product innovation within four classes of innovation networks. *Information Systems Journal*, 26(1), 47–75.

Makadok, R. (2001). Toward a synthesis of the resource-based and dynamic-capability views of rent creation. *Strategic Management Journal*, 22(5), 387–401.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. *Harvard Business Review*, 90(10), 60–68.

Mikalef, P., Framnes, V. A., Danielsen, F., & Krogstie, J. (2018). *Big Data Analytics Capability: Antecedents and Business Value* (Vol. 16). Springer-Verlag GmbH Germany.

Mikalef, P., & Gupta, M. (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. *Elsevier: Information & Management*, 58, 1–20.

Miloslavskaya, N., & Tolstoy, A. (2016). Big data, fast data and data lake concepts. *Procedia Computer Science*, 88, 300–305.

Podesta, J., Pritzker, P., Moniz, E., Holdren, J., & Zients, J. (2014). Big Data: Seizing Opportunities, Preserving Values. *Executive Office of the President*.

Ransbotham, S., Kiron, D., & Prentice, P. K. (2015). *Minding the analytics gap* (Vol. 56). MIT Sloan Management Review.

Rosenthal, D. S. H., Rosenthal, D. C., Miller, E. L., Adams, I. F., Storer, M. W., & Zadok, E. (2012). The Economics of Long -Term Digital Storage. *The Memory of the World in the Digital Age: Digitization and Preservation*, 1–17.

Ross, J. W., Beath, C. M., & Goodhue, D. L. (1996). Develop long-term competitiveness through IT assets. *Sloan Management Review*, 38(1), 31–42.

Russom, P. (2011). Big Data Analytics. *The Data Warehousing Institute*.

Sabherwal, R., & Kirs, P. (1994). The alignment between organizational critical success factors and information technology capability in aca-

demic institutions. *Decision Sciences*, *25*(2), 301–330.

Santhanam, R., & Hartono, E. (2003). Issues in linking information technology capability to firm performance. *MIS quarterly*, 125–153.

Seddon, P. B., Constantinidis, D., Tamm, T., & Dod, H. (2017). How does business analytics contribute to business value? *Information Systems Journal*, *27*(3), 237–269.

Soltanpoor, R., & Sellis, T. (2016). *Prescriptive Analytics for Big Data*. Springer International Publishing.

Soto Setzke, D., Opderbeck, L., Böhm, M., & Krcmar, H. (2020). Pathways to Successful Business Model Innovation in the Context of Digital Transformation. *24th Pacific Asia Conference on Information Systems*.

Tallon, P. P., Ramirez, R. V., & Short, J. E. (2013). The information artifact in IT governance: toward a theory of information governance. *Journal of Management Information Systems*, *30*(3), 141–178.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, *18*(7), 509–533.

Tiguint, B., & Hossari, H. (2020). Big data analytics and artificial intelligence: a meta-dynamic capability perspective. *changes*.

Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research*, *261*(2), 626–639.

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J.-F., & Dubey, R. (2017). Big data analytics and firm performance: effects of dynamic capabilities. *Journal of Business Research*, *70*, 356–365.

Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Elsevier: Technological Forecasting & Social Change*, *126*, 3–13.

Webster, J., & Watson, R. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, *26*(2), 13–13.

West, B. C., Battleson, D. A., Kim, J., & Ramesh, B. (2014). Achieving dynamic capabilities with cloud computing. *IT Professional*, *16*(6), 18–24.

Westerman, G., Bonnet, D., & McAfee, A. (2012). The digital capabilities your company needs. *MIT Sloan Management Review*, 1–5.

White, T. (2012). *Hadoop: The definitive guide*. O'Reilly Media, Inc.

Wiener, M., Saunders, C., & Marabelli, M. (2020). Big-data business models: A critical literature review and multiperspective research framework. *Journal of Information Technology*, *35*(1), 66–91.

Winegar, J. (2021). *What is "Big Data" – Understanding the History*. https://towardsdatascience.com/what-is-big-data-understanding-the-history-32078f3b53ce. (Accessed: 2021-06-15)

Yasmina, M., Tatoglua, E., Kilicb, H. S., Zaimc, S., & Delen, D. (2020). Big data analytics capabilities and firm performance: An integrated MCDM approach. *Elsevier: Journal of Business Research*, *114*, 1–15.

Effects of Internal Auditing on the Quality of Financial Reporting – A Critical Analysis

Auswirkungen der internen Revision auf die Qualität der Rechnungslegung – Eine kritische Analyse

Jette Fabian

Universität Hamburg

Abstract

In order to uncover the weaknesses in corporate governance revealed by both accounting scandals and the financial crisis, one possible solution is for internal auditing, as a process-independent in-house auditing institution, to continuously monitor and improve the effectiveness of measures taken within the framework of governance and risk management. In addition to presenting both definitional basics and the normative and conceptual framework, this paper critically analyzes whether and to what extent internal auditing has an impact on the quality of financial reporting and how it should be designed in order to optimally fulfill its mandate to the respective company management. As a result, a high-quality internal audit, as one of the four pillars of corporate governance, can represent a significant success factor in reducing earnings management, uncovering accounting manipulation and, thus, avoiding accounting scandals. The decisive factor for positively influencing financial reporting is the consideration of the key drivers of internal audit quality as well as the alignment of audit procedures with predefined frameworks.

Zusammenfassung

Um die sowohl von Bilanzskandalen als auch der Finanzkrise aufgedeckten Schwächen in der Corporate Governance rechtzeitig unternehmensintern aufzudecken, soll als eine mögliche Lösung die Interne Revision als prozessunabhängige Prüfungsinstitution die im Rahmen der Governance und des Risikomanagements getroffenen Maßnahmen hinsichtlich ihrer Wirksamkeit laufend überwachen und verbessern. Nebst der Darstellung definitorischer Grundlagen und des normativen sowie konzeptionellen Rahmens wird in dieser Arbeit kritisch analysiert, ob und inwiefern die Interne Revision Auswirkungen auf die Qualität der Rechnungslegung hat und wie diese ausgestaltet sein sollte, um ihrem Auftrag gegenüber der jeweiligen Unternehmensführung optimal gerecht zu werden. Im Ergebnis kann festgehalten werden, dass eine qualitativ hochwertige Interne Revision, als einer von vier Grundpfeilern der Corporate Governance, einen wesentlichen Erfolgsfaktor zur Reduktion von Bilanzpolitik, zur Aufdeckung von Bilanzmanipulationen und somit zur Vermeidung von Bilanzskandalen darstellen kann. Entscheidend für die positive Einflussnahme auf die Rechnungslegung ist die Berücksichtigung der wesentlichen Treiber der Qualität der Internen Revision sowie die Orientierung der Prüfungshandlungen an vorgegebenen Rahmenwerken.

Keywords: Interne Revision; Corporate Governance; Rechnungslegungsqualität; Bilanzpolitik.

1. Einleitung

Das Bewusstsein für die Legitimität und Relevanz der Internen Revision als Teil der Corporate Governance[1] und des Risikomanagements wurde seit mehreren S&P 500[2] Unternehmenszusammenbrüchen und den damit verbundenen regulatorischen Anpassungen geschaffen.[3] Bei zahlreichen

---

[1] „Unter Corporate Governance wird der rechtliche und faktische Ordnungsrahmen für die Leitung und Überwachung eines Unternehmens verstanden." DCGK (Hrsg.) (2019), S. 2.Weiterführend hierzu: Mallin (2019), S. 17-27).

[2] Der Aktienindex S&P 500 umfasst die Aktien der größten börsennotierten US-amerikanischen Unternehmen.

[3] Vgl. Bajra und Cadez (2018, S. 152); Behrend und Eulerich (2019, S. 104); Davidson, Goodwin-Stewart und Kent (2005, S. 242); Hölscher und Rosenthal (2007, S. 9).

Bilanzskandalen der Vergangenheit haben mehrere der von Unternehmen eingerichteten Maßnahmen im Zuge der Corporate Governance gleichzeitig versagt.[4] Um die sowohl von Bilanzskandalen als auch der Finanzkrise aufgedeckten Schwächen in der Corporate Governance rechtzeitig unternehmensintern aufzudecken, soll als eine mögliche Lösung die Interne Revision als prozessunabhängige Prüfungsinstitution die im Rahmen der Governance und des Risikomanagements getroffenen Maßnahmen hinsichtlich ihrer Wirksamkeit laufend überwachen und verbessern.[5] Im Rahmen dieser Masterarbeit soll kritisch analysiert werden, ob und inwiefern die Interne Revision Auswirkungen auf die Qualität der Rechnungslegung hat. Es stellt sich die Frage, ob und inwiefern Bilanzskandale durch eine qualitativ hochwertige Interne Revision verhindert oder vorzeitig aufgedeckt werden können. Zudem gilt es zu klären, wie eine Interne Revision ausgestaltet sein sollte, damit diese ihrem Auftrag gegenüber der jeweiligen Unternehmensführung optimal gerecht werden kann. Zur Klärung dieser Ausgangsfragen erfolgt eingangs eine Darstellung der definitorischen Grundlagen sowie des normativen und konzeptionellen Rahmens der Internen Revision. Im Anschluss werden die normativen und konzeptionellen Grundlagen der Rechnungslegung dargelegt. Zudem erfolgt eine Einführung in Rechnungslegungspolitik und -qualität. Im Anschluss wird mittels Betrachtung sowohl nationaler als auch internationaler wissenschaftlicher Literatur erläutert, ob und inwiefern die Interne Revision einen Einfluss auf die Qualität der Rechnungslegung hat und welche Faktoren diesen determinieren. Eine kritische Würdigung der Thematik sowie ein Fazit schließen diese Masterarbeit ab.

2. Interne Revision

2.1. Definitorische Grundlage

Jedes Unternehmen verfolgt in seinem täglichen Bestreben selbst gesetzte Ziele und kann bei deren Verfolgung auf Hindernisse stoßen, die der Erreichung der gesetzten Ziele im Wege stehen.[6] Diese Hindernisse schaffen Risiken, die von den Unternehmen erkannt, analysiert und mit entsprechender Reaktion angegangen werden müssen. Zur Minimierung von Risiken stehen den Unternehmen diverse Methoden zur Verfügung. Abgesehen von Überwachungsprozessen und Risikomanagementsystemen zählt insbesondere die Einrichtung eines Internen Kontrollsystems (IKS)[7] zu den wesentlichen Maßnahmen. Wie die nachfolgend aufgeführten Definitionen darstellen, setzt die Interne Revision genau an dieser Stelle an, indem sie die Wirksamkeit der von Unternehmen eingerichteten Maßnahmen zur Reduktion und frühzeitigen Erkennung von Risiken überwacht und verbessert. Zudem dient die Interne Revision der Reduktion

von Informationsasymmetrien und Interessenkonflikten wie sie gemäß der Prinzipal-Agenten-Theorie in Unternehmen häufig zu finden sind.[8] Da die wissenschaftliche Literatur für die Darstellung der Aufgaben der Internen Revision keine einheitliche Definition bietet, wird im Folgenden eine Auswahl der gängigsten Definitionen aufgeführt. Von großer Bedeutung ist die deutsche Übersetzung der Definition des International Professional Practices Framework (IPPF) des Institute of Internal Auditors (IIA) als internationale Vertretung des Berufsstandes, da diese als Selbstverständnis interpretiert werden kann. Sie kann zudem als definitorische Grundlage der Aufgaben und Tätigkeitsfelder der Internen Revision interpretiert werden. Das IIA definiert die Interne Revision als Institution, die „unabhängige und objektive[9] Prüfungs- und Beratungsdienstleistungen [erbringt – Anm. d. Verf.], welche darauf ausgerichtet sind, Mehrwerte zu schaffen und die Geschäftsprozesse zu verbessern. Sie unterstützt die Organisation bei der Erreichung ihrer Ziele, indem sie mit einem systematischen und zielgerichteten Ansatz die Effektivität des Risikomanagements, der Kontrollen und der Führungs- und Überwachungsprozesse bewertet und diese verbessern hilft".[10] Eine ähnliche Definition stellt das Institut der Wirtschaftsprüfer e. V. (IDW) bereit, denn gemäß dessen „[umfassen – Anm. d. Verf.] Ziele und Aufgabenbereich der Internen Revision (...) i.d.R. Prüfungs- und Beratungstätigkeiten, die dazu konzipiert sind, die Wirksamkeit der Überwachungsprozesse, des Risikomanagements und des Internen Kontrollsystems der Einheit zu beurteilen und zu verbessern".[11] Das IDW betont im Zuge der weiteren Ausführungen des Prüfungsstandards, dass unter anderem Risikopotenziale identifiziert, die Wirksamkeit getroffener Maßnahmen beurteilt, Verbesserungsempfehlungen ausgesprochen sowie Wirtschaftlichkeit, Effizienz und Wirksamkeit der laufenden Geschäftstätigkeit bewertet werden sollen.[12] Außerdem gehört die Untersuchung finanzieller und betrieblicher Informationen sowie die Überwachung der Einhaltung von Rechtsnormen gemäß dem Verständnis des IDW zum Aufgabengebiet der Internen Revision. Eine weiterführende und deutlich detailliertere Definition bietet der IDW durch den Prüfungsstandard *IDW PS 321*, demnach „unter der [eigenständigen organisatorischen Einheit, genannt Interne Revision, – Anm. d. Verf.] (...) eine unternehmenseigene prozessunabhängige, prüfende, beurteilende und beratende Tätigkeit verstanden [wird – Anm. d. Verf.], die innerhalb eines Unternehmens oder Konzerns durchgeführt wird. Sie unterstützt das (Mutter-)Unternehmen bei der Erreichung seiner unternehmerischen Ziele. Dazu bewertet sie mit einem systematischen und zielgerichteten Ansatz die Wirksamkeit des Risikomanagementsystems, des Steuerungs- und Überwachungssystems einschließlich der Kontrollen und hilft

---

[4]Vgl. Bantleon et al. (2017, S. 684); ECIIA and FERMA (2014, S. 5); Kirkpatrick (2009, S. 2-3).

[5]Vgl. DIIR (Hrsg.) (2020b, S. 3); Kirkpatrick (2009, S. 20 & 23).

[6]Vgl. hier und im Folgenden Anderson und Eubanks (2015, S. 1).

[7]Weiterführend zum IKS Hamacher (2015, S. 25-29).

[8]Weiterführend zur Prinzipal-Agenten-Theorie im Zusammenhang mit der Internen Revision siehe Eulerich und Velte (2013, S. 146-147 & 150).

[9]Weiterführend zur Definition von Unabhängigkeit und Objektivität der Internen Revision siehe DIIR (Hrsg.) (2018), S. 24.

[10]DIIR (Hrsg.) (2018), S. 13.

[11]ISA [DE] 610.A1 (Revised 2013).

[12]Vgl. hier und im Folgenden ISA [DE] 610.A1 (Revised 2013).

diese zu verbessern. Insbesondere umfasst die Funktion der internen Revision die Untersuchung, Bewertung und Überwachung der Angemessenheit und Wirksamkeit des internen Kontrollsystems einschließlich des Rechnungslegungssystems (...) und der in diesem System ablaufenden Prozesse".[13] Die nachfolgende Abbildung veranschaulicht den im Rahmen der definitorischen Grundlagen dargestellten Aufgabenbereich der Internen Revision.

Die Abbildung illustriert, welche Aufgaben in den Zuständigkeitsbereich der Internen Revision fallen und welche Tätigkeiten klar von der Internen Revision abgegrenzt sein sollten. Hierbei wird deutlich, dass klassische Assurance-Aufgaben in den Kerntätigkeitsbereich gehören, wohingegen beratende Aktivitäten eher eine untergeordnete Rolle spielen und alle darüberhinausgehende Verantwortung eindeutig bei der Unternehmensleitung bleiben sollte.[14] Die Interne Revision bildet als objektive, unabhängige und unternehmensinterne Institution einen integralen Bestandteil der Überwachung des Governance- und Risikomanagementsystems von Unternehmen.[15] Die Ergebnisse der durchgeführten Prüfungstätigkeiten sowie daraus abgeleitete Handlungsempfehlungen werden in der Regel direkt an die Unternehmensführung gemeldet. Die Einordnung der Internen Revision in den gesamtunternehmerischen Kontext kann auf Basis von Rahmenwerken und Modellen vorgenommen werden. Zwei wesentliche Ansätze, auf die sich die wissenschaftliche Literatur bezieht, sind das COSO Internal Control – Integrated Framework und das Three Lines of Defense Modell. Das COSO Internal Control – Integrated Framework ist das bis dato am häufigsten verwendete Rahmenwerk für die Implementierung eines internen Kontrollsystems und besteht aus fünf IKS Komponenten und 17 Grundsätzen, wie diese Komponenten miteinander verbunden sind.[16] In der Regel arbeiten bezüglich des Risikomanagements unterschiedliche Fachabteilungen eines Unternehmens zusammen und die anfallenden Aufgaben der im Rahmen des IKS implementierten Kontrollen müssen koordiniert und überwacht werden, um die Funktionsfähigkeit sicherzustellen.[17] Als ein sehr verbreitetes und effizientes Modell zur Koordination des Risikomanagements hat sich das Three Lines of Defense Model bewährt, das im Bankenwesen sogar verpflichtend anzuwenden ist.[18] Die Relevanz und Akzeptanz des Three Line of Defense Modells können anhand der Rezeption und Bezugnahme durch zahlreiche Standardsetzer hergeleitet werden.[19] Das Modell bietet einen effizienten und systematischen Ansatz zur Koordination der anfallenden Aufgaben, bestehend aus drei voneinander unabhängigen[20] Verteidi-

gungslinien.[21] Jede Verteidigungslinie spielt eine separate Rolle in der Überwachung des Risikomanagement- und Kontrollumfeldes von Unternehmen. Das Three Lines of Defense Modell verhindert im optimalen Fall die doppelte Ausführung gleicher Aufgaben durch die Regelung klarer Verantwortlichkeiten mit deutlich abgegrenzten Tätigkeitsbereichen und zudem das Übersehen von Kontrolllücken und bedeutsamen Risiken.[22] Der Standardsetzer COSO hat das von zahlreichen Standardsetzern adaptierte Three Lines of Defense Modell in der neuesten Überarbeitung des COSO Internal Control – Integrated Framework allerdings nicht berücksichtigt.[23]

2.2. Normativer und konzeptioneller Rahmen

Abgesehen von Ausnahmen in spezifischen Branchen[24] liegt keine Rechtsnorm vor, die deutsche Unternehmen explizit dazu verpflichtet, eine Interne Revision als Institution einzurichten.[25] Dennoch finden sich in den deutschen, europäischen und internationalen Rechtsquellen zahlreiche Passagen, aus denen sowohl die Legitimation der Internen Revision als auch die Einordung in das Corporate-Governance-System hergeleitet werden kann. Zur besseren Übersicht werden in der nachfolgenden Abbildung die im Rahmen dieses Kapitels erläuterten Normen grafisch dargestellt.[26]

Regelungen bezüglich der Internen Revision beziehen sich in den deutschen Normen insbesondere auf die Geschäftsleitung bzw. den Vorstand sowie dessen Kontrollgremium, also den Aufsichtsrat.[27] Allem voran ist die Ergänzung des Aktiengesetzes um *§ 91 Abs. 2 AktG* zu nennen, genauer durch den Erlass des Gesetzes zur Kontrolle und Transparenz im Unternehmensbereich (KonTraG) im Jahr 1998.[28] Ziel dieses Artikelgesetzes[29] mit Änderungen sowohl im Handelsrecht als auch im Aktiengesetz ist die Verbesserung der Corporate Governance in deutschen Unternehmen.[30] Gemäß *§ 91 Abs. 2 AktG* hat „der Vorstand [von Aktiengesellschaften – Anm. d. Verf.] (...) geeignete Maßnahmen zu treffen,

---

[13]IDW PS 321.1.

[14]Vgl. IIA (2009, S. 4-6).

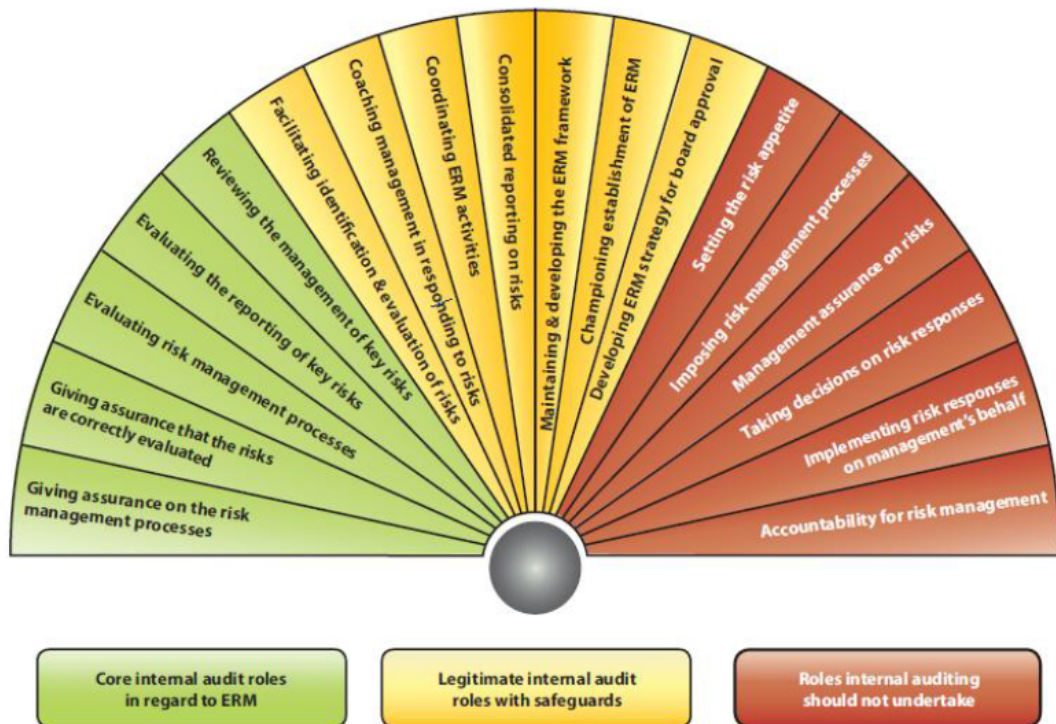[15]Vgl. hier und im Folgenden DIIR (Hrsg.) (2020a, S. 6).

[16]Vgl. Anderson und Eubanks (2015, S. 13); ECIIA and FERMA (2011, S. 15).

[17]Vgl. IIA (2013b, S. 1).

[18]Vgl. Bantleon et al. (2021, S. 59); Bantleon et al. (2017, S. 683).

[19]Eine Auflistung der Bezugnahmen von Standardsetzern bezüglich des Three Line of Defense Modells finden sich unter Bantleon et al. (2017, S. 685).

[20]Die Auffassung inwiefern die Aufgaben der Verteidigungslinien vonein-

ander abgegrenzt und unabhängig sein müssen variiert in der vorhandenen Literatur. Einigkeit besteht aber darüber, dass die jeweilige Abgrenzung der Zuständigkeiten und Tätigkeitsfelder der Verteidigungslinien klar kommuniziert sein muss.Siehe hierzu beispielsweise DIIR (Hrsg.) (2020b, S. 3).

[21]Vgl. hier und im Folgenden IIA (2013a, S. 1).

[22]Weiterführend zur Definition bedeutsamer Risiken siehe IDW PS 261.65-67.

[23]Vgl. Bantleon et al. (2017, S. 686).

[24]Eine gesetzliche Verpflichtung zur Einrichtung einer Internen Revision besteht für Kreditinstitute gemäß §25a Abs. 1 Satz 3 Nr. 3 KWG und für Versicherungsunternehmen gemäß §64a Abs. 1 Satz 4 Nr. 4 VAG.Weiterführend über Notwendigkeit der Internen Revision in Kredit- und Finanzdienstleistungsunternehmen siehe BaFin (Hrsg.) (2017, AT 4.4.3 und für Versicherungsunternehmen BaFin (Hrsg.) (2017, TZ 9.4.Da es sich bei Kreditinstituten und Versicherungsunternehmen um besondere Geschäftsmodelle handelt, werden diese nur der Vollständigkeit halber erwähnt, aber im Rahmen der Arbeit nicht näher betrachtet.

[25]Vgl. Berwanger und Kullmann (2012, S. 89); Hamacher (2015, S. 17 & 19); Marten, Quick und Ruhnke (2015, S. 27).

[26]Der Fokus der Ausführungen dieses Kapitels liegt auf den wesentlichen Normen, die Aufzählung ist nicht vollzählig.

[27]Vgl. Hamacher (2015, S. 18).

[28]Vgl. Art. 1 Nr. 9 Buchstabe c) KonTraG.

[29]Ein Artikelgesetz ändert zeitgleich unterschiedliche Gesetze.

[30]Vgl. Wengert und Schittenhelm (2013, S. 9).

Abbildung 1: Aufgabenbereich der Internen Revision

insbesondere ein Überwachungssystem einzurichten, damit den Fortbestand der Gesellschaft gefährdende Entwicklungen früh erkannt werden".[31] Die durch den Vorstand gemäß *§ 91 Abs. 2 AktG* durchzuführenden Maßnahmen werden als Risikofrüherkennungssystem[32] bezeichnet und untergliedern sich zum einen in die Etablierung von Maßnahmen, um Risiken frühzeitig zu erkennen, und zum anderen in die Überprüfung, ob entsprechende Regelungen eingehalten wurden.[33] Eine Konkretisierung der bestandgefährdenden Risiken und Ausgestaltung des geforderten Überwachungssystems wird im Rahmen der gesetzlichen Regelung nicht vorgenommen. Die Intention der Gesetzgebung wird erst durch die Regierungsbegründung zum Gesetzesentwurf deutlicher hervorgehoben, denn durch den Erlass des *§ 91 Abs. 2 AktG* soll „die Verpflichtung des Vorstands, für ein angemessenes Risikomanagement und für eine angemessene interne Revision zu sorgen, (...) verdeutlicht werden".[34] Zudem präzisiert die Regierungsbegründung, dass die Regelungen des *§ 91 Abs. 2 AktG* unter bestimmten Voraussetzungen auch für Unternehmen anderer Rechtsformen[35] Ausstrahlungs-

wirkung entfalten.[36] Gemäß *§ 317 Abs. 4 HGB*[37] ist das von börsennotierten Aktiengesellschaften eingerichtete Überwachungssystem von einem Abschlussprüfer[38] daraufhin zu beurteilen, „ob (...) die (...) obliegenden Maßnahmen in einer geeigneten Form getroffen [wurden – Anm. d. Verf.] und ob das danach einzurichtende Überwachungssystem seine Aufgaben erfüllen kann".[39] Der den gesetzlichen Prüfungsauftrag präzisierende Prüfungsstandard, IDW PS 340 n.F. weist auf die Notwendigkeit der unternehmensinternen und prozessunabhängigen Überwachung der Maßnahmen nach *§ 91 Abs. 2 AktG* hin, mit Verweis auf eine potenziell vorhandene Interne Revision und empfiehlt die Aufnahme der Beschreibung der Aufgaben der Internen Revision in das Risikohandbuch[40].[41] Der IDW PS 340 n.F. zeigt auf, an welchen Maßnahmen und Prozessen eine Interne Revision im Rahmen eines Risikofrüherkennungssystems beteiligt sein kann und inwiefern diese für die Prüfung relevante Erkenntnisse

---

[31] §91 Abs. 2 AktG.

[32] Gemäß IDW PS 340 n.F. umfasst das Risikofrüherkennungssystem: Risikokultur, Ziele und Organisation der Maßnahmen, Risikoidentifikation, Risikobewertung, Risikosteuerung, Risikokommunikation, Überwachung und Verbesserung.

[33] Vgl. Hamacher (2015, S. 19).

[34] Deutscher Bundestag (1998), S. 15.

[35] Gemeint ist hiermit insbesondere die Rechtsform der GmbH. Zum einen

wird in der Regierungsbegründung festgehalten, dass für die Rechtsform der GmbH nichts anderes gilt als in § 91 Abs. 2 AktG bestimmt ist (vgl. Deutscher Bundestag (1998), S. 15) und zum anderen haben die Sorgfaltspflichten der Geschäftsführung einer GmbH (vgl. §43 Abs. 1 GmbHG), denen des Vorstandes einer AG (vgl. §93 Abs. 1 Satz 1 AktG) zu entsprechen.

[36] Vgl. Deutscher Bundestag (1998), S. 15.

[37] Wurde durch den Beschluss des TransPuG im Jahr 2002 in Kraft gesetzt.

[38] Weiterführend zur Auswahl der Abschlussprüfer siehe §319 Abs. 1 HGB.

[39] §317 Abs. 4 HGB.

[40] Weiterführend zum Aufbau des Risikohandbuchs siehe Wengert und Schittenhelm (2013, S. 16-18 & 89).

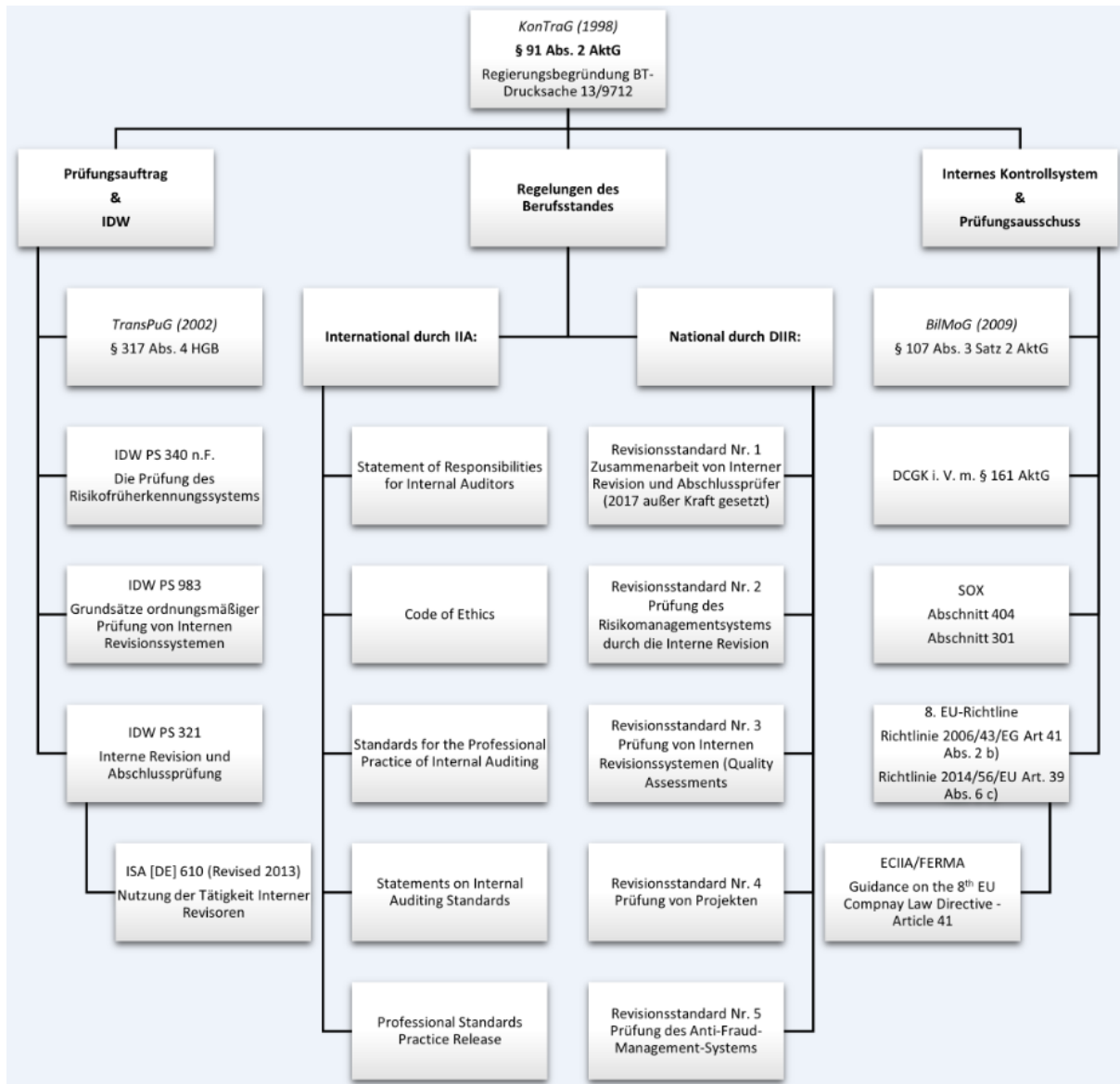[41] Vgl. IDW PS 340.A27 n.F.; IDW PS 340.A30 n.F.

Abbildung 2: Grafische Darstellung der einschlägigen Normen

Quelle: eigene Darstellung.

liefern kann.[42] Es ist ersichtlich, dass die Interne Revision, wenn vorhanden, ein zentrales Element des Risikofrüherkennungssystems bildet. Gemäß Prüfungsstandards des IDW können Wirtschaftsprüfer im Rahmen der Jahresabschlussprüfung[43] Arbeiten bzw. Ergebnisse der Internen Revision eines Unternehmens verwerten, wenn dessen Wirksamkeit als positiv eingeschätzt wird.[44] Für die Vornahme dieser Einschätzung sieht *IDW PS 321*[45] die folgenden vier Kriterien

vor, gemäß derer die Abschlussprüfer die organisatorische Einordnung, den Umfang der Tätigkeit, die fachliche Kompetenz und die berufliche Sorgfalt einschätzen sollen.[46]

Da derzeit keine expliziten gesetzlichen Vorgaben zur Ausgestaltung der Internen Revision existieren, können sowohl die IDW PS als auch die darin enthaltenen Verweise auf entsprechende Rahmenwerke als Orientierung herangezogen werden. Regelungen zur freiwilligen Prüfung des Internen Revisionssystems finden sich in *IDW PS 983*, der

---

[42]Vgl. hier und im Folgenden IDW PS 340.35 n.F.; IDW PS 340.A27 n.F.; IDW PS 340.A37 n.F.; IDW PS 340.A45 n.F.; IDW PS 340.A48 n.F.; IDW PS 340.A50 n.F.

[43]Die Pflicht zur Prüfung des Jahresabschlusses ergibt sich gemäß §316 HGB.

[44]Vgl. IDW PS 321.11-17.

[45]In Deutschland wird für die Prüfung von Berichtszeiträumen die nach

dem 15.12.2021 beginnen, von IDW PS 321 auf ISA [DE] 610 (Revised 2013) übergangen, wodurch sich jedoch keine neuen Anforderungen für den Abschlussprüfer ergeben.Vgl. ISA [DE] 610 (Revised 2013), S. 1; Wirtschaftsprüferkammer (WPK) (Hrsg.) (2020, S. 1).

[46]Vgl. IDW PS 321.17.

insbesondere auf die Relevanz der verbindlichen Elemente[47] des International Professional Practices Framework (IPPF) des Institute of Internal Auditors (IIA), als international anerkanntes Rahmenwerk für die berufliche Praxis der internen Revision, verweist.[48] Das IPPF ist demnach das konzeptionelle Rahmenwerk der internen Revision und enthält sowohl verbindliche Leitlinien als auch Empfehlungen für die berufliche Praxis der Internen Revision.[49] Zwar muss sich nicht zwingend an diesen Regelungen orientiert werden, aber die Mitglieder des IIA sowie assoziierte Verbände haben sich sowohl dem IPPF als Rahmenwerk als auch den vom IIA herausgegebenen International Standards for the Professional Practice of Internal Auditing verpflichtet.[50] Auch die Einhaltung der vom Deutschen Institut für Interne Revision e. V. (DIIR)[51], als Vertretung des deutschen Berufsstandes der Internen Revision, veröffentlichten Revisionsstandards Nr. 1-5 sind insbesondere auf nationaler Ebene dringend empfohlen, allerdings werden diese auf internationaler Ebene eher als praktische Empfehlung angesehen.[52]

Regelungen bezüglich der Internen Revision, in Verbindung mit dem Aufsichtsrat, finden sich zudem in dem als unternehmerischer Verhaltenskodex ins Leben gerufenen Deutschen Corporate Governance Kodex (DCGK). „Sowohl die Regelungen des Gesetzgebers als auch die Richtlinien für gute Corporate Governance haben das gemeinsame Ziel, die Transparenz für die Investoren zu erhöhen, die Interessen der Stakeholder angemessen zu berücksichtigen und die Rolle des Aufsichtsrates zu stärken."[53] In klassischen Aktiengesellschaften sind mit dem Corporate Governance Management die Unternehmensführung sowie, in kontrollierender Funktion, Vorstand und Aufsichtsrat betraut. Gemäß Präambel des DCGK „enthält [dieser – Anm.d.Verf.] Grundsätze, Empfehlungen und Anregungen für den Vorstand und den Aufsichtsrat, die dazu beitragen sollen, dass die Gesellschaft im Unternehmensinteresse geführt wird, (...) unter Berücksichtigung der Belange der Aktionäre, der Belegschaft und der sonstigen mit dem Unternehmen verbundenen Gruppen (...)".[54] Eine gesetzliche Verpflichtung zur Einhaltung der Empfehlungen des DCGK gibt es nicht, allerdings sind börsennotierte Unternehmen gemäß *§ 161 AktG* zu einer jährlichen Entsprechenserklärung verpflichtet.[55] In dieser müssen sie darstellen, inwieweit den Empfehlungen im Geschäftsjahr entsprochen wurde bzw. im folgenden Geschäftsjahr entsprochen wird und mit einschlägiger Erläuterung, welche nicht angewendet wurden bzw. werden. Gemäß Empfehlung des DCGK soll durch den Aufsichtsrat ein Prüfungsausschuss eingerichtet werden, der sich mit „der Prüfung der Rech-

nungslegung, der Überwachung des Rechnungslegungsprozesses, der Wirksamkeit des internen Kontrollsystems, des Risikomanagementsystems und des internen Revisionssystems sowie der Abschlussprüfung und der Compliance befasst".[56] Diese Empfehlung knüpft an das Wahlrecht gemäß *§ 107 Abs. 3 Satz 2 AktG* an, demnach kann der Aufsichtsrat aus seiner Mitte einen Prüfungsausschuss bestellen, der unter anderem mit der Überwachung der Wirksamkeit des Internen Revisionssystems betraut ist.[57] Sollte auf die Bestellung eines Prüfungsausschusses verzichtet werden, muss der Aufsichtsrat diese Aufgabe selbst wahrnehmen. Die ausdrückliche Nennung der Internen Revision im Gesetzeswortlaut signalisiert den Aufsichtsräten die Erwartung der Gesetzgebung, dass in entsprechenden Unternehmen eine Interne Revision vorhanden ist.[58] Auch die von der Europäischen Union erlassene 8. EU-Richtlinie verweist auf die Aufgabe des Prüfungsausschusses, die Wirksamkeit eines vorhandenen Internen Revisionssystems zu überwachen, bleibt allerdings, sowohl in der ursprünglichen Fassung als auch in der überarbeiteten Version einer detaillierteren Ausführung, wie diese Überwachung ausgeführt werden soll, schuldig.[59] Allerdings haben die europäischen berufsständischen Verbände, Federation of European Risk Management Associations (FERMA) und European Confederation of Institutes of Internal Auditing (ECIIA), im Rahmen einer gemeinsamen Richtline die Frage der Umsetzung der Überwachung des Internen Revisionssystems durch den Prüfungsausschuss näher konkretisiert.[60] Zu nennen ist ebenfalls der im Jahr 2002 in den USA implementierte Sarbanes-Oxley Act of 2002 (SOX), mit Gültigkeit für alle Unternehmen, die den dortigen Kapitalmarkt nutzen. Somit haben die darin beschlossenen Regelungen auch für deutsche Unternehmen, „deren Aktien an US-Börsen oder außerbörslich in den USA gehandelt oder angeboten werden und die der Aufsicht der Securities and Exchange Commission (SEC) unterstehen"[61], Relevanz. Auch der Sarbanes-Oxley Act of 2002 enthält Regelungen bezüglich des Prüfungsausschusses und der Internen Revision. So verweist insbesondere der Abschnitt 404 auf die Notwendigkeit eines internen Kontrollsystems sowie die unternehmensinterne Beurteilung der Funktionsfähigkeit des internen Kontrollsystems.[62] Der Abschnitt 301 empfiehlt die Einrichtung eines Prüfungsausschusses und auch wenn es nicht explizit erwähnt, ist sich die wissenschaftliche Literatur einig, dass der Prüfungsausschuss auch gemäß dieser Vorschrift für die Überwachung der Wirksamkeit der Internen Revision zuständig ist.[63] Abgesehen von den vorab aufgeführten normativen Regelungen

---

[47]Vgl. IDW PS 983.11; IDW PS 983.12; IDW PS 983.27.
[48]Vgl. IDW PS 983.2; IDW PS 983.11.
[49]Vgl. DIIR (Hrsg.) (2018), S. 7-8.
[50]Vgl. DIIR (Hrsg.) (2011, S. 4).
[51]Weiterführend zum DIIR siehe Berwanger und Kullmann (2012, S. 29-31).
[52]Vgl. Bungartz (2021, S. 1).
[53]Welge und Eulerich (2014, S. 2).
[54]DCGK (Hrsg.) (2019), S. 2.
[55]Vgl. hier und im Folgenden §161 Abs. 1 AktG.

[56]DCGK (Hrsg.) (2019), S. 10.
[57]Vgl. hier und im Folgenden §107 Abs. 3 Satz 2 AktG.
[58]Vgl. Amling und Bantleon (2007, S. 17); Eulerich und Velte (2013, S. 146).
[59]Vgl. Bantleon et al. (2017, S.684); Europäisches Parlament und Rat der EU (Hrsg.) (2006), Art. 41 Abs. 2 b); Europäisches Parlament und Rat der EU (Hrsg.) (2014a), Art. 39 Abs. 6 c).
[60]Weiterführend hierzu siehe ECIIA and FERMA (2014, S. 7-13).
[61]Heinzelmann (2020, S. 1).
[62]Vgl. SOX (2002), Sec. 404.
[63]Vgl. Berwanger und Kullmann (2012, S. 163-164); SOX (2002), Sec. 301.

spielen insbesondere die intern von Unternehmen festgelegten Vorschriften eine große Rolle für die Ausgestaltung und Tätigkeit der Internen Revision, da diese Richtlinien und Handlungsanweisungen als Grundlage für die operative Revisionsarbeit gelten.[64]

3. Rechnungslegung

3.1. Normativer und konzeptioneller Rahmen

Der Begriff Rechnungslegung definiert insbesondere die „gesetzliche oder freiwillige Übermittlung unternehmensbezogener Informationen (...) an Stakeholder eines Unternehmens"[65] und wird in Deutschland primär durch das Handelsgesetzbuch (HGB) geregelt. Die zur Verfügung gestellten Informationen basieren im Wesentlichen auf im Rahmen des betrieblichen Rechnungswesens ermittelten Werten, wobei das betriebliche Rechnungswesen verschiedene Funktionen umfasst.[66] Durch die mengen- und wertmäßige Erfassung aller in Unternehmen auftretenden Finanz- und Leistungsströme wird der Dokumentations- und Überwachungsfunktion nachgekommen. Die Dispositionsfunktion kann erfüllt werden, da mittels des betrieblichen Rechnungswesens eine Unternehmensplanung und Wirtschaftlichkeitskontrolle durch Vergleiche von Bestands- und Erfolgsgrößen ermöglicht werden. Die Aufbereitung und Darstellung der Vermögens-, Finanz- und Ertragslage gegenüber internen und externen Adressaten erfüllt die Rechenschaftslegungs- und Informationsfunktion. Der Umfang und Inhalt der von Unternehmen veröffentlichten Unterlagen variiert stark und ist abhängig von der Gesellschaftsform[67], Unternehmensgröße[68], Branche sowie der von Unternehmen verfolgten Publizitätspolitik. Bei der unternehmerischen Publizitätspolitik soll mittels der freiwilligen Bereitstellung von Zusatzinformationen zum Jahresabschluss die Meinungsbildung beziehungsweise die Verhaltensweise externer Abschlussadressaten zugunsten des Unternehmens beeinflusst werden.[69] Der handelsrechtliche Jahresabschluss von Unternehmen hat eine Informations- und Ausschüttungsfunktion sowie bedingt durch die Maßgeblichkeit der Handelsbilanz für die Steuerbilanz[70] auch eine Steuerbemessungsfunktion.[71] Dahingegen gilt die wesentliche Funktion des Abschlusses nach international anerkannten Rechnungslegungsnormen, wie der International Financial Reporting Standards (IFRS) bzw. vormals International Accounting Standrads (IAS), insbesondere der Informationsfunktion.[72] Dementsprechend ist die Normierung

der Rechnungslegung aus handels-, steuerrechtlicher und internationaler Sicht zu differenzieren.

Je nach Unternehmensgröße, -form und Branche umfassen die gesetzlich geforderten Jahresabschlussunterlagen Bilanz, Gewinn- und Verlustrechnung, Anhang, Eigenkapitalspiegel, Kapitalflussrechnung, Segmentberichterstattung sowie Lagebericht und darüber hinaus Veröffentlichungen ohne gesetzliche Normierung wie beispielsweise Nachhaltigkeitsberichte oder Aktionärsbriefe.[73] Die Grundlage für die Erstellung der unterschiedlichen Berichtselemente ist insbesondere die Finanzbuchhaltung auf Basis der handelsrechtlichen[74] Buchführungspflicht.[75] Letztere ergibt sich aus der Verpflichtung, gemäß *§ 238 Abs. 1 HGB* Bücher zu führen[76], verbunden mit der Aufstellung des Inventars[77] sowie der Fertigung des Jahresabschlusses[78], da diese Elemente miteinander verknüpft sind und aufeinander aufbauen.[79] Grundsätzlich haben Jahresabschlüsse gemäß *§ 264 Abs. 2 HGB* unter der „Beachtung der Grundsätze ordnungsmäßiger Buchführung ein den tatsächlichen Verhältnissen entsprechendes Bild der Vermögens-, Finanz- und Ertragslage (...) zu vermitteln".[80] Das dritte Buch des Handelsgesetzbuches umfasst die wesentlichen Normen zur handelsrechtlichen Rechnungslegung und besitzt insbesondere Gültigkeit[81] für Einzelkaufleute, Kapital- und Personenhandelsgesellschaften.[82] Der erste Abschnitt des HGB umfasst die allgemeinen Grundsätze und Vorschriften zur Buchführung und der Erstellung des Jahresabschlusses. Der zweite Abschnitt beinhaltet ergänzende Regelungen für Kapitalgesellschaften, insbesondere bezüglich des Jahresabschlusses, Lagebericht, Spezifizierung für Konzerne, Prüfung und Offenlegung. Der durch handelsrechtliche Rechnungslegungsnormen geforderte Inhalt und Umfang der Jahresabschlussunterlagen hängt somit unmittelbar von der Rechtsform der Unternehmen ab. Die Grundsätze ordnungsmäßiger Buchführung (GoB) sind obligatorische Regelungen zur Führung der Handelsbücher und Erstellung des Jahresabschlusses.[83] Sie ergänzen bestehende Gesetze und finden insbesondere bei Gesetzeslücken oder Auslegungssachverhalten Berücksichtigung. Die GoB untergliedern sich in Klarheit, Vollständigkeit, Stetigkeit, Vorsicht, Abgrenzung, Richtigkeit und Willkürfreiheit.[84] Für Konzern-

---

[64]Vgl. Berwanger und Kullmann (2012, S. 91-92).

[65]Freidank und Velte (2013, S. 4).

[66]Vgl. hier und im Folgenden Freidank und Velte (2013, S. 3-4).

[67]Weiterführend zu Charakteristika der Grundtypen privatrechtlicher Unternehmensformen siehe Freidank und Velte (2013, S. 7-12).

[68]Die Größenkriterien finden sich für Kleinstkapitalgesellschaften in §267a HGB und für kleine bis große Kapitalgesellschaften in §267 HGB.

[69]Vgl. Küting und Weber (2015, S. 34); Coenenberg, Haller und Schultze (2018, S. 1374-1375).

[70]Vgl. §5 Abs. 1 Satz 1 EstG.

[71]Vgl. Coenenberg et al. (2018, S. 18-24 & 1024); Pfaff und Ising (2010, S. 293).

[72]Vgl. Küting und Weber (2015, S. 35).

[73]Vgl. Freidank und Velte (2013, S. 4 & 14).

[74]Steuerrechtlich ergibt sich die Buchführungspflicht originär gemäß §141 AO oder derivativ gemäß §141 AO, steuerrechtliche Aspekte werden im Rahmen dieser Arbeit nicht näher betrachtet.

[75]Vgl. Freidank und Velte (2013, S. 51).

[76]Vgl. §238 Abs. 1 HGB.

[77]Vgl. §240 Abs. 1 & 2 HGB.

[78]Vgl. §242 Abs. 1 bis 3 HGB.

[79]Vgl. Freidank und Velte (2013, S. 51-53).

[80]§264 Abs. 2 HGB.

[81]Das dritte Buch des HGB besitzt insbesondere mit den Regelungen der Abschnitte drei bis fünf, Gültigkeit für Kreditinstitute, Versicherungsunternehmen, Genossenschaften und Unternehmen nach dem Publizitätsgesetz, allerdings werden diese im Rahmen dieser Arbeit nicht näher betrachtet.

[82]Vgl. hier und Folgenden Coenenberg et al. (2018, S. 25-55).

[83]Vgl. hier und im Folgenden Coenenberg et al. (2018, S. 38-46); Freidank und Velte (2013, S. 295-309).

[84]Weiterführend zu den GoB siehe Coenenberg et al. (2018, S. 38-46); Freidank und Velte (2013, S. 295-309).

abschlüsse gilt es zudem, die Deutschen Rechnungslegungs-standards (DRS), Interpretationen und Anwendungshinweise des gemäß *§ 342 HGB*[85] tätigen Deutschen Rechnungslegungs Standards Committee (DRSC) zu berücksichtigen.

Die wesentlichen internationalen Rechnungslegungsnormen sind beispielsweise die United States Generally Accepted Accounting Principles (US-GAAP) und die International Financial Reporting Standards. Auch für Jahresabschlüsse, die gemäß IFRS aufgestellt werden, gilt der Grundsatz, dass „die Vermögens-, Finanz- und Ertragslage (...) eines Unternehmens den tatsächlichen Verhältnissen entsprechend [dargestellt werden soll – Anm. d. Verf.]"[86].[87] Gemäß § 24 des IASB Rahmenkonzeptes, haben Jahresabschlüsse nach IFRS zudem den Grundsätzen der Verständlichkeit, Relevanz, Verlässlichkeit und Vergleichbarkeit zu folgen.[88] Insbesondere kapitalmarktorientierte Konzernmuttergesellschaften mit Sitz innerhalb der EU müssen seit dem Jahr 2005 ihren Konzernabschluss nach den IFRS erstellen. Trotz des eingeräumten Wahlrechts[89] für nicht-kapitalmarktorientierte Unternehmen, den Konzernabschluss nach IFRS zu erstellen, wählt die überwiegende Mehrheit die Aufstellung des Konzernabschlusses nach handelsrechtlichen Vorschriften.[90] Grundsätzlich gilt, dass je nach zugrunde liegender Rechnungslegungsnorm sowie der gewählten Auslegung von Ermessenspielräumen, Ansatz- und Bewertungswahlrechten im Zuge der von Unternehmen verfolgten Rechnungslegungspolitik, stark voneinander abweichende Unternehmensbilder gezeigt werden können.[91]

3.2. Rechnungslegungspolitik und Earnings Management

Eine Teildisziplin der Unternehmenspolitik von Unternehmen ist die Rechnungslegungspolitik, durch die unter anderem mittels zweckorientierter, bilanzpolitischer Maßnahmen übergeordnete Unternehmensziele oder individuelle Ziele der Geschäftsführung verfolgt werden.[92] Die Bilanzpolitik von Unternehmen umfasst im weitesten Sinne „die zielgerichtete Gestaltung der externen Rechnungslegung durch das Management im Rahmen der Möglichkeiten, die unter Einhaltung der Regeln des jeweils zur Anwendung kommenden Normensystem bestehen".[93] Mittels der zur Verfügung gestellten Informationen sollen dem Unternehmensziel dienliche Effekte bewirkt werden, indem Stakeholder ihr Verhalten auf Basis der Erkenntnisse aus dem Jahresabschluss ausrichten.[94] Je nach Unternehmensgröße, -form und Branche umfassen die gesetzlich geforderten

Jahresabschlussunterlagen Bilanz, Gewinn- und Verlustrechnung, Anhang, Eigenkapitalspiegel, Kapitalflussrechnung, Segmentberichterstattung sowie Lagebericht.[95] Darüber hinaus können Veröffentlichungen ohne gesetzliche Normierung wie beispielsweise Nachhaltigkeitsberichte, Aktionärsbriefe oder Pressemitteilungen der Rechnungslegungspolitik von Unternehmen unterliegen.[96] Da es praktisch unmöglich ist, jeden denkbar möglichen Geschäftsvorfall sowie die damit verbundene buchhalterische Erfassung im Gesetz zu definieren, und da die Bewertung der Sachverhalte zudem von der subjektiven Einschätzung des Erfassers, dem Zeithorizont und potenziellen Zukunftszenarien abhängt, bestehen in der Auslegung und Anwendung der Normierung gewisse Wahlmöglichkeiten.[97] In der wissenschaftlichen Literatur werden mögliche bilanzpolitische Handlungsspielräume sowie die damit verbundene, opportunistische Auslegung und Anwendung unter dem Begriff Earnings Management diskutiert.[98] Die eingangs dargestellte Rechnungslegungspolitik umfasst lediglich die Anwendung gesetzlich zulässiger Bilanzierungsspielräume, wohingegen das Earnings Management auch die unzulässige Bilanzmanipulation durch die missbräuchliche Anwendung von Bilanzierungsvorschriften umfasst. Rechnungslegungspolitik entspricht demnach nur einem Teilgebiet des Earnings Managements. Gemäß Healy und Wahlen (1999) spricht man von Earnings Management wenn „(...) managers use judgment in financial reporting and in structuring transactions to alter financial reports to either mislead some stakeholders about the underlying economic performance of the company or to influence contractual outcomes that depend on reported accounting numbers".[99] Die nachfolgende Abbildung zeigt auf, welche Möglichkeiten bestehen, durch den Einsatz bilanzpolitischer Mittel den ausgewiesenen Gewinn eines Unternehmens zu verschlechtern oder zu verbessern.[100]

Zu unterscheiden ist im Wesentlichen in sachverhaltsgestaltende Maßnahmen (Real Earnings Management), die vor dem Bilanzstichtag vorgenommen werden, und sachverhaltsabbildende Maßnahmen (Accounting Earnings Management), die die Ergebnislage des abgelaufenen Geschäftsjahres eines Unternehmens durch rückwirkende Eingriffe beeinflussen.[101] Bei sachverhaltsgestaltenden Maßnahmen stehen die Entscheidungen über die Durchführung sowie den Zeitpunkt von Geschäftsvorgängen vorrangig im Fokus

---

[85]Vgl. §342 HGB.

[86]IAS 1.15.

[87]Vgl. IAS 1.15; §46 IASB Rahmenkonzept.

[88]Vgl. §24 IASB Rahmenkonzept.

[89]Vgl. §315e Abs. 3 HGB.

[90]Vgl. hier und im Folgenden Freidank und Velte (2013, S. 263-266); Küting und Weber (2015, S. 16-23).

[91]Vgl. Küting und Weber (2015, S. 18); Pfaff und Ising (2010, S. 292-293).

[92]Vgl. Freidank und Velte (2013, S. 13-14); Ismael und Kamel (2021, S. 951); Küting und Weber (2015, S. 33).

[93]Coenenberg et al. (2018, S. 1023).

[94]Vgl. Coenenberg et al. (2018, S. 1024); Küting und Weber (2015, S. 33).

[95]Es handelt sich hierbei um keine abschließende Auflistung, da unter anderem auch Zwischenabschlüsse, Sonder- und Konzernbilanzen sowie branchenspezifische Veröffentlichungen gefordert sein können.

[96]Vgl. Freidank und Velte (2013, S. 4 & 14); Küting und Weber (2015, S. 33).

[97]Vgl. Dechow und Schrand (2004, S. 37); Pfaff und Ising (2010, S. 292); Teets (2002, S. 355).

[98]Vgl. hier und im Folgenden Davidson et al. (2005, S. 243); Pfaff und Ising (2010, S. 294-303); Teets (2002, S. 356).

[99]Healy und Wahlen (1999, S. 368).

[100]Weiterführend zu den bilanzpolitischen Instrumenten siehe die tabellarischen Übersichten von Coenenberg et al. (2018, S. 1030-1033); Küting und Weber (2015, S. 42-45).

[101]Vgl. hier und im Folgenden Coenenberg et al. (2018, S. 1026); Küting und Weber (2015, S. 40); Pfaff und Ising (2010, S. 295).
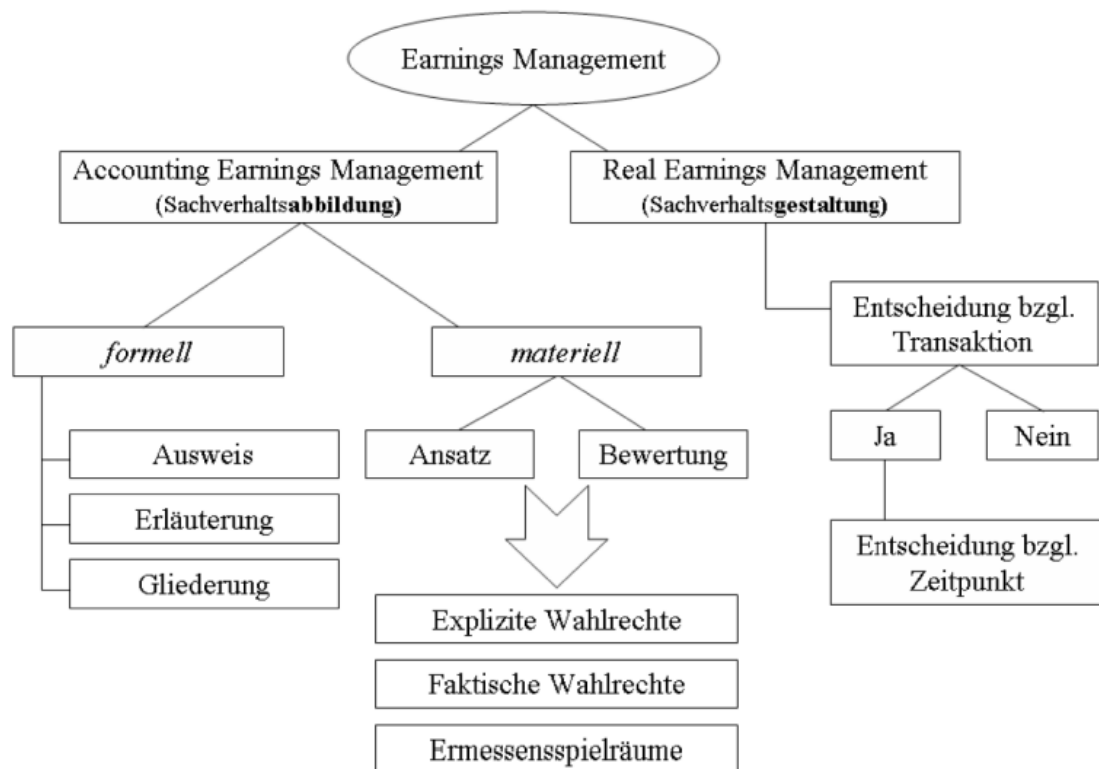
Abbildung 3: Instrumente des Earnings Management

Quelle: Pfaff und Ising (2010, S. 295).

und nicht deren Abbildung. Dahingegen werden bei sachverhaltsabbildenden Maßnahmen die Berichtsinstrumente der Finanzbuchhaltung unterteilt in formelle und materielle Maßnahmen, entsprechend der gewählten Konzern-, Handels- und Steuerbilanzpolitik. Die materielle Sachverhaltsabbildung umfasst Ansatz- und Bewertungswahlrechte sowie Ermessenspielräume wie beispielsweise Schätzungen oder Auslegungen. Die formelle Sachverhaltsabbildung nutzt gesetzliche Spielräume bezüglich Ausweis, Erläuterung und Gliederung im Einklang mit den unternehmenspolitischen Zielen aus. Der zielgerichtete Einsatz der soeben erläuterten Maßnahmen verstößt der bereits dargestellten Generalnorm, dass Jahresabschlüsse, sowohl nach IFRS als auch handelsrechtlich, ein den tatsächlichen Verhältnissen entsprechendes Bild der Vermögens-, Finanz- und Ertragslage vermitteln sollen. Demnach besteht das Risiko, dass die Qualität von Jahresabschlüssen und somit der entscheidungsrelevante Informationsgehalt für Stakeholder durch Earnings Management verringert wird. Wie bereits erläutert, können durch die gewährten gesetzlichen Spielräume, je nach Motivation des Jahresabschlusserstellers, trotz identisch geltender Rechnungslegungsnormen unterschiedliche und nicht direkt miteinander vergleichbare Unternehmensbilder aufgezeigt

werden.[102] Die Qualität der Rechnungslegung bzw. eines Jahresabschlusses hängt somit nicht primär von den geltenden Rechnungslegungsnormen ab, sondern von den individuellen, unternehmensspezifischen und finanziellen Anreizstrukturen des Jahresabschlusserstellers und der dementsprechend verfolgten Rechnungslegungspolitik.[103]

3.3. Qualität der Rechnungslegung

In der wissenschaftlichen Literatur wird bei Betrachtung der Rechnungslegungsqualität (Accounting Quality/Financial Reporting Quality) regelmäßig von der Ergebnisqualität (Earnings Quality) gesprochen. Eine einheitliche Definition beziehungsweise Abgrenzung der Begrifflichkeiten liegt nicht vor und die Auslegung der Termini hängt von der jeweils betrachteten Zielsetzung der Rechnungslegung ab.[104] Da die Ergebnisqualität als Summenindikator die Qualität der Rechnungslegung widerspiegelt[105], wird im Folgenden vorrangig auf diese Begrifflichkeit Bezug genommen.[106]

---

[102]Vgl. Eulerich und Fligge (2020, S. 3); Küting und Weber (2015, S. 18); Pfaff und Ising (2010, S. 292-293).

[103]Vgl. Pfaff und Ising (2010, S. 292-293).

[104]Vgl. Dechow, Ge und Schrand (2010, S. 344); Gegenfurtner (2010, S. 26-27); Wagenhofer und Dücker (2007, S. 264); Weiterführend siehe Gegenfurtner (2010, S. 26-27).

[105]Vgl. Gegenfurtner (2010, S. 27).

[106]Weiterführend hierzu siehe Gegenfurtner (2010, S. 28).

Die Herangehensweisen zur Definition von Earnings Quality variieren in der wissenschaftlichen Literatur und hängen stark von den betrachteten Parametern ab, demnach werden nachfolgend zur besseren Übersicht lediglich zwei wesentliche Definition vorgestellt.[107] Einen in der wissenschaftlichen Literatur häufig zitierten Definitionsansatz liefern Dechow und Schrand (2004), die Earnings Quality aus analytischer Sicht wie folgt definieren: „A high-quality earnings number is one that accurately reflects the company's current operating performance, is a good indicator of future operating performance, and is a useful summary measure for assessing firm value. We define earnings to be of high quality when the earnings number accurately annuitizes the intrinsic value of the firm."[108] Wie im vorangegangenen Kapitel 3.1. bereits erläutert, soll gemäß internationaler und nationaler Standardsetter die Rechnungslegung eines Unternehmens möglichst die tatsächlichen Verhältnisse der Vermögens-, Finanz- und Ertragslage darstellen, um eine hohe Entscheidungsnützlichkeit zu generieren und dadurch die optimale Allokation von Kapital auf Kapitalmärkten zu bestärken.[109] Dechow et al. (2010) nehmen unter diesem Gesichtspunkt eine Definition wie folgt vor: „Higher quality earnings more faithfully represent the features of the firm's fundamental earnings process that are relevant to a specific decision made by a specific decision-maker."[110]

Da die Rechnungslegungs- bzw. Ergebnisqualität unter anderem von der von Unternehmen verfolgten Rechnungslegungspolitik bzw. dem betriebenen Earnings Management abhängt, werden im Folgenden ausgewählte Möglichkeiten zur Aufdeckung von Earnings Management dargestellt. Ähnlich wie die voneinander abweichenden Definitionen der Rechnungslegungs- bzw. Ergebnisqualität variieren auch die Ansätze und Methoden zur Ermittlung von Earnings Quality.[111] Aufgrund der Komplexität der Messung von Earnings Quality soll im Nachfolgenden lediglich eine kurze Übersicht über ausgewählte Maßgrößen und Messmethoden gegeben werden. Grundsätzlich wird es als schwierig erachtet, die Messung der Ergebnisqualität eines Unternehmens auf Basis von nur einer einzigen Methode vorzunehmen, dementsprechend werden in der Regel mehrere Ansätze miteinander kombiniert.[112] Um die Ergebnisqualität im Rahmen der empirischen Forschung zu modellieren, werden regelmäßig Modelle auf Basis von diskretionären Periodenabgrenzungen[113] (Abnormal bzw. Discretionary Accruals) wie das Jones Mo-

del[114] bzw. Modified Jones Model[115] verwendet, aber auch Messmethoden wie The Penman Index[116], Leuz Index[117], Dichev and Dechow Model[118], Red Flags[119] und zahlreiche weitere Methoden kommen zum Einsatz.[120] Zu den prominentesten Modellen zur Modellierung der diskretionären Periodenabgrenzungen zählen das Jones Model sowie das Modified Jones Model.[121] Bei der Aufdeckung von Earnings Management mittels diskretionärer Periodenabgrenzungen ist „(...) die Abgrenzung zwischen obligatorischen (sog. normalen) Periodenabgrenzungen, die in den Vorschriften eines Rechnungslegungsstandards begründet sind, und strategischen (sog. diskretionären) Periodenabgrenzungen, die allein dem Earnings Management dienen"[122], relevant.[123] Da jedoch nur die gesamten Periodenabgrenzungen eines Unternehmens beobachtbar sind, muss mittels Hypothesenbildung über den Betrachtungszeitraum eine Schätzung vorgenommen werden, welcher Teil der Periodenabgrenzungen normal ist.[124] So wird angenommen, dass sich der Anteil der normalen Periodenabgrenzungen aus einem Durchschnitt der gesamten Periodenabgrenzungen mehrerer vergangenen Betrachtungsperioden ergibt. Da die angewendeten Modelle den normalen Anteil der Periodenabgrenzungen schätzen, spiegelt sich der abnormale Teil in der Differenz wider. Abschließend bleibt festzuhalten, dass eine „(...) Bewertung über die Qualität des Gewinns (...) nur als Konsensusschlussfolgerung aus der Betrachtung verschiedener Dimensionen und verschiedener Maßgröße erfolgen [kann – Anm. d. Verf.] und der subjektiven Interpretation des Betrachters [unterliegt – Anm. d. Verf.]. Dies impliziert, dass die aus den Kennzahlen abgeleiteten Schlussfolgerungen grundsätzlich kritisch und kontextbezogen zu bewerten sind".[125]

---

[107]Weiterführend zu ausführlichen Analysen der Definitionsansätze in der wissenschaftlichen Literatur siehe Gegenfurtner (2010, S. 28-31); Menicucci (2020, S. 6-12); Wagenhofer und Dücker (2007, S. 264-267).

[108]Dechow und Schrand (2004, S. 5).

[109]Vgl. §264 Abs. 2 HGB; §15 IASB Rahmenkonzept; Wagenhofer und Dücker (2007, S. 266-267); Weiterführend siehe Wagenhofer und Dücker (2007, S. 266-267).

[110]Dechow et al. (2010, S. 344).

[111]Vgl. Menicucci (2020, S. 6).

[112]Vgl. Menicucci (2020, S 43).

[113]Der Grundsatz der Periodenabgrenzung ist in §252 Abs. 1 Nr. 5 HGB geregelt und besagt, dass Aufwendungen und Erträge unabhängig von den jeweiligen Zahlungszeitpunkten im Jahresabschluss zu berücksichtigen sind. Beispiele für nicht zahlungswirksame Aufwendungen und Erträge sind der

Abgang von Anlagevermögen, Abschreibungen, Rückstellungen, Forderungen aus Lieferungen und Leistungen, Verbindlichkeiten aus Lieferungen und Leistungen sowie Änderungen im Vorratsvermögen.Vgl. Dechow und Schrand (2004, S. 40-43); Pfaff und Ising (2010, S. 303).

[114]Weiterführend siehe Jones (1991, S. 210-223); Menicucci (2020, S 43-45).

[115]Weiterführend siehe Dechow, Sloan und Sweeny (1995, S. 199); Weiterführend zu den Modifikationen des Jones Model siehe Ronen und Yaari (2008, S. 433-460).

[116]Weiterführend siehe Menicucci (2020, S. 43-45).

[117]Weiterführend siehe Gegenfurtner (2010, S. 39-48); Leuz, Dhananjay und Wysocki (2003, S. 511-524); Menicucci (2020, S. 43-45).

[118]Weiterführend siehe Dechow und Dichev (2002, S. 36-41); Menicucci (2020, S. 45-46).

[119]Vgl. Dechow und Schrand (2004, S. 34-35); Pfaff und Ising (2010, S. 305).

[120]Es handelt sich um keine vollständige Aufzählung der möglichen Maßgrößen und Modelle.

[121]Vgl. Dechow und Schrand (2004, S. 20); Weiterführend zu Messmethoden auf Basis diskretionärer Periodenabgrenzungen siehe Dechow et al. (1995, S. 197-200); Keßler (2015, S. 31-70).

[122]Pfaff und Ising (2010, S. 303).

[123]Vgl. Davidson et al. (2005, S. 249-251); Pfaff und Ising (2010, S. 303).

[124]Vgl. hier und im Folgenden Bajra und Cadez (2018, S. 156); Gegenfurtner (2010, S. 48-50); Ismael und Kamel (2021, S. 959); Pfaff und Ising (2010, S. 304).

[125]Gegenfurtner (2010, S. 52).

4. Auswirkungen der Internen Revision auf die Qualität der Rechnungslegung

In den nachfolgenden Kapiteln werden die wesentlichen Forschungsergebnisse, die den Zusammenhang zwischen der Internen Revision und den die Qualität der Rechnungslegung beeinflussenden Faktoren dargelegt. Die nächstfolgend, tabellarisch dargestellten Studienergebnisse fokussieren sich unter Zuhilfenahme variierender Forschungsansätze insbesondere auf den Zusammenhang zwischen Interner Revision auf die Qualität der Rechnungslegung, wohingegen die in Kapitel 4.5. zusammengefassten Studienergebnisse eher den indirekten Einfluss untersuchen.[126]

4.1. Vorhandensein einer Internen Revision

Das American Institute of Certified Public Accountants (AICPA) argumentiert, dass die Angst erwischt zu werden ein starkes Hindernis darstellt und demnach abschreckend wirkt.[127] Daran anknüpfend umfasst der Untersuchungsgegenstand der nachfolgend aufgeführten Studien unterschiedliche Corporate Governance Mechanismen und deren Auswirkungen auf die Qualität der Rechnungslegung. Die Besonderheit der nachfolgenden Studien ist die Untersuchung auf Basis der Existenz von Internen Revisionen in Unternehmen, ohne eine Betrachtung von Qualitätsaspekten. Demnach stellt sich die Frage, ob sich die bloße Existenz einer Internen Revision bereits positiv auf die Qualität der Rechnungslegung auswirkt.

Davidson et al. (2005) untersuchen den Zusammenhang zwischen Corporate Governance Strukturen von 434 kapitalmarktorientierten Unternehmen in Australien und Earnings Management.[128] Zu den untersuchten Corporate Governance Mechanismen der Unternehmen gehören der Vorstand, der Prüfungsausschuss, die Interne Revision und der gewählte Jahresabschlussprüfer. Zum Untersuchungszeitpunkt im Jahr 2000 bestand für australische Unternehmen keine gesetzliche Verpflichtung zur Einrichtung einer Internen Revision oder eines Prüfungsausschusses. Die Autoren argumentieren auf Basis vorangegangener wissenschaftlicher Literatur, dass eine Interne Revision per se mit einer Reduktion von Earnings Management zusammenhängt. Eine Formulierung von Qualitätskriterien erfolgt nicht, demnach wird die Hypothese untersucht, dass das bloße Vorhandensein einer Internen Revision Earnings Management reduziert.[129] Earnings Management wurde mittels des absoluten Wertes der Discretionary Accruals auf Basis des Modified Jones Model sowie des Earnings Benchmark Test[130], zur Messung der Wahrscheinlichkeit, einen kleinen Verlust oder Ertrag zu erfassen, gemessen.[131] Die Forschungsergebnisse von Davidson et al. (2005)

zeigen einen signifikanten Zusammenhang zwischen dem Vorhandensein eines Prüfungsausschusses sowie überzähliger nicht-exekutiver Direktoren als Mitglieder des Vorstandes und des Aufsichtsrates und einer niedrigeren Ausprägung von Discretionary Accruals.[132] Die Autoren konnten jedoch, analog zur Wahl eines Big-5-Jahresabschlussprüfers[133], keinen signifikanten Zusammenhang zwischen dem Vorhandensein oder Nichtvorhandensein einer Internen Revision und einer niedrigeren Ausprägung von Discretionary Accruals feststellen.[134] Allerdings weisen Davidson et al. (2005) in ihren Ausführungen auf die Notwendigkeit weiterer Forschungstätigkeiten hin, denn aufgrund der Größe der Unternehmen in der Stichprobe verfügten nahezu alle über eine Interne Revision und dementsprechend wäre der Einfluss der Qualität der Internen Revision eher untersuchenswert gewesen.[135] Beasley et al. (2000), Sierra García et al. (2012) und Gebrayel et al. (2018) liefern Forschungsergebnisse, die sich ebenfalls auf die Untersuchung der Existenz von Internen Revisionen fokussieren, aber entgegengesetzte Erkenntnisse hervorbringen. Beasley et al. (2000) analysieren 200 Fälle von Bilanzbetrug auf Basis von Unternehmen der Geschäftszweige Technologie, Gesundheits- und Finanzwesen in den Jahren 1987 bis 1997.[136] Die Ergebnisse zeigen, dass im Falle von Bilanzbetrug in allen drei Branchen keine Unterstützung durch eine Interne Revision gegeben war und im Bereich des Gesundheitswesens und der Technologieunternehmen eine geringere Anzahl an Sitzungen des Prüfungsausschusses stattfanden.[137] Sierra García et al. (2012) untersuchen auf Basis von 108 spanischen Unternehmen, die im Zeitraum 2003 bis 2006 an der Madrid Stock Exchange gehandelt wurden, den Zusammenhang zwischen ausgewählten Corporate Governance Mechanismen und Earnings Management beziehungsweise der Qualität des Jahresabschlusses.[138] Auf Grundlage einzelner Merkmale des Prüfungsausschusses sowie der Internen Revision soll der Zusammenhang zu Earnings Management, gemessen durch Ermittlung der Abnormal Accruals mittels des Modified Jones Models, untersucht werden. Sierra García et al. (2012) formulieren die Hypothesen, dass der Zusammenhang zwischen einem wirksamen Prüfungsausschuss[139], dem Vorhandensein einer Internen Revision sowie der Ausprägung der Beziehung zwischen den beiden Corporate Governance Mechanismen zu einer Reduktion der Möglichkeiten der Bilanzmanipulation durch die Geschäftsleitung führt. Die Forschungsergebnisse der Autoren zeigen, dass eine größere Anzahl an Mitgliedern

---

[126]Alle nachfolgend ausgewerteten Studien exkludieren Finanzdienstleistungsunternehmen, da es sich um besondere Geschäftsmodelle handelt.

[127]Vgl. AICPA, IIA und ACFE (2008, S. 30).

[128]Vgl. hier und im Folgenden Davidson et al. (2005, S. 241-242).

[129]Vgl. Davidson et al. (2005, S. 247-248).

[130]Weiterführend siehe Holland und Ramsay (2003, S. 42-62).

[131]Vgl. Davidson et al. (2005, S. 249-250 & 260).

[132]Vgl. Davidson et al. (2005, S. 243 & 262-263).

[133]Zu dieser Zeit bestehend aus: Arthur Andersen, Deloitte Touche Tohmatsu, Ernst & Young, PricewaterhouseCoopers und KPMG.

[134]Vgl. Davidson et al. (2005, S. 243 & 262-263).

[135]Vgl. Davidson et al. (2005, S. 263-264); Prawitt et al. (2009, S. 1258).

[136]Vgl. Beasley et al. (2000, S. 445-446).

[137]Vgl. Beasley et al. (2000, S. 453).

[138]Vgl. hier und im Folgenden Sierra García et al. (2012, S. 306).

[139]Gemäß wissenschaftlicher Literatur umfassen die Merkmale eines wirksamen Prüfungsausschusses unter anderem die Unabhängigkeit des Prüfungsausschusses, mindestens vier jährliche Sitzungen und mindestens drei Mitglieder.Vgl. Sierra García et al. (2012, S. 312); The Blue Ribbon Committee (1999, S. 1089-1094).

des Prüfungsausschusses, eine höhere Anzahl an jährlichen Sitzungen des Prüfungsausschusses und die Existenz einer Internen Revision jeweils und unabhängig voneinander mit einer höheren Wahrscheinlichkeit zur Reduktion von Earnings Management führen.[140] Sierra García et al. (2012) stellen zudem fest, dass Unternehmen, in denen eine Interne Revision gepaart mit einem wirksamen Prüfungsausschuss existiert, mit einer geringeren Wahrscheinlichkeit Earnings Management betreiben. Einen sehr ähnlichen Forschungsansatz haben auch Gebrayel et al. (2018) gewählt.[141] Die Autoren untersuchen ebenfalls den Zusammenhang zwischen der Qualität des Jahresabschlusses und unterschiedlichen Ausprägungen der Corporate Governance Mechanismen, Interne Revision und Prüfungsausschuss. Auf Basis der Beobachtungen von 71 Unternehmen, die im Zeitraum 2013 bis 2014 am Muscat Securities Market gehandelt wurden, stellen die Autoren fest, dass die Höhe der Anzahl der jährlichen Prüfungsausschusssitzungen sowie die Existenz einer Internen Revision jeweils unabhängig voneinander positiv mit Rechnungslegungsqualität korrelieren.[142] Gebrayel et al. (2018) konkludieren analog zu Sierra García et al. (2012), dass die Interne Revision durch die Überwachung organisatorischer Risiken, die Bewertung der internen Kontrollen von Unternehmen und der Fähigkeit, Bilanzmanipulationen zu erkennen, ein wichtiger Mechanismus zur Verbesserung der Qualität der Rechnungslegung ist.[143] Aufgrund der Bewertung rechnungslegungsbezogener Sachverhalte und der Aufbereitung relevanter Informationen für den Prüfungsausschuss kann die Interne Revision wesentlich zur Reduktion von Informationsasymmetrien beitragen. Darüber hinaus existieren diverse Studien, die zeigen, dass sich die Konsequenzen der Tätigkeit des Prüfungsausschusses von Unternehmen insbesondere dann positiv auf die Rechnungslegungsqualität auswirken, wenn sie mit einer Internen Revision zusammenarbeiten.[144] Wie bedeutsam dabei ein qualitativ hochwertiger und wirksamer Prüfungsausschuss ist, zeigen hierbei die Forschungsergebnisse von Raghunandan et al. (1998) und Abbott et al. (2004). Ein sich regelmäßig treffender Prüfungsausschuss kann die Wahrscheinlichkeit für Fehlerfeststellungen im Jahresabschluss, insbesondere bei regelmäßigen Zusammenkünften mit der Internen Revision, reduzieren.

4.2. Qualität der Internen Revision

Wie im vorangegangenen Kapitel bereits dargestellt, ist abgesehen von der bloßen Existenz der Internen Revision auch deren Qualität von Interesse. Nachfolgend werden demnach die Studien betrachtet, die sich mit der Untersuchung

des Einflusses der Qualität der Internen Revision auf die Qualität der Rechnungslegung befassen. Beispielsweise untersuchen Prawitt et al. (2009) den Zusammenhang zwischen der Qualität der Internen Revision und Earnings Management.[145] Das Maß zur Darstellung der Qualität der Internen Revision setzt sich aus den sechs Komponenten des vom AICPA verfassten Statement on Auditing Standard (SAS) No. 65 zusammen und wurde mittels der Umfrageergebnisse von Chief Audit Executives (CAE)[146] von 218 Unternehmen ermittelt. Diese abgefragten Komponenten umfassen die durchschnittliche Berufserfahrung der Internen Revisoren, den prozentualen Anteil der Certified Internal Auditors (CIA) oder Certified Public Accountans (CPA), die jährliche Anzahl an Fortbildungsstunden, die relative Fokussierung der Internen Revision auf finanzielle Prüfungsschwerpunkte, die Berichtsbeziehung des Leiters der Internen Revision[147] sowie die Größe der Internen Revision in Relation zur Unternehmensgröße.[148] Earnings Management wird mittels der Proxy-Variablen des, durch das Performance-Adjusted Modified Jones Model ermittelten, absoluten Wertes der Abnormal Accruals[149] sowie der Tendenz von Unternehmen, die Earnings-Vorhersagen von Bilanzanalysten zu treffen oder zu verfehlen, gemessen.[150] Prawitt et al. (2009) gelangen zu dem Ergebnis, dass eine qualitativ hochwertige Interne Revision im Zusammenhang mit einem niedrigeren Grad an Earnings Management steht.[151] Je höher die Qualität einer Internen Revision, desto geringer fallen die gemessenen Abnormal Accruals aus. Demnach reduziert eine qualitativ hochwertige Interne Revision die übermäßige und zielgerichtete Bildung diskretionärer Periodenabgrenzungen und stiller Reserven sowie sonstiges opportunistisches Verhalten seitens der Geschäftsführung. Die Forschungsergebnisse zeigen, dass Unternehmen, die die von Analysten vorhergesagten Gewinnschwellen treffen bzw. übertreffen, mit einer höheren Wahrscheinlichkeit Earnings Management betreiben.[152] Die Qualität der Internen Revision ist insbesondere bei den Unternehmen hoch, die die vorhergesagten Gewinnschwellen unterschreiten, was sich mit dem vorherigen Ergebnis deckt, dass eine hochwertige Interne Revision Earnings Management einschränkt.

Überraschende Ergebnisse liefern hingegen Johl et al. (2013), die den Zusammenhang zwischen der Qualität einer Internen Revision und der Rechnungslegungsqualität untersuchen, genauer auf Basis von 128 durch CIAs oder Controller ausgefüllte Fragebögen von 64 an der Bursa Malaysia gelisteten Unternehmen in den Jahren 2009-2010.[153] Das Maß zur Darstellung der Qualität der Internen Revision setzt

---

[140]Vgl. hier und im Folgenden Sierra García et al. (2012, S. 321-322).
[141]Vgl. hier und im Folgenden Gebrayel et al. (2018, S. 197-207).
[142]Vgl. Gebrayel et al. (2018, S. 207-211).
[143]Vgl. hier und im Folgenden Gebrayel et al. (2018, S. 210).
[144]Vgl. hier und im Folgenden Abbott, Parker und Peters (2004, S. 69-87); Gebrayel et al. (2018, S. 202); Raghunandan, Rama und Scarbrough (1998, S. 181-194).

[145]Vgl. hier und im Folgenden Prawitt et al. (2009, S. 1255-1257 & 1259-1262).
[146]Weiterführend zu den Aufgaben eines CAE siehe Deloitte (2017, S. 1-11).
[147]Berichtet die Leitung der Internen Revisison an die Geschäftsführung oder an den Aufsichtsrat/Prüfungsausschuss.
[148]Weiterführend siehe AICPA (Hrsg.) (1997), S. 1805-1812.
[149]Vgl. Prawitt et al. (2009, S. 1255-1257 & 1262-1267).
[150]Vgl. Prawitt et al. (2009, S. 1255-1257 & 1267-1268).
[151]Vgl. hier und im Folgenden Prawitt et al. (2009, S. 1269-1270).
[152]Vgl. hier und im Folgenden Prawitt et al. (2009, S. 1271-1272).
[153]Vgl. Johl et al. (2013, S. 787).

sich aus den folgenden fünf weitgefassten Charakteristika zusammen: Berufserfahrung, Unabhängigkeit, Qualitätssicherungsmaßnahmen, relative Fokussierung auf finanzielle Prüfungsschwerpunkte und finanzielle Ausstattung.[154] Die Rechnungslegungsqualität modellieren die Autoren mittels Abnormal Accruals auf Basis des Cross-Sectional Modified Jones Model. Johl et al. (2013) stellen, entgegen ihrer ursprünglichen Hypothesen, eine positive Korrelation zwischen der Qualität der Internen Revision und Abnormal Accruals in Abhängigkeit der politischen Verbindungen und der potenziellen Ausgliederung der Tätigkeiten der Internen Revision fest.[155] Dementsprechend betreiben politisch verbundene Unternehmen mit ausgegliederten Internen Revisionen bei steigender Qualität der Internen Revision mit einer höheren Wahrscheinlichkeit Earnings Management und verfügen demnach über eine geringere Rechnungslegungsqualität. Die finanzielle Ausstattung einer Internen Revision ist negativ korreliert mit Abnormal Accruals, wohingegen Unabhängigkeit und Berufserfahrung schwach positiv mit Abnormal Accruals korreliert sind. Johl et al. (2013) konkludieren, dass im Allgemeinen das Resultat und der Einfluss der Tätigkeit von Internen Revisionen davon abhängen, ob es sich um eine ausgegliederte oder interne Abteilung handelt und wie diese politisch verbunden sind. Die Autoren stellen zudem fest, dass die Qualität des Vorstandes und die Qualität der Internen Revision substitutiv zusammenwirken, um die Qualität der Rechnungslegung auf einem bestimmten Niveau zu halten. Die Autoren haben sich bewusst dazu entschieden, Interne Revisionen im Kontext des Schwellenlandes Malaysia zu untersuchen, da dieses sich signifikant von der ökonomischen Entwicklung anderer Länder wie beispielsweise den USA unterscheidet. Aus der gleichen Motivation heraus, aber mit anderen Forschungsergebnissen konkludierend, untersuchen im Jahr 2009 Hutchinson und Zain (2009) im Rahmen einer Preliminary Study auf Basis von 60 ausgefüllten Fragebögen von an der Bursa Malaysia gelisteten Unternehmen, ob Wachstumschancen und die Unabhängigkeit des Prüfungsausschusses den Zusammenhang zwischen der Qualität der Internen Revision und der Performanz des Unternehmens beeinflussen.[156] Die Autoren verweisen auf die Vorläufigkeit ihrer Ergebnisse, können aber feststellen, dass ein Zusammenhang zwischen wirksamer Corporate Governance, insbesondere durch die Einrichtung einer kompetenten Internen Revision, und der Aufdeckung und Vermeidung von Earnings Management, existiert.[157]

Ege (2015) untersucht auf Basis proprietärer Umfragedaten[158], welchen Einfluss die Qualität, die Kompetenz und die Objektivität der Internen Revision auf die Wahrscheinlichkeit von Fehlverhalten durch die Geschäftsführung haben.[159] Unter Fehlverhalten der Geschäftsführung subsummiert der Au-

tor Bilanzmanipulation, Bestechung und irreführende Offenlegungspraktiken. Das Maß zur Darstellung der Qualität der Internen Revision setzt sich aus den folgenden Komponenten zusammen: Erfahrung, Zertifikation, Schulungsstunden, Management Training Ground[160], Beziehung zum Prüfungsausschuss sowie Abteilungsgröße. Um ein entsprechendes Maß für die Kompetenz und Objektivität von Internen Revisionen zu kreieren, führt Ege (2015) eine Faktorenanalyse durch. Ege (2015) untersucht 1.398 Geschäftsjahre von 617 Unternehmen, gegen die glaubwürdige Vorwürfe vorsätzlicher, rechtswidriger Handlungen durch die Geschäftsführung im Zeitraum von 2000 bis 2009 vorliegen. Hierbei kontrolliert der Autor für andere Determinanten wie Vorstand, Prüfungsausschuss, Qualität des IKS und Qualität des Jahresabschlussprüfers.[161] Unter Anwendung des Differenz-von-Differenzen-Ansatzes ermittelt der Autor, dass Unternehmen, gegen die glaubwürdige Vorwürfe vorsätzlicher, rechtswidriger Handlungen durch die Geschäftsführung vorliegen, in den entsprechenden Jahren des Fehlverhaltens über eine qualitativ niedrige und vergleichsweise wenig kompetente Interne Revision verfügen.[162] Ege (2015) stellt fest, je höher die Qualität der Internen Revision ist, desto geringer ist die Wahrscheinlichkeit für das oben definierte Fehlverhalten durch die Geschäftsführung. Zudem findet Ege (2015) einen negativen, signifikanten Zusammenhang zwischen der Kompetenz einer Internen Revision und der Wahrscheinlichkeit für Fehlverhalten durch die Geschäftsführung. Für die Objektivität der Internen Revision kann kein signifikanter Zusammenhang festgestellt werden. Der Autor hält fest, dass Unternehmen in den Folgejahren eines rechtswidrigen Vorfalls die Qualität ihrer Internen Revisionen deutlich erhöhen. In der Regel erfolgt die Steigerung der Qualität durch die Erhöhung der Schulungsstunden und die Neueinstellung zertifizierter Interner Revisoren.

Abbott et al. (2016) möchten die Determinanten einer effektiven Internen Revision ermitteln und untersuchen dazu den Interaktionseffekt der Kompetenz und Unabhängigkeit von Internen Revisionen als Qualitätsmerkmal und dessen Effekt auf die Rechnungslegungsqualität.[163] Die Autoren stützen ihre Hypothesen auf die Definition externer Prüfungsqualität, wie formuliert durch DeAngelo (1981): „The quality of audit services is defined to be the market-assessed joint probability that a given auditor will both (a) discover a breach in the client's accounting system and (b) report the breach".[164] Die Autorin verweist demnach auf die wahrgenommene Kompetenz der Jahresabschlussprüfer und deren Bereitschaft, Feststellungen zu melden, als Faktor der Unabhängigkeit. Abbott et al. (2016) nehmen dementsprechend an, dass sich auch die Qualität der Internen Revision auf

---

[154]Vgl. hier und im Folgenden Johl et al. (2013, S. 788).
[155]Vgl. hier und im Folgenden Johl et al. (2013, S. 809-810).
[156]Vgl. Hutchinson und Zain (2009, S. 50).
[157]Vgl. Hutchinson und Zain (2009, S. 57-59).
[158]Eigentümer der Daten ist das IIA.
[159]Vgl. hier und im Folgenden Ege (2015, S. 496).

[160]Eine Personalstrategie, die angewendet wird, um Fach- und Führungskräfte auf Managementpositionen vorzubereiten, indem diese mehrjährige Revisionserfahrung sammeln.Weiterführend siehe Eulerich und van Uum (2014, S. 132-138).
[161]Vgl. Ege (2015, S. 497).
[162]Vgl. hier und im Folgenden Ege (2015, S. 497-498).
[163]Vgl. Abbott et al. (2016, S. 4).
[164]DeAngelo (1981, S. 186).

das Zusammenspiel von Kompetenz und Unabhängigkeit jener zurückführen lässt.[165] Auf Basis von 189 durch CIAs oder CAEs ausgefüllte Fragebögen von ausschließlich Fortune 1000 gelisteten Unternehmen[166] im Jahr 2009, stellen Abbott et al. (2016) ein Zwei-Faktoren-Modell für die Qualität der Rechnungslegung als Funktion aus der Fähigkeit der Internen Revision, Bilanzmanipulationen zu verhindern beziehungsweise zu erkennen und der Tendenz, diese an den Prüfungsausschuss oder externen Jahresabschlussprüfer zu melden, auf.[167] Rechnungslegungsqualität wird mittels Abnormal Accruals unterteilt in gewinnsteigernd und -minimierend, sowie in der Tendenz von Unternehmen, die Earnings-Vorhersagen von Bilanzanalysten zu treffen oder zu verfehlen, gemessen. Abbott et al. (2016) stellen fest, dass das gemeinsame Vorhandensein von Kompetenz und Unabhängigkeit eine notwendige Voraussetzung für die Wirksamkeit der Überwachung der Finanzberichterstattung durch die Interne Revision darstellt und insbesondere in Zusammenarbeit mit dem Prüfungsausschuss die Gewinnsteuerung sowohl nach oben als auch nach unten eingeschränkt wird. Die Besonderheit der Studie von Abbott et al. (2016) im Vergleich zu anderen Studien besteht in der Darstellung von Unabhängigkeit und Kompetenz als separate Charakteristika, die in Interaktion miteinander die Qualität der Internen Revision darstellen.

Bajra und Cadez (2018) untersuchen den Zusammenhang zwischen ausgewählten Corporate Governance Mechanismen von 127 europäischen Unternehmen, die in den USA börsennotiert sind, und Earnings Management.[168] Zu den untersuchten Corporate Governance Mechanismen der Unternehmen gehören die Qualität des Vorstandes sowie die Qualität der Internen Revision in den Jahren 2000 bis 2013. Eine Auswertung dieser Daten ist besonders interessant, da der Zeitraum vor und nach der Implementierung des Sarbanes-Oxley Act of 2002 und der 8. EU-Richtlinie abgedeckt wird. Earnings Management messen die Autoren mittels Discretionary Accruals, zur Darstellung der Qualität der Internen Revision nutzen Bajra und Cadez (2018) neuartige und umfassende Messansätze. Die multidimensionale Operationalisierung der Qualität der Internen Revision erfolgt mittels der Qualitätsdimensionen Existenz beziehungsweise Einführung einer Internen Revision im Betrachtungszeitraum, Expertise[169], Unabhängigkeit und Einbezug in die externe Jahresabschlussprüfung.[170] Bajra und Cadez (2018) stellen für alle fünf Qualitätsdimensionen fest, dass eine gesteigerte Qualität der Internen Revision negativ mit Earnings Management korreliert.[171] Bajra und Cadez (2018) finden heraus, dass eine gesteigerte Qualität des Vorstandes negativ mit Earnings Management korreliert. Bei der Untersuchung des Interaktions-

effektes zwischen der Qualität der Internen Revision und der Qualität des Vorstandes stellen die Autoren entgegen ihrer Hypothese fest, dass sich dieser positiv auf Earnings Management auswirkt. Allerdings weisen Bajra und Cadez (2018) darauf hin, dass der Earnings Management begünstigende Effekt im Vergleich zum Earnings Management reduzierenden Effekt sehr klein ist. Außerdem zeigen die Studienergebnisse, dass sich sowohl die Qualität der Internen Revision als auch die Qualität des Vorstandes im Laufe der Zeit steigern.

Ismael und Kamel (2021) untersuchen im Rahmen ihrer Forschung den Zusammenhang zwischen Earnings Management und der Qualität Interner Revisionen auf Basis von 115 britischen, an der London Stock Exchange (LSE) gehandelten Unternehmen.[172] Die Qualität der Internen Revision bilden Ismael und Kamel (2021) mittels Fragebögen, die die Revisionsleiter der jeweiligen Unternehmen ausgefüllt haben, ab. Die Revisionsleiter beurteilen ihre Abteilungen auf Basis der vier Qualitätskomponenten des IIA, Größe der Internen Revision (Size), Unabhängigkeit der Internen Revision (Independence), Fachkompetenz (Skills and Knowledge) und Arbeitspraktiken beziehungsweise Methoden (Methodology and Management).[173] Earnings Management messen die Autoren mittels Abnormal Accruals auf Basis des Modified Jones Model. Ismael und Kamel (2021) stellen fest, dass Unternehmen mit einer qualitativ hochwertigen Internen Revision mit einer geringeren Wahrscheinlichkeit gewinnerhöhendes Earnings Management betreiben.[174] Zudem untersuchen die Autoren, analog zu Abbott et al. (2016), ob eine hohe Kompetenz oder Unabhängigkeit der Internen Revision zur Reduktion von Earnings Management führt, finden jedoch nur signifikante Ergebnisse für eine der beiden Qualitätskomponenten. Demnach betreiben Unternehmen mit einer kompetenteren Internen Revision zu einer geringeren Wahrscheinlichkeit gewinnerhöhendes Earnings Management. Dieses Ergebnis deckt sich mit den Ausführungen von und Ege (2015), der einen negativen, signifikanten Zusammenhang zwischen der Kompetenz einer Internen Revision und der Wahrscheinlichkeit für Fehlverhalten durch die Geschäftsführung feststellt.[175] Grundsätzlich deckt sich dieses Ergebnis mit den Forschungsergebnissen von Abbott et al. (2016), allerdings sehen diese das gemeinsame Vorhandensein von Kompetenz und Unabhängigkeit der Internen Revision als Erfolgsfaktor zur Steigerung der Rechnungslegungsqualität.[176]

Grundsätzlich kann festgehalten werden, dass der überwiegende Teil der vorliegenden Studien einen Zusammenhang zwischen der Qualität von Internen Revisionen und der Qualität der Rechnungslegung feststellt. Abgesehen von den Ergebnissen von Johl et al. (2013) und dem festgestellten Interaktionseffekt zwischen Vorstand und Interner Revision bei Bajra und Cadez (2018) zeigen alle anderen, dass eine qualitativ hochwertigere Interne Revision mit einer höheren Wahr-

---

[165]Vgl. Abbott et al. (2016, S. 4).

[166]Listet die umsatzstärksten amerikanischen Unternehmen.

[167]Vgl. hier und im Folgenden (Abbott et al., 2016, S. 5-6 & 13-33).

[168]Vgl. hier und im Folgenden Bajra und Cadez (2018, S. 152-158).

[169]Messung mittels der Subkomponenten Qualifikation, Erfahrung und Zertifikation.Weiterführend siehe Bajra und Cadez (2018, S. 157).

[170]Vgl. Bajra und Cadez (2018, S. 156-157).

[171]Vgl. hier und im Folgenden Bajra und Cadez (2018, S. 161-163).

[172]Vgl. hier und im Folgenden Ismael und Kamel (2021, S. 956-961).

[173]Vgl. Ismael und Kamel (2021, S. 955 & 959).

[174]Vgl. hier und im Folgenden Ismael und Kamel (2021, S. 961-973).

[175]Vgl. Ege (2015, S. 497-498).

[176]Vgl. Abbott et al. (2016, S. 13-33).

scheinlichkeit Earnings Management reduziert und somit positiv zur Qualität der Rechnungslegung beiträgt. Wesentliche Treiber der Qualität von Internen Revisionen scheinen gemäß der Studienergebnisse von Ege (2015), Abbott et al. (2016) und Ismael und Kamel (2021) insbesondere die Qualitätskomponenten Kompetenz und Unabhängigkeit zu sein.

4.3. Wahrnehmung der Internen Revision

Aufbauend auf den Erkenntnissen bezüglich der Relevanz von qualitativ hochwertigen Internen Revisionen, soll im Folgenden ein Einblick in die Forschungsergebnisse bezüglich der Wahrnehmung von Internen Revisionen und deren Auswirkungen gegeben werden. Schneider und Wilner (1990) untersuchen anhand eines experimentellen Forschungsdesigns mittels randomisierter Antwort-Technik die Auswirkungen auf Unregelmäßigkeiten in der Rechnungslegung in Abhängigkeit der Wahrnehmung der internen und externen Rechnungslegungsprüfung durch Manager.[177] Insgesamt 264 Probanden, zu überwiegender Mehrheit bestehend aus erfahrenen Managern, beantworten Fragen zu drei Rechnungslegungsfällen mit unterschiedlichen Merkmalsausprägungen[178] und abweichendem Kontrollumfeld.[179] Schneider und Wilner (1990) stellen im Rahmen ihrer Forschung fest, dass es bezüglich der abschreckenden Wirkung keine Unterschiede zwischen der Wahrnehmung der Tätigkeit der Internen Revision oder des externen Jahresabschlussprüfers gibt.[180] Die Forschungsergebnisse zeigen, dass die Interne Revision analog der Prüfung durch einen externen Jahresabschlussprüfer einen eindeutig abschreckenden Effekt hat, wenn es sich um einen wesentlichen Geldbetrag, die Überbewertung von Vermögensgegenständen, eindeutige Verstöße gegen Regelungen des US-GAAP und einen geringen Anreiz zur Bilanzmanipulation handelt. Die Forschungsergebnisse von Schneider und Wilner (1990) sind inkonsistent zu den vorangegangenen Erkenntnissen von Uecker et al. (1981), die im Rahmen ihres experimentellen Forschungsdesigns in Form einer Postkorb-Fallstudie die Auswirkungen der wahrgenommenen Aggressivität von internen und externen Rechnungslegungsprüfern testen.[181] Uecker et al. (1981) stellen auf Basis vorangegangener Forschungen die Hypothese auf, dass ein Interner Auditor, der als aktive und starke Persönlichkeit wahrgenommen wird und direkt an den Vorstand berichtet, mit einer höheren Wahrscheinlichkeit Unregelmäßigkeiten in der Rechnungslegung erkennt und aufdeckt.[182] Zudem testen die Autoren die Hypothese, dass eine höhere wahrgenommene Aggressivität der Prüfungstätigkeit das Auftreten von Unregelmäßigkeiten in der Rechnungslegung

reduziert. Keine der aufgestellten Hypothesen konnte im Zuge der Forschung bestätigt werden und Uecker et al. (1981) konkludieren, dass die reine Existenz einer Internen Revision und deren Auftreten die Geschäftsführung nicht davon abhält, Bilanzmanipulation zu betreiben oder betrügerische Handlungen in Betracht zu ziehen.[183]

Besonders spannend ist die Veränderung der Wahrnehmung der Internen Revision und die damit verbundenen Auswirkungen auf die Qualität der Rechnungslegung im Verlauf der vergangenen Jahrzehnte. Zwar wurde bereits in den 1970er Jahren vermehrt darauf hingewiesen, dass eine starke Interne Revision mit freiem Zugang zu allen relevanten Informationen als eine der wesentlichen Kontrollempfehlungen zur Vermeidung und Aufdeckung von Bilanzmanipulationen und dolosen Handlungen gilt, dieser Zusammenhang konnte jedoch erst deutlich später in Studien belegt werden.[184] Während das Vorhandensein und die Wahrnehmung einer Internen Revision in den 1980er Jahren noch keinen wesentlichen Einfluss auf das Verhalten von Managern und damit verbunden auf die Vermeidung von Bilanzmanipulationen hatten, sehen aktuellere Studien andere Forschungsergebnisse. Dies zeigt sich insbesondere dadurch, dass die überwiegende Mehrheit der in Kapitel 4.1. und 4.2. ausgewerteten Studien zeigt, dass die Existenz einer Internen Revision per se entweder mindestens zu einer gleichbleibenden, meist aber eher verbesserten Qualität der Rechnungslegung führt. Die Ergebnisse zeigen eine Entwicklung und Steigerung des Bewusstseins über die Relevanz und Tätigkeitsfelder der Internen Revision und somit die abschreckende Wirkung Bilanzmanipulation oder dolose Handlungen zu betreiben.

4.4. Ausgliederung der Internen Revision

Die wissenschaftliche Literatur beschäftigt sich neben dem Einfluss der bloßen Existenz von Internen Revisionen auch mit den Vor- und Nachteilen der Ausgliederung der entsprechenden Tätigkeiten. Grundsätzlich lässt sich feststellen, dass das Ausgliedern der Tätigkeiten der Internen Revision an externe Dienstleister zunehmend an Relevanz gewinnt.[185] Generell werben die externen Anbieter von Revisionsdienstleistungen damit, dass sie kosteneffiziente Arbeitsergebnisse liefern und darüber hinaus gegenüber unternehmensinternen Revisionen auf überlegenes Fachwissen zurückgreifen können.[186] Dementsprechend gilt es, auch diesen Aspekt näher zu betrachten und der wissenschaftlichen Literatur zu entnehmen, ob eine ausgegliederte Interne Revision einer unternehmenseigenen Revisionsabteilung überlegen ist und demnach einen größeren positiven Einfluss auf die Qualität der Rechnungslegung haben könnte.

Baatwah et al. (2021) untersuchen, anknüpfend an die Untersuchungen von Baatwah, Al-Ebel und Amrah (2019), auf Basis der Paneldaten von 928 omanischen Unternehmen

---

[177]Vgl. hier und im Folgenden Schneider und Wilner (1990, S. 669-677).
[178]Wesentlichkeit des Geldbetrages, Art des Rechnungslegungsverstoßes, Wahrgenommenes Ausmaß der Verstöße gegen Regelungen des US-GAAP und Stärke des Anreizes zur Bilanzmanipulation.
[179]Prüfung der Rechnungslegung durch den externen Jahresabschlussprüfer, die Interne Revision oder ohne Prüfung.
[180]Vgl. hier und im Folgenden Schneider und Wilner (1990, S. 677-681).
[181]Vgl. Schneider und Wilner (1990, S. 680); Uecker et al. (1981, S. 465).
[182]Vgl. hier und im Folgenden Uecker et al. (1981, S. 467).

[183]Vgl. Uecker et al. (1981, S. 465 & 476-478).
[184]Weiterführend siehe Uecker et al. (1981, S. 466).
[185]Vgl. Carey, Subramaniam und Ching (2006, S. 26).
[186]Vgl. Caplan und Kirschenheiter (2000, S. 392).

im Zeitraum 2005 bis 2017, den Zusammenhang zwischen der Expertise einer ausgegliederten Internen Revision und echtem Earnings Management, also sachverhaltsgestaltenden, bilanzpolitischen Maßnahmen.[187] Die Autoren haben bewusst den omanischen Markt für ihre Untersuchungen gewählt, da zum Untersuchungszeitpunkt 58 % der omanischen Unternehmen ihre Interne Revision ausgegliedert haben.[188] Der hohe Anteil ausgegliederter Interner Revisionen resultiert im Wesentlichen aus der Schwierigkeit, kompetente Mitarbeiter auf dem omanischen Arbeitsmarkt zu finden und die Objektivität Interner Revisionen zu gewährleisten.[189] Gemäß Barr-Pulliam (2016) kann diese Problematik jedoch auch auf andere Länder übertragen werden, da es generell als schwierig erachtet wird, entsprechend kompetente Mitarbeiter zur Besetzung von unternehmensinternen Revisionen zu finden.[190] Davon abgesehen ist Oman eine besonders interessante Forschungsregion, da Geschäftsführungen in Oman im Vergleich mit 38 anderen Ländern am meisten Bilanzmanipulation zu betreiben scheinen.[191] Baatwah et al. (2021) stellen im Rahmen ihrer Forschung fest, dass die Branchenexpertise einer ausgegliederten Internen Revision nicht mit der Reduktion von echtem Earnings Management zusammenhängt, das firmenspezifische Wissen hingegen schon.[192] Die Autoren argumentieren, dass die entsprechende Branchenexpertise auf Managementebene der ausgegliederten Internen Revisionen durchaus vorhanden sein könnte, die Weitergabe an das Personal jedoch nur begrenzt stattfindet. Außerdem vermuten die Autoren, dass die Aufführung von branchenspezifischen Fachkenntnissen auf der Homepage eines Anbieters lediglich Marketingzwecken dienen könnte. Das firmenspezifische Wissen einer ausgegliederten Internen Revision ist gemäß Baatwah et al. (2021) zur Vermeidung und Aufdeckung von echtem Earnings Management naturgemäß von größerer Bedeutung. Laut diesen Erkenntnissen weisen die Autoren darauf hin, dass eine längere Mandatsdauer mit einer gesteigerten Qualität der ausgegliederten Internen Revision einhergehen könnte. Bereits 2019 untersuchen Baatwah et al. (2019) auf Basis von 711 Beobachtungen omanischer Unternehmen im Zeitraum 2005 bis 2014, wie effizient ausgegliederte Interne Revisionen sind und wie diese im Zusammenhang mit qualitativ hochwertigen externen Jahresabschlussprüfern stehen.[193] Die Autoren stellen fest, dass die ausgegliederten Internen Revisionen insbesondere dann effizient arbeiten, wenn die Dienstleistung von einer Big-4-Prüfungsgesellschaft[194] durchgeführt wird.[195] Zudem ist die Zusammenarbeit zwischen den externen Jahresabschlussprüfern von Unternehmen und deren ausgegliederten Internen Revisionen signifikant höher, wenn die ausgegliederte Dienstleistung von einer Big-4-Prüfungsgesellschaft durchgeführt wird. Dieses Ergebnis deckt sich mit Prawitt et al. (2012), die auf Basis von 334 Beobachtungen von 159 Unternehmen aus 37 Ländern den Zusammenhang zwischen der Ausgliederung von Internen Revisionen und Rechnungslegungsrisiken untersuchen.[196] Prawitt et al. (2012) subsummieren, dass die Jahresabschlussqualität von Unternehmen steigt, wenn die Tätigkeiten der Internen Revision zumindest partiell von externen Jahresabschlussprüfern erbracht werden.[197] Zudem können die Autoren feststellen, dass die Existenz einer Internen Revision, unabhängig davon, ob unternehmensintern oder ausgegliedert, Rechnungslegungsrisiken minimiert. Carey et al. (2006) fügen dem noch hinzu, dass insbesondere kleinere Unternehmen die Ausgliederung der Internen Revision als sinnvollen Schritt erachten, wohingegen größere Unternehmen vorrangig aus Gründen der Kosteneinsparung den Wechsel von einer unternehmensinternen zu einer ausgegliederten Revision in Betracht ziehen.[198]

Im Zusammenhang mit der Internen Revision wird in der wissenschaftlichen Literatur häufig von der Notwendigkeit der Unabhängigkeit der Internen Revision gesprochen.[199] Caplan und Kirschenheiter (2000) argumentieren daher im Rahmen ihrer Forschung, dass insbesondere das Angestelltenverhältnis von unternehmensinternen Revisoren dazu führt, dass diese deutlich weniger Verantwortung tragen und nicht haftbar gemacht werden können, im Vergleich zu Dienstleistern.[200] Und tatsächlich zeigen die Ergebnisse der Autoren, dass ausgegliederte Interne Revisionen mindestens eine gleichwertige, oft aber auch eine qualitativ höherwertige Prüfung durchführen, allerdings resultiert dies in einem deutlich höheren Honorar.[201] Interessanterweise widerspricht dies den Forschungsergebnissen von Carey et al. (2006), die argumentieren, dass insbesondere große Firmen aus Kosteneinsparungsgründen die Ausgliederung der Internen Revision in Betracht ziehen.[202] Caplan und Kirschenheiter (2000) stellen allerdings ebenfalls fest, dass die Ausgliederung der Internen Revision gleichzeitig zur Erhöhung von Risiken führen kann und insbesondere das Risiko besteht, dass Kontrollrisiken unentdeckt bleiben. Durch Skaleneffekte können externe Dienstleister teilweise überlegene Audit-Ressourcen bieten, wohingegen unternehmensinterne Revisionen meistens über mehr firmenspezifisches Wissen verfügen. James (2003) stellen im Zuge einer durchgeführten Umfrage fest, dass sowohl unternehmensinterne als auch ausgegliederte Interne Revisionen als kompetent genug wahrgenommen werden, um grundsätzlich betrügerische Handlungen aufzudecken, vorausgesetzt sie berichten direkt

---

[187]Vgl. Baatwah et al. (2021, S. 206).

[188]Vgl. Baatwah et al. (2021, S. 211).

[189]Weiterführend siehe Baatwah et al. (2019, S. 425).

[190]Vgl. Barr-Pulliam (2016, S. 1-2).

[191]Vgl. Baatwah et al. (2021, S. 211). Weiterführend siehe EY (2015, S. 22).

[192]Vgl. hier und im Folgenden Baatwah et al. (2021, S. 224-225).

[193]Vgl. Baatwah et al. (2019, S. 429).

[194]Heute bestehend aus: Deloitte, EY, PricewaterhouseCoopers und KPMG.

[195]Vgl. hier und im Folgenden Baatwah et al. (2019, S 430-441).

[196]Vgl. Prawitt et al. (2012, S. 1110-1112).

[197]Vgl. hier und im Folgenden Prawitt et al. (2012, S. 1129-1130).

[198]Vgl. Carey et al. (2006, S. 26-27).

[199]Vgl. Mubako (2019, S. 539).

[200]Vgl. Caplan und Kirschenheiter (2000, S. 392-393).

[201]Vgl. hier und im Folgenden Caplan und Kirschenheiter (2000, S. 411-412).

[202]Vgl. Carey et al. (2006, S. 26-27).

an den Prüfungsausschuss.[203] Allerdings stellt James (2003) fest, dass die Umfrageteilnehmer davon ausgehen, dass ausgegliederte Interne Revisionen über weniger branchen- und insbesondere firmenspezifisches Wissen verfügen. Diese Erkenntnis kollidiert teilweise mit den Forschungsergebnissen von Baatwah et al. (2021), die feststellen, dass insbesondere das firmenspezifische Wissen zur Aufdeckung von betrügerischen Handlungen eine große Rolle bei Internen Revisionen spielt.

Coram, Ferguson und Moroney (2008) untersuchen, ob Unternehmen, in denen eine Internen Revision existiert, betrügerische Handlungen mit einer höheren Wahrscheinlichkeit aufdecken und melden als Unternehmen ohne eine Interne Revision. Ihre Forschung basiert auf den 491 Unternehmen aus Australien und Neuseeland, die Teil des KMPG Fraud Survey[204] im Jahr 2004 waren.[205] Die relevanten Informationen bezüglich der Internen Revision der entsprechenden Unternehmen erhielten die Autoren mittels Fragebögen. Coram et al. (2008) stellen fest, dass Unternehmen mit einer Internen Revision mit einer höheren Wahrscheinlichkeit betrügerische Handlungen aufdecken.[206] Zudem zeigen die Forschungsergebnisse, dass dies insbesondere für Unternehmen gilt, die zumindest partiell eine unternehmenseigene Interne Revision haben. Dieses Forschungsergebnis deckt sich mit der in Kapitel 4.1. dargestellten Erkenntnis, dass die Existenz einer Internen Revision zur Steigerung der Rechnungslegungsqualität beiträgt. Gemäß Coram et al. (2008) ist die Wahrscheinlichkeit zur Aufdeckung und Meldung von betrügerischen Handlungen bei Unternehmen mit einer ausgegliederten Internen Revision niedriger.[207] Dieses Forschungsergebnis bildet gegenüber der allgemeinen Meinung der wissenschaftlichen Literatur eine Ausnahme, denn diese zeigt auf, dass eine ausgegliederte Interne Revision eine mindestens gleichbleibende Rechnungslegungsqualität bietet, wenn nicht sogar eine gesteigerte gegenüber unternehmensinternen Revisionen.[208]

Zusammenfassend kann festgestellt werden, dass die Ausgliederung der Internen Revision eine überlegenswerte Maßnahme für Unternehmen darstellen kann. Wenn Unternehmen auf qualitativ hochwertige Anbieter der Revisionsdienstleistung zurückgreifen, Maßnahmen einleiten, um dem potenziell gestiegenen Risiko des Übersehens von Kontrollrisiken entgegenzuwirken und höhere Honorare in Kauf nehmen, kann die Ausgliederung der Internen Revision durch eine mindestens gleichbleibende, im besten Fall aber gesteigerte Qualität der Dienstleistung, zur Steigerung der Qualität der Rechnungslegung beitragen. Allerdings ist sich die wissenschaftliche Literatur einig, dass die bisherigen Erkenntnisse zwar eine erste Tendenz, für eine abschließende Beurteilung der Vor- und Nachteile der Ausgliederung der

Internen Revision erkennen lassen, jedoch dass auch noch deutlich mehr Forschungsbedarf besteht.[209]

4.5. Sonstige Einflussfaktoren der Internen Revision
4.5.1. Ausgestaltung einer qualitativ hochwertigen Internen Revision

Wie bereits dargestellt, wird die Interne Revision in der wissenschaftlichen Literatur als einer von vier Grundpfeilern der Corporate Governance gesehen und ergänzt demnach das System, bestehend aus dem externen Jahresabschlussprüfer, dem Prüfungsausschuss und der Geschäftsführung.[210] Gramling et al. (2004) tragen in ihren Ausführungen einige der wesentlichen Treiber der Qualität der Internen Revision zusammen. Demnach spielen insbesondere zur Beurteilung der Kompetenz die Kenntnisse der Internen Revision über die unternehmensspezifischen Betriebsabläufe, Prozesse und Verfahren eine große Rolle.[211] Zwar zeigen zahlreiche der in den vorherigen Kapiteln zusammengetragenen Studien, dass die bloße Existenz einer Internen Revision bereits einen positiven Einfluss auf die Qualität der Rechnungslegung hat[212], dennoch verdeutlichen die in Kapitel 4.2. dargestellten Forschungsergebnisse, dass eine Internen Revision darüber hinaus einer ausreichend hohen Qualität entsprechen muss[213], damit diese einen wünschenswerten und positiven Einfluss auf die Qualität der Rechnungslegung haben kann. Dementsprechend werden im Folgenden die gemäß aktuellem Forschungsstand häufig genannten und die Qualität der Internen Revision beeinflussenden Faktoren aufgeführt. Für die Messung und Beurteilung der Qualität der Leistungen und Effektivität der Funktionen der Internen Revision stehen diverse Methoden zur Verfügung, eine Messung der Prüfungsqualität unter Zuhilfenahme einer einzelnen direkt beobachtbaren Variable ist nicht möglich. Bei dem von Internen Revisionen erbrachten Leistungsspektrum handelt es sich in der Regel um ein Vertrauensgut, da „es nicht möglich ist, die Verlässlichkeit der finanziellen Berichterstattung oder die Einhaltung von Gesetzen und Vorschriften usw. eindeutig auf die Revisionstätigkeit zurückzuführen".[214] Bedingt durch die Besonderheit des Vetrauensgutes und, wie in Kapitel 4.2. bereits dargestellt, der Menge an unterschiedlichen Überprüfungsmethoden ist die Messung der Qualität der Internen Revision problematisch.[215] Gemäß wissenschaftlicher Literatur gibt es die nachfolgend dargestellten, wesentlichen Treiber zur Steigerung und Beurteilung der Qualität einer Internen Revision. Schneider (1985) stellt fest, dass externe Jahresabschlussprüfer, die die Arbeitsergebnisse von Internen Revisionen gemäß SAS No. 65[216] in deren Arbeit integrieren möchten, die Qualität der Arbeitsergebnisse mit 41,7 % gewichten,

---

[203]Vgl. hier und im Folgenden James (2003, S. 315 & 323).
[204]Weiterführend siehe KPMG (2004, S. 5-26).
[205]Vgl. hier und im Folgenden Coram et al. (2008, S. 543).
[206]Vgl. hier und im Folgenden Coram et al. (2008, S. 557-558).
[207]Vgl. Coram et al. (2008, S. 557-558).
[208]Vgl. Mubako (2019, S. 538).

[209]Vgl. Mubako (2019, S. 542-543).
[210]Vgl. Gramling, Maletta, Schneider und Church (2004, S. 194).
[211]Vgl. Gramling et al. (2004, S. 200).
[212]Vgl. Beasley et al. (2000, S. 453); Coram et al. (2008, S. 557-558); Gebrayel et al. (2018, S. 211); Sierra García et al. (2012, S. 321-322).
[213]Vgl. Bajra und Cadez (2018, S. 161-163); Ege (2015, S. 497-498); Ismael und Kamel (2021, S. 961-973); Prawitt et al. (2009, S. 1269-1270).
[214]Zehnder und Brönnimann (2005, S. 449).
[215]Vgl. Hölscher und Rosenthal (2007, S. 15-16).
[216]Weiterführend siehe AICPA (Hrsg.) (1997), S. 1805-1812.

gefolgt von der Kompetenz der Internen Revision mit 35,2 % und abschließend mit deren Objektivität mit 23,1 %.[217] Zwar variiert die Gewichtung in der wissenschaftlichen Literatur, jedoch besteht durchaus Einigkeit darüber, dass es sich bei diesen drei Faktoren um wesentliche Treiber handelt.[218] Anknüpfend an die soeben genannten Faktoren führen Ismael und Kamel (2021) in Anlehnung an die Qualitätscharakteristiken gemäß des IIA[219] vier, teilweise übereinstimmende Faktoren auf.[220] Eine Interne Revision sollte in Abhängigkeit der Art und Größe des Unternehmens eine gewisse Mindestgröße umfassen, da davon ausgegangen wird, dass somit pro Person mehr Zeit für Prüfungshandlungen zur Verfügung steht, unterschiedliche berufliche Hintergründe vorhanden sind und verschiedene Prüfungstätigkeiten durchgeführt werden.[221] Die Unabhängigkeit und Objektivität einer Internen Revision gilt als gewährleistet, wenn die funktionale beziehungsweise direkte Berichtspflicht gegenüber dem Prüfungsausschuss gilt und gegenüber der Geschäftsführung nur die administrative beziehungsweise operative Berichtsbeziehung.[222] Die Zuverlässigkeit der Arbeitspraktiken und Methoden, die eine Interne Revision anwendet, um Prüfungssicherheit zu erlangen, sowie deren Einklang mit den berufsständischen Standards spielen ebenfalls eine elementare Rolle.[223] Der letzte Faktor umfasst die individuellen Eigenschaften und Kompetenzen der Internen Revisoren, wie deren Persönlichkeit, technisches Geschick, Bildungsstand, Berufserfahrung, Teilnahme an Schulungsmaßnahmen oder die professionellen Zertifizierung.[224] Prawitt et al. (2009) nehmen an, je kompetenter ein Interner Revisor ist, desto größer ist die Wahrscheinlichkeit, dass er Faktoren, die Earnings Management beeinflussen, versteht und Indikatoren für eine Verzerrung des Jahresergebnisses durch die Geschäftsführung erkennt und zu überprüfen weiß.[225] Die Anforderungen an die Kenntnisse und Kompetenzen von Mitarbeitern der Internen Revision werden zum einen im IPPF erläutert und zum anderen ausführlich im Kompetenzrahmen des IIA, dem Internal Audit Competency Framework dargestellt.[226] Der Kompetenzrahmen stellt mittels seiner definierten Wissensbereiche und aufeinander aufbauenden Kompetenzstufen eine Orientierung für die Weiterbildung und -entwicklung von Internen Revisoren dar und soll dadurch Qualifikationslücken identifizieren und beheben.[227] Regelmäßige Fortbildungen gehören zu den Anforderungen des IPPF an die Mitarbeiter der Internen Revision, zudem sollen Qualitätssicherungsmaßnahmen, interne und externe Beurteilungen zur Sicherung und Verbesserung der Mitarbeiter und Tätigkeiten der Internen Revision beitragen.[228]

4.5.2. Berücksichtigung kultureller Unterschiede

Bedingt durch die Globalisierung und Digitalisierung steigt zum einen die Anzahl an multinationalen Unternehmen und zum anderen arbeiten vermehrt Menschen aus verschiedenen Kulturen und Ländern zusammen.[229] Zur vollständigen Darstellung des Einflusses der Internen Revision auf die Qualität der Rechnungslegung gilt es daher, kulturelle Unterschiede in die Betrachtung miteinzubeziehen. Während in entwickelten Ländern die Relevanz und Notwendigkeit der Tätigkeiten und Maßnahmen einer qualitativ hochwertigen Internen Revision bereits erkannt wurden, ergeben sich in Entwicklungsländern teilweise große Hürden bei der Erfüllung der Aufgaben der Internen Revision aufgrund fehlender Akzeptanz durch differierende kulturelle Anschauungen.[230] Die Betrachtungen der im Rahmen dieses Kapitels ausgewerteten Studien, die den Effekt von kulturellen Unterschieden auf die Qualität der Internen Revision untersuchen, lehnen sich an die Kulturdimensionen von Hofstede (1980) an.[231] Die Dimensionen gliedern sich in Machtdistanz[232], Kollektivismus vs. Individualismus[233], Maskulinität vs. Femininität[234], Langezeit- vs. Kurzzeitorientierung[235], Genuss vs. Zurückhaltung[236] und Unsicherheitsvermeidung[237]. Die jeweiligen Dimensionen geben Tendenzen an, wie die Werte und Einstellungen von Menschen in verschiedenen Ländern im beruflichen Kontext typologisiert werden können. Die Forschungsergebnisse von Alzeban (2014) zeigen, dass Unternehmen in einem Umfeld mit hoher Machtdistanz sowie hoher Unsicherheitsvermeidung durch eine niedrigere Qualität der Internen Revision charakterisiert werden, wohingegen ein positiver Zusammenhang zwischen Individualismus und der Qualität der Internen Revision zu bestehen scheint.[238] Sarens und Abdolmohammadi (2010) stellen im Rahmen ihrer Forschung fest, dass ein Umfeld mit hoher Unsicherheitsvermeidung und Kollektivismus die Professionalität von Internen Revisionen negativ beeinflusst.[239] Zu-

---

[217]Vgl. Schneider (1985, S. 300).

[218]Vgl. Glover, Prawitt und Wood (2008, S. 195).

[219]Weiterführend siehe IIA (Hrsg.) (2017), S. 1-25.

[220]Vgl. zu diesem Absatz Arena und Azzone (2009, S. 43-60); Hass, Abdolmohammadi und Burnaby (2006, S. 835-844); Ismael (2019, S. 113-143); Ismael und Kamel (2021, S. 955); Sarens (2009, S. 1-7).

[221] Weiterführend siehe IIA (2010), S. 50-53.

[222] Weiterführend siehe Eighme und Cashell (2002, S. 3-10); Sarens (2009, S. 3-4).

[223]Weiterführend siehe Sarens (2009, S. 4).

[224]Weiterführend siehe Sarens (2009, S. 4).

[225]Vgl. Prawitt et al. (2009, S. 1260).

[226]Vgl. DIIR (2018), S. 29-31; DIIR (Hrsg.) (2020b), S. 1-7.

[227]Vgl. DIIR (Hrsg.) (2020b), S. 1-7.

[228]Vgl. DIIR (2018), S. 32-35.

[229]Vgl. Cizmas, Feder, Maticiuc und Vlad-Anghel (2020, S. 1-2).

[230]Vgl. Alzeban (2014, S. 57-58).

[231]Vgl. Alzeban (2014, S. 57); Eulerich und Ratzinger-Sakel (2018, S. 218); Sarens und Abdolmohammadi (2010, S. 5).

[232]Gibt an bis zu welchem Ausmaß weniger mächtige Menschen erwarten und akzeptieren, dass Macht ungleich verteilt ist. Bei hoher Machtdistanz ist die Macht sehr ungleich verteilt.

[233]Wertesystem, in dem die Bedürfnisse eines Individuums, vor dem der Gemeinschaft stehen vs. Wertesystem, in dem die Bedürfnisse eines Individuums, den Interessen der Gemeinschaft untergeordnet werden.

[234]Gibt an, in welchem Ausmaß eine Gesellschaft von Idealen wie Erfolg oder Macht vs. Fürsorge und Beziehungen geprägt ist.

[235]Ausrichtung einer Gesellschaft auf langanhaltende Lösungen vs. temporäre kurzfristige Erfolge.

[236]Gibt an, wie wichtig Menschen das Ausleben eigener Bedürfnisse ist.

[237]Grad, zu dem sich Menschen durch ungewisse oder unbekannte Situationen bedroht fühlen.

[238]Vgl. Alzeban (2014, S. 73).

[239]Vgl. hier und im Folgenden Sarens und Abdolmohammadi (2010, S. 22-25).

dem finden sie einen marginalen negativen Zusammenhang zwischen dem Durchsetzungsvermögen und der Professionalität von Internen Revisionen, was die Autoren darauf schließen lässt, dass Interne Revisionen am professionellsten vorgehen, wenn ein unaufdringliches, nicht aggressives und sanftmütiges Auftreten vorliegt. Im Allgemeinen finden Sarens und Abdolmohammadi (2010) heraus, dass nicht die geografische Lage eines Unternehmens über die Professionalität einer Internen Revision entscheidet, sondern die vorherrschenden Kulturdimensionen sowie das vorliegende Rechtssystem. So stellen die Autoren fest, dass ein Common Law-Umfeld im Vergleich zu einem Civil Law-Umfeld in der Regel mit einer besseren Corporate Governance verknüpft ist, was sich wiederum positiv auf die Professionalität von Internen Revisionen auswirkt. Die Forschungsergebnisse von Eulerich und Ratzinger-Sakel (2018) zeigen, dass die Dimensionen Unsicherheitsvermeidung und Langzeitorientierung die Unabhängigkeit und Objektivität von Internen Revisionen verbessern, wohingegen ein Umfeld mit hohem Individualismus einen negativen Effekt aufweist.[240] Es gilt, diese Forschungsergebnisse und den Einfluss der Internen Revision im Kontext mit dem jeweiligen kulturellen Umfeld auf die Qualität der Rechnungslegung zu berücksichtigen.

4.5.3. Whistleblowing

Nur allzu häufig hat die Vergangenheit gezeigt, dass nicht Behörden, Wirtschaftsprüfer oder eingerichtete interne Kontrollen Missstände oder Straftaten aufdecken, sondern Whistleblower.[241] Whistleblower legen potenzielle Risiken, illegale, unmoralische oder nicht legitimierte Praktiken gegenüber unternehmensinternen oder externen Parteien, die das entsprechende Fehlverhalten beeinflussen können, offen.[242] Sollten solche Sachverhalte in Unternehmen vorliegen, kann dies bei Nichtaufdeckung die Qualität der Rechnungslegung auf unterschiedliche Weisen negativ beeinflussen. Neben Maßnahmen der Unternehmen, die ergriffen werden, um betrügerische Handlungen von vornherein zu verhindern, gehört es zu einem erfolgreichen Corporate Governance Management dazu[243], ein gut funktionierendes Hinweisgebersystem für Whistleblower zu implementieren.[244] Hierbei soll die Möglichkeit gegeben werden, vermeintliche Missstände oder illegale Aktivitäten im Optimalfall unternehmensintern zu melden. Häufig handelt es sich bei Whistleblowern um Mitarbeiter der betroffenen Unternehmen.[245] Von zentraler Bedeutung für Unternehmen ist es daher, dass sie versuchen, mittels eines gut funktionierenden internen

Whistleblowing-Systems das externe Whistleblowing[246] und somit eventuelle Reputationsschäden, finanzielle Einbußen und Strafen[247] zu vermeiden.[248] Die Interne Revision kann in diesem Zusammenhang unterschiedliche Rollen einnehmen. Zum einen kann sie in den Prozess der Hinweissichtung und -untersuchung eingebunden sein, zum anderen muss sie im Rahmen der Risikoeinschätzung entsprechend evaluieren, ob das Unternehmen Rahmenbedingungen geschaffen hat, die das Hinweisen auf Missstände oder Straftaten so attraktiv wie möglich machen.[249] Zudem hat sich im Rahmen einer Umfrage unter Leitern von Revisionsabteilungen gezeigt, dass die Abteilung der Internen Revision regelmäßig als Anlaufstelle für Whistleblower genutzt wird, vorausgesetzt sie genießt ein entsprechend hohes Ansehen im Unternehmen.[250] Ein zentraler Erfolgsfaktor für die Nutzung der Internen Revision als Anlaufstelle für Hinweise ist die direkte Berichtsbeziehung der Internen Revision zum Prüfungsausschuss und somit die Unabhängigkeit von der Geschäftsführung.[251] Gemäß der Forschungsergebnisse von Kaplan und Schultz (2007) scheint die Qualität der Internen Revision, entgegen ihrer ursprünglichen Hypothese, keinen Einfluss auf die Meldewahrscheinlichkeit von Missständen zu haben.[252] Die Autoren begründen dies damit, dass in Unternehmen teilweise ein sehr eingeschränktes Wissen über die Aufgaben und Tätigkeitbereiche der Internen Revision vorherrscht. Die bisherigen Forschungsergebnisse zeigen, dass die Existenz einer Internen Revision ein entscheidender Erfolgsfaktor im Umgang mit Whistleblowern und damit verbunden der Aufdeckung von, die Qualität der Rechnungslegung beeinflussenden Missständen sein kann.

4.5.4. Rotation des Revisionsleiters

Die von der EU eingeführte Rotationspflicht[253] für Jahresabschlussprüfer von Unternehmen von öffentlichem Interesse[254] wirft die Frage auf, ob auch die Rotation des Leiters der Internen Revision zu einer höheren Prüfungsqualität und damit verbunden zu einer höheren Rechnungslegungsqualität führen könnte. Auf die hohe Relevanz der Objektivität der Internen Revision wird im Rahmen des praktischen Ratschlages der internationalen Grundlagen für die berufliche Praxis der Internen Revision Nr. 1120 verwiesen.[255] Die Autoren

---

[240]Vgl. hier und im Folgenden Eulerich und Ratzinger-Sakel (2018, S. 224).
[241]Vgl. ACFE (2020, S. 19); Dworkin und Baucus (1998, S. 1282); Dyck, Morse und Zingales (2010, S. 2215); EY (2016, S. 2).
[242]Vgl. Near und Miceli (1985, S. 508).
[243]Empfehlung des DCGK: „(...) Beschäftigten soll auf geeignete Weise die Möglichkeit eingeräumt werden, geschützt Hinweise auf Rechtsverstöße im Unternehmen zu geben (...)". DCGK (Hrsg.) (2019), S. 4.
[244]Vgl. hier und im Folgenden Schemmel, Ruhmannseder und Witzigmann (2012, S. 5).
[245]Vgl. ACFE (2020, S. 19); Dyck et al. (2010, S. 2214); Stubben und Welch (2020, S. 474).

[246]Wenn sich Whistleblower mit ihren Beobachtungen an externe Dritte, wie beispielsweise Medienvertreter oder Behörden der Strafverfolgung wenden.
[247]Beispielsweise straf- und arbeitsrechtliche Konsequenzen, Wettbewerbsverbote, Strafzahlungen und Verbot der Vergabe von öffentlichen Aufträgen.
[248]Vgl. Rothschild und Miethe (1999, S. 120).
[249]Vgl. Schneider (2008, S. 74-75).
[250]71% der befragten Revisionsleiter gaben an Hinweise von Whistleblowern erhalten zu haben.Vgl. Read und Rama (2003, S. 356).
[251]Vgl. Chambers (1995, S. 195); IIA (2013b, S. 1-3); Schneider (2008, S. 74-75).
[252]Vgl. hier und im Folgenden Kaplan und Schultz (2007, S. 121).
[253]Vgl. Europäisches Parlament und Rat der Eulerich und van Uum (2014, Art. 17).
[254]Weiterführend zu den Begriffsbestimmungen der Unternehmen von öffentlichem Interesse siehe Europäisches Parlament und Rat der EU (Hrsg.) (2006), Art. 2 Nr. 13.
[255] Vgl. DIIR (Hrsg.) (2016), S. 87-89.

d'Arcy und Hoos (2012) stellen im Rahmen einer Umfrage unter 118 Internen Revisoren fest, dass die Rotation des Revisionsleiters als eine mögliche Maßnahme zur Steigerung der Objektivität und Unabhängigkeit der Internen Revision gesehen wird, allerdings mit dem Nachteil des Verlustes des firmenspezifischen Wissens und der Erhöhung der Kosten.[256] Daugherty, Dickins, Hatfield und Higgs (2012) verweisen in ihren Forschungsergebnissen ebenfalls darauf, dass die Rotation von Prüfungsleitern mit dem Verlust von kundenspezifischem Wissen einhergeht, was gemäß den Forschungsergebnissen von Baatwah et al. (2021) als ein wesentlicher Faktor zur Reduktion von Earnings Management betrachtet werden sollte.[257] Daugherty et al. (2012) zeigen zudem auf, dass durch Prüferrotation die Prüfungsqualität, bedingt durch die gesteigerte Unabhängigkeit des Prüfers, zwar als gesteigert wahrgenommen wird, es tatsächlich aber insbesondere in der Startphase nach der Rotation eher zur Reduktion der Prüfungsqualität kommt.[258] Diese Ergebnisse decken sich mit Geiger und Raghunandan (2002) und St. Pierre und Anderson (1984), die ebenfalls feststellen, dass die Wahrscheinlichkeit für Fehler im Prüfungsverfahren insbesondere in den ersten Jahren nach der Rotation deutlich höher ist.[259] Demgegenüber stehen Forschungsergebnisse, die einen Zusammenhang zwischen der Rotation des Abschlussprüfers und der Reduktion von Earnings Management[260] beziehungsweise der Verschlechterung der Prüfungsqualität mit laufender Prüfungsmandatsdauer[261] feststellen. Abschließend bleibt festzuhalten, dass in der wissenschaftlichen Literatur zum aktuellen Zeitpunkt noch Uneinigkeit bezüglich der Vor- und Nachteile der Rotation von externen Jahresabschlussprüfern herrscht und dass insbesondere im Hinblick auf die niedrige Forschungsdichte bezüglich der Rotation von Prüfungsleitern in der Internen Revision derzeit noch keine abschließende Aussage getroffen werden kann.

5. Kritische Würdigung

Die wissenschaftliche Literatur zum Thema Interne Revision ist zwar zahlreich, jedoch aufgrund der Abhängigkeit des Einflusses der Internen Revision auf die Qualität der Rechnungslegung von vielfältigen Faktoren sehr widersprüchlich und stellenweise veraltet. Dies begründet sich unter anderem darin, dass die ausgewerteten Daten und das Umfeld, aus dem diese stammen, stark voneinander abweichen. Beispielsweise verweisen die ausgewerteten Studien im Rahmen ihrer Hypothesenbildung auf die jeweils geltenden Normen im Hinblick auf die Interne Revision. Demnach sind die rechtlichen Rahmenbedingungen und das kulturelle Umfeld der

Studien, die über USA, Australien, Neuseeland, Europa, Spanien und UK bis hin zu Malaysia und Oman reichen, sehr unterschiedlich. Wie in den vorangegangenen Kapiteln dargestellt, sind insbesondere die rechtlichen Rahmenbedingungen und Anforderungen an die Tätigkeit von Internen Revisionen sehr unterschiedlich, da zahlreiche Rahmenwerke und Stellungnahmen von relevanten Institutionen auf nationaler und internationaler Ebene existieren. Hierbei gilt es zudem, im Blick zu behalten, in welchem strukturellen und sozioökonomischen Umfeld die betrachteten Unternehmen tätig sind, welchen Wirtschaftszweigen sie angehören und was für eine Art von Personen in dem Bereich tätig sind. Dementsprechend müssen die Studienergebnisse insbesondere im Hinblick auf die Übertragbarkeit auf deutsche Unternehmen mit Vorsicht interpretiert werden.

Das gleiche gilt auch für die Betrachtung des betriebenen Earnings Managements und dessen Einfluss auf die Qualität der Rechnungslegung. Bei der Rechnungslegungsqualität handelt es sich um ein komplexes Konstrukt, dass sich aus vielen unterschiedlichen Einflussfaktoren, wie beispielsweise der Qualität des externen Jahresabschlussprüfers[262], weiterer Corporate Governance Mechanismen, den rechtlichen und demografischen Rahmenbedingungen[263] sowie der Herkunft der Umsätze[264] zusammensetzt. Beispielsweise hat sich im Rahmen empirischer Forschungen gezeigt, dass Earnings Management in Ländern mit starken Corporate Governance Strukturen, großen Aktienmärkten, starken Anlegerrechten, Streubesitz und einer strenger Rechtsdurchsetzung am niedrigsten ist.[265] Die Messung der Qualität von Corporate Governance Mechanismen, insbesondere der Internen Revision, erfolgt in den meisten Studien mit Konstrukten, die sich in der Regel nur auf eine oder eine kleine Anzahl von Qualitätsdimensionen konzentrieren.[266] In der wissenschaftlichen Literatur wird darauf verwiesen, dass eine Erweiterung und Aktualisierung der etablierten Methoden zur Messung der Qualität der Internen Revision sinnvoll ist.[267] Insbesondere haben sich die notwendigen Fähigkeiten und Kompetenzen zur Durchführung von Prüfungshandlungen in der Internen Revision verändert, so dass, um ein Beispiel zu nennen, zukünftig die Fähigkeit IT-Landschaften zu prüfen, von großer Bedeutung sein könnte. Aber auch der bislang eingeschränkte Zugang zu unternehmensinternen Dokumenten bezüglich Interner Revisionen, die beispielsweise Informationen über die durchgeführten Prüfungshandlugen, Budgets, Zeitvorgaben und Vergütungsstrukturen enthalten, erschwert die Einschätzung der Qualität der Internen Revisionen. Darauf aufbauend ist limitierend zu berücksichtigen, dass insbesondere die häufige Erfassung der Qualitätsaspekte der Internen Revision auf Basis von Befragungen mit der Repräsentation von

---

[256]Vgl. d'Arcy und Hoos (2012, S. 124-131).

[257]Vgl. Baatwah et al. (2021, S. 224-225); Daugherty et al. (2012, S. 111-112).

[258]Vgl. Daugherty et al. (2012, S. 111).

[259] Vgl. Geiger und Raghunandan (2002, S. 75-76); St. Pierre und Anderson (1984, S. 256).

[260]Weiterführend siehe Nagy (2005, S. 57-66).

[261]Weiterführend siehe Copley und Doucet (1993, S. 29-34).

[262]Weiterführend siehe Becker, DeFond, Jiambalvo und Subramanyam (1998, S. 5-21).

[263]Weiterführend siehe Francis und Wang (2008, S. 158-162 & 185-186); Leuz et al. (2003, S. 505-527).

[264]Weiterführend siehe Dyreng, Hanlon und Maydew (2012, S. 683-684).

[265]Vgl. Francis und Wang (2008, S. 185); Leuz et al. (2003, S. 506).

[266]Vgl. Bajra und Cadez (2018, S. 153).

[267]Vgl. hier und im Folgenden Gramling et al. (2004, S. 235-236).

subjektiven Sichtweisen einhergeht.

Im Rahmen der Forschung kann nur der finale veröffentlichte Jahresabschluss für die Untersuchung herangezogen werden und es ist nahezu unmöglich, klar abzugrenzen, welche Faktoren in welchem Maß zur Qualität des Jahresabschlusses beigetragen haben. Außerdem ist die Verwendung von Abnormal beziehungsweise Discretionary Accruals als Maß zur Berechnung von Earnings Management und damit verbunden, dem Ausmaß der Rechnungslegungsqualität, kritisch zu betrachten, da die Höhe der ermittelten Abnormal Accruals unter anderem davon abhängt, welche Methode zur Ermittlung zum Einsatz kommt.[268] Des Weiteren kann sich das ermittelte Ausmaß des Earnings Managements auch in einem vollständig legalen Rahmen ohne jegliche bewusste Manipulation bewegen. Zwar ist das Vorgehen in der wissenschaftlichen Literatur gängige Praxis und zum Großteil bedingt durch mangelnde Datenverfügbarkeit, aber ein genaueres und individuelleres Ermittlungsverfahren wäre für zukünftige Forschungen wünschenswert. Es stellt sich demnach die Frage, ob Abnormal beziehungsweise Discretionary Accruals für die Beurteilung des Einflusses der Internen Revision auf die Qualität der Rechnungslegung ein optimales Maß sind.

Im Allgemeinen geben nahezu alle Studien als wesentliche Limitation einen kleinen Stichprobenumfang an. Zudem finden kaum Untersuchungen bzw. Differenzierung zwischen unternehmensinternen Revisionen und ausgegliederten Revisionen statt. Während zahlreiche Untersuchungen zu den potenziellen Risikofaktoren einer eingeschränkten Unabhängigkeit der externen Jahresabschlussprüfer vorliegen, bleibt die wissenschaftliche Literatur solchen Untersuchungen im Hinblick auf die Interne Revision schuldig.[269] Da die in Kapitel 4.2. ausgewerteten Studien keine signifikanten Ergebnisse für den Zusammenhang mit der Objektivität der Internen Revision liefern, in der wissenschaftlichen Literatur aber dennoch Einigkeit darüber herrscht, dass die Objektivität einer Internen Revision zu den wesentlichen Treibern der Qualität gehört, sollte dieser Faktor in einem anderen Setting erneut erforscht werden.

Aufgrund der voranschreitenden Globalisierung ist ein vermehrter Fokus auf die Interne Revision als Teil des Internen Kontrollsystems zu setzen. Sowohl die Prozesse als auch die organisatorischen Strukturen in Unternehmen werden zunehmend komplexer und machen ein wirksames und effizientes IKS unverzichtbar. Dementsprechend muss der Aufgabe der unternehmensinternen Überprüfung der Wirksamkeit und Optimierung des IKS durch die Interne Revision auch auf politischer Ebene mehr Rechnung getragen werden. Zwar bietet die vage Formulierung des Gesetzgebers in § 91 Abs. 2 AktG den Unternehmen die Möglichkeit möglichst flexibel und individuell auf unternehmensinterne Besonderheiten und Bedürfnisse einzugehen, aber auf der anderen Seite fehlt der unternehmensübergreifende Nachdruck zur Einrichtung

einer einheitlichen Internen Revision durch eine explizite gesetzliche Forderung. Demnach ist es grundsätzlich problematisch, dass die fehlende Konkretisierung der Notwendigkeit, Ausgestaltung und Aufgaben der Internen Revision in Unternehmen zu differierenden Meinungsauffassungen führt. Eine Erläuterung, welche Maßnahmen der Gesetzgeber als geeignet auffasst, um die Anforderungen des § 91 Abs. 2 AktG zu erfüllen, bleibt aus, wodurch die Auslegungen der Rechtsnorm stark variieren.[270] Klar definierte und einheitliche Anforderungen an Interne Revisionen von Unternehmen wären insbesondere zur Verbesserung der Qualität von Internen Revisionen und der Entscheidungsfindung zur Ausgliederung von Revisionstätigkeiten an externe Dienstleister wünschenswert. Insbesondere da die Forschungsergebnisse zeigen, dass eine Interne Revision positiv zur Qualität der Rechnungslegung beitragen kann, wäre die explizite Formulierung der Mindestanforderungen an eine gesetzlich explizit geforderte Interne Revision sinnvoll.

Trotz der soeben aufgeführten Limitationen lassen sich zur Beantwortung der Ausgangsfrage dieser Masterarbeit einige Schlussfolgerungen bezüglich des Einflusses der Internen Revision auf die Qualität der Rechnungslegung ziehen. Da eine effektive Interne Revision als einer von vier Grundpfeilern qualitativ hochwertiger Corporate Governance gewertet wird, ist dessen Rolle, Ausgestaltung und Funktionalität von hoher Bedeutung[271] und die Interne Revision kann dazu beitragen, Informationsasymmetrien zwischen dem Prüfungsausschuss und der Geschäftsführung zu reduzieren.[272] Die überwiegende Mehrheit der im Rahmen dieser Masterarbeit betrachteten Studien zeigt einen Zusammenhang zwischen Internen Revisionen und der Qualität der Rechnungslegung von Unternehmen. Zwar variieren die Untersuchungsansätze sowie die zugrunde gelegten Annahmen und Datensätze, aber zusammenfassend kann festgestellt werden, dass die Qualität der Rechnungslegung durch die Existenz einer Internen Revision zumindest gleichbleibt, oft aber auch verbessert werden kann. Einen Einfluss auf die Stärke und Richtung des Einflusses der Internen Revision auf die Qualität der Rechnungslegung hat die Qualität der Internen Revision. Auch wenn die Gewichtungen der qualitätsbeeinflussenden Faktoren von den relevanten Institutionen und Forschern unterschiedlich dargestellt werden, so können dennoch ein paar Schlussfolgerungen zur optimalen Ausgestaltung der Internen Revision gezogen werden. Wie die Auswertung der Studien gezeigt hat, stellen Kompetenz, Objektivität und qualitativ hochwertige Prüfungsergebnisse der Internen Revisoren die zentralen Erfolgsfaktoren für die wirksame Eindämmung von Earnings Management und somit für die Steigerung bzw. Sicherung der Qualität der Rechnungslegung dar. Dies beinhaltet insbesondere die direkte Berichtsbeziehung zum Prüfungsausschuss, die gemäß

---

[268]Weiterführend siehe Keßler (2015, S. 31-70).
[269]Vgl. Abbott et al. (2016, S. 6).

[270]Vgl. Bantleon et al. (2017, S. 684); Berwanger und Kullmann (2012, S. 93-94 & 96).
[271]Vgl. Vgl. Prawitt et al. (2009, S. 1256 & 1258); Sierra García et al. (2012, S. 306).
[272]Vgl. Gebrayel et al. (2018).

berufsständischer Rahmenwerke angewandten Prüfungsmethoden und die als positiv wahrgenommenen persönlichen Eigenschaften der Internen Revisoren. Wie jedoch in den Limitationen bereits aufgeführt, variieren auch hier die Gewichtungen und das inhaltliche Verständnis der Faktoren, dementsprechend ist weiterer Forschungsbedarf zur optimalen Ausgestaltung von Internen Revisionen notwendig.

Zudem gilt es für jedes Unternehmen, individuell zu entscheiden, ob die Tätigkeiten der Internen Revision vollständig unternehmensintern durchgeführt werden sollen, oder ob eine partielle beziehungsweise vollständige Ausgliederung an externe Dienstleister sinnvoll ist. Wenn auf ein entsprechendes Niveau der Qualität des externen Dienstleisters geachtet wird und aufgrund genutzter Skaleneffekte die Honorare nicht zu hoch sind, kann auch die Ausgliederung einer Internen Revision gegenüber einer unternehmenseigenen Revision von Vorteil sein. Da der aktuelle Forschungsstand noch keine abschließende Empfehlung bezüglich der Vor- und Nachteile der Ausgliederung von Revisionsabteilungen machen kann, muss jedes Unternehmen mittels Kosten-Nutzen-Analyse eine individuelle Entscheidung treffen.

Wie die vorangegangenen Ausführungen zeigen, kann die Existenz einer qualitativ hochwertigen Internen Revision zur Reduktion von Informationsasymmetrien und Interessenkonflikten, insbesondere bezüglich des Prüfungsausschusses, beitragen und darüber hinaus einen wesentlichen Beitrag zur Verbesserung der Qualität der Rechnungslegung leisten, jedoch besteht im Allgemeinen weiterer Forschungsbedarf.

6. Fazit

Abschließend bleibt festzuhalten, dass zahlreiche Faktoren determinieren, welchen Einfluss eine Interne Revision auf die Qualität der Rechnungslegung von Unternehmen hat. Grundsätzlich kann eine qualitativ hochwertige Interne Revision, als einer von vier Grundpfeilern der Corporate Governance, einen wesentlichen Erfolgsfaktor zur Reduktion von Bilanzpolitik, zur Aufdeckung von Bilanzmanipulationen und somit zur Vermeidung von Bilanzskandalen darstellen. Entscheidend für die positive Einflussnahme auf die Rechnungslegung ist die Berücksichtigung der wesentlichen Treiber der Qualität der Internen Revision sowie die Orientierung der Prüfungshandlungen an vorgegebenen Rahmenwerken.

Eine Auswertung des Einflusses der Internen Revision auf Basis von mehrjährig und international angelegten Studien wäre wünschenswert. Weitere Erkenntnisse zu diesem Thema sind für zahlreiche nationale und internationale Akteure von Interesse, da die Erkenntnisse über die Erfolgsfaktoren einer qualitativ hochwertigen Internen Revision unabhängig von der Unternehmensart, -größe oder -branche ein entscheidender Erfolgsfaktor zur Reduktion von Bilanzmanipulationen sein kann. Es sollte daher weitere Forschung zum Einfluss der Internen Revision auf die Qualität der Rechnungslegung betrieben werden.

Tabelle 1: Zusammenfassung ausgewählter Studienergebnisse

Autor (en)	Stichprobe/ Zeitraum	Studiendesign/ Methode(n)	Untersuchungsgegenstand	Qualität der IR	Earnings Management	Forschungsergebnisse bezüglich Interner Revision	Limitationen
Beasley, Carcello, Hermanson und Lapides (2000)	N=300; n=200 Unternehmen im Zeitraum 1987-1997	Deskriptive Statistik; Analyse	Analyse von mutmaßlichen Fällen von Bilanzbetrug in den Branchen Gesundheit, Finanzen und Technologie und der Existenz von IR sowie der Anzahl an Zusammenkünften des Prüfungsausschusses	Nur Analyse des **Vorhandenseins** gegenstand	Nicht Untersuchungs-	(1) In allen Unternehmen der untersuchten Branchen, mit Fällen von mutmaßlichem Bilanzbetrug, gab es im Vergleich zu Unternehmen ohne Bilanzbetrug häufig keine Interne Revision (2) In den Branchen Gesundheitswesen und Technologie fanden bei Unternehmen mit mutmaßlichem Bilanzbetrugsfällen eine geringere Anzahl an Sitzungen des Prüfungsausschusses statt	Keine Verwendung eines Diskriminanzmodells zur Unterscheidung der Stichprobe, mit und ohne fraudulente Handlungen; ein verfeinertes Modell zur Vorhersage von Betrugsfällen zu empfehlen
Davidson et al. (2005)	n=434 Australische Unternehmen im Jahr 2000	Empirische Forschung mit Archivdaten; Querschnittstudie	Zusammenhang zwischen Corporate Governance Strukturen und Earnings Management (Prüfungsausschuss & IR) und der Existenz einer IR	Nur Analyse des **Vorhandenseins**	(1) Absoluter Wert der Discretionary Accruals auf Basis des Modified Jones Model (2) Earnings Benchmark Test	Kein Zusammenhang zwischen dem Vorhandensein oder Nichtvorhandensein einer IR und einem niedrigeren Level an Earnings Management	Nur mittel- bis große kapitalmarktorientierte Unternehmen in der Stichprobe, daher Einfluss der Qualität der IR eher untersuchenswert als das Vorhandensein oder nicht Vorhandensein einer IR
Sierra García, Barbadillo und Orta Pérez (2012)	n=432 Beobachtungen; 108 an der Madrid Stock Exchange gelistete, spanische Unterter	Empirische Forschung mit Archivdaten; deskriptive Statistik, Multivariate Regression	Zusammenhang zwischen Corporate Governance Mechanismen (Prüfungsausschuss & IR) und Earnings Management/Financial Reporting Quality	Nur Analyse des **Vorhandenseins**	Abnormal Accruals auf Basis des Modified Jones Model	(1) Eine größere Anzahl an Mitgliedern des Prüfungsausschusses reduziert die Wahrscheinlichkeit für Earnings Management (2) Eine höhere Anzahl an jährlichen Sitzungen des Prüfungsausschusses reduziert die Wahrscheinlichkeit für Earnings Management (3) Die Existenz einer Internen Revision reduziert die Wahrscheinlichkeit für Earnings Management (4) Die Existenz einer Internen Revision, gepaart mit einem wirksamen Prüfungsausschuss, reduziert die Wahrscheinlichkeit für Earnings Management	Kleiner Stichprobenumfang; Einfluss der Qualität der Internen Revision eher untersuchenswert als das Vorhandensein oder nicht Vorhandensein einer Internen Revision, ggf. andere Proxy-Variablen[a] zur Messung der Wirksamkeit des Prüfungsausschusses sinnvoll
Gebrayel, Jarrar, Salloum und Lefebvre (2018)	n=139 Beobachtungen von 71 am Muscat Stock Securities Market gehandelten Unternehmen	Empirische Forschung; Paneldatenanalyse; deskriptive Statistik, Multivariate Regression	Zusammenhang zwischen Corporate Governance Mechanismen (Prüfungsausschuss & IR) und Financial Reporting Quality	Nur Analyse des **Vorhandenseins** einer IR und der wechselseitigen Beziehung zum Prüfungsausschuss	(1) Accruals Quality auf Basis des Modifizierten Dechow and Dichey Models (2) Absoluter Wert der Discretionary Accruals auf Basis des Performance-Adjusted Modified Jones Model	(1) Höhe der Anzahl der jährlichen Prüfungsausschusssitzungen ist positiv korreliert mit der Rechnungslegungsqualität (2) Die Existenz einer Internen Revision ist positiv korreliert mit der Rechnungslegungsqualität	Mangelnde Datenverfügbarkeit, da bestimmte Informationen nicht öffentlich zugänglich sind, daher keine Auswertung des Einflusses des gegliederten IR und keine Auswertung der Agenda und Dauer von Prüfungsausschusssitzungen mit der IR; Eingeschränkte Überprüfbarkeit der Wirksamkeit des Prüfungsausschusses
Prawitt, Smith und Wood (2009)	n=528 Daten von 218 kapitalorientierten Unternehmen für die Jahre 2000-2005	Empirische Forschung mit Archivdaten; Fragebogen; Multivariate Regression; Sensitivitätsanalyse	Zusammenhang zwischen der **Qualität** der IR und Earnings Management	Komponenten des SAS No. 65 mittels Befragung der Chief Audit Exekutives	(1) Absoluter Wert der Abnormal Accruals auf Basis des Modified Jones Model (2) Tendenz von Unternehmen die Earnings-Vorhersagen von Bilanzanalysten zu treffen oder zu verfehlen	(1) Je höher die Qualität der IR, desto weniger Earnings Management wird betrieben (2) Unternehmen, die die von Analysten vorhergesagten Gewinnschwellen treffen bzw. übertreffen, betreiben mit einer höheren Wahrscheinlichkeit Earnings Management (3) Unternehmen mit einer qualitativ hochwertigen IR unterschreiten die von Analysten vorhergesagten Gewinnschwellen mit einer höheren Wahrscheinlichkeit	Keine randomisierte Stichprobe, daher limitierte Generalisierbarkeit; eher kleiner Stichprobenumfang; Potenzielle Endogenität der Qualität der IR betreffende Entscheidungen mit anderen Corporate Governance Entscheidungen; nur zwei Proxy-Variablen zur Untersuchung von Earnings Management

(Continued)

Table 1—continued

Studie	Stichprobe	Methodik	Untersuchungsgegenstand	Komponenten / Qualitätsdimensionen	Modell	Ergebnisse	Limitationen
Johl, Kaur Johl, Subramaniam und Cooper (2013)	n=128 durch CIA oder Controlling ausgefüllte Fragebögen von 64 an der Bursa Malaysia gelisteten Unternehmen in den Jahren 2009-2010	Empirische Forschung mit Archivdaten & Fragebögen; Deskriptive Statistik; Multivariate Regression; Sensitivitätsanalyse	Zusammenhang zwischen der **Qualität** der IR und der Rechnungslegungsqualität in dem Schwellenland Malaysia	Komponenten: Erfahrung, Unabhängigkeit, Qualitätssicherungsmaßnahmen, Prüfungsschwerpunkte, finanzielle Ausstattung	Abnormal Accruals auf Basis des Cross-Sectional Modified Jones Model	(1) Positive Korrelation zwischen der Qualität der IR und Abnormal Accruals, in Abhängigkeit der politischen Verbindungen und der potenziellen Ausgliederung der Tätigkeiten der IR (2) Politisch verbundene Unternehmen mit ausgegliederten IR, betreiben bei steigender Qualität der IR, mit einer höheren Wahrscheinlichkeit Earnings Management und verfügen demnach über eine geringere Rechnungslegungsqualität (3) Die finanzielle Ausstattung einer Internen Revision ist negativ korreliert mit Abnormal Accruals (4) Unabhängigkeit und Berufserfahrung sind schwach positiv mit Abnormal Accruals korreliert. (5) Die Qualität des Vorstandes und die Qualität der IR wirken substitutiv zusammen, um die Qualität der Rechnungslegung auf einem bestimmten Level zu halten	Kleiner Stichprobenumfang mit einer großen Anzahl an Kontrollvariablen, daher geschmälerte Aussagekraft; Mangelnde Signifikanz einzelner Qualitätsattribute der IR könne dem Forschungsdesign oder der Stichprobengröße geschuldet sein; ggf. Maß für Rechnungslegungsqualität nicht optimal
Ege (2015)	n=1.398 Geschäftsjahre von 617 Unternehmen, gegen die glaubwürdige Vorwürfe vorsätzlicher, rechtswidriger Handlungen durch die Geschäftsführung vorliegen; Zeitraum 2000-2009	Empirische Forschung mit proprietären Umfragedaten; Differenzen-von-Differenzen-Ansatz; Faktorenanalyse	Zusammenhang zwischen der **Qualität** der IR und der Wahrscheinlichkeit von Fehlverhalten (Bilanzmanipulation, Bestechung, irreführende Offenlegungspraktiken) durch die Geschäftsführung	(1) Komponenten: Erfahrung, Zertifikation, Schulungsstunden, Führungskräftetraining, Beziehung zum Prüfungsausschuss, Größe (2) Kompetenz (3) Objektivität	Nicht Untersuchungsgegenstand	(1) Unternehmen, gegen die glaubwürdige Vorwürfe vorsätzlicher, rechtswidriger Handlungen durch die Geschäftsführung vorliegen, verfügen in den entsprechenden Jahren des Fehlverhaltens über eine qualitativ niedrige und vergleichsweise wenig kompetente Interne Revision (2) Je höher die Qualität der IR ist, desto geringer ist die Wahrscheinlichkeit für Fehlverhalten durch die Geschäftsführung. (3) Negativen, signifikanter Zusammenhang zwischen der Kompetenz einer IR und der Wahrscheinlichkeit für Fehlverhalten durch die Geschäftsführung (4) Unternehmen steigern in den Folgejahren eines rechtswidrigen Vorfalls, die Qualität ihrer IR durch die Erhöhung der Schulungsstunden und die Neueinstellung zertifizierter Interner Revisoren	Fehleranfälligkeit von Fragebögen; Die Qualität der IR und die Wahrscheinlichkeit des Fehlverhaltens durch die Geschäftsführung sind gemeinsam bestimmt; Das Ergebnis, dass die Objektivität keinen signifikanten Zusammenhang mit der Wahrscheinlichkeit des Fehlverhaltens durch die Geschäftsführung hat, sollte mit Vorsicht interpretiert werden
Abbott, Daugherty und Parker Peters (2016)	n=189 durch CIAs ausgefüllte Fragebögen von Fortune 1000 Unternehmen im Jahr 2009	Empirische Forschung; Fragebögen; Multivariate Regression; Zwei-Faktoren-Modell	Ermittlung der Determinanten einer effektiven IR und Untersuchung des Interaktionseffektes der Kompetenz und Unabhängigkeit von IR, als **Qualitätsmerkmal** und dessen Effekt auf die Rechnungslegungsqualität.	Interaktionseffekt von Unabhängigkeit und Kompetenz als separate Charakteristika	(1) Gewinnsteigernde und -minimierende Abnormal Accruals auf Basis des Performance-Adjusted Cross-Sectional Modified Jones Model -> Kothari Model (2) Tendenz von Unternehmen die Earnings-Vorhersagen von Bilanzanalysten zu treffen oder zu verfehlen	Das gemeinsame Vorhandensein von Kompetenz und Unabhängigkeit stellt eine notwendige Voraussetzung für die Wirksamkeit der Überwachung der Finanzberichterstattung durch die Interne Revision dar und schränkt insbesondere in Zusammenarbeit mit dem Prüfungsausschuss die Gewinnsteuerung sowohl nach oben als auch nach unten ein	Eingeschränkte Generalisierbarkeit aufgrund der Verwendung von Fragebögen und der Entwicklung der Weltwirtschaftskrise[b]; Potenzielle Endogenität[c]; Potenzielle Qualität der IR betreffende Entscheidungen mit anderen Corporate Governance Entscheidungen
Bajra und Cadez (2018)	n=127 europäische Unternehmen, die in den USA börsennotiert sind; Zeitraum 2000-2013	Empirische Forschung mit Archivdaten & Datenerhebung; Multivariate Regression; Sensitivitätsanalyse	Zusammenhang zwischen Corporate Governance Strukturen (**Qualität** der IR & Qualität des Vorstandes) und Earnings Management	Qualitätsdimensionen: Existenz / Einführung IR, Expertise, Unabhängigkeit, Einbezug in die externe Jahresabschlussprüfung	(1) Total Discretionary Accruals auf Basis der Modified Jones Model (2) ROA-adjusted Discretionary Accruals auf Basis des Performance-Adjusted Cross-Sectional Modified Jones Model -> Kothari Model	(1) Die Qualität der Internen Revision ist negativ korreliert mit Earnings Management (2) Alle fünf Qualitätsdimensionen sind negativ korreliert mit Earnings Management (3) Die Qualität des Vorstandes ist negativ korreliert mit Earnings Management (4) Der Interaktionseffekt zwischen der Qualität der IR und der Qualität des Vorstandes ist positiv korreliert mit Earnings Management, der Effekt ist gegenüber dem direkten Einfluss der Qualität der IR und des Vorstandes jedoch sehr klein (5) Sowohl die Qualität der IR als auch die Qualität des Vorstandes steigt im Laufe der Zeit	Discretionary Accruals ist nur eine Proxy-Variable für Earnings Management; Erhöhte Wahrscheinlichkeit für Messfehler

Table 1—continued

Autor (Jahr)	Stichprobe	Methodik	Untersuchungsgegenstand	Messung / Anmerkung	Zentrale Ergebnisse	Kritische Würdigung
Ismael und Kamel (2021)	n=115 Fragebögen (Ausgefüllt von Leitern der IR britischer, an der LSE gehandelten Unternehmen)	Empirische Forschung zwischen der **Qualität** der IR und Earnings Management; Multivariate Regression; Sensitivitätsanalyse	Zusammenhang zwischen der **Qualität** der IR und verschiedenen Qualitätskomponenten des Earnings Management	Zusammengesetztes Abnormal Accruals Maß auf Basis des Modified Jones Model	(1) Ceteris paribus, betreiben Unternehmen mit einer qualitativ hochwertigen IR zu einer geringeren Wahrscheinlichkeit Earnings Management (2) Ceteris paribus, betreiben Unternehmen mit einer kompetenteren IR zu einer geringeren Wahrscheinlichkeit Earnings Management	Kleiner Stichprobenumfang; Reliabilität der Daten bezüglich IR Qualität aufgrund der Datengewinnung mittels Fragebögen eventuell eingeschränkt; Eingeschränkte Generalisierbarkeit aufgrund der Konzentration auf UK; keine Untersuchung von ausgegliederten IR
Uecker, Brief und Kinney Jr. (1981)	n=86 verwertbare Rückmeldungen von Managern	Experimentelles Forschungsdesign mittels Postkorb-Fallstudie	Untersuchung der **Wahrnehmung** des persönlichen Auftretens Interner Revisoren und der Aggressivität der Prüfungshandlungen auf die Wahrscheinlichkeit Bilanzmanipulation zu betreiben	Nur Analyse der **Wahrnehmung** einer IR	(1) Ablehnung der Hypothese, dass ein Interner Auditor, der als aktive und starke Persönlichkeit wahrgenommen wird und direkt an den Vorstand berichtet, mit einer höheren Wahrscheinlichkeit Unregelmäßigkeiten in der Rechnungslegung erkennt und aufdeckt (2) Ablehnung der Hypothese, dass eine höhere wahrgenommene Aggressivität der Prüfungstätigkeit das Auftreten von Unregelmäßigkeiten in der Rechnungslegung reduziert (3) Die Anwesenheit und Wahrnehmung einer IR hält die Geschäftsführung nicht davon ab, Bilanzmanipulation zu betreiben oder betrügerische Handlungen in Betracht zu ziehen	Kleiner Stichprobenumfang; Mangelnde interne Validität zwar nicht bestätigt, aber dennoch Befürchtung, dass Probanden die Aufgabe nicht vollständig verstanden oder nicht gewissenhaft geantwortet haben; fehlende erkennbare Trends bei der Beantwortung der Fragen durch die Manager
Schneider und Wilner (1990)	n=264 Probanden (Großteil erfahrene Manager)	Experimentelles Forschungsdesign; randomisierter Antwort-Technik	Untersuchung der abschreckenden Wirkung auf Unregelmäßigkeiten in der Finanzberichterstattung in Abhängigkeit der **Wahrnehmung** der internen und externen Rechnungslegungsprüfung durch Manager	Nicht Untersuchungsgegenstand	(1) Bezüglich der abschreckenden Wirkung liegen keine Unterschiede zwischen der Wahrnehmung der Tätigkeit der IR oder des externen Jahresabschlussprüfers vor (2) IR hat einen eindeutig abschreckenden Effekt bei folgenden Merkmalsausprägungen: wesentlicher Geldbetrag, Überbewertung von Vermögensgegenständen, eindeutige Verstöße gegen US-GAAP, geringer Anreiz zur Bilanzmanipulation	Nur hypothetische finanzielle Anreize; trotz anonymisierter Umfrage besteht das Risiko des Phänomens der Sozialen Erwünschtheit bei der Beantwortung der Fragen; subjektive Antworten; keine Analyse der kognitiven Verarbeitung der gegebenen Informationen durch die Probanden
Prawitt, Sharp und Wood (2012)	n=334 Beobachtungen (von 159 Unternehmen aus 37 Ländern); Zeitraum 2000-2002	Empirische Forschung mit proprietären Daten; Multivariate Regression	Untersuchung des Zusammenhangs zwischen **ausgegliederten IR** und Rechnungslegungsrisiken	Nur Untersuchung der Rechnungslegungsrisiken (Messung gemäß dem Konzept von Audit Integrity[d])	(1) Die Jahresabschlussqualität von Unternehmen steigt, wenn die Tätigkeiten der Internen Revision zumindest partiell von externen Jahresabschlussprüfern erbracht wird (2) Die Existenz einer Internen Revision, egal ob unternehmensintern oder ausgegliedert, führt zu einer Reduktion von Rechnungslegungsrisiken	Validität der Forschungsergebnisse in Abhängigkeit der Validität der Messung der Rechnungslegungsrisiken gemäß Audit Integrity; Ergebnisse ggf. durch alternative Mechanismen gesteuert (wie beschrieben durch Kinney Jr., Palmrose und Scholz (2004)[e]
Baatwah, Omer und Al-jaaidi (2021)	n=928 Beobachtungen (Paneldaten von 928 omanischen Unternehmen); Zeitraum 2005-2017	Empirische Forschung; Pooled Panelregression; descriptive & univariate Analyse	Zusammenhang zwischen der Expertise einer **ausgegliederten IR** und Real Earnings Management	Real Earnings Management: (1) Abnormal Sales (2) Abnormal Production (3) Abnormal Discretionary Expenses; (1) Branchenexpertise (2) Firmenspezifische Expertise	(1) Die Branchenexpertise einer ausgegliederten IR hängt mit der Reduktion von echtem Earnings Management zusammen (2) Die firmenspezifische Expertise einer ausgegliederten IR steht im Zusammenhang mit niedrigerem echtem Earnings Management	Trotz großem Stichprobenumfang begrenzte Generalisierbarkeit; mangelnde Datenverfügbarkeit über Honorare ausgegliederter IR; Endogenität kann nicht ausgeschlossen werden

IR=Interne Revision. N=Gesamtzahl der Grundgesamtheit. n=Stichprobenumfang.
Quelle: eigene Darstellung.[f]

[a] Proxy-Variablen kommen zum Einsatz, um nicht direkt beobachtbare bzw. messbare Variablen zu ersetzen, vorausgesetzt, dass diese miteinander hoch korreliert sind. Vgl. Albers, Klapper, Konradt, Walter und Wolf (2007, S. 237).
[b] Weiterführend siehe Arnold (2009, S. 803-809).
[c] Korrelation der unabhängigen Variable und der Störgröße.
[d] Audit Integrity verwendet proprietäre Daten zur Messung des „overall risk of potentially fraudulent or misleading financial reporting". Prawitt et al. (2012, S. 1110). Weiterführend siehe Audit Integrity (2005, S. 2-12).
[e] Weiterführend siehe Kinney Jr et al. (2004, S. 561-586).

Literatur

Abbott, L. J., Daugherty, B., Parker, S. & Peters, G. F. (2016). Internal Audit Quality and Financial Reporting Quality: The Joint Importance of Independence and Competence. *Journal of Accounting Research*, *54* (1), 3–40.

Abbott, L. J., Parker, S. & Peters, G. F. (2004). Audit Committee Characteristics and Restatements. *Auditing: A Journal of Practice & Theory*, *23* (1), 69–87.

ACFE. (2020). *2020 Report of the Nations.* https://acfepublic.s3-us -west-2.amazonaws.com/2020-Report-to-the-Nations.pdf.

AICPA, IIA & ACFE. (2008). *Managing the Business Risk of Fraud: A Practical Guide.*

Albers, S., Klapper, D., Konradt, U., Walter, A. & Wolf, J. (2007). *Methodik der empirischen Forschung.*

Alzeban, A. (2014). The Impact of Culture on the Quality of Internal Audit: An Empirical Study. *Journal of Accounting, Auditing & Finance*, *30* (1), 57–77.

Amling, T. K. & Bantleon, U. (2007). *Handbuch der Internen Revision: Grundlagen, Standards, Berufsstand, 1. Aufl.*

Anderson, D. J. & Eubanks, G. (2015). *Leveraging COSO Across the Three Lines of Defense.* coso-2015-3lod.pdf. (Zuletzt abgerufen: 26.08.2021)

Arena, M. & Azzone, G. (2009). Identifying Organizational Drivers of Internal Audit Effectiveness. *International Journal of Auditing*, *13* (1), 43–60.

Arnold, P. J. (2009). Global Financial Crisis – The Challenge to Accounting Research. *Accounting, Organizations and Society*, *34* (6), 803–809.

Audit Integrity. (2005). The Audit Integrity Bankruptcy Risk Model – An Enhanced Approach to Predicting Bankruptcies and Identifying Companies in Severe Financial Distress. *The Audit Integrity AGR Model.*

Baatwah, S. R., Al-Ebel, A. M. & Amrah, M. R. (2019). Is the Type of Outsourced Internal Audit Function Provider Associated With Audit Efficiency? *International Journal of Auditing*, *23* (3), 424–443.

Baatwah, S. R., Omer, W. K. & Aljaaidi, K. S. (2021). Outsourced Internal Audit Function and Real Earnings Management: The Role of Industry and Firm Expertise of External Providers. *International Journal of Auditing*, *25* (1), 206–232.

Bajra, U. & Cadez, S. (2018). The Impact of Corporate Governance Quality on Earnings Management: Evidence from European Companies Cross-listed in the US. *Australian Accounting Review*, *28* (2), 152–166.

Bantleon, U., d'Arcy, A., Eulerich, M., Hucke, A., Knoll, M., Köhler, A. & Pedell, B. (2017). Das Three-Lines-of-Defence-Modell: ein Beitrag zu einer besseren Corporate Governance? Entstehung und Rezeption durch Standardsetzer und Regulatoren. *Die Wirtschaftsprüfung*, *28* (2), 152–688.

Bantleon, U., d'Arcy, A., Eulerich, M., Hucke, A., Pedell, B. & Ratzinger-Sakel, N. V. S. (2021). Coordination Challenges in Implementing the Three Lines of Defense Model. *International Journal of Auditing*, *25* (1), 59–74.

Barr-Pulliam, D. (2016). *Engaging Third Parties for Internal Audit Activities – Strategies for Successful Relationships.* IIARFCBOKEngagingThirdPartiesForIAActivitiesJan2016_0 .pdf. (Zuletzt abgerufen: 02.10.2021)

Beasley, M., Carcello, J., Hermanson, D. & Lapides, P. (2000). Fraudulent Financial Reporting: Consideration of Industry Traits and Corporate Governance Mechanisms. *Accounting Horizons*, *4* (4), 441–454.

Becker, C. L., DeFond, M. L., Jiambalvo, J. & Subramanyam, K. R. (1998). The Effect of Audit Quality on Earnings Management. *Contemporary Accounting Research*, *15* (1), 1–24.

Behrend, J. & Eulerich, M. (2019). The Evolution of Internal Audit Research: A Bibliometric Analysis of Published Documents (1926-2016). *Accounting History Review*, *29* (1), 103–139.

Berwanger, J. & Kullmann, S. (2012). *Interne Revision, 2. Aufl.*

Bungartz, O. (2021). DIIR Revisionsstandard Nr. 3 – Qualitätsmanagement in der Internen Revision. *IPPF Nr. 1300.*

Caplan, D. H. & Kirschenheiter, M. (2000). Outsourcing and Audit Risk for Internal Audit Services. *Contemporary Accounting Research*, *17* (3), 387–428.

Carey, P., Subramaniam, N. & Ching, K. C. W. (2006). Internal Audit Outsourcing in Australia. *Accounting and Finance*, *46* (1), 11-30.

Chambers, A. (1995). Whistleblowing and the Internal Auditor. *Business Ethics*, *4* (4), 192–198.

Cizmas, A., Feder, E., Maticiuc, M. D. & Vlad-Anghel, S. (2020). Team Management, Diversity, and Performance as Key Influencing Factors of Organizational Sustainable Performance. *Sustainability*, *12* (18), 1-33.

Coenenberg, A. G., Haller, A. & Schultze, W. (2018). *Jahresabschluss und Jahresabschlussanalyse: betriebswirtschaftliche, handelsrechtliche, steuerrechtliche und internationale Grundlagen - HGB, IAS/IFRS, US-GAAP, DRS, 25. Aufl.*

Copley, P. & Doucet, M. (1993). Auditor Tenure, Fixed Fee Contracts, and the Supply of Standard Single Audits. *Public Budgeting & Finance*, *13* (3), 23–35.

Coram, P, Ferguson, C. & Moroney, R. (2008). Internal Audit, Alternative Internal Audit Structures and the Level of Misappropriation of Assets Fraud. *Accounting and Finance*, *48*, 543–559.

d'Arcy, A. & Hoos, F. (2012). Welche Faktoren beeinflussen Unabhängigkeit und Objektivität der Internen Revision? *Zeitschrift Interne Revision* (3), 124–131.

Daugherty, B. E., Dickins, D., Hatfield, R. C. & Higgs, J. L. (2012). An Examination of Partner Perceptions of Partner Rotation: Direct and Indirect Consequences to Audit Quality. *Auditing: A Journal of Practice & Theory*, *31* (1), 97–114.

Davidson, R., Goodwin-Stewart, J. & Kent, P. (2005). Internal Governance Structures and Earnings Management. *Accounting and Finance*, *45*, 241–267.

DeAngelo, L. E. (1981). Auditor Size and Audit Quality. *Journal of Accounting and Economics*, *3* (3), 183–199.

Dechow, P. M. & Dichev, I. D. (2002). The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors. *The Accounting Review*, *77*, 35–59.

Dechow, P. M., Ge, W. & Schrand, C. M. (2010). Understanding Earnings Quality: A Review of the Proxies, their Determinants and their Consequences. *Journal of Accounting and Economics*, *50* (2), 344–401.

Dechow, P. M. & Schrand, C. M. (2004). *Earnings Quality, Research Foundation of CFA Institute.*

Dechow, P. M., Sloan, R. G. & Sweeny, A. P. (1995). Detecting Earnings Management. *The Accounting Review*, *70* (2), 193–225.

Deloitte. (2017). *Chief Audit Executives.* us-deloitte-chief-audit -executives-trends.pdf. (Zuletzt abgerufen: 20.09.2021)

DIIR (Hrsg.). (2011). Stellungname des DIIR – Deutsches Institut für Interne Revision e.V. zum EU-Grünbuch "Europäischer Corporate Governance Rahmen". *Deutsches Institut für Interne Revision e. V..*

DIIR (Hrsg.). (2020a). Das Drei-Linien-Modell des IIA – Eine Aktualisierung der Three Lines of Defense. *Deutsches Institut für Interne Revision e. V..*

DIIR (Hrsg.). (2020b). Internal Audit Competency Framework. *Deutsches Institut für Interne Revision e. V..*

Dworkin, T. M. & Baucus, M. (1998). Internal vs. External Whistleblowers: A Comparison of Whistleblowering Processes. *Journal of Business Ethics*, *17*, 1281–1298.

Dyck, A., Morse, A. & Zingales, L. (2010). Who Blows the Whistle on Corporate Fraud? *The Journal of Finance*, *65* (6), 2213–2253.

Dyreng, S. D., Hanlon, M. & Maydew, E. L. (2012). Where do Firms Manage Earnings? *Review of Accounting Studies*, *17*, 649–687.

ECIIA and FERMA. (2011). *URL falsch, BITTE NACHFRAGEN: Guidance on the 8th EU Company Law Directive – Article 41.* (Zuletzt abgerufen: 15.08.2021)

ECIIA and FERMA. (2014). *Audit and Risk Committees – News from EU Legislation and Best Practices.* eciia_ferma_brochure_v8.pdf. (Zuletzt abgerufen: 25.08.2021)

Ege, M. S. (2015). Does Internal Audit Function Quality Deter Management Misconduct? *The Accounting Review*, *90* (2), 495–527.

Eighme, J. E. & Cashell, J. D. (2002). Internal Auditors' Roles in Overcoming the Financial Reporting Crisis. *International Auditing*, *17*, 3–10.

Eulerich, M. & Fligge, B. (2020). Aggressive Berichterstattung in deutschen Unternehmen – Der Einfluss der Mitbestimmung auf die Ausnutzung von Bilanzierungs- und Steuergestaltungsspielräumen. In *Mitbestimmungsreport Nr. 62.* Hans-Böckler-Stiftung.

Eulerich, M. & Ratzinger-Sakel, N. V. S. (2018). The Effects of Cultural Dimensions on the Internal Audit Function – A Worldwide Comparison

of Internal Audit Characteristics. *Corporate Ownership & Control*, *15* (3), 217–229.

Eulerich, M. & van Uum, C. (2014). Die Interne Revision als Management Training Ground – Chancen und Herausforderungen bei der Weiterentwicklung von Führungskräften. *Zeitschrift Interne Revision*, *49* (3), 132–138.

Eulerich, M. & Velte, P. (2013). Theoretische Fundierung der Internen Revision – Zur Ökonomischen Notwendigkeit einer Internen Revision. *Zeitschrift Interne Revision*, *48* (3), 146–151.

EY. (2015). *Fraud and Corruption – The Easy Option for Growth – Europe, Middle East, India and Africa Fraud Survey 2015 (verfügbar unter: EMEIA-FRAUD-SURVEY-2015-FINAL-060515.pdf (eyfinancial-servicesthoughtgallery.ie))*. (Zuletzt abgerufen: 27.09.2021)

EY. (2016). *Whistleblowing – a Tick in the Compliance Box or a Vital Source of Business Intelli-gence?* . `https://acfe.de/wp-content/uploads/0067f20160429_009_Broschuere_2016_EY_Whistleblowing.pdf`. (Zuletzt abgerufen: 02.10.2021)

Florea, R. & Florea, R. (2009). Internal Audit and Risk Management - ISO 3100 and ERM Approaches. *Economy Transdisciplinary Cognition*, *19* (1), 72–77.

Francis, J. R. & Wang, D. (2008). The Joint Effect of Investor Protection and Big 4 Audits on Earnings Quality around the World. *Contemporary Accounting Research*, *25* (1), 157–191.

Freidank, C. C. & Velte, P. (2013). *Rechnungslegung und Rechnungslegungspolitik, 2. Aufl.*

Gebrayel, E., Jarrar, H., Salloum, C. & Lefebvre, Q. (2018). Effective association between audit committees and the internal audit function and its impact on financial reporting quality: empirical evidence from Omani listed firms. *International Journal on of Auditing*, *22* (2), 197–213.

Gegenfurtner, B. (2010). Kapitalmarkteffekte der Rechnungslegung unter Berücksichtigung von Eigentümerstrukturen - eine empirische Studie des deutschen Kapitalmarkts. *zugl. Dissertation Technische Universität München*.

Geiger, M. A. & Raghunandan, K. (2002). Auditor Tenure and Audit Reporting Failures. *Auditing: A Journal of Practice & Theory*, *21* (1), 67–78.

Glover, S. M., Prawitt, D. F. & Wood, D. A. (2008). Internal Audit Sourcing Arrangements and the External Auditor's Reliance Decision. *Contemporary Accounting Research*, *25* (1), 193–213.

Gramling, A. A., Maletta, M. J., Schneider, A. & Church, B. K. (2004). The Role of Internal Audit Function in Corporate Governance: A Synthesis of the Extant Internal Auditing Literature and Directions for Future Research. *Journal of Accounting Literature*, *23*, 194–244.

Hamacher, A. H. (2015). *Anforderungen an die Interne Revision: Corporate Governance und Internes Kontrollsystem (IKS), 1. Aufl.*

Hass, S., Abdolmohammadi, M. J. & Burnaby, P. (2006). The Americas Literature Review on Internal Auditing. *Managerial Auditing Journal*, *21* (8), 835–844.

Healy, P. M. & Wahlen, J. M. (1999). A Review of the Earnings Management Literature and its Implications for Standard Setting. *Accounting Horizons*, *13* (4), 365–383.

Heinzelmann, R. (2020). *SOX Compliance: Das sind die Anforderungen.* `https://www.haufe.de/compliance/recht-politik/sox-compliance_230132_457488.html`. (Zuletzt abgerufen: 22.05.2021)

Hofstede, G. H. (1980). *Culture's Consequences: International Differences in Work-Related Value.*

Holland, D. & Ramsay, A. (2003). Do Australian Companies Manage Earnings to Meet Simple Earnings Benchmarks? *Accounting and Finance*, *43*, 41–62.

Hölscher, L. & Rosenthal, J. (2007). *Leistungsmessung der Internen Revision (verfügbar unter: Arbeits86 (core.ac.uk))*. (Zuletzt abgerufen: 03.09.2021)

Hutchinson, M. & Zain, M. M. (2009). Internal Audit Quality, Audit Committee Independence, Growth Opportunities and Firm Performance. *Corporate Ownership & Control*, *7* (2), 50–65.

IIA. (2009). The Role of Internal Auditing in Enterprise-wide Risk Management. *Institute of Internal Auditors - IIA Position Paper*.

IIA. (2010). Effective Sizing of Internal Audit Departments. *Institute of Internal Auditors*.

IIA. (2013a). Internal Audit and Whistleblowing. *Institute of Internal Auditors*.

IIA. (2013b). The Three Lines of Defense in Effective Risk Management and Control. *Institute of Internal Auditors*.

Ismael, H. R. (2019). The Association Between Firm Characteristics and the Quality Characteristics of the Internal Audit Function in the UK – An Agency Perspective. *International Journal of Accounting, Auditing and Performance Evaluation*, *15* (2), 113–143.

Ismael, H. R. & Kamel, H. (2021). Internal Audit Quality and Earnings Management: evidence from the UK. *Managerial Auditing Journal*, *36* (7), 951–978.

James, K. L. (2003). The Effects of Internal Audit Structure on Perceived Financial Statement Fraud Prevention. *Accounting Horizons*, *17* (4), 315–327.

Johl, K. S., Kaur Johl, S., Subramaniam, N. & Cooper, B. (2013). Internal Audit Function, Board Quality and Financial Reporting Quality: Evidence from Malaysia. *Managerial Auditing Journal*, *28* (9), 780–814.

Jones, J. J. (1991). Earnings Management During Import Relief Investigations. *Journal of Accounting Research*, *29* (2), 193–228.

Kaplan, S. E. & Schultz, J. J. (2007). Intentions to Report Questionable Acts: An Examination of the Influence of Anonymous Reporting Channel, Internal Audit Quality, and Setting. *Journal of Business Ethics*, *71* (2), 109–124.

Keßler, S. (2015). Die Messbarkeit von Gewinnsteuerung mit Hilfe von Periodenabgrenzungen. In W. Ballwieser, C. Kuhn & D. Ordelheide (Hrsg.), *Betriebswirtschaftliche Studien, Rechnungs- und Finanzwesen, Organisation und Institution, Band 98*. Dissertation Ludwig-Maximilians-Universität München.

Kinney Jr., W. R., Palmrose, Z. V. & Scholz, S. (2004). Auditor Independence, Non-Audit Services, and Restatements: Was the U.S. Government Right? *Journal of Accounting Research*, *42* (3), 561–588.

Kirkpatrick, G. (2009). *The Corporate Governance Lessons from the Financial Crisis (verfügbar unter: Microsoft Word - The Corporate Governance Lessons from the Finan-cial Crisis.doc (oecd.org))*. (Zuletzt abgerufen: 25.08.2021)

KPMG. (2004). *Fraud Survey 2004 Report.*

Küting, K. & Weber, C. P. (2015). *Die Bilanzanalyse – Beurteilung von Abschlüssen nach HGB und IFRS, 11. Aufl.*

Leuz, C., Dhananjay, N. & Wysocki, P. D. (2003). Earnings Management and Investor Protection: An International Comparison. *Journal of Financial Economics*, *69*, 505–527.

Mallin, C. A. (2019). *Corporate Governance, 6. Aufl.*

Marten, K. U., Quick, R. & Ruhnke, K. (2015). *Wirtschaftsprüfung – Grundlagen des betriebswirtschaftlichen Prüfungswesens nach nationalen und internationalen Normen.*

Menicucci, E. (2020). *Earnings Quality – Definitions, Measures, and Financial Reporting, 1. Aufl.*

Mubako, G. (2019). Internal Audit Outsourcing: A Literature Synthesis and Future Directions. *Australian Accounting Review*, *29* (3), 532–545.

Nagy, A. L. (2005). Mandatory Audit Firm Turnover, Financial Reporting Quality, and Client Bargaining Power: The Case of Arthur Andersen. *Accounting Horizons*, *19* (2), 51–68.

Near, J. P. & Miceli, M. P. (1985). Whistle-Blowing: Myth and Reality. *Journal of Management*, *22* (3), 507–526.

Pfaff, D. & Ising, P. (2010). Earnings Management: Erscheinungsformen und Aufdeckungsmöglichkeiten. In G. Seicht (Hrsg.), *Jahrbuch für Controlling und Rechnungswesen* (S. 291–312).

Prawitt, D. F., Sharp, N. Y. & Wood, D. A. (2012). Internal audit outsourcing and the risk of misleading or fraudulent financial reporting: Did Sarbanes-Oxley get it wrong? *Contemporary Accounting Research*, *29* (4), 1109–1136.

Prawitt, D. F., Smith, J. L. & Wood, D. A. (2009). Internal Audit Quality and Earnings Management. *The Accounting Review*, *84* (4), 1255–1280.

Raghunandan, K., Rama, D. V. & Scarbrough, D. P. (1998). Accounting and Auditing Knowledge Level of Canadian Audit Committees – Some Empirical Evidence. *Journal of International Accounting, Auditing and Taxation*, *7* (2), 181–194.

Read, W. J. & Rama, D. V. (2003). Whistle-blowing to Internal Auditors. *Managerial Auditing Journal*, *18* (5), 354–362.

Ronen, J. & Yaari, V. (2008). *Earnings Management – Emerging Insights in Theory, Practice, and Research, 1. Aufl.*

Rothschild, J. & Miethe, T. D. (1999). Whistle-Blower Disclosures and Management Retaliation: The Battle to Control Information about Organization Corruption. *Work and Occupations*, *26* (1), 107–128.

Sarens, G. (2009). Internal Auditing Research: Where Are We Going? Editorial. *International Journal of Auditing*, *13* (1), 1–7.

Sarens, G. & Abdolmohammadi, M. J. (2010). *Cultural Dimension and Professionalism Uniformity of Internal Auditing Practice*.

Schemmel, A., Ruhmannseder, F. & Witzigmann, T. (2012). *Hinweisgebersysteme – Implementierung in Unternehmen*.

Schneider, A. (1985). Consensus Among Auditors in Evaluating the Internal Audit Function. *Accounting and Business Research*, *15*, 297–301.

Schneider, A. (2008). The Roles of Internal Audit in Complying With the Sarbanes-Oxley Act. *International Journal of Disclosure and Governance*, *6* (1), 69–79.

Schneider, A. & Wilner, N. (1990). A Test of Audit Deterrent to Financial Reporting Irregularities Using the Randomized Response Technique. *The Accounting Review*, *65* (3), 666–681.

Sierra García, L., Barbadillo, E. R. & Orta Pérez, M. (2012). Audit Committee and Internal Audit and the Quality of Earnings: Empirical Evidence from Spanish Companies. *Journal of Management & Governance*, *16*, 305–331.

St. Pierre, K. & Anderson, S. (1984). An Analysis of the Factors Associated With Lawsuits Against Public Accountants. *The Accounting Review*, *59* (2), 242–263.

Stubben, S. R. & Welch, K. T. (2020). Evidence on the Use and Efficacy of International Whistleblowing Systems. *Journal of Accounting Research*, *58* (2), 473–518.

Teets, W. R. (2002). Quality of Earnings: An Introduction to the Issues in Accounting Education. *Issues in Accounting Education*, *17* (4), 355–360.

The Blue Ribbon Committee. (1999). Report and Recommendations of the Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees. *The Business Lawyer*, *54* (3), 1067–1095.

Uecker, W. C., Brief, A. P. & Kinney Jr., W. R. (1981). Perception of the Internal and External Auditor as a Deterrent to Corporate Irregularities. *The Accounting Review*, *56* (3), 465–478.

Wagenhofer, A. & Dücker, H. (2007). Die Messung von "Earnings" Qualität. *Journal für Betriebswirtschaft/Management Review Quarterly*, *57* (3), 263–297.

Welge, M. K. & Eulerich, M. (2014). *Corporate-Governance-Management – Theorie und Praxis der guten Unternehmensführung*.

Wengert, H. & Schittenhelm, F. A. (2013). *Corporate Risk Management*.

Wirtschaftsprüferkammer (WPK) (Hrsg.). (2020). *Anwendung der International Standards on Auditing (ISA) in Deutschland (verfügbar unter: Anwendung der International Standards on Auditing (ISA) in Deutschland (wpk.de))*. (Zuletzt abgerufen: 18.08.2021)

Zehnder, M. & Brönnimann, T. (2005). Interne Revision als Vertrauensgut - Transaktionskostentheorie - ein Beitrag zur Erhöhung der Wertschöpfung der Internen Revision. *Der Schweizer Treuhänder*, *79* (6), 448–454.

Junior Management Science

Understanding the Impact of Future Social Self-Concepts on Newcomer Adjustment

Jonas Franz Buerkner

Technische Universität München

Abstract

The purpose of this study is to investigate the impacts of future social self-concepts on newcomer short-term adjustment. Employing a qualitative longitudinal design based on interviews, this work aims to uncover how the shape of identities before and three weeks after entry, as well as the identity processes between them, impact adjustment success. This is important because adjustment is a precursor for job outcomes, such as performance, satisfaction, and intentions to remain. In the first part, the relevance of identities for job outcomes is carved out and major theoretical contributions to identity and socialization are identified and presented. The thesis then discusses a fitting methodology for studying identity and describes key methodological choices. Three newcomers participated in the narrative-based interviews. The first interview was conducted shortly before the second interview three weeks after organizational entry. The interviews were audio-recorded, transcribed, and coded employing an abductive coding procedure. The results support the view that identity plays a key role in newcomer socialization and illustrate currently discussed identity processes. The complexity of self-concept phenomena involved in newcomer socialization calls for further research efforts.

Keywords: Newcomer socialization; Newcomer adjustment; Self-concept; Possible selves; Identity partnership.

1. Introduction

Imagining our future is a capability unique to the human species. From today's psychological point of view, imagination is our capacity to leave the present behind. Our capability to imagine the future 'as it could be' complements our ability to experience the present 'as it is'. Imagining oneself in the future marks a specific type of imagination: that of future 'possible selves'. This thesis examines future possible selves in the context of newcomer socialization in organizations.

From a theoretical point of view, two main research conversations matter. First, the way newcomers imagine themselves in the future organization is described by the 'future social self-concept'. The self-concept is the cognitive component of the self. It is "a knowledge representation that contains knowledge about us, including our beliefs about our personality traits, physical characteristics, abilities, values, goals, and roles, as well as the knowledge that we exist as individuals" (Stangor, Jhangiani, & Tarry, 2014, p. 21). These cognitive components of the current self are linked to those of the future self. They include fears, hopes, dreams and possibilities directed to ourselves in the future. The newcomer thinking "I could be a team leader in organization ABC" is in fact exploring their possible selves (Markus & Nurius, 1986). Because possible selves link self-concept and motivation, they

are critical to understanding how newcomer socialization efforts are incentivized (de Place & Brunot, 2019).

Second, newcomer socialization literature has conceived adjustment as the key factor impacting job outcomes, such as performance, satisfaction, and organizational commitment. Adjustment is defined as the newcomer's ability to transition toward becoming an organizational insider (Bauer, Bodner, Erdogan, Truxillo, & Tucker, 2007). Although research has progressed in both fields separately, knowledge about the links between the two is still in its infancy. Scholars seeking to fill this gap will contribute to understanding how newcomer future social self-concepts in the organization will impact their success on the job. Because most studies explore the impacts on newcomer mid-term adjustment, little is known about adjustment shortly after the newcomer's entry. However, short-term adjustment success is important for understanding longer-term follow-on consequences for the success of newcomer socialization.

Importantly, future selves often must undergo a 'reality check' once the newcomer begins integrating into the organization. 'Reality checking' will determine the degree of overlap between the future possible self before entry–i.e., how the newcomer had imagined him- or herself in the new job–and the self he or she experiences after having entered the orga-

nization. The self is constructed by several features, such as structure, complexity, and clarity. I suggest that these characteristics will individually affect adjustment as they are scrutinized in the process of 'reality checking'.

Furthermore, one relevant research trajectory in this realm is how the self-concept determines how effectively individual and organizational socialization efforts impact newcomer adjustment.

From a practical standpoint, this contribution is paramount, as workers continue to switch jobs frequently. US labor market data shows that individuals change jobs on average 12.3 times in their first 34 years at work. Interestingly, most of those changes occur in the first 6 years in the labor market (Bureau of Labor Statistics, 2019).[1] Hence, this work aims to provide guidance for businesses and newcomers alike targeted toward successful adjustment in the socialization process.

In essence, I aim to shed light on the interconnection of newcomer's imagination of themselves in the organization, and their adjustment to the future firm. The puzzle I am trying to illuminate is the following: *"What is the impact of future social self-concepts on newcomer short-term adjustment?"*.

2. Theory

2.1. Newcomer Socialization

2.1.1. The Core Model of Newcomer Adjustment

Entering a new organization is often accompanied by feelings of uncertainty and anxiety. Berger and Calabrese's (1975) 'Uncertainty Reduction Theory' (URT) is illustrative of the processes underlying first encounters among strangers– such as when a newcomer joins an organization. Uncertainty is defined as being faced with multiple possible behaviors of both the other person and oneself (Berger & Calabrese, 1975). According to URT, newcomers are faced with high levels of uncertainty when entering an organization. The authors theorize that the effort of the newcomer to reduce the level of uncertainty is the core driver of interactions with other members of the organization. Hence, the newcomer seeks to predict the others' likely actions to react appropriately. To effectively do so, the others' response alternatives need to be minimized to the most likely actions. Here is where the process of uncertainty reduction comes into play: to generate accurate predictions of the other's action, information is sought and consequently uncertainty reduced.

With ongoing interaction, typically the accuracy of predictions about the situations increases. Berger and Calabrese (1975) describe this way of reducing uncertainty as 'predictive'. Additionally, they identified 'explanatory' processes as a second form of uncertainty reduction–the individual's attempts to explain the other's behavior retroactively. Linking these insights back to the newcomer socialization process, both predictive and explanatory uncertainty reduction

are likely to be at play. A newcomer aiming to anticipate a coworker's contribution in a meeting to respond appropriately engages in predictive uncertainty reduction. Alternatively, retroactive uncertainty reduction may consist of the newcomer explaining to him- or herself after the meeting why the coworker acted the way he or she did.

While uncertainty reduction theory is conducive to understanding the dynamics of individual interaction situations, a higher-level framework is imperative to grasping newcomer socialization processes at large. Generally, "organizational socialization refers to the process by which newcomers make the transition from being organizational outsiders to being insiders" (Bauer et al., 2007). For assessing the success of the socialization process, theorists and practitioners alike rely on job outcome variables. These include the positive outcomes of newcomer performance, job satisfaction, organizational commitment, intentions to remain, and the negative outcome of intentions to quit. Accordingly, the focus has been put into researching the relationships of outcome variables and their causal constructs. These efforts have produced two main strings of research. First, organizational socialization tactics emphasize the role of the organization in facilitating newcomer adjustment. Second, newcomer proactivity or information seeking underscores the relevance of newcomer personality traits and behavior in supporting adjustment. Newcomer adjustment in turn connects those antecedents to the job outcome variables of interest (Bauer et al., 2007; Figure 1).

In their meta-analytic review, the authors conceptualize newcomer adjustment as a multifaceted construct consisting of the dimensions of (1) role clarity, (2) self-efficacy, and (3) social acceptance. This typology is grounded in theoretical insights of prior socialization research. Feldman (1981) first described adjustment as a threefold concept, including (1) role definition – "an implicit or explicit agreement with the work group on what tasks one is to perform and what the priorities and time allocation for those tasks is to be" (p. 380), (2) initiation to the task – "the extent to which the individual feels competent and treated as a full work partner" (p. 380), and (3) initiation to the group – "the extent to which the individual feels trusted by coworkers and accepted personally by them" (p. 380). One of the core tenets of this concept is that a newcomer may need to resolve multiple demands to facilitate adjustment in the new organization.

2.1.2. Newcomer Information Seeking

Newcomer information seeking is understood as being necessitated by the newcomer's perception of not receiving enough or inadequate information from the organization. URT suggests that if this happens, newcomers experience uncertainty. "This uncertainty is reflected in the levels of role ambiguity and role conflict which newcomers experience" (Miller & Jablin, 1991, p. 93). Langer, Blank, and Chanowitz (1978) opened an elucidating perspective on information seeking as interpersonal interaction leading to "mindless" or "mindful" behavior. They proposed that when the participants know the structure of communication due

---
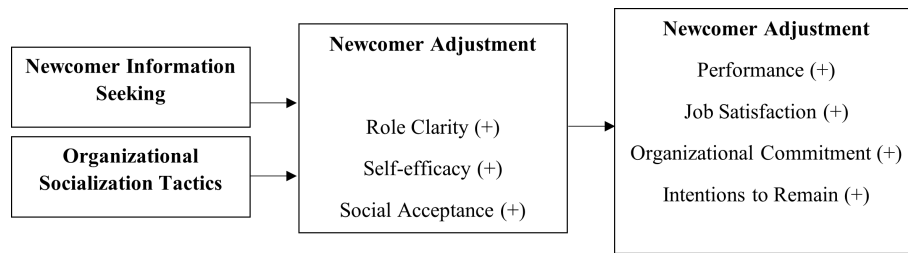
Figure 1: Current Model of Newcomer Adjustment

Note. From *Journal of Applied Psychology*, *92*(3), pp. 707–721, by Bauer et al. (2007)

to previous similar experiences, they may behave in a more "mindless" fashion. That is, they may overlook relevant details. On the contrary, when prior knowledge of the communication situation does not exist – such as if one of the participants is a newcomer to the organization – they are likely to behave in a more "mindful" fashion.

Such prior knowledge is often stored in "scripts". A script is understood as "a hypothesized cognitive structure that when activated organizes comprehension of event-based situations" (Abelson, 1981, p. 717). As such, a script offers an "expectation bundle" – it provides guidance toward making the 'right' inference from the situation (Abelson, 1981). Importantly, newcomers during organizational entry, usually lack scripts for a variety of organizational interactions – for instance, for weekly meetings. They must therefore engage in "mindful" processing and more intentional information seeking (Miller & Jablin, 1991). This implies that newcomers are likely to actively seek out the information most relevant to their organizational adjustment.

Miller & Jablin identify five main factors affecting newcomer information-seeking behavior: (1) perception of uncertainty, (2) perception of social cost, (3) type of information, (4) key sources of information, and (5) individual differences and contextual factors.

First, as touched upon, uncertainty plays a substantial role in understanding newcomer information seeking. To conclude, Berger and Calabrese (1975) posit in their URT seminal work, that information-seeking behavior increases with the level of uncertainty.

Second, social costs imply that individuals involved in interaction consciously weigh up rewards and costs of social exchange. According to Roloff (1981), social exchange includes individuals mutually and voluntarily transferring objects or activities. Miller and Jablin (1991) specify these exchanges to include "resources that may be cognitive, affective, or material in nature" (p. 95). Newcomers may perceive high social costs for information seeking, for instance, when they fear to ask their supervisor for information that they think incumbents would see as trivial. As a consequence, newcomers may abstain from information seeking if social costs are perceived as too high.

Third, the sources of information are relevant to information-seeking behavior. Supervisors and co-workers are considered the main sources of information for the newcomer. Hanser

and Muchinsky (1978) suggest that the sources of information differ in terms of their psychological distance from newcomers "Generally the informativeness of the sources increased as they moved from psychologically distant (i.e., the formal organization) to psychologically nearer sources (i.e., personal thoughts and feelings)" (p. 49). Interestingly, the authors discovered that supervisors are rated as psychologically closer, and thus more informative sources of information to newcomers.

Fourth, the content of information, i.e., the type of information sought, is an important factor affecting information seeking. Miller and Jablin (1991) develop a typology of information types, condensing previous typologies into three categories: (1) referent, (2) appraisal, and (3) relational information. Referent information refers to what is required to function on the job. Appraisal information is sought to assess the degree of functioning successfully on the job. Finally, relational information tells the newcomer about the nature of relationships with others. It appears logical to suspect that depending on which type of information is required, the newcomer may approach the interaction situation differently.

Finally, individual differences and contextual factors may also impact how newcomers seek information. Miller and Jablin (1991) posit that newcomers' self-esteem, tolerance for ambiguity, cognitive complexity, self-efficacy, and prior experience may direct their information seeking.

Morrison (1993) offers another useful approach addressing the impacts of information seeking on socialization. She proposes that the socialization process consists of four newcomer's tasks: (1) task mastery, (2) role clarification, (3) acculturation – "gaining an understanding of their organization's culture" (p. 174), and (4) social integration. Importantly, the author theorizes that different types of information will differentially contribute to the achievement of socialization tasks. First, the findings corroborate the impact of information-seeking behavior on three dimensions of newcomer socialization tasks: task mastery, role clarity, and social integration. Second, the results indeed evidence differential influences of information types. Whereas task mastery was positively correlated to seeking technical information, role clarity was associated with referent information and performance feedback. Social integration was related to seeking normative information – insights on "the behaviors and attitudes that [the] firm values and expects" (p. 183). Interest-

ingly, the author found no influence of information seeking on acculturation. In sum, these results support the notion that information seeking – as captured in the described socialization model Bauer et al. (2007) – critically affects newcomer adjustment. Furthermore, it adds the perspective that information seeking is not universally effective toward adjustment. Rather, newcomers should expend efforts directed to seeking specific information to adjust successfully.

To conclude, it is noteworthy that, while current research acknowledges that newcomer individual differences play a role in information seeking and adjustment, the precise role of the self-concept remains largely unexplored. I suggest that investigating the situational-perceptual factors affecting adjustment (Miller & Jablin, 1991) provides only a superficial view. Instead, I argue that the latent construct of the self-concept will reveal a more holistic perspective of the influences on information seeking. As a consequence, a gap in research exists around the impacts of subcategories of the self-concept on newcomer information seeking. Open questions invite exploratory research along the following lines: 1) "How does the self-concept influence when newcomers seek information?", 2) "What is the impact of the self-concept on the newcomer's choice of information source?", and 3) "How does the self-concept direct which specific type of information is sought?". Exploring the links between self-concept and information seeking will benefit the conversation by opening new directions for future research to embark on.

2.1.3. Organizational Socialization Tactics

Next to the newcomer, the organization plays a role in facilitating newcomer adjustment. The organizational efforts targeted at integrating newcomers into the organization are summarized under the concept of organizational socialization tactics. "Socialization tactics refer to organization-driven or more or less formalized means of socializing individuals" (Ashforth, Sluss, & Saks, 2007, p. 448). Early socialization researchers have coined the term "people processing" to refer to these strategies of newcomer socialization (Van Maanen, 1978). Under this terminology, contrary to research stressing the proactive role of the newcomer, the novice assumes a relatively passive role.

Socialization scholars have modeled tactics in a variety of ways. Van Maanen and Schein (1979) inductively derived a model of six organizational socialization tactics in bipolar dimensions (Figure 2). They argue that the tactics organizations employ range somewhere on a continuum on any of the six dimensions. Furthermore, the tactics are not considered mutually exclusive. Rather, a particular organizational tactic can be understood as the sum of scores on each of the dimensions. In their argument, the chosen tactic influences the newcomer's role orientation.

Later research on socialization tactics has largely built upon the work by Van Maanen and Schein (1979). Jones (1986) argued that the polar outcomes on the six dimensions can be categorized into two major types of socialization tactics: He termed collective, formal, sequential, fixed, serial, and investiture approaches "institutionalized tactics". At the

other end of the scale, individual, informal, random, variable, disjunctive, and divestiture, form "individualized tactics". According to Jones, the two prototypical tactics differ in terms of how the newcomer takes on her role in the organization. While institutionalized tactics "encourage newcomers to passively accept preset roles and thus maintain the status quo" (p. 150), individualized tactics "encourage newcomers to question the status quo and develop their own approach to their roles" (Ashforth & Saks, 1996, p. 150).

Importantly, Ashforth and Saks modeled person change as a key indicator of newcomer adjustment. However, they noted that because the self is a deeply rooted concept, "socialization primarily influences what Schein (1971) referred to as the more labile' self" (p. 153). Thus, an influence on the self is viewed as possible mainly on a surface level.

Crucially, while research has attempted to investigate the impacts of socialization tactics on the newcomer's self, it has largely neglected relationships in the opposite direction. Hence, researchers would do well in exploring how the self-concept directs how a particular socialization tactic is received by the newcomer. Such efforts would contribute to exploring precisely which factors of the self-concept influence whether a chosen socialization tactic promotes newcomer adjustment. Furthermore, insights are needed on the processes by which these impacts play out.

In another vein, scholars have conceived of tactics as antecedents of newcomer information seeking (Mignerey, Gorden, & Rubin, 1995). They argue that "the type of tactic chosen by the organization to socialize the newcomer determines the extent to which information will be made available to the newcomer" (p. 58) and thus the newcomer's information-seeking behavior. Surprisingly, scholars have not considered the self-concept in this realm. It remains yet to be explored how the newcomer's self-concept influences when and how the newcomer seeks information in response to organizational socialization tactics.

2.2. Self-Concept and Identities
2.2.1. The Self-Concept

The self-concept is of central importance to a variety of disciplines as evidenced by the vast body of work, not only in psychology but also in neighboring disciplines such as sociology and cultural studies. Possessing a self-concept and an identity is critical to humans. It answers the fundamental question "Who am I?" and thus provides a sense of direction to human life at large. "A consistent sense of self" (Selenko et al., 2018, p. 5) is needed for humans to function and effectively control the environment. The self-concept is the cognitive representation of the self that stores knowledge about us, "including our beliefs about our personality traits, physical characteristics, abilities, values, goals and roles, as well as the knowledge that we exist as individuals" (Stangor et al., 2014, pp. 109-110).

The self-concept is unique and private; it is different from other selves and no one else can completely grasp how it looks like (Oyserman, 2001). "Improving oneself, knowing

Institutionalized	Individualized
Collective: Newcomers go through a common set of experiences designed to produce standardized responses to situations	Individual: Each newcomer receives unique training in isolation from others.
Formal: Newcomers are segregated from other organizational members and put through experiences tailored to newcomers.	Informal: Newcomers receive on-the-job training to learn their roles.
Sequential: Newcomers are given a clear sequence of experiences or stages they will go through	Random: The sequence of stages isn't communicated in advance.
Fixed: Newcomers are given a fixed timetable about when they will move through stages	Variable: The timing of role transitions is variable.
Serial: Newcomers observe and get training from experienced role models, who give newcomers a clear view of the experiences they will encounter in the organization	Disjunctive: Newcomers must develop their own definition of the situation and do not have more senior people to observe.
Investiture: Newcomer receive positive feedback confirming their prior identity	Divestiture: Newcomers receive negative feedback expressing organizational disapproval of their prior identity.

Figure 2: *Organizational Socialization Tactics*

Note. From *Research in Organizational Behavior*, pp. 209–264, by Van Maanen and Schein (1979) and *Academy of Management Journal*, *29*(2), pp. 262–279, by Jones (1986)

oneself, discovering oneself, creating oneself anew, expressing oneself, taking charge of one's self, being happy with oneself, being ashamed of oneself, are all essential self-projects, central to our understanding of what self-concept and identity are and how they work" (Oyserman, 2001, p. 499). Thus, while the self-concept is at core a cognitive concept, it is critically involved in organizing behavior, which can–but not must–in turn, be directed at the self.

Investigating how the self-concept is constructed is critical to understanding its' workings and the implications that follow from it. Importantly, the self-concept does not exist in isolation. Rather, as posited in Social Identity Theory (SIT) (Tajfel & Turner, 1979) individuals define their identities based on their social group memberships. While this central notion of SIT is largely accepted by scholars, a debate has flourished around the precise role of social identities within the framework of individual-level identities. Researchers have historically put differing emphasis on social identity as a means of purely intrapersonal differentiation vs. as a means of extending the individual-level self (Brewer, 1991).

Turner and Oakes (1986) note that "the self-concept exists at different levels of abstraction" (p. 242), hence arguing for the latter position. Similarly, following Brewer's (1991) conception, the social identities entail categorizing the self into social units where it becomes depersonalized. The author illustrates this approach through a schematic model: a central point representing personal identity is surrounded by concentric circles representing social identities. Going from inward circles to further outward ones, the respective social identities become tied to higher-order frames of reference. One possible application to the context of newcomer socialization could be the following example: A male biology teacher has recently joined a new school. At the center is his identity as an individual teacher. Here, he considers mainly how he as a teacher is different from his fellow colleagues in terms of his teaching experience, approaches and accomplishments, etc. The first concentric circle may represent his social identity as a teacher within his school's department of science. The frame of reference is set by the department, and

hence he considers how he is similar to other biology teachers in his department and distinct from colleagues in physics and chemistry. Higher-order social identities could, for instance, include the school- or city-level and further extend to identifications with the nationwide teaching body or worldwide educators. I suggest that the outcomes of comparison and follow-on behaviors may differ depending on the respective level of self-definition that is currently activated. Hence, this case underlines the importance of considering different levels of social identity when evaluating newcomer socialization outcomes.

The existence of a multitude of theories on the self requires making choices on the type of self-concept approach to utilize. A useful self-concept theory should 1) provide direction for linking self-concepts to socialization outcomes in line with the research question, 2) depict the time dimension of different self-concepts in the present and future, and 3) allow depiction of the social dimension of the self as the newcomer enters into an organization. I argue that an intrapersonal identity network approach (Ramarajan, 2014) integrating the notion of possible selves (Markus & Nurius, 1986) and social selves (Brewer, 1991) will be best capable to serve these purposes.

2.2.2. A Multiple Identities Framework

Building on his extensive review of multiple identity research, Ramarajan (2014) provides a holistic framework of multiple identities following an intrapersonal network approach. He argues that "because preliminary evidence suggests that multiple identities shape important outcomes in organizations" (p. 590) it is important to look at those multiple identities instead of only singular or pairwise considerations. Selenko et al. (2018) explain that multiple identities are not in contrast with feeling as a single person. It is widely accepted by identity scholars that people can feel as one while possessing multiple identities (Ramarajan, 2014; Selenko et al., 2018). Because newcomers' multiple identities are interlinked, we need to understand not just their singular identities but also other identities and their interconnectedness. This means that to understand when certain job-relevant identities are enacted we must reveal the person's other identities and investigate the links between them. Therefore, a multiple identity framework appears most suitable in the context of newcomer socialization.

To disentangle self-concept and identities, I adopt the view of the author in defining the self-concept as the broad term to include all identities of a person, while conceptualizing identities as the more concrete term: Identities are thus more specific targets organized under the umbrella of the self-concept. These multiple identities can then be mapped in a network representation with nodes and arcs connecting the nodes. Nodes denote identities and may differ in number, importance, and meaning. The connections between them show relationships, for instance, that two identities enhance each other, are in conflict, or overlap and in which relation across time they stand, etc.

Imagine a female nurse who defines herself as a newcomer to nursing, a woman, a mother, a soccer player, and a dependable teammate. She also has a future identity as a caring nurse who takes on responsibility for her team and has an open ear for her patients. One could now map those identities in a network fashion and investigate the links between the nodes. First, you could imagine different relationships between the identities. The dependable teammate identity might have enhancing links with her soccer player, as well as her current and future nurse identities because team skills help her on the job and in the soccer club, and those, in turn, provide opportunities to improve these skills. Due to the demanding work environment with irregular working hours including night shifts she might lack time for her family. Hence, because the expression of her nurse-identity limits the extent to which she can express her mother-identity (she cannot bring her small children to work) these identities are conflicting. Moreover, her now newcomer-self and future nurse-self will be related temporally. Whereas some of the aspects or meanings of her current self are likely to be included in her future self, some aspects may be rather aspirational.

2.2.3. Possible Selves

In this vein, possible selves (Markus & Nurius, 1986) are a useful concept to encompass these thoughts about possibilities and potentials in the future. Markus & Nurius propose: "an individual's repertoire of possible selves can be viewed as the cognitive manifestation of enduring goals, aspirations, motives, fears, and threats. Possible selves provide the specific, self-relevant form, meaning, organization, and direction to these dynamics" (1986, p. 954).

Markus and Nurius (1986) further elaborate on important properties of possible selves: Linking back to the consideration of the social self, the plethora of possible selves is portrayed as a product of social comparison. The individual may derive possible selves building on the realization that "What others are now, I could become" (1986, p. 954). As these cognitive forms represent future selves, they reflect self-images that often have not been previously experienced by the individual. In other words, while building possible selves, the individual does not simultaneously consider whether they would become reality.

Through their aspirational nature, possible selves can act as "incentives for future behavior" (Markus & Nurius, 1986, p. 955). Possible selves are understood as the link between self-concept and motivation (de Place & Brunot, 2019). According to de Place and Brunot, as such, possible selves can guide behavior by selecting suitable action plans and spurring efforts in achieving those possible selves. However, the authors further argue that not all possible selves equally motivate and encourage behavior. Instead, they show that the effectiveness of possible selves in motivating and initiating behavior depends on a range of different factors. For instance, Oyserman, Bybee, and Terry (2006), identify that possible selves are not equally effective toward behavior because the "self is multidimensional and includes multiple potentially competing goals" (p. 188). In their reasoning, possible selves

may not effectively guide action due to conflict with other parts of the self-concept. Additionally, and perhaps least contested, the salience and ease with which a possible self can be accessed determine its self-regulatory effectiveness (de Place & Brunot, 2019). In this vein, the social context cueing a possible self, i.e., making it salient, also plays an important role because cued information is more likely to affect judgment and decision-making (Oyserman et al., 2006).

Because the newcomer initially lacks a current self-concept in the organization, possible selves are likely to substantially direct the newcomer's experience in the initial entry phase. How might possible selves guide behavior in workplace entry depending on the characteristics outlined? Strauss, Griffin, and Parker (2012) argue that future-oriented work selves allow individuals to be ambitious, take risks, and think creatively about their possibilities. Building on these arguments, the authors link future work selves to proactive career decision-making behavior. In the socialization context, too, I contend that those future work selves can incite proactivity, for instance, in encouraging the newcomer to seek information. Furthermore, it is likely that whether or not the work context of the particular organization cues certain possible work selves, will direct their salience and in turn their impact on behavior. Similarly, the newcomer's possible selves may be less effective if the action plans derived from it seem unlikely to be achieved within the context of the organization. In particular, I suggest that in this case the newcomer's endurance in following the goals derived from possible selves will be undermined.

Moreover, as the newcomer builds up his or her current self in the organization, the proposed reality-checking mechanism comes into play. There, it is suggested that in an intrapersonal comparison the overlap of possible and current self is scrutinized. The outcomes of this comparison are suggested to steer the newcomer's future behavior toward reaching possible self-relevant goals.

3. Method

3.1. Context

Management literature articulates the pressing need for addressing anomalies and surprising events with adequate theory (Sætre & Van de Ven, 2021). Scholars have identified abductive reasoning as a fruitful approach in theoretically explaining the increasingly rapid changes to organizations and social contexts (Sætre & Van de Ven, 2021). Hansen (2008) states that abduction forms explanatory hypotheses, and thus is contemporarily seen as the driving force in much-needed theory creation in organizational and management research. The author further distinguishes abduction from an inductive approach: "While abduction is more of a leap than induction, induction is actually a stronger form of inference" (Hansen, 2008, p. 4).

According to Sætre and Van de Ven (2021), the process of theory creation involves three steps. It starts with an abductive step aimed at developing "ideas or hunches" (Sætre &

Van de Ven, 2021, p. 3) to explain incomplete observations, puzzles, or anomalies. Next, deduction serves to create a theory based on the discovered hunches, and finally, induction allows for testing and verification of the theory.

In newcomer socialization, a key puzzle yet to be solved is how similar future social selves can lead to opposing adjustment outcomes, depending on a wide range of factors yet to be discovered. Seeking to solve this puzzle, hunches are formed abductively and attempts are made at deductive theory creation and inductive testing. While the scope of this bachelor's thesis certainly does not allow for a holistic abductive study, it should nevertheless be emphasized that an abductive approach has impacted some of the methodological choices of this thesis. For instance, whereas inductive coding for data analysis would encompass deriving codes from data only, here codes are created both from data and from theory.

Next to the research method, the design choice is a key decision affecting the ability to gain relevant research insights. Previously, scholars from multiple perspectives have relied on both qualitative and quantitative methods to study self-concept and identity (Ramarajan, 2014). Because the self-concept is a complex, multi-faceted phenomenon, which needs to be interpreted in context, this thesis adopts a qualitative design based on interviews. Also, inherently due to the nature of the self-concept as a latent variable, this thesis had to rely on self-report data for gaining insights. In particular, possible selves and their impact on behavior during newcomer socialization are best investigated using a qualitative design.

Furthermore, the real-world complexity of the investigated phenomena calls for original data collection in field research. Also, because theory on the links between self-concept and short-term newcomer adjustment is relatively nascent, a qualitative interview approach appears to be a good methodological fit (Edmondson & McManus, 2007). In particular, Edmondson and McManus (2007) state that for qualitative approaches to nascent theory, relatively open-ended research questions are typically explored using rather open-ended data collection methods. This approach furthers knowledge by interpreting the rich open-ended qualitative findings for meaning (Edmondson & McManus, 2007).

3.2. Data Collection
3.2.1. Theoretical Sampling

To best investigate emerging theory, theoretical sampling appears to be a useful approach to data collection (Draucker, Martsolf, Ross, & Rusk, 2007). According to Draucker et al. (2007), theoretical sampling is "the process of data collection directed by evolving theory rather than by predetermined population dimensions" (p. 1137). Hence, sampling decisions are primarily guided by the research question under investigation.

When selecting the sample to explore self-concept and newcomer short-term adjustment, I considered findings from my review of the literature in both fields. For instance, past work experience might impact adjustment outcomes and is

typically controlled for (Ashforth et al., 2007). Hence, because past similar roles might critically guide role adjustment in the new organizational context, efforts were made at obtaining a sample with comparable prior work experience. Specifically, the sample includes three newcomers, all of age 22, at entry-level positions, who typically have little to no previous work experience in the specific field of work.

Insights on the proposed links between self-concept and adjustment would benefit from generalizable findings to establish strong theoretical connections. In my view, a central trade-off within the frame of resources afforded by a bachelor's thesis lies between the chosen depth and breadth of elaboration. The strength of this work lies in its deep real-life accounts while limiting the volume of cases investigated and thus curbing generalizability. Increased resources, for instance, multiple trained investigators, a longer time frame, and access to larger samples would promote more generalizable results and come closer to theoretical saturation.

3.2.2. Procedure

This thesis adopts a longitudinal design consisting of two interviews. They were conducted shortly before entering the organization and at three weeks tenure in the organization. Longitudinal designs are useful for investigating changes in variables across time. As outlined above, the newcomer enters the organization with future possible selves for the job, and during initial socialization starts building present identities in the organization. Because this thesis is interested in the interplay of incoming self and self in socialization, as well as the processes between them, measuring both concepts at the same point in time is not appropriate as it would introduce hindsight bias.

Conversely, the key methodological strength of using this type of design with two separate interviews at different points in time is that it effectively eliminates the influence of hindsight bias. Roese and Vohs (2012) define this type of bias as the "belief that an event is more predictable after it becomes known than it was before it became known." (p. 411). It involves being unable to recall uncertainty before a past event and thus disturbs appraisal of these events (Roese & Vohs, 2012). For instance, if both self-concepts were measured at three weeks tenure and the newcomer reports positive adjustment outcomes, he or she might conclude that it was very likely all along that adjustment would be positive. This might involve stressing his or her abilities and fit for the position–which were clear from the onset–and ignoring the uncertainties almost inevitably involved in crafting present identities in the new organization. On the other hand, the newcomer would be likely to also overestimate the chances of negative adjustment outcomes in hindsight.

Based on my literature review on both self-concept and newcomer adjustment, I created two interview protocols targeted at uncovering the links between the two fields (*see Appendices 1 and 2*). Specifically, the interviews were aimed at elucidating the nature of the newcomers' self-concept and socialization context and recalling critical incidents important for newcomer adjustment outcomes. Interviews were designed to trace newcomers throughout their organizational entry phase. Whereas the pre-entry interview looked at the newcomer's future self-concept for the coming entry phase, the post-entry interview examined the present self-concept three weeks later. Additionally, the second interview was aimed at finding self-reports of 1) organizational socialization tactics, 2) newcomer proactivity, and 3) adjustment outcomes in line with the core model of newcomer adjustment (Bauer et al., 2007). Similar structuring of the self-concept questions in both interviews allowed for increased comparability of the two, as important for examining the suggested 'reality checking' mechanism.

The interview protocols created made use of narrative methods of inquiry. Researchers interested in personality have relied on and developed narrative identity methods–methods aimed at eliciting the story of one's self as it is subjectively constructed across time (Adler et al., 2017). According to Adler et al. (2017), "narratives of personal experiences are dynamically reconstructed representations of events [...] [and] are deeply idiographic" (p. 520). Because the present research is rather exploratory, narrative methods appear to be well suited (Adler et al., 2017). The authors further highlight a few key points to consider when employing narrative methods: First, narrative methods are suited for finding meaning in investigating data. Second, narratives depict identity in that they show how the individual makes sense of the surroundings and experiences, which is strongly correlated with the perception of who he or she is, and hence the self-concept. In essence, "narrative methods are particularly well situated to examine meaning-making processes that concern the self." (Adler et al., 2017, p. 520).

Hence, the interviews utilized semi-structured interviews employing narrative methods. Inspiration was obtained from Atkinson's (2002) contribution on the Life Story Interview (LSI) as a narrative method. Designing the study material thus involved creating narrative questions and narrative prompts. While a narrative question is an open-ended inquiry about a story regarding its explicit, implicit, and structural aspects (Adler et al., 2017), narrative prompts are a strong tool for bringing participants into 'story-telling mode'. Prompts nudge participants into mentally recalling specific critical incidents, promising to yield informative insights. A narrative prompt in the present study takes, for instance, the following form:

> *"Please describe a scene, special moment, or event that is symbolic of a beginning in a new organization. Please describe this scene in detail. What happened, when and where did it happen, who was there? What were you thinking and feeling? Also, please describe why this particular moment stands out and what it reveals about you as a person."*
> *(see Appendices 1-3 for Study Material)*

At the start of each interview, the introduction was read out to the participant verbatim, to make sure that he or she felt well-informed, and to establish that the situation needed

to be taken seriously (Hermanowicz, 2002). During the interviews, emphasis was put on achieving flow throughout the conversation to facilitate detailed narration and the finding of meaningful data. Moreover, a key concern was to establish similar conditions to maximize the extent to which the results can be intersubjectively reproduced. Two of the three participants were interviewed in person, one using video-calling. Because the interviews were being performed during the Covid19 pandemic, with mandatory social distancing requirements, the choice of interview mode was left to the participants to ensure they would feel comfortable in the interview situation. While before conducting the interviews, uncertainty surrounded the point of whether participants would "open up" to equal degrees in in-person and online interviewing, these concerns turned out to be unwarranted, as both variants produced comparably rich accounts.

During all face-to-face interviews, a neutral location was chosen, and in video calls, it was, for instance, made sure that a neutral virtual background was displayed. The interviewer attempted to take an eager, genuinely interested, impartial stance throughout the entire interview process. These considerations informed the preparation for interviews and were key factors involved in minimizing bias–such as through verbal and nonverbal cues including posture, voice pitch, and fluency of speech.

Interviews lasted between 50 and 80 minutes and were conducted in German. They were audio-recorded and transcribed word-for-word. Additionally, the transcripts attempt to capture speaking breaks and signals of contemplation ("Uhmm", "ahhh", "hmmm", etc). I argue that these cues help data evaluation as they signify when the interviewees engage in effortful reflection. Because inquiries about deeply rooted, personal phenomena such as the self-concept typically pose "difficult questions" to the participants, the mentioned cues can serve as signposts to self-relevant revelations.

Additionally, to further reduce the potential influence of biases in interviewing, participants were asked to fill out the Twenty Statement Test (TST) prior to the first interview for cross-validation (see Appendix 3) (Baumann, Mitchell, & Hodges Persell, 1989).

3.3. Data Analysis

Data were analyzed following qualitative content analysis approaches developed by Mayring (2000). Inferences were made from transcribed data only, which included the spoken word with speaking breaks. Following Mayring, inductive and deductive data analysis utilizes different coding procedures. Because this thesis takes an abductive research approach, both inductive and deductive coding were used to draw conclusions. While Mayring states that both processes are based on the formulated research question, he also highlights some key differences between the two. Inductive analysis appeared useful where little prior theory could guide coding, for instance, when looking at individual newcomer entry expectations in the first interview or identity processes in the second interview. On the other hand, deduction was used when theory was readily available and could

guide coding decisions. For example, newcomer adjustment appeared suitable for deductive coding, because a comprehensive model was available (Bauer et al., 2007).

First, for inductive coding, according to Mayring (2000), codes were iteratively derived from the interview transcripts. Codes were obtained directly from data and later revised, summarized, and cross-checked for redundancies. Through subsequent iterations, these codes became increasingly structured hierarchically and could be summarized into higher-order categories. After roughly 30% of the material, the emergent categories were again revised and coding was finalized.

Second, deductive coding is based on defining aspects and categories to analyze based on the research question under investigation (Mayring, 2000). Again, a stepwise process was followed to arrive at codes fitting to the higher-order categories. Codes were deduced from data with the theoretical categories in mind and later cross-checked and summarized so that they would explain the categories to the best possible extent. Again, revision took place when roughly 30% of the material was coded. Out of inductive and deductive coding techniques, a comprehensive coding scheme emerged (see Appendix 4).

Finally, TST responses were used as background information to validate interview results.

4. Results

4.1. Pre-Entry

4.1.1. The Self-Concept

Centrally, interviewees revealed rich accounts of their self-concept, including their social selves, and general personality traits. Participant names were pseudonymized using name occurrence distributions of the participants' years of birth to ensure confidentiality. I will refer to the participants as Taylor, Jacob, and Naomi.

Taylor. said that he is open, willing to help, and peace-loving. Taylor placed importance on making an impression on others, enjoyed being in groups, and particularly appreciated the valuation and interest of other group members.

Jacob. identified as being agreeable, spontaneous, creative, flexible, good at improvising, perfectionist, and open.

Naomi. Most prominent were Naomi's accounts of her lack of self-confidence and self-doubt with 13 occurrences. The following quote highlights this aspect: *"But sometimes as I said, I doubt my abilities. And then I would not see myself as the first contact person if there is anything that needs to be solved now or so in a .... If there is a problem and you need a solution for it."* Additionally, Naomi said that she is conscientious, structured (and talented at organizing) agreeable, extraverted, creative, willing to help, and peace-loving. Naomi placed importance on making an impression on others, enjoyed being in groups, and appreciated the valuation and interest of other group members.

4.1.2. Uncertainty

Participants further described feelings of uncertainty regarding three categories: purpose and orientation uncertainty, social uncertainty, and task and role uncertainty.

All participants expressed that they did not know central information regarding their task and role and were uncertain about those aspects in the future organization. This is not surprising as all newcomers are in entry-level positions and have limited previous similar role experiences. Taylor reported considering being provided a certain degree of task and role freedom as a valid strategy for resolving task and role uncertainty proactively.

Social uncertainty also mattered to participants. All reported feelings of uncertainty about their future social environment, in particular regarding harmony in working with colleagues. As only Naomi worked from home, she reported feeling uncertain about the nature of interactions with colleagues restricted to a purely virtual setting. Interestingly, whereas she stressed the potential difficulties resulting from online interactions, she also considered it an advantage not to have to face colleagues in real life as she reported a significant lack of self-confidence and self-doubts (*see Self-Concept*). Taylor and Naomi stated that they placed a high value on their supervisor expressing interest in working together; consistent with their expressed importance for being valued and shown outside interest (*see Self-Concept*).

Only Taylor explicitly reported requiring purpose and orientation information to resolve his uncertainty regarding those aspects. This especially mattered to him as he considers being provided these insights as an expression of appreciation and valuation for his work:

"Yes, but that definitely makes a difference, how valued I think my work is and how I think they treat interns there or how they deal with interns, whether that's just cheap labor for them or whether that's potential later employees for them, that they like to take good care of so they have a good image of the company

4.1.3. The Organization

The newcomers considered several aspects of their future organization. First, participants all supported the notion of gaining support for their entry into the new organization. Being supported in understanding job tasks and roles in the face of the new environment played a key role for newcomers. Jacob explained the role of individual support in two polar scenarios: *"Either you're just... you're just taken by the hand, and you're just shown what you have to do or you're just completely thrown in at the deep end and you just have to see that you can swim somehow yourself or you just can't get it together at all."* Next to general support in entering the organization, all three newcomers stressed that it is important to them that the organization would tolerate their initial mistakes and consider their learning progress individually. The following quote from Jacob illustrates this point: *"So... it's absolutely justified, so it's always a question of how much puppy protection you have overall. So, you're new, you don't have any experience yet, you have time to learn the whole*

thing, that's about it. And depending on how well this puppy protection simply works, the better it is."

Naomi underscored that she highly values social support as she repeatedly mentioned potential shared after-work experiences. Furthermore, she considered social support a possibility for mitigating her shyness in social interaction.

Taylor imagined possible learnings from his team, including, for instance, orientation for future job choices.

Next to that, Taylor and Naomi articulated that they imagined engaging in comparisons with other organizational members and perceiving differences between other interns of their age and other colleagues, which would affect their interactions with them.

4.1.4. The Individual

On the individual level, too, insights relevant to newcomer adjustment were obtained. Regarding the imagined role, Naomi articulated the wish to develop self-confidence and assertiveness in the new organization. Also, she imagined playing the role of a communicator and mediator, as well as a reliable colleague. Furthermore, she, and in particular also Taylor, mentioned that they find it important that they perceive their work as being valued by the organization. Jacob had clear valence for his imagined role: enjoying working with people, and being supported while doing so, and negative valence for inadequate support. While imagining initially missing orientation, he also had clear conceptions of becoming productive on the job and pictured missing routine and strategies for building up routine. He imagined following colleagues' demands initially, but had a strong vision for taking on responsibility, as illustrated in the following: *"...Until I eventually get into a position where I can explain to other people what they can do best and how, because I have simply gained competence behind it."* Furthermore, participant two articulated having pronounced demands on himself:

Naomi imagined not being proactive, Jacob could picture himself proactively looking for tasks after his initial entry. Taylor gave a detailed account of his proactivity fluctuating depending on how he feels received:

"In the past, I myself have noticed that it depends very much on how I feel received. That doesn't mean that I'm always waiting for someone else to welcome me. But if I notice that someone is not at all interested in dealing with me or in who I actually am and what I actually do here, then I don't feel like making an effort to say "Hello, here I am". But on the contrary, when I notice that people are also interested, then I love to slide into new groups and discover new groups. But somehow... I find it very different, depending on how the mindset is.?...So, in one scenario I also have a real desire uh, myself ... no idea, to change the group, to do something different, to do new things, to establish new things in the group and so on and in the other scenario just not so."

Furthermore, Jacob mentioned multiple paths through which he might learn the job, including learning by following experienced colleagues, experimenting with tasks, and learning by pretending to know ("Fake it, till you make it."). He also stressed considering adaptive capabilities, for instance, to handle novelty, manage expectations, balance different task demands, or adapt to a high workload.

All newcomers stated that they anticipated the new job, with Naomi also reporting feeling frightened in the face of novelty.

4.1.5. Background Information

Newcomers reported job-related information to the extent that they knew what could be expected. Unsurprisingly, job tasks varied as the participants entered different jobs at various organizations. All stated the overarching goals or purposes of the organization. Taylor reported possessing most previous knowledge as compared to the other interviewees. Taylor and Jacob carried out their jobs in person, while Naomi worked from home during the first three weeks.

Next to general job information, newcomers also reported their imaginations and valences for the first day in the organization. Overall, participants stated that they would miss orientation, feel nervous, and imagine scenarios of failure. Tied to these possibilities were behavioral strategies newcomers imagined adopting on the first day. In particular, all newcomers placed importance on preparing well for their entry day. Also, newcomers planned to counter nervousness by seeking distraction and social support and evaluated options for self-regulation regarding imaginations of failure. Furthermore, Jacob already pictured that he could find orientation on the job due to the limited complexity of tasks combined with his ability for developing routine for job tasks.

4.2. Post-Entry

4.2.1. The Self-Concept

Personality Traits and Social Self. Participants provided accounts of their self-concept and social self.

Taylor. reported being adaptable, for instance, in accepting inadequate information provided by the organization and evaluated this personality trait as too pronounced. He showed tendencies to justify his lack of questioning for relevant information Associated with this aspect, he stated that he aims at avoiding a negative impression as being too "intrusive" when asking for information.

Jacob. noted being encouraging with colleagues, for instance, checking on their wellbeing, or providing considerate gestures. Furthermore, he described possessing a social team identity in metaphorical terms:

> *Sums it up pretty well actually, "Part of the ship, part of the crew." Because it just works best when we all work together when we all effectively pull together. As trite as that may sound now. But if we simply pay attention to each other and, um, pay attention to what the other people are doing and what's happening there, then it all just works*

> *best, because it's a very social interaction, where you simply have to work with people. And then just not that it's impossible that you just don't work together with them, with your fellow human beings, because that just causes problems.*

He also reported identifying as an egoistic altruist. With this concept, he described helping colleagues to maximize his enjoyment of work and learning outcomes. He stated making colleagues enjoy their work because he enjoys working along and interacting with satisfied colleagues. He also mentioned that this personality aspect would affect his choices of relationships with colleagues. Additionally, Jacob perceived a social hierarchy within his hierarchical level in the organization, according to their job experience. He reported behaving differently toward those categories of people–for example, helping colleagues with less experience, while approaching experienced "mentor" colleagues for his learning progress,

Naomi. reflected being extraverted, agreeable, and willing to help. She also stressed having a creative side, which she mentioned utilizing for generating suggestions for improving job performance. She mentioned that her supervisor had confirmed this personality aspect to her.

Identity Processes. The newcomers described a variety of identity processes, encompassing identity aspirations, possibilities, and implementations of pre-entry wished identities.

Taylor. reported being dependent on consideration by his colleagues. He stressed the importance of being given attention and time from colleagues – for instance, being explained his tasks in detail and being provided background information. He mentioned that the amount of consideration given would determine his motivation to acquire task-relevant learnings. Taylor said that he wished pre-entry to take on responsibility and be proactive in asking for task-relevant information. He assessed having been only partially successful in implementing his desired identity.

Jacob. described enjoying being himself on the job and feeling completed through his job, exemplified by the following quote:

> *I, I am effectively me again. I can work with people again. I'm getting to know a bunch of new people again, I can just. I'm meeting new people every day. I can talk to a bunch of new people again. I can be around people again. I can just be with people again! So that's, I've noticed what I was missing before, that I'm just back in a room with a lot of people, and I can talk to a lot of people.*

He stated being able to enact a part of his identity again on the current job, that he previously could not. Associated with this, he stressed repeatedly a need to be in control over his job, job routines, outcomes, and interactions. The newcomer mentioned being proud of being independent. Regarding performing tasks delegated to him, he outlined the importance of being given reasons for individual task assignments. On the contrary, Jacob expressed disliking being

given orders that deprive him of his freedom of choice on how to perform tasks. He stated finding it challenging to handle situations where he is facing outside factors out of his control, for instance, the Covid19 pandemic limiting his possibilities to interact with people on the job. Participant two also mentioned enjoying being in control not only of his own but also of other's outcomes. He reported liking to influence colleagues to make them perform better at their tasks. Along with being in control he described feeling less uncertain and capable to act as opposed to previous feelings of surrendering to outside factors.

Moreover, Jacob stated that he desired an identity as an assertive authority being able to *"yell at colleagues to do their work better"*. His motivation for this wished identity connects to his desire to improve job performance and dealing with people appropriately: he described colleagues' mistakes in the bigger picture including their follow-on effects, such as delays in performance and treating customers unkindly and disrespectfully. Finally, he assessed hierarchical structures to be blocking the enactment of this desired identity.

Naomi. said that it is central to her that she has become more self-confident:

> *Yeah. So, I feel more confident as I said. And like I said, yeah, and honestly, when this job is done, I just look forward to more jobs because I'm always learning new things. So now I don't feel so scared or anything. I feel more like "YOLO experience is good like that." Hm. But I really think it's always my self-concept, in the beginning, I worry too much, I think too much or, or sometimes in the beginning... Oh yeah, what I also noticed is in the beginning I thought about messages a thousand times before... so I thought about the message carefully before I sent it. And then, in retrospect, I thought to myself, "Oh no, that comes across badly, what will they think of me? They're definitely going to think I'm rude." But now I just send messages like that, and I think to myself like that and when I think about it I think to myself like "oh no, it doesn't matter how it should come across". Yeah, and I just try to worry less now. Exactly.*

In this realm, she mentioned being able to increasingly tolerate her own mistakes. She also mentioned using downward comparison to increase self-confidence and implementing individual regulation for self- doubts.

4.2.2. Newcomer Adjustment

Adjustment Cues. Participants frequently gave cues about the success of their socialization experience. Newcomers often described adjusting as feeling at home and having arrived and reflecting less about what to do and how.

Taylor. displayed uncertainty regarding his adjustment success. For him, the freedom of structuring his workday relatively independently was key to feeling at home in the organization. He stated that initial adjustment was not very difficult because the division deals with a wide variety of tasks

that do not require specialized previous knowledge. Furthermore, he again stressed the importance of being individually considered by colleagues for feeling at home in the organization and questioned colleagues' efforts in introducing him. Conversely, he cued that his initial entry experience was mediocre and that he felt neglected by his supervisor. Associated with the perceived ease of initial entry and his supervisor's absence in the first two weeks, newcomer two described having neglected key learnings that became apparent when interacting with his supervisor in week three.

Jacob. gave positive cues about his adjustment outcomes. He reported feeling socially accepted and welcome and stated that he enjoyed getting to know colleagues more deeply. While he initially felt overwhelmed and confused, he evaluated it as chiefly important for feeling at home that he understood the structure and functioning of the organization after three to four days. His associated valence was positive, as he felt more motivated and clearer about his job and tasks:

> *And, on the other hand, it was also simply easier, simply because the work as a whole felt more gripping. Because you know exactly, I'm doing exactly this now. Then this happens, and then that happens. That you understood the individual work steps.*

Jacob also understood his job within the larger organizational structure, from which he could derive purpose and orientation.

Naomi. also gave signs of positive adjustment. The newcomer felt initially overwhelmed by the amount of information but felt motivated to start learning for her progress. While she was initially neglected by her supervisor, she perceived his later apologies and promises of support as credible. Thus, after initial difficulties, she felt well supported by her supervisor and reported that this helped her feel more at ease, comfortable, and at home in the organization. This coincided with feeling at home starting in her third week in the organization. She stated that she felt socially accepted within her team. The newcomer associated being at home in the organization with feeling secure and self-confident. She felt effective at learning her tasks and role, perceiving decreasing dependence on others to explain tasks to her. Furthermore, she enjoyed being in contact with customers and had a positive valence for the job enjoying working in the organization. Finally, a positive comforting valence was reported regarding the fact that she entered the organization together with other newcomers. She felt that she would not be the only one to make mistakes and that individual mistakes would not attract as much attention.

Role Clarity.

Taylor. appeared to be unsure about his role: while stating that he knew his role – as rather a role assigned to him from the outside organizational environment – he displayed uncertainty regarding various role aspects. For instance, he was not sure whether his role clarity mattered for him feeling at home in the organization and did not know how he

acquired his role. He described his role as encompassing the performance of simple, individual supportive tasks, while not understanding the purpose of his work in the organizational framework. The newcomer reported being dependent on colleagues to provide tasks and expressed a negative valence toward his role. He stated that he took over a predetermined role, which he had little influence in shaping, and was uncertain regarding his future valence toward his role.

Jacob. on the other hand, could relatively easily recall aspects of his role, including dealing with people and making them feel comfortable while doing so. He perceived having freedom over the development of his role and had a positive valence for notion. Understanding his role in the organizational structure appeared important to the newcomer. While reporting that he initially did not understand the structures, he later grasped important structures relevant to his role. Finally, structure embedding his role was important because it provided orientation.

Naomi. was certain about her role in being a reliable partner and being fast at executing tasks.

Self-Efficacy.

Taylor. stressed that he did not feel demanded on the job. He said he contributed small parts of information for colleagues but that this did not challenge him in his skillset. He stated having tried to take on more challenging tasks but found it difficult to do so. The newcomer mentioned that he was passively waiting for colleagues to approach him rather than be proactive in searching for adequate tasks and that his proactivity depended on being given more challenging tasks.

Jacob. found it not very difficult to perform his tasks, and said that they require attention and motivation. The novice stated that he knew which tasks he particularly enjoyed performing and managed to make his job performance a habit. He adopted individualized approaches to dealing with his customers and viewed this as a key aspect of learning his job. Coupled with his *"sure-fire success"* performing main job tasks, he mentioned that he could deal with other tasks and contribute to organizational structure. The newcomer imagined optimizing processes, contributing to structures, and making a difference through his work. Regarding his proactivity in learning the job, he stated that he would engage in *"learning by doing"* and proactively look for more experienced colleagues for obtaining job-relevant information. He mentioned that he got praise for his work and started taking on responsibility for other colleagues after one to one and a half weeks. Colleagues relied on him and approached him when they encountered any problems and he stated that he engaged as a problem-solver in these instances. Furthermore, Jacob said that he approached problems proactively and was perceived as a considerate colleague interested in his coworkers' needs and problems. In line with taking on responsibility, the newcomer reported keeping an overview over his team. He felt joyfully surprised and valued for being able to take on responsibility. The newcomer felt increasingly capable of handling difficult, unexpected situations.

Naomi. reported that she felt hard to replace and motivated to distinguish herself from other colleagues and learn her tasks quickly. She said that she compared positively with other interns and felt confident despite initial self-doubts. The newcomer identified strongly as a fast and disciplined learner and was also perceived as such by her colleagues, which she thinks they view positively. She said that she was diligent at learning the tasks on her own starting in her first week and that this helped her gain orientation in the organization. She received repeated praise for her work from colleagues. The newcomer dared to try challenging tasks, which she felt helped her build up self-confidence. For instance, she outlined taking on customer calls in her first week at work–a challenging situation associated with social uncertainty for her, which she said she mastered using a variety of acquired techniques. Moreover, Naomi stated picking her work times so that colleagues were available for her to proactively approach. She also reported providing possible solutions for problems, which impressed her colleagues. Taking on more challenging tasks was perceived as enjoyable and appreciative She felt valued for taking on responsibility for other colleagues, too, and could imagine performing a leading role in the future.

Social Acceptance.

Taylor. reported that feeling accepted by colleagues was associated with reflecting less about his actions. While he stated that he became less reflexive throughout his entry phase, he gave mixed signs of social acceptance and described feelings of separation when working: *"On the social level?....... hmmm...Difficult..... well,..... it's just rather ... well, it's an open-plan office. But so, the work itself is already yes ....very, very separate."* He noted that little collaboration took place and that he usually performed delegated tasks alone. His work was not part of discussions in team meetings, which revolved around overarching topics, which he felt he was not contributing to. Regarding the social atmosphere, the newcomer said that he got along well with open colleagues and that the work topics are generally suited for starting conversations with other coworkers. He cued that he perceived differences between interns and permanent employees, which he said both have separate group dynamics. However, he noted that the general atmosphere was harmonious and collegial, where he felt he could talk about beyond-work topics and felt hierarchies were unimportant.

Jacob. gave no more specific cues regarding his social acceptance, beyond feeling accepted and welcome as he got to know his colleagues better (*see Adjustment Cues*).

Naomi. reported that her mistakes were tolerated and that others knew who she is and what she does and felt valued by them in those aspects. While she enjoyed being seen as competent, she also wanted to avoid making a too strong or arrogant impression. A sense of belonging fostered by her supervisor and a large-scale organizational meeting was recalled. The newcomer felt more confident feeling as part of the organization. Social support from another intern also

played a role for her: the other intern admitted her own mistakes, which helped the participant feel accepted in making mistakes herself and encouraged her learning experience. The intern provided multiple functions for her: she took over aspects that her supervisor had neglected, buffered her uncertainties and insecurities when approaching her supervisor by being readily approachable in place of her supervisor. She also enjoyed the close contact with her and felt that they got well along. The newcomer described a difference when interacting with colleagues of her age and older colleagues, but also stated that this difference did not matter a lot and that she could get along with both perceived categories. A positive valence of feeling more at ease was reported regarding her many similar-age colleagues. The general work atmosphere was described as harmonious and on eyes-sight and reflected as contributing to her motivation to learn and feeling at ease with less need for reflection. Naomi said that she gets along well with her supervisor and another working student she has contact with.

5. Discussion

5.1. Adjustment Outcomes and Identity Work

5.1.1. Naomi's Identity Partnerships and Identity Work

Naomi gave clear signs of positive adjustment outcomes with high role clarity, self-efficacy, and social acceptance. I suggest that the identity work she performed during the three weeks under investigation played a key role in reaching these advantageous results. She provided insights into interactions with incumbents that helped her in her initial socialization experience. She described that another intern took over socialization tasks normally expected by her supervisor and provided a variety of functions for her: She reported that the incumbent acted humbly–admitting her own mistakes and insecurities–and that this helped her relate and feel secure about making mistakes herself. Because she stressed enjoying interacting on eyes sight, I suggest that the humble behavior of her colleague made the newcomer consider the incumbent as a more likely partner to open up to and discuss her organizational identity with. Furthermore, the implicit message to the newcomer was likely targeted (whether intentionally or not) at boosting the newcomer's self-confidence, which she said she had doubts about pre-entry. Her colleague is described as her first approachable point of access for issues at work:

> Exactly, so the intern, of course, she doesn't always have so much to do. That's why she said I can always ask her and then I ask her and if she can't help me or if she doesn't know because she's still new, then she says, "Ask the supervisor." or something. And then I do that and then it gets sorted out. And that's exactly. So first I always go to her. And then if she can't help me, then I go to him. And still, it's very relaxed.

Naomi's enjoyment of talking about non-work topics and frequently being on calls with her colleague hints that the interaction was not only focused on work and task structure but also involved overarching concerns, for instance about her identity in the firm.

Coupled with the interaction, the newcomer described a feeling of belongingness or identification with the organization. Identification is a key predictor of socialization success because it predicts contributions to the team and team performance (Sluss & Ashforth, 2007). Thus, the newcomer's identification with the organization is likely to constitute one reason for her successful adjustment.

The depth of exchange between the incumbent and the newcomer points to discussing the processes underlying interactions leading to identification and adjustment success. While it is impossible to ascertain the precise contents of interaction, it is plausible to suspect that her peer communicated a lot of relevant information "in narrative form, that is, as event-driven stories, myths, and folklore that vividly highlight both positive and negative cultural cues in action" (Ashforth, Sluss, & Harrison, 2008, p. 7). Because these narratives are so engaging and relatable, the authors suggest that they provide memorable background and (tacit) knowledge upon which newcomer identification can be built.

Processes such as the one described are also part of the scholarly conversation seeking to gain missing knowledge on how newcomers obtain identity resources in organizations. It is currently debated how newcomers and incumbents interact to exchange identity resources important for newcomer socialization. Cooper, Rockmann, Moteabbed, and Thatcher (2021) make a point for identity partnerships–"a relationship in which identity resources are exchanged between an incumbent team member and a team newcomer" (p. 2).

Judging by the rich descriptions of Naomi's interaction with the incumbent, she might likely be describing an identity partnership in Cooper's (2021) sense, which is outlined in the following. An identity partnership is a dyadic level concept, which is made up of one newcomer and one incumbent for mutual exchange of identity resources to satisfy identity needs. While there are partnerships with which all identity needs can be satisfied, it is likely that newcomers and incumbents may seek multiple identity partnerships that each satisfy individual identity needs. The identity resources are twofold: team identity cues, which transmit information of descriptive and normative nature about the team identity, and acceptance cues, which convey a sense of belonging. The partnership may be initiated either by the newcomer or the incumbent. On the one hand, because newcomers enter teams generally lacking a team-based identity, they may experience identity ambiguity. "Identity ambiguity in a new team is likely to trigger two specific identity needs–uncertainty reduction and belonging" (p. 11). Individuals may seek information to understand how they fit into the social system and to grasp their team identity. Next to that, newcomers want to feel belonging to people around them and feel like organizational insiders, as opposed to outsiders. Thus, the experience of identity ambiguity may trigger the

Table 1: *Overview of Newcomer Self-Concepts, Identity Processes, and Adjustment Outcomes*

	Taylor	Jacob	Naomi
Pre-Entry Self-Concept	Importance of Making an Impression	Agreeableness	Importance of Making an Impression
		Spontaneity	
	Enjoying Being in Groups	Creativity	Enjoying Being in Groups
	Being Oneself in Groups	Flexibility	Importance of Outside Interest and Valuation
		Improvisation	
	Importance of Outside Interest and Valuation	Perfectionism	Conscientiousness
			Structuredness
	Willingness to Help	Openness	Lack of Self-Confidence, Self-Doubt
			Agreeableness
	Peace-Loving		Extraversion
	Openness		Creativity
			Willingness to Help
			Peace-Loving
Post-Entry Self-Concept	Adaptability	Encouraging Colleagues	Extraversion
	Lack of Social Team Identity	Social Team Identity	Agreeableness
		Egoistic Altruist	Willingness to Help
		Categorization into Experience Level	Creativity
Identity Work	Dependence on Consideration	Being Oneself Again	Increasing Self-Confidence
	Wished Pre-Identity: Taking on Responsibility	Wished Identity: Assertive Authority	
Adjustment			
• **Role Clarity**	Low	High	High
• **Self-Efficacy**	Low	High	High
• **Social Acceptance**	Low – Medium	Probably High	High

newcomer to proactively seek information and resources including team identity and acceptance cues. On the other hand, incumbents usually possess a team-based identity and thus have different identity needs. Those who strongly identify with the team are likely to proactively approach newcomers to initiate an identity partnership. "Because strong identifiers value their connection to the team, they feel that they belong to that team and will seek out positive interpersonal relationships with team members–including the newcomer–to maintain that feeling of belonging within the team" (p.

13). While strong identifiers (*integrators*) are likely to support newcomers in the socialization process, incumbents who strongly disidentify with the team (*gremlins*) search for confirmation for their negative team identities in the newcomer and may be detrimental to newcomer socialization.

From the data, the degree of identity ambiguity experienced by the newcomer can only be roughly approximated. Naomi signaled to feel lost and neglected by her supervisor initially, although this may alternatively just be a sign of task and role uncertainty: "*Because I mean, seriously, if*

there hadn't been an intern, that would have been really bad. None of us would have known what we were supposed to do." Moreover, precisely which identity resources the incumbent sought from the newcomer is not apparent from the data. However, it appears unlikely that the incumbent had strong team disidentification, as Naomi did not report any cues of anti-team identity on the part of the intern. Rather, one possible explanation for the described integrative potential of the incumbent is that she recently has been a newcomer in the team herself, allowing her to empathize with Naomi's socialization needs. Furthermore, Naomi's positive valence toward colleagues of a similar age may have further simplified the initiation and development of the partnership.

Thus, who made the initiating step for the identity partnership is not clear, but I suggest that this is also not very important – what matters in the realm of the research question is the adjustment success resulting out of the identity partnership. Naomi not only reported feeling more at ease and self-confident having built the partnership, but she even took the incumbent as a role model for her future development in the firm–a possible self providing an incentive to take on a leading role with responsibility for other interns in the future. The perspective of possible selves (Markus & Nurius, 1986), in particular, future-oriented work selves (Strauss et al., 2012) allowing newcomers to be ambitious and take risks is significant in this respect. Specifically, I suggest that the ability to take risks aids in identifying role aspects and developing one's role and that a certain self-confidence is a prerequisite to engage in risk-taking behavior. Naomi reported that the partnership helped her feel more self-confident. Hence, the formation of the identity partnership may not only have contributed to the formation of a future possible work self but also supported its enactment by lowering self-confidence barriers.

This example points at an important aspect of identification in teams and the construction of identity partnerships: the need to not only consider newcomer's identity needs but also further interpersonal and communicational factors. For example, Naomi preferred colleagues of a similar age and enjoyed being in frequent exchange with the incumbent, including topics that exceed the work context. Next to that, it is imaginable that other interpersonal factors played a role in the development of the partnership.

Naomi also described that her supervisor signaled to her a feeling of belongingness to the organization: *"He was so, so confident and also so really.... he made us feel like we're all friends, so even though he's my supervisor. He makes us feel like we can all still be very, very close to each other or communicate with each other."* Naomi might thus be pointing at another identity partnership with her supervisor, mostly for obtaining acceptance cues. The point in time that her supervisor provided these cues also coincides with when she reported feeling at home in the organization–a strong indication of adjustment success.

Taken together, next to the identity needs satisfied in the two identity partnerships, Naomi's increases in self-confidence facilitated by the relationships appear to have

supported adjustment in another way. Self-confidence boosts allowed Naomi to explore identity aspects she would otherwise not have dared to. Thus, the identity partnerships also supported her in performing identity work: constructing and altering her identity over time (Bataille & Vough, 2020).

5.1.2. Taylor's Possible Self, Identity Work, and Reality Checking

Hm... So, I wanted to be like that... so... I just wanted to be, um, someone who just cares a lot that people show you something, so I ask a lot. is interested... Um...maybe even asks if he can be part of things where he actually would not be there um.... partly I have implemented it that way, but partly not because I noticed when just people have no desire to deal with additional things. Umm but just want to do their thing, want to do their work. Um, yes, then it just felt a bit out of place and that's why I just partly put it back, so to speak. Uhm...So... yeah... that would actually be my role...

Taylor gave negative signs of his adjustment success, as he stated not understanding his role in the larger frame of the organization, felt little demanded in his tasks, and perceived separation from his colleagues. However, Taylor did possess a wished identity of taking on responsibility and being proactive before entering but said he was largely unsuccessful in implementing it. This observation raises the question: why did Taylor's possible self before entering not perform well as an incentive for proactive behavior to facilitate adjustment?

Two perspectives appear useful in explaining Taylor's identity processes: possible selves (Markus & Nurius, 1986) and identity work (Bataille & Vough, 2020).

First, scholars have argued that the effectiveness of possible selves to function as incentives depends on a wide variety of factors. Critically, Taylor's organizational context is unlikely to have cued proactivity and taking responsibility, as he said he is viewed in the preformed role of the "typical intern": *"Yes, and that is actually how one imagines an internship in the classic sense, a bit of internship work if you want to say it pejoratively. Yes, so my position in the company is intern."* In this role, he said he was expected to perform small supportive jobs that he felt little demanded with. Due to the contrary role orientation, Taylor's organization does not make his possible self salient, which in turn diminishes its likelihood of impacting his behavior (de Place & Brunot, 2019).

Second, Taylor may not have seen identity opportunities that could have made him engage in identity work. Bataille and Vough (2020) conceptualize that newcomers respond to identity opportunities and threats with identity work. According to Bataille and Vough (2020), identity opportunities trigger identity work that makes the individual grow to achieve a desired identity. However, Taylor described scarce identity opportunities in taking on responsibility and becoming proactive. He only mentioned one task assignment that was challenging to him where he felt more responsible. Next to that, the lack of role flexibility may pose another barrier to his performance of identity work. Given his "preformed" role, he may have assessed it unlikely that he could implement his

identity within the context of the organization. Moreover, as he described depending on outside valuation and consideration by his colleagues, he may also have required this feedback with regard to his identity work. I suggest that missing affirmation in the process of identity work has posed another obstacle to successful identity work toward his desired identity.

Finally, the sum of processes described may be understood as Taylor engaging in reality-checking with a negative outcome: the perceived gap between wished and actual identity disincentivizing identity work.

5.1.3. Jacob's Fulfillment of an Identity

Jacob stated that he perceived a sense of completeness through his job. He said that he could be himself again and that he could not fulfill his identity of working with and helping people previously due to external circumstances. I suggest that Jacob's case chiefly underscores the importance of work for identity processes. Work "offers an important place to learn about what one is able to do; which is fundamental to enhancing self-efficacy" (Selenko et al., 2018, p. 2). Jacob thus managed to use work in alignment with his identity to explore how he could advance his skillset to better implement his identity in the organization. The identity itself may have served as a key motivator in this process (Selenko et al., 2018).

5.2. Limitations

Because research is interested in newcomers of all ages, positions, and career stages, the sample in itself might allow for only limited generalizability to the entire population. Future research would do well to explore identity processes with larger samples more representative of the population of newcomers. However, the sample was built under the resource constraints inherent in a bachelor's thesis and thus did not allow for investigating a larger-sized sample. Larger samples could be explored with multiple trained investigators. Furthermore, the research question guided sampling decisions in an important way: by studying student newcomers, previous job experience is not likely to impact findings.

Next to sampling, the interview itself carries aspects worth discussing in the realm of methodological limitations. As stated by Alvesson (2003) a certain interpretation of the interview situation on part of the participant always takes place: "Intensive interpretation of what the researcher is after – before, during, and perhaps after the interview (before a repeat interview) – and the forming of work assumptions of what the entire exercise is about and how specific themes addressed should be understood guide interviewee responses" (p. 19). Thus, despite extensive efforts to ensure questions are clear and easily understandable, it cannot be avoided that there exists a certain range of different interpretations and that answers vary based on that.

Furthermore, social desirability issues are inherent to interview situations. Hence, participants might base their replies on what they think is socially acceptable instead of on their "true" thoughts. The interviewer paid particular attention to minor cues exposing answers biased by social desirability.

Regarding the interview technique, efforts were made at avoiding priming participants (see Section 3.2.2). One potential limitation in this respect could be that newcomers understand key concepts differently than scholars typically define them. Alternatively, could defining terms scientifically also bias participants' responses? Here, the definitions of key concepts such as the self-concept were withheld until participants asked for a clear definition, or the facilitator felt it was necessary to define terms during the flow of the interview. In future research with larger samples, scholars could experiment with explaining terms in researchers' definitions, using a definition put in context ("This is how researchers define the self-concept"), or using no definition at all.

Finally, data analysis could also bear potential limitations. In particular, coding decisions were performed by a single investigator. Employing multiple coders with diverse perspectives could serve to uncover more aspects of the data.

5.3. Summary and Outlook

In sum, the accounts of the three newcomers Taylor, Jacob, and Naomi provide informative insights into identity processes in the organizational entry phase and how they affect short-term adjustment. The participants significantly differed in their identity work processes and adjustment outcomes and thus provide a background against which to discuss (un)successful adjustment processes. The newcomers' identity processes illustrate current research on identity partnerships (Cooper et al., 2021) and identity work (Bataille & Vough, 2020) and designate avenues for future research.

As underlined by Cooper et al. (2021), "our understanding of how newcomers come to define their own identity vis-à-vis the team is still relatively unknown" (p. 3). Identity partnerships as proposed by the authors are one possible path through which newcomers come to build their identity in the team, however, key questions remain. While identity partnerships are also a form of association of individuals, they serve different purposes than, for instance, a friendship or romantic partnership. Scholars would do well in exploring more precisely how identity partnerships differ from associations of individuals not meant for the exchange of identity resources within organizations and what this means for newcomer adjustment. This work provides some cues of success factors of identity partnerships, yet this can only be the beginning. Researchers should thoroughly explore which concepts determine whether identity partnerships become successful guides of newcomer adjustment.

While factors influencing the initiation of identity partnerships are cautiously suggested (Cooper et al., 2021), future research should also investigate how identity partnerships *develop* based on interpersonal factors. Furthermore, there is a gap in knowledge regarding the dynamics that follow after a mutual exchange of identity resources. If both partners' identity needs are satisfied, does the partnership

lose its core purpose and consequently dissolve? Alternatively, when do partners decide to continue investing in their partnership? One suggestion may be that when initial demands are satisfied, partners may still expect having to satisfy identity needs in the future and thus engage in farsighted behavior to continue the partnership.

Another aspect worth exploring is how multiple identity partnerships may interact. As outlined by Cooper et al. (2021), identity partnerships may provide complementary identity resources–could there conversely also be competition for the provision of the same or similar identity resources for the newcomer? For example, one could imagine that an *integrator* (with strong team identification) and a *gremlin* (with strong team disidentification) both build an identity partnership with the newcomer. While both may seek to provide the newcomer with identity and acceptance cues, they apparently may want to satisfy competing interests and needs from the relationship. Which dynamics would follow such competitive scenarios? Could they even lead to larger-scale detrimental outcomes, such as endangering the overarching team identity or igniting team conflict?

Additionally, research has largely focused on investigating identity work on individual identities. Yet, identity researchers have theorized that individuals possess multiple identities and that they relate to one another (Bataille & Vough, 2020; Ramarajan, 2014). Hence, it needs to be studied how identity work on a single identity influences other identities. Beyond that, could identity work alter multiple identities at the same time? What could be the implications of such processes, depending on the relationship between the identities being changed?

From a practical perspective, this work highlights aspects of newcomer adjustment relevant to both incoming newcomer and socializing organization. First, newcomers could monitor their identity needs more closely to help them more effectively take steps toward satisfying them. Second, organizations need to recognize that newcomer adjustment success, and consequently job success factors important to the organization, critically depend on newcomer identity processes. Raising awareness for the importance of newcomers' identity needs could aid the organization in best meeting these needs with fitting identity partnerships enabling newcomer identity work. Thinking beyond, organizations could take efforts to develop more structured approaches to managing individual identities and embedding them into standardized management processes. With the precise design and implementation yet to discuss, could there be a way toward an integrated "Identity Management System"?

Finally, it remains to be stressed that the self-concept is central to a wide array of fields not just in organizational research, but also in psychology at large. Great strides have been made at unveiling the various roles of the self-concept, yet the quest for discovery has just begun. The sheer complexity inherent in self-concept research will continue to challenge scholars in advancing theory toward solving a key puzzle: *Understanding the Impact of Future Social Self-Concepts on Newcomer Adjustment.*

References

Abelson, R. P. (1981). Psychological status of the script concept. *American Psychologist*, *36*(7), 715–729.

Adler, J. M., Dunlop, W. L., Fivush, R., Lilgendahl, J. P., Lodi-Smith, J., McAdams, D. P., . . . Syed, M. (2017). Research Methods for Studying Narrative Identity. *Social Psychological and Personality Science*, *8*(5), 519–527.

Alvesson, M. (2003). Beyond Neopositivists, Romantics, and Localists: A Reflexive Approach to Interviews in Organizational Research. *The Academy of Management Review*, *28*(1), 13.

Ashforth, B. E., & Saks, A. M. (1996). Socialization Tactics: Longitudinal Effects on Newcomer Adjustment. *Academy of Management Journal*, *39*(1), 149–178.

Ashforth, B. E., Sluss, D. M., & Harrison, S. H. (2008). Socialization in Organizational Contexts. *International Review of Industrial and Organizational Psychology*, *22*, 1–70.

Ashforth, B. E., Sluss, D. M., & Saks, A. M. (2007). Socialization tactics, proactive behavior, and newcomer learning: Integrating socialization models. *Journal of Vocational Behavior*, *70*(3), 447–462.

Atkinson, R. (2002). *Handbook of Interview Research: Context & Method*. SAGE Publications Inc.

Bataille, C. D., & Vough, H. C. (2020). More than the Sum of My Parts: An Intrapersonal Network Approach to Identity Work in Response to Identity Opportunities and Threats. *Academy of Management Review*.

Bauer, T. N., Bodner, T., Erdogan, B., Truxillo, D. M., & Tucker, J. S. (2007). Newcomer adjustment during organizational socialization: A meta-analytic review of antecedents, outcomes, and methods. *Journal of Applied Psychology*, *92*(3), 707–721.

Baumann, E. A., Mitchell, R. G., & Hodges Persell, C. (1989). *Encountering Society: Student Resource Manual to accompany Persell, Understanding Society* (Third ed.). New York, Harper & Row. (Exercise 5, Twenty Statement Test,)

Berger, C. R., & Calabrese, R. J. (1975). Some Explorations in Initial Interaction and beyond: Toward a Developmental Theory of Interpersonal Communication. *Human Communication Research*, *1*(2), 99–112.

Brewer, M. B. (1991). The Social Self: On Being the Same and Different at the Same Time. *Personality and Social Psychology Bulletin*, *17*(5), 475–482.

Cooper, D., Rockmann, K. W., Moteabbed, S., & Thatcher, S. M. B. (2021). Integrator or Gremlin? Identity Partnerships and Team Newcomer Socialization. *Academy of Management Review*, *46*(1), 128–146.

de Place, A. L., & Brunot, S. (2019). Motivational and Behavioral Impact of Possible Selves: When Specificity Matters. *Imagination, Cognition and Personality*, *39*(4), 329–347.

Draucker, C. B., Martsolf, D. S., Ross, R., & Rusk, T. B. (2007). Theoretical Sampling and Category Development in Grounded Theory. *Qualitative Health Research*, *17*(8), 1137–1148.

Edmondson, A. C., & McManus, S. E. (2007). Methodological fit in management field research. *Academy of Management Review*, *32*(4), 1246–1264.

Feldman, D. C. (1981). The Multiple Socialization Of Organization Members. *Academy of Management Review*, *6*(2), 309–318.

Hansen, H. (2008). Abduction. In *The SAGE handbook of new approaches in management and organization* (pp. 454–463). SAGE Publications Ltd.

Hanser, L. M., & Muchinsky, P. M. (1978). Work as an information environment. *Organizational Behavior and Human Performance*, *21*(1), 47–60.

Hermanowicz, J. C. (2002). The Great Interview: 25 Strategies for Studying People in Bed. *Qualitative Sociology*, *25*(4), 479–499.

Jones, G. R. (1986). Socialisation Tactics, Self-Efficacy, and Newcomers' Adjustments to Organizations. *Academy of Management Journal*, *29*(2), 262–279.

Langer, E. J., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of „placebic" information in interpersonal interaction. *Journal of Personality and Social Psychology*, *36*(6), 635–642.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*(9), 954–969.

Mayring, P. (2000). Qualitative Content Analysis [28 paragraphs]. *Forum: Qualitative Social Research*, *1*(2).

Mignerey, J., Gorden, W., & Rubin, R. (1995). Organizational Entry: An Investigation of Newcomer Communication Behavior and Uncertainty. *Communication Research*, *22*(1), 54–85.

Miller, V. D., & Jablin, F. M. (1991). Information Seeking during Organizational Entry: Influences, Tactics, and a Model of the Process. *The Academy of Management Review*, *16*(1), 92.

Morrison, E. W. (1993). Longitudinal study of the effects of information seeking on newcomer socialization. *Journal of Applied Psychology*, *78*(2), 173–183.

Oyserman, D. (2001). Self-concept and identity. In *The blackwell handbook of social psychology* (pp. 499–517). A. Tesser & N. Schwarz.

Oyserman, D., Bybee, D., & Terry, K. (2006). Possible selves and academic outcomes: How and when possible selves impel action. *Journal of Personality and Social Psychology*, *91*(1), 188–204.

Ramarajan, L. (2014). Past, Present and Future Research on Multiple Identities: Toward an Intrapersonal Network Approach. *The Academy of Management Annals*, *8*(1), 589–659.

Roese, N. J., & Vohs, K. D. (2012). Hindsight Bias. *Perspectives on Psychological Science*, *7*(5), 411–426.

Roloff, M. E. (1981). *Interpersonal Communication: The Social Exchange Approach (Commtext Series)*. SAGE Publications Inc.

Sætre, A. S., & Van de Ven, A. H. (2021). Generating Theory by Abduction. *Academy of Management Review*.

Schein, E. H. (1971). The Individual, the Organization, and the Career: A Conceptual Scheme. *The Journal of Applied Behavioral Science*, *7*(4), 401–426.

Selenko, E., Berkers, H., Carter, A., Woods, S. A., Otto, K., Urbach, T., & de Witte, H. (2018). On the dynamics of work identity in atypical employment: setting out a research agenda. *European Journal of Work and Organizational Psychology*, *27*(3), 324–334.

Sluss, D. M., & Ashforth, B. E. (2007). Relational Identity and Identification: Defining Ourselves Through Work Relationships. *Academy of Management*, *32*(1), 9–32.

Stangor, C., Jhangiani, R., & Tarry, H. (2014). *Principles of Social Psychology* (1st International ed.). Victoria, B.C, BCcampus.

Strauss, K., Griffin, M. A., & Parker, S. K. (2012). Future work selves: How salient hoped-for identities motivate proactive career behaviors. *Journal of Applied Psychology*, *97*(3), 580–598.

Tajfel, H., & Turner, J. (1979). An Integrative Theory of Intergroup Conflict. In S. Worchel & W. Austin (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA, Brooks / Cole.

Turner, J. C., & Oakes, P. J. (1986). The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence. *British Journal of Social Psychology*, *25*(3), 237–252.

Van Maanen, J. (1978). People processing: Strategies of organizational socialization. *Organizational Dynamics*, *7*(1), 19–36.

Van Maanen, J., & Schein, E. H. (1979). Toward a theory of organizational socialization. *Research in Organizational Behavior*, 209–264.

Junior Management Science

How Can Fairness Tools Impact the Understanding of Fairness and the Processes Within a Machine Learning Development Team?

Cosima Friedle

London School of Economics and Political Science

Abstract

Over the last years, a wide spread of Machine Learning in increasingly more, especially sensitive areas like criminal justice or healthcare has been observed. Popular cases of algorithmic bias illustrate the potential of Machine Learning to reproduce and reinforce biases present in the analogous world and thus lead to discrimination. The realisation of this potential has led to the creation of the research stream on fair, accountable and transparent Machine Learning. One aspect of this research field is the development of fairness tools, algorithmic toolkits that aim to assist developers of Machine Learning in identifying and eliminating bias in their models and thus ensuring fairness. The literature review on fairness tools has revealed a research gap in the impact of these on the understanding of fairness and the processes within a development team. Thus, the aim of this research was to investigate the impact that fairness tools can have on the notion of fairness and the processes in a development team. Therefore, a case study with a development team of a large, globally operating corporation has been conducted. Applying Kallinikos' theory of technology as a regulative regime and Oudshoorn and Pinch's idea of the co-construction of users and technologies on the empirical findings revealed two important conclusions. Firstly, it shows that fairness tools act as regulative regimes by shaping the understanding of fairness and the processes within a development team. Secondly, this character of fairness tools as regulative regimes needs to be understood as part of the co-construction process between the technology and the developer.

Keywords: Machine learning; Fairness; Fairness tools; Regulative regime of technology; Co-construction of user and technology.

1. Introduction

„There's software used across the country to predict future criminals. And it's biased against blacks." (Mattu, Kirchner, & Surya, 2016). This quote illustrates the potential inherent in Machine Learning (ML) to reproduce and reinforce biases and thus create discrimination (Lepri, Oliver, Letouzé, Pentland, & Vinck, 2018). Popular examples like the COMPAS, a criminal recidivism algorithm that discriminated against black people (Saxena et al., 2019) or Amazon's recruiting algorithm that was biased against women show the severe consequences that unfair AI can have (Holstein & Vaughan, 2019).

Against the backdrop of these popular examples of algorithmic bias, there have been growing concerns in both academia and practice over the deployment of ML applications. Especially since the use of ML applications has spread to increasingly more areas over the last years (Adadi & Berrada, 2018), including to contexts in which decision-

making has critical consequences (Saxena et al., 2019), such as healthcare and recruitment processes, this topic has gained momentum. It is argued that given this wide spread and the areas in which these models are employed (Srivastava, Heidari, & Krause, 2019), they can have an enormous, negative impact on the life of many people (Holstein & Vaughan, 2019).

This realisation of the negative potential of ML has created the FAT ML research community which is concerned with Fairness, Accountability, and Transparency in Machine Learning (Gade, Geyik, Kenthapadi, Mithal, & Taly, 2019; Pasquale, 2015). Research within this stream is mainly focused on advancing fair, explainable and accountable algorithmic decision-making models (Adadi & Berrada, 2018). One major theme within the literature on fair ML is the development of fairness tools, which are algorithmic tools that support developers in detecting and eliminating unfairness in ML (Holstein & Vaughan, 2019).

To date, the research on fairness tools has been mainly driven by two theoretical lenses. The deterministic, technical-rational perspective focuses on the technical functionalities of the tools and assumes that they can be easily implemented and will directly lead to more fairness in ML (Berk, Heidari, Jabbari, Kearns, & Roth, 2021; Calmon, Wei, Vinzamuri, Ramamurthy, & Varshney, 2017). Authors analysing fairness tools from a contingent, socio-technical perspective, criticise them for being too simplified and poorly suited to ensure fairness and call for an increased focus on the broader social context of algorithmic systems (Binns, 2018; Holstein & Vaughan, 2019).

One aspect that has been studied very little is the impact that fairness tools can have on the developers that are interacting with them and, an aspect that has been neglected so far, on the processes within a development team. Some papers suggest that fairness tools impact the knowledge of developers on fairness and bias mitigation, however, this has not been investigated empirically to date (Bellamy et al., 2019; Yan, Gu, Lin, & Rzeszotarski, 2020).

In order to fill this research gap, a case study has been conducted with a development team to evaluate the impact that fairness tools have on the understanding of fairness of the developers and the processes within the development team. The aim of the paper is to answer the research question "How can fairness tools impact the understanding of fairness and the processes within a machine learning development team?".

The essay will be structured as follows. In the literature review, current key debates and approaches to the topic will be critically evaluated and juxtaposed. Following, the theoretical framework used for sense-making of the empirical data and the methodology are outlined. In the next part, the empirical findings from the data collection are presented. The discussion embeds the findings in the context of the current literature and presents the contributions of this study. The conclusion synthesises the central aspects of this study and identifies limitations of this work as well as possibilities for future research.

2. Literature Review

2.1. Machine Learning and Algorithmic Bias

In order to embed the issue of fairness in ML in its context, a short overview of ML and algorithmic bias will be provided in the following.

ML can be described as one of the most important subsets of AI (Rossi, 2018) that enables computers to learn something without being explicitly programmed to fulfil this specific task. It is the development of a mathematical model by the use of training data (Zhang, 2020) with the aim to make predictions and classifications (Binns, 2018). Algorithmic-decision making processes supported by ML are frequently described as potentially leading to fairer decisions because they are able to prevent human bias (Lepri et al., 2018). However, this has not proven correct.

With the increasing spread of ML applications in sensitive areas, the concern over the potential of these systems to reproduce and reinforce existing biases (Chouldechova & Roth, 2020) has received a lot of attention (Holstein & Vaughan, 2019). During the learning process, ML models are likely to adopt discriminatory correlations which means that these patterns are transferred from the analogous to the digital world (Binns, 2018). Given the nature of ML models, these biases are scaled and reinforced when translated to the digital world (Garcia, 2016). Bias in ML can have a number of different sources, such as the variables, the size of the training set or the decision to deploy an algorithmic system in a certain context itself (Lepri et al., 2018).

2.2. Definition of Fairness

A significant amount of effort within the literature on fair ML has been dedicated to the elaboration of a unified definition of fairness (Holstein & Vaughan, 2019). To date, there exists a number of different fairness definitions, however, no one is commonly accepted (Binns, 2018). It can be argued that the reason for the different definitions of fairness lies in the distinct theoretical lenses that are used (Lepri et al., 2018). These approaches can be divided into two main theoretical lenses.

2.2.1. Reductionist Approaches

Authors representing the reductionist perspective focus their work on the definition of fairness on mathematical considerations. Fairness in ML is understood as a static concept (Liu, Dean, Rolf, Simchowitz, & Hardt, 2018); and various definitions of fairness are considered and compared with regards to their limitations and shortcomings (Corbett-Davies, Pierson, Feller, Goel, & Huq, 2017). However, these are limited to a theoretical, mathematical level.

All authors within the reductionist lens agree on the fact that it is impossible to simultaneously satisfy all different fairness definitions or requirements (Kleinberg, Mullainathan, & Raghavan, 2017; Srivastava et al., 2019). On this basis, it is attempted to reduce and simplify various definitions in order to achieve a combined concept that satisfies as many requirements as possible. This is done by calculating the trade-offs between the individual definitions (Corbett-Davies et al., 2017; Kleinberg et al., 2017) with the aim to minimise the disadvantages (Chouldechova & Roth, 2020). This leads to a high level of abstraction and simplification that is performed by the authors.

Within this theoretical perspective, the importance of the specific context in which fairness in ML should be applied for the definition as well as other environmental and dynamic factors are neglected. It is assumed that one definition of fairness is applicable to all ML contexts (Agarwal, Beygelzimer, Dudík, Langford, & Wallach, 2018).

2.2.2. Contingent Approaches

Within this view, the authors share the assumption that since ML decisions affect people's lives, the definition of fairness should match people's perception of fairness (M. K. Lee

& Baykal, 2017; Saxena et al., 2019; Srivastava et al., 2019). This is underpinned by the idea that for fairness in ML to have a positive impact, it needs to be informed by the common sense of justice (Srivastava et al., 2019). By conducting empirical research, some authors discover that most people perceive demographic parity as the most suitable fairness notion (Srivastava et al., 2019), other authors find calibrated fairness to be the one that represents layperson's perception in the best way (Saxena et al., 2019).

The underlying assumption of this perspective can be described as inspired by the contingency theory. It is argued that fairness is highly contingent (Srivastava et al., 2019) and by investigating people's perceptions of fairness, a better understanding of the appropriateness of different definitions depending on the context can be generated (Saxena et al., 2019). Authors within this approach emphasise the importance of moving beyond the attempts to mathematically define fairness. Instead, they argue that it is imperative to acknowledge the contingency of fairness and establish a process of profoundly investigating the context in order to evaluate the best suitable definition of fairness for a specific problem (Lepri et al., 2018). They frequently criticise that the reductionist approaches to fairness don't reflect the highly context-specific reality of fairness in ML (M. S. A. Lee, Floridi, & Singh, 2021) and that instead, fairness needs to be negotiated every time according to the specific use case (Binns, 2018; Holstein & Vaughan, 2019).

The previous paragraphs have shown that the debate around the definition of fairness is controversial. While authors following a reductionist approach present fairness as easy to define and reduce it to a mathematical notion, authors within the contingent school of thought clearly argue for taking the context of the respective ML system into account when defining fairness.

2.3. Fairness Tools

A major theme within the literature on fair ML is the development of fairness tools. Against the backdrop of algorithmic bias and the necessity to ensure fairness in ML, researchers have developed these tools aimed at helping developers to detect and eliminate cases of bias in their models (M. S. A. Lee & Singh, 2021). These tools are typically classified according to the stage in the development process in which they can be deployed. They are typically divided into pre-processing tools, which eliminate bias from the dataset (Binns, 2018), in-processing tools that build fairness requirements directly into the algorithm (Calmon et al., 2017), and post-processing tools that evaluate and adjust the performance of a model after it has been developed (Berk et al., 2021).

The last years have seen a surge in the development of integrated toolkits in order to facilitate the implementation in development teams (Holstein & Vaughan, 2019). Companies can purchase these toolkits and integrate them into their processes. A commercial example is the AI 360 Toolkit that has been developed by IBM ('The AI 360 Toolkit', IBM Developer, 2021). Having reviewed and critically analysed the

broad and interdisciplinary literature on fairness tools, two main theoretical directions could be identified.

2.3.1. Deterministic, Technical-Rational View

Authors within the deterministic, technical-rational view focus on the technical functionalities of the tools and assume that they can be easily implemented and will automatically lead to more fairness in ML (Berk et al., 2021; Feldman, Friedler, Moeller, Scheidegger, & Venkatasubramanian, 2015; Hajian, Bonchi, & Castillo, 2016). In their works, they study the technical functionalities of fairness tools (Calders & Verwer, 2010), the shortcomings of individual tools (Hajian et al., 2016) and how tools can optimise the trade-off between accuracy and fairness (Kamiran, Calders, & Pechenizkiy, 2010; Kearns, Neel, Roth, & Wu, 2019).

One identified limitation is that most tools are not able to account for indirect discrimination that is introduced via proxies that are correlated with sensitive attributes. Therefore, they are not suitable to eliminate systemic discrimination (Berk et al., 2021). The shortcomings that are studied within this perspective are limited to technical disadvantages, such as a loss of accuracy within a tool as compared to another tool (Kamiran et al., 2010). Shortcomings in a wider, socio-technical sense, such as implementation challenges, are not addressed within this approach.

Another aspect studied within this approach is the fairness-accuracy trade-off. It is widely acknowledged that attempts to increase fairness in ML will always be a balance between the performance of the model and fairness (Menon & Williamson, 2018); there will be no technical solution that is able to maximise both fairness and accuracy (Berk et al., 2021). However, the deterministic approach does not address the fact that in particularly sensitive contexts, such as healthcare or criminal justice, this trade-off will need to be considered under another light than in less sensitive contexts. Instead, it is assumed that an optimised balance between fairness and accuracy is applicable for every model and every context (Chen, Johansson, & Sontag, 2018).

The assumption underpinning this approach is that fairness tools are fixed objects that are implemented in a development team and then automatically eliminate the bias from ML models; regardless of the context, the perception and understanding of fairness that is present in the development team, and the general environment of the system. By studying fairness tools from a deterministic lens, they are abstracted from their implementation in real-world development teams. The fairness definition underlying these approaches is a reductionist view. It is assumed that fairness can be formalised independent of the context and then assured through the use of a tool that can be implemented regardless of the context. This means that one tool represents a certain definition of fairness, but no guidance is provided as to which tool should be applied in which context.

2.3.2. Contingent, Socio-Technical View

Within the contingent, socio-technical view, authors study the implementability of fairness tools in practice and

the interaction between tool and developer adopting a socio-technical viewpoint. These works are partly located in the realm of human-computer-interaction research.

A major critic within this approach is the focus of the deterministic view on the technical development of fairness toolkits and an evaluation based on the functional shortcomings. Authors claim that the deterministic lens is narrowly focused on mathematical models and aims at modifying complex models without taking into account the human aspect of the issue and how their solutions can be successfully translated into practice (Abdul, Vermeulen, Wang, Lim, & Kankanhalli, 2018). Issues like the usability of these toolkits in commercial contexts (M. S. A. Lee & Singh, 2021), the challenges faced by practitioners when ensuring fairness in ML (Veale, Van Kleek, & Binns, 2018) or a developer team's general challenges and needs evolving around fairness (Holstein & Vaughan, 2019) are neglected in the field.

In the last years, some papers made an effort in bridging these gaps and answering open questions. M. S. A. Lee and Singh (2021) studied the gaps between the needs of practitioners and the functionalities of fairness toolkits. These include the user-friendliness and the contextualisation of these tools. The work from Richardson, Garcia-Gathright, Way, Thom, and Cramer (2021) focusses on implementation barriers of fair ML faced by practitioner teams and the effectiveness of fairness toolkits. One of the barriers was described as the lack of a contextualisation of a fairness definition within the tools which hindered a successful implementation (Richardson et al., 2021). A study with ML practitioners is conducted in the work of Holstein and Vaughan (2019) in order to investigate their challenges and needs in the work with fairness tools.

Adopting a contingent view on fairness in ML, authors highlight that a socio-technical issue like fairness cannot be tackled only with technological tools. Instead, the socio-technical environment of ML systems and organisational processes need to be studied in order to ensure fairness (Holstein & Vaughan, 2019). The conclusion that is drawn by all the authors investigating fairness tools in ML from a socio-technical perspective is that there is a "disconnect" between the real-world situation of ML teams and the literature (Holstein & Vaughan, 2019; M. S. A. Lee & Singh, 2021; Veale et al., 2018).

One example is the call for a more holistic approach. Interviewees emphasise the importance of having a better communication between different teams that develop a ML model, such as data scientists and developers, and call for the development of tools that facilitate this coordination (Holstein & Vaughan, 2019). Another challenge expressed by practitioners is the amount of time that is required to monitor the fairness of the ML models and to work with the tools. The lack of sufficient time and human resources constitutes a major challenge in ensuring fairness (Holstein & Vaughan, 2019). Apart from the lack of time, the lack of knowledge in fairness compared to tools that require high-level expertise knowledge is identified as a challenge (M. S. A. Lee & Singh, 2021). Researchers also identify the scepticism toward a

full automation of fairness processes due to quantifying challenges of fairness as a major disconnect between practice and academia. While the literature proposes a full automation with the support of fairness tools, interviewees voice concerns over this elimination of a human element (Holstein & Vaughan, 2019).

A smaller research strand within the papers that investigate fairness tools through a contingent, socio-technical lens is focused on the way that the use of fairness tools impacts the fairness understanding of developers that interact with them. While there are many papers that study the perception of fairness of lay people or users of the systems, very few studies focus on developers (Woodruff, Fox, Rousso-Schindler, & Warshaw, 2018).

Yan et al. (2020) examine the way tools enable developers to understand the sources of bias in datasets and draw meaningful conclusions. The specific toolkit used in the paper enables the developers to assess the fairness of the ML model and also leads them to reflect about systemic biases inherent in the system. The way the practitioners use the tool to explore different explanations and improve their understanding of fairness is analysed by using the sensemaking theory. The authors conclude that through interaction, fairness tools have the potential to support developers in their "sensemaking process" of assessing fair ML (Yan et al., 2020).

The work from Bellamy et al. (2019) explore how the specific toolkit AI Fairness 360 impacts the understanding of fairness of developers. The authors conclude that the toolkit shapes the knowledge of the users in terms of fairness, detection of unfairness and mitigation of bias (Bellamy et al., 2019).

2.4. Research Question

The literature review has revealed a research gap on the impact that fairness tools have on the understanding of fairness and the processes within a development team. Specifically, to the best of my knowledge, this aspect has not been investigated empirically to date. Given the potential consequences it might have for the implementation of fairness tools, it is imperative to understand the impact they can have when used in a development team. Therefore, this dissertation aims to answer the research question "How can fairness tools impact the understanding of fairness and the processes within a machine learning development team?". In order to answer this research question, a case study has been conducted. The theoretical framework used to analyse the findings of this case study as well as the details of this case study and the methodology used will be elaborated in the following parts.

3. Theoretical Framework

During the sense-making process of the empirical findings, two theories have been deployed in order to distil the full meaning of the data.

The first theory is Kallinikos' idea of the "regulative regime of technology". He states that technology can be

considered a "distinctive regulative regime" that forms the operations within a company in important ways and therefore influences social practice. Technology is described as an objectified system consisting of processes and forces that have an impact on tasks carried out in the environment it is implemented into. One aspect of that impact is the forming of perceptions, professional rules and routines. It is important to note that regulation in this sense is not defined as a rigid, constraining force, but rather used in a broader sense (Kallinikos, 2009, 2010).

The second theory that has been used is the theory of the co-construction of users and technologies by Oudshoorn and Pinch. The theory can be understood as an augmentation of the social construction of technological systems theory (SCOT). In the co-construction theory the main idea is that with their creative capacity, users shape technology in important ways throughout the whole development and implementation process. It examines the ways users impact technology by using, altering, resisting and reconfiguring it, and the ways technology impacts users by transforming and defining them (Oudshoorn & Pinch, 2003).

4. Research Design and Methodology

4.1. Case Study Method and Case Selection

In order to answer the research question and to create a fit between the research question and the data collection, I have decided to conduct a qualitative data collection. The objective of the data collection was to understand the impact of fairness tools on the understanding of fairness and on the processes within the development team. The purpose was to generate hypotheses about the interaction between developers and fairness tools in ML development teams as well as on their definition of fairness. Therefore, a qualitative data collection was the most suitable method for answering my research question.[1]

The specific research method that has been chosen is a single case study. Since the research question aimed at explaining the impact of fairness tools on development teams, a case study was the most suitable research method (Benbasat, Goldstein, & Mead, 1987). The role that fairness tools play and the impact they have needed to be understood in depth in order to answer the research question. Conducting a case study allowed me to obtain a "holistic and real-world perspective", which was needed to address the previously identified research gap. The case study as a data collection method matched the research question because of its focus on explaining the how and its focus on contemporary events (Yin, 1994). Furthermore, since the phenomenon that was to be studied is highly complex, it was not possible to remove it from its context and study it separately (Gerring, 2004).

The unit of analysis of the case study was the ML development team of a large, globally operating company. Due to

legal reasons, it has been agreed with the interviewees that the name of the company will not be revealed in this work. The company is a multinational technology company that is headquartered in the U.S. and operates in approximately 170 countries. The development team that was used for the case study is embedded in the branch of the company that develops, sells and implements AI tools, solutions and applications. The team consults its clients with the aim to ensure a successful adoption of AI and an ethical, responsible use of the technology.

The tool that is being used by the development team is a fairness toolkit that integrates features for explainability and fairness. It aims at giving the developer a comprehensive view of the model and to evaluate the biases inherent in it. Therefore, it examines the data and retraces how the model reached a decision, and on basis of that evaluates whether the model is biased or not. If the tool finds that the model is biased, it alerts the developer, who can then decide how to proceed with this information. The tool also indicates ways to alter the model in order to eliminate the bias; these include protecting sensitive features in the training data so that these features will not be taken into account for the decision-making.

The specific development team has been selected as a unit of analysis for various reasons. Firstly, the operational range spans the whole development cycle of ML. This means that the team develops the models, selects the training data and trains the algorithm before selling it to their customers. Secondly, the team uses a fairness toolkit on a daily basis since many years. This enabled me to investigate and understand the impact that this toolkit has on the team. Since an open-source version of the toolkit is available online, I was able to study the tool itself in greater detail which prepared me for the interviews and enabled me to understand the details of the functionality. Lastly, the company has agreed to grant access to developers for interviews. Considering the high sensitivity of the topic and legal requirements from corporates, this proved to be difficult to find. Many development teams have been contacted for this case study; however, many declined a cooperation for different reasons. Some claimed not to experience any fairness issues in their AI applications, others are aware of the topic but do not employ any specific fairness tools. Others, as mentioned before, use fairness tools but were not able to disclose any information due to company guidelines.

4.2. Data Collection Strategy

The case study consisted of six semi-structured, in-depth interviews that have been conducted with members from the development team. The selection of interviewees was made based on availability, tenure and position within the team. It was important for this research to achieve diversity among the interviewees in terms of tenure and position since these factors probably influence the experience of working with the fairness tool. Given the limited timeframe of this study and the current size of the team, it was not feasible to interview

---

[1]Taken from 2021, "How can development teams effectively use fairness tools in order to ensure fairness in Machine Learning?", Research Proposal for MY401 Course, LSE

more developers. The aim of the interviews was to understand how different team members view the role and the impact of the fairness toolkit on their work and on their understanding of fairness.[2]

The interviews were conducted over Zoom and each lasted approximately 35-45 minutes. The interview guideline included questions on the interviewees' understanding of fairness, their experience of working with the tool and how the tool impacts their understanding of fairness and the processes within the team. The participants were informed about the research beforehand with an information sheet and gave their consent to the interview being recorded. After the interviews, they were fully transcribed by using the interview recordings.

Apart from the interviews, an analysis of documents detailing the fairness toolkit and its functionality was conducted.

With the aim to cross-validate certain findings, an additional interview has been conducted with an employee from a company that develops and sells fairness toolkits. This interview enabled me to receive a second perspective to some points mentioned in the interviews with the developers and contextualise them. This interview was also conducted via Zoom and was fully transcribed afterwards.

One of the advantages of interviews in this case is that they are suited for exploring perspectives, which was one of the aims for the data collection. Conducting interviews allowed the collection of relevant and focused data. One major weakness is that the interviewees might not be able to answer all the questions or that given the artificial setting, the answers don't always reflect the reality. This challenge was sought to be addressed with the diversity among the interviewees in terms of position and experience within the team.

4.3. Data Analysis Strategy

According to the qualitative data collection, the data has been analysed through a thematic analysis as proposed by Attride-Stirling (2001). In the first step, a coding framework was developed based on the literature review and the research question, and also on salient issues from within the data. Using this coding framework, the text has then been divided into separate sections (Attride-Stirling, 2001). These sections have then been read and analysed to identify common themes across the different interviews (Fereday & Muir-Cochrane, 2006). These themes have then been rearranged into a network of themes and clustered into basic themes, organising themes and global themes. In a next step, these thematic networks have then been further explored to detect the "underlying patterns". Consequently, the main themes within the networks have been summarised and interpreted with regard to the research question (Attride-Stirling, 2001). The

thematic analysis has helped answering the research question by identifying the most important and salient aspects of the collected data and enabled me to interpret them. [3]

5. Findings

After having evaluated the interviews through a thematic analysis, three main themes emerged: Definition of Fairness, Limitations and Implementation Challenges of Fairness Tools, and the Impact of Fairness Tools. In the following, the findings corresponding to the themes will be presented.

5.1. Definition of Fairness

All the respondents agree on the difficulty of defining fairness in ML and struggle to pin the different notions they are aware of down to one unequivocable definition. They frequently equate it to the concepts of equal opportunity and equal outcome or associate it with transparency, explainability and accountability. All of the interviewees mention the importance of defining fairness differently for separate stakeholder groups.

One respondent also acknowledges that the working definition of fairness that is used in the industry will most likely differ from the one agreed on in academia.

> *"From how we work in the industry, is we have a what we call an enterprise design thinking approach for data and AI, where we look at the personas and user groups for that specific AI solution."*

A ML solution will then be judged in terms of its fairness by evaluating whether it is fair to all the different users and personas that use the system. This evaluation is conducted by ethics experts in order to then assess whether a solution is fair or not. The result of this evaluation process can then be considered as the definition of fairness.

The interviewees also emphasise that fairness needs to be defined every time depending on the context and the specific use case. In some cases, a certain form of bias might be acceptable or even wanted, which requires developers to define bias and fairness individually for every model. Therefore, as mentioned by the interviewees, the choice of the appropriate fairness definition for a specific context is not made by the fairness tool, but by the developers that then configure the tool accordingly.

It is also highlighted that the importance does not lie in choosing a particular concept of fairness, but in properly defining, understanding and operationalising one notion intentionally and transparently. For every use case, it is imperative to reveal the model of fairness that underpins the development of the model.

5.2. Limitations and Implementation Challenges of Fairness Tools

One major theme present in the interviews are the limitations and implementation challenges evolving around the use of fairness tools. These can be clustered into the three sub-themes Holistic Approach, Accuracy Trade-Offs, and Implementation Challenges.

5.2.1. Holistic Approach

Interviewees emphasise the importance of having a holistic approach when talking about fairness in ML. Instead of only focussing on the part of the development team that is concerned with the building of the model, it is fundamental to take into account the whole process of development, which includes the ML team as well as other teams that share responsibility for the development. Given the dependencies that frequently exist between various different processes, the necessity arises to monitor the whole development process to detect "opportunities for debiasing" and also weak points that can render the fairness efforts useless.

One interviewee highlighted in particular the importance of ensuring a tight connection between the business and the development area. Since the business problems that are to be solved with a ML solution and the requirements for the ML model itself are created within the business team and then transferred to the development team, it is imperative to involve the business department in the fairness efforts. This strong connection between business and development is described as success-defining, however, it is often not enabled by the use of fairness tools, as reported by interviewees.

5.2.2. Accuracy Trade-Offs

In terms of accuracy and the overall performance of a model that might be reduced when fairness tools are applied on it, all of the interviewees reject the idea of always prioritising one of the two parameters and highlight the need to consider the respective context. The process of developing a model always involves a "tweaking" of the individual parameters to achieve the desired performance and accuracy rate. The factors that guide, or should guide, this process are industry standards, the company and the corporate culture as well as the nature of the ML solution that shall be deployed, such as the objective and the data that is available.

It became evident during the interviews that the difficulty hereby mainly lies in the lack of a clear definition of the terms fairness and accuracy.

> "What is accurate then? Is it your historical data that is accurate? Or any assumption on how the future should look like? The golden standard that you have created artificially? Is that accurate?"

This quote clearly illustrates this challenge in labelling a model as fair or accurate. The respondent concludes that a fairness tool is not able to deal with this problematic. Instead, best practices, design methods and an ethics board need to be implemented in order to evaluate the specific context and decide which factors should be prioritised over others. This means that the decision of the trade-off is not left to the tool, but negotiated beforehand and then the tool is configured accordingly.

One of the respondents highlights the importance of moving beyond the dichotomy of accuracy and fairness. When bias is detected in a model, and it is decided to remove this bias, this has to be done regardless of the accuracy, since an unfair model is not accurate either.

> "There is this habit of some people in the field to become very focussed on improving this one number, sometimes accuracy, and then they become tunnel-visioned on that. I don't think that is a very good way to develop models, fairness-aside".

This quote illustrates that the formulation of a fairness-accuracy trade-off can be obsolete since accuracy is not the best and only indicator for a good model and if fairness shall be added to the set of requirements of a model, it should be treated equal to the other requirements, such as accuracy.

5.2.3. Implementation Challenges

During the interviews, various challenges to the implementation of fairness tools in everyday work practices have been stated. One of them is the lack of time in development that can be dedicated to fairness efforts. This is exacerbated by the fact that in many cases, when a tool indicates that a model is biased, further data needs to be collected in order to eliminate the bias from the model. This lack of time presents a hurdle for development teams in the implementation of fairness tools. This can be described as a fairness-cost trade-off, cost in that sense meaning both money and time and it advantages big companies compared to smaller companies which typically have less resources at their disposal.

Another major challenge is the lack of skilled talent. Properly understanding the tools and their functionality requires high-level math skills which are, according to the interviewees, not always represented in a development team. This is connected with the issues that companies who are about to develop a ML model often experience difficulties in finding developers and data scientists. Once they have the necessary resources, they want to start developing their model immediately without taking fairness into account. Also the fact that some toolkits are aligned with a traditional waterfall-development process makes it hard to implement them in today's iterative and agile development environments.

5.3. Impact of Fairness Tools

The theme impact of fairness tools consists of the three sub-themes Understanding of Fairness, Processes and Practices, and Fairness Automation.

5.3.1. Understanding of Fairness

The respondents state that the use of fairness tools impacts and shapes the understanding of fairness within a development team. The use of a certain tool shows the team

what is fair by indicating certain models as biased and others not.

> "Yeah, it contributes [to the understanding of fairness], because it just visibly shows how the data is biased. And sometimes we do not even recognise as human beings that our data that we have (..) is biased."

The interviewees explain how, by making visible the hidden bias in the data, the model contributes to a change in the understanding of fairness because of its functionality. One respondent states that sometimes, the results of the fairness tools are surprising because the development team did not expect any bias to occur in the respective model. If the model then indicates an instance of unfairness, it shows to the developers that their assessment was not correct and adds to their definition and understanding of which models are fair and which are not. Thus, the interviewees agree on the fact that the technology contributes to the awareness of fairness among the people who are working with it.

5.3.2. Processes and Practices

Apart from shaping the understanding of fairness, tools deployed in development teams to ensure the fairness of ML also limit the scope of actions of developers and the possibilities in their work, as reported by the interviewees. By giving an alert when the model seems to be biased or behaves in an unwanted way, the model gives the human the sign to pause the development process and take one step back to evaluate and, if needed, correct the bias. By giving the developers this insight, the tool enables the team to engage in a discussion and pivot the approach. This means that the developer needs to evaluate the bias instance evaluated by the model and, if it is found to be a bias that should be eliminated, take action to ensure the fairness of the model. Thus, the interviewees view the tool as limiting their possibilities of work – however not in a negative sense, but rather understood as additional support.

One respondent draws an analogy between fairness tools and compliance guidelines within a company. Given that supervising and controlling nature of fairness tools, some interviewees describe these as regulating the ML development process within a team.

> "If you are (. . .) introducing a fairness tool, then at one point, the people or the data scientist is not able to do it, like he did before. And then of course, it's something like a regulation point."

The fairness tool does not automatically change the development process or force actions to be taken, it rather makes the development team reflect about the model and decide whether the biases indicated by the tool shall be eliminated or not. This means that the fairness tool does not solve the bias issue automatically and independently, but requires the interaction from the developer side.

Interviewees also report that the usage of a tool impacts the development process by introducing additional checkpoints. They create a certain structure around the development and testing process and introduce new steps to it, so that the development process will be different than it was before the implementation of the tool.

5.3.3. Fairness Automation

The interviewees agree on the fact that fairness tools should not automatically eliminate bias in the future, instead of alerting a human to then make a decision. Some state that, given the sensitive and highly contingent context of ML systems, the process of eliminating bias will and should not be automated in the future. Therefore, human oversight will always be needed to make sense of the results given by a fairness tool. Others claim that the tools are still in a nascent state of development which is why they are currently not able to perform actions automatically. They agree on the importance of augmenting the regulative impact of fairness tools with human controls. It is emphasised that a fairness tool should not be used as a standalone indication of bias, but rather in addition to human judgement.

In contrast, one employee of a company which develops and sells fairness tools has been interviewed in order to cross-validate the findings and he emphasises that the tools are planned to automatically conduct the process of bias elimination in the future and thus replace the human element in the loop.

6. Discussion

In the following, it will be reflected on the empirical results, connecting them to the findings from the literature review and the research question on how fairness tools can impact the understanding of fairness and the processes within a machine learning development team. Consequently, the contributions to research and implications for practice will be presented.

6.1. Reductionist and Contingent Approaches to Fairness

The findings from the interviews show the impossibility of pinning fairness down to one single definition and that fairness needs to be treated as a contingent, context-specific concept. This finding falls in line with the idea of contingent approaches to fairness (Holstein & Vaughan, 2019; M. S. A. Lee et al., 2021; Srivastava et al., 2019). Similar to these ideas, the findings reject the idea of following a reductionist approach to fairness. Instead of attempting to analyse the trade-offs between several definitions of fairness and combine them into one simplified notion (Corbett-Davies et al., 2017; Kleinberg et al., 2017), the findings support the idea of defining fairness according to the specific use case and context. Conclusively, the empirical findings support the contingent lens on the definition of fairness and reject reductionist approaches.

6.2. Rejection of Deterministic / Technical-Rational Approaches to Fairness Tools

Analysing the empirical findings, it can be noted that they align with the socio-technical, contingent perspective, rejecting the deterministic view on fairness tools. Instead of viewing these tools as fixed solutions that can be implemented easily and directly lead to fairness, the interview results confirm the challenges encountered when implementing them and the complexity involved in the interaction between developer and tool.

The implementation challenges mentioned in the literature have been mostly confirmed by the empirical findings, namely the lack of time in development teams that can be dedicated to fairness efforts, the lack of human resources and skilled talent and the difficulties in collecting additional data Holstein and Vaughan (2019). One additional challenge that has not been mentioned in the literature to the best of my knowledge is the misfit between fairness toolkits that are sometimes designed for a waterfall development process and today's agile, iterative processes.

These implementation challenges described by the interviewees serve as a support to the claim made in the contingent approach that tools need to take into account the socio-technical environment of the ML model in order to be successfully implemented. It also shows that the claim from the deterministic perspective that tools do not face any difficulties when being implemented in a team (Berk et al., 2021; Hajian et al., 2016) does not hold true.

In general, the empirical findings emphasise the importance of considering the context of ML systems. Examples are the trade-off between accuracy and fairness which needs to be negotiated according to the respective context, as well as the holistic approach that is requested in order to consider the whole development context of a system. This falls in line with the socio-technical approaches in the literature (Holstein & Vaughan, 2019; M. S. A. Lee & Singh, 2021) which argue that fairness toolkits are too focused on the technical aspects and are neglecting the context of the respective systems.

6.3. Fairness Tools as Regulative Regimes

Concerning the impact that fairness tools have in development teams, the empirical findings reveal that fairness tools impact and shape the understanding of fairness within a team in important ways. In this sense, the technology of fairness tools contributes to both the awareness and the understanding of fairness by indicating the developers instances of unfairness. Fairness tools can also impact the processes within a team by creating a structure around the development and testing process and introducing new steps and checkpoints. What makes these findings significant is that they can be seen as an enactment of Kallinikos' theory on technology as a regulative regime (Kallinikos, 2009).

By alerting biased decisions made by an algorithm and thus indicating the developer to modify the model to eliminate this bias, the tool clearly shows the developer instances of bias and thus shapes their understanding of fairness. The output of the technology determines the next steps for the developers and therefore forms the way they define and understand fairness. By limiting the scope of actions for developers, the tools regulate the development process and enable the team to engage in a discussion and pivot their approach. Thus, the technology of fairness tools can be understood as a "regulating practice" that shapes the operations and development processes within a team and governs the social practice by raising the awareness for fairness and shaping its understanding (Kallinikos, 2010). In future investigations, it would have to be investigated in a comparative study how the implementation of different tools in one development team impacts the understanding of fairness.

The application of Kallinikos' theory on the impact of fairness tools can be regarded as an augmentation to the current, limited research on the impact of fairness tools on the understanding of fairness of developers. The empirical findings confirm the ideas present in the literature on how fairness tools enable developers to better understand unfairness and fairness and shape their knowledge in terms of bias mitigation (Bellamy et al., 2019; Yan et al., 2020). The sensemaking theory used in the existing literature cannot be confirmed since the methodological setting used in this study did not allow for a close observation of sensemaking loops that are supported by fairness tools. Thus, the analysis of the impact of fairness tools through the lens of Kallinikos' theory reveals the regulating impact these tools have on developers interacting with them.

Another aspect of the role that fairness tools play in development teams is the degree of automation and independence of these tools. Here, the empirical findings reflect the results from the literature review. While the interviewee who works in a company that develops fairness tools emphasises that these tools will function automatically in the future, without including a human in the loop, the interviewees are more concerned about a full automation. Given the sensitive and context-dependent area, they highlight the importance of human control, and this scepticism is also reflected in the literature (Holstein & Vaughan, 2019). This raises questions about the future role of fairness tools in development teams and their desired degree of independence and automation. This will need to be investigated in more detail in future research.

6.4. Co-Construction of Developers and Fairness Tools

The empirical findings have revealed that, apart from the fairness tool impacting the understanding of fairness of the developer and the processes within the team, the tools are also shaped by the choices made by the developers. This becomes evident at the choices developers make about the definition of fairness and desired fairness-accuracy trade-offs depending on the context. Instead of the tool having these choices already pre-programmed, the developers shape them by making choices and configuring the tools accordingly. This rejects the deterministic view on fairness tools that assumes that choices around fairness, such as the definition of fairness and fairness-accuracy trade-offs can be simplified and built into a fairness tool.

Applying Oudshoorn and Pinch's theory on the co-construction of users and technologies, this process of mutual shaping between developer and fairness tool can be understood as a process of co-construction. The character of fairness tools as regulative regimes can then be viewed as an embedded part of this process. By acting as a regulative regime and shaping the understanding of fairness and development processes, the fairness tools transform and influence the developers. Simultaneously, developers shape fairness tools by making choices and configuring the tools accordingly.

This finding also supports the socio-technical perspective and challenges the deterministic perspective on fairness tools because it confirms that tools cannot be simply implemented into a development team and expected to yield in more fairness. Fairness tools are no fixed objects that function independently from the development team and the socio-technical environment. Instead, they are actively shaped by the developers through choices and then configured accordingly; and the developers and development processes are also shaped and influenced by the fairness tools. Rather than focussing on the technical functionalities and shortcomings of fairness tools, this process of mutual shaping, of co-construction between user and technology, needs to be brought into focus.

6.5. Contributions to Theory and Implications for Practice

These findings contribute to the existing literature by explaining the impact of tools on the understanding of fairness with the use of the theory on regulative regimes which challenges the deterministic, technical-rational view on fairness tools which assumes that choices on fairness are built into a tool that then directly eliminates bias in ML and leads to fairness. Simultaneously, it supports the contingent, socio-technical perspective by showing the importance of the context and environment of ML systems.

Furthermore, the findings contribute to the understanding of the interaction between developer and fairness tool by describing the mutual shaping of developer and fairness tool as a co-construction process, of which the regulative character of fairness tools can be understood as one part.

The empirical findings also entail important implications for the practice of development teams. Having analysed how developers construct fairness tools and vice versa, development teams should make an increased effort to educate all members of the team on fairness in ML. Since the choices, both conscious and subconscious, taken by developers, influence the functionality of fairness tools, they should understand for instance what consequences different definitions of fairness have and how accuracy and fairness can be traded off against each other, depending on the context. Apart from that, companies that implement fairness tools in their teams should be aware of this co-construction process and actively shape it. The findings also show that companies cannot simply implement fairness tools and expect them to eliminate algorithmic bias from their models. Instead, they need to actively construct the functionality of the tool.

For companies that develop and sell fairness toolkits, the findings signify that they should educate their clients in more detail about the impact of fairness tools and how they are co-constructed and influenced by the choices made by developers.

7. **Conclusion**

The aim of this research was to investigate the impact that fairness tools can have on developers and on the processes within a development team.

The literature review has revealed the tensions between the deterministic, technical-rational and the socio-technical, contingent approach to fairness tools. It has been analysed that the existing, but very limited literature on the impact of fairness tools argues that fairness tools impact the knowledge of developers on fairness tools. The empirical findings from the case study conducted with the ML development team has confirmed this and revealed the extent of this impact. The application of Kallinikos' theory on the regulative regime of technology brought to light the character of fairness tools as regulative regimes in development teams.

The empirical findings on the impact of fairness tools on developers, against the backdrop of the current literature, led to a novel understanding of the relation between developer and tool. The use of Oudshoorn and Pinch's theory enabled an understanding of the interaction between the developer and the technology, the fairness tool, as a process of co-construction in which both elements shape and construct the other.

With these two main conclusions, the research question can be answered as follows. By shaping the understanding of fairness and the processes within a development team, fairness tools act as a regulative force. This process of regulating has to be understood in the wider context of a co-construction process between the technology, in this case the fairness tool, and the user, who is the ML developer in this case.

These findings contribute to the limited literature on the impact of fairness tools by confirming and extending the impact of tools on the knowledge of developers and offering a theoretical framework to analyse the impact of this technology. It also adds to the contingent, socio-technical perspective on fairness tools by highlighting the importance of the context for the functionality of the tools. Lastly, it challenges the deterministic perspective present in the current literature by showing how the impact of fairness tools is not determined by their technical characteristics, but by the way they are constructed through conscious choices of the developers.

7.1. Limitations & Further Studies

Although this research contributes to both research and practice in significant ways, it also exhibits important limitations.

Firstly, it is imperative to note that the limited scope of this work constrains it to the most relevant and significant findings that answer the research question. In future work,

the impact of fairness tools on developers and the concrete processes of co-construction will have to be studied further and in more detail.

Another limitation is the possibly constrained eagerness of the respondents to talk openly in the interviews. Given the high sensitivity of the topic, it is possible that certain thoughts have not been expressed freely. This could be overcome in future works by the conduction of an anonymised study, like it is also used by Binns (2018).

Regarding the general reliability of the data, it can also be noted that through an observation of the development team over a longer period of time, richer data could have been collected concerning the interaction with the tool. Due to the COVID-19 Pandemic, the majority of employees is currently still working from home, which is why this has not been possible and should be conducted in future works on the topic.

References

Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., & Kankanhalli, M. (2018). Trends and Trajectories for Explainable, Accountable and Intelligible Systems: An HCI Research Agenda. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–18).

Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, *6*, 52138–52160.

Agarwal, A., Beygelzimer, A., Dudík, M., Langford, J., & Wallach, H. (2018). A reductions approach to fair classification. In *International Conference on Machine Learning* (pp. 60–69).

Attride-Stirling, J. (2001). Thematic networks: An analytic tool for qualitative research. *Qualitative Research*, *1*(3), 385–405.

Bellamy, R. K. E., Dey, K., Hind, M., Hoffman, S. C., Houde, S., Kannan, K., … Mojsilović, A. (2019). AI Fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias. *IBM Journal of Research and Development*, *63*(4/5), 4–1.

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The case research strategy in studies of information systems. *MIS Quarterly*, 369–386.

Berk, R., Heidari, H., Jabbari, S., Kearns, M., & Roth, A. (2021). Fairness in Criminal Justice Risk Assessments: The State of the Art. *Sociological Methods & Research*, *50*(1), 3–44.

Binns, R. (2018). Fairness in Machine Learning: Lessons from Political Philosophy. *Conference on Fairness, Accountability and Transparency*, 149–159.

Calders, T., & Verwer, S. (2010). Three naive bayes approaches for discrimination-free classification. *Data Mining and Knowledge Discovery*, *21*(2), 277–292.

Calmon, F. P., Wei, D., Vinzamuri, B., Ramamurthy, K. N., & Varshney, K. R. (2017). Optimized pre-processing for discrimination prevention. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 3995–4004).

Chen, I., Johansson, F. D., & Sontag, D. (2018). Why Is My Classifier Discriminatory? *Advances in Neural Information Processing Systems*(31), 3543–3554.

Chouldechova, A., & Roth, A. (2020). A snapshot of the frontiers of fairness in machine learning. *Communications of the ACM*, *63*(5), 82–89.

Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, A. (2017). Algorithmic Decision Making and the Cost of Fairness. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 797–806).

Feldman, M., Friedler, S. A., Moeller, J., Scheidegger, C., & Venkatasubramanian, S. (2015). Certifying and Removing Disparate Impact. In *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 259–268).

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, *5*(1), 80–92.

Gade, K., Geyik, S. C., Kenthapadi, K., Mithal, V., & Taly, A. (2019). Explainable AI in Industry. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 3203–3204.

Garcia, M. (2016). Racist in the Machine: The Disturbing Implications of Algorithmic Bias. *World Policy Journal*, *33*(4), 111–117.

Gerring, J. (2004). What is a case study and what is it good for? *American Political Science Review*, *98*(2), 341–354.

Hajian, S., Bonchi, F., & Castillo, C. (2016). Algorithmic Bias: From Discrimination Discovery to Fairness-aware Data Mining. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 2125–2126).

Holstein, K., & Vaughan, J. W. (2019). Opportunities for Machine Learning Research to Support Fairness in Industry Practice. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–16.

IBM Developer. (2021). *The AI 360 Toolkit: AI models explained*. https://developer.ibm.com/articles/the-ai-360-toolkit-ai-models-explained/.

Kallinikos, J. (2009). The regulative regime of technology. In *ICT and innovation in the public sector* (pp. 66–87). Springer.

Kallinikos, J. (2010). *Governing through technology: Information artefacts and social practice*. Springer.

Kamiran, F., Calders, T., & Pechenizkiy, M. (2010). Discrimination aware decision tree learning. In *2010 IEEE International Conference on Data Mining* (pp. 869–874). IEEE.

Kearns, M., Neel, S., Roth, A., & Wu, Z. S. (2019). An Empirical Study of Rich Subgroup Fairness for Machine Learning. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 100–109). ACM.

Kleinberg, J., Mullainathan, S., & Raghavan, M. (2017). Inherent Trade-Offs in the Fair Determination of Risk Scores. *Proceedings of Innovations in Theoretical Computer Science (ITCS)*.

Lee, M. K., & Baykal, S. (2017). Algorithmic Mediation in Group Decisions: Fairness Perceptions of Algorithmically Mediated vs. Discussion-Based Social Division. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 1035–1048). ACM.

Lee, M. S. A., Floridi, L., & Singh, J. (2021). Formalising trade-offs beyond algorithmic fairness: lessons from ethical philosophy and welfare economics. *AI and Ethics*, 1–16.

Lee, M. S. A., & Singh, J. (2021). The landscape and gaps in open source fairness toolkits. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–13).

Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2018). Fair, Transparent, and Accountable Algorithmic Decision-making Processes: The Premise, the Proposed Solutions, and the Open Challenges. *Philosophy & Technology*, *31*(4), 611–627.

Liu, L. T., Dean, S., Rolf, E., Simchowitz, M., & Hardt, M. (2018). Delayed impact of fair machine learning. In *International Conference on Machine Learning* (pp. 3150–3158). PMLR.

Mattu, J. L., Kirchner, L., & Surya, J. A. (2016). Machine Bias. *ProPublica*. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Menon, A. K., & Williamson, R. C. (2018). The cost of fairness in binary classification. In *Conference on Fairness, Accountability and Transparency* (pp. 107–118). PMLR.

Oudshoorn, N., & Pinch, T. (2003). How Users Matter: The Co-Construction of Users and Technologies. *The MIT Press*.

Pasquale, F. (2015). *The black box society*. Harvard University Press.

Richardson, B., Garcia-Gathright, J., Way, S. F., Thom, J., & Cramer, H. (2021). Towards Fairness in Practice: A Practitioner-Oriented Rubric for Evaluating Fair ML Toolkits. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–13).

Rossi, F. (2018). Building trust in artificial intelligence. *Journal of International Affairs*, *72*(1), 127–134.

Saxena, N. A., Huang, K., DeFilippis, E., Radanovic, G., Parkes, D. C., & Liu, Y. (2019). How do fairness definitions fare? Examining public attitudes towards algorithmic definitions of fairness. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 99–106).

Srivastava, M., Heidari, H., & Krause, A. (2019). Mathematical Notions vs. Human Perception of Fairness: A Descriptive Approach to Fairness for Machine Learning. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2459–2468).

Veale, M., Van Kleek, M., & Binns, R. (2018). Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14.

Woodruff, A., Fox, S. E., Rousso-Schindler, S., & Warshaw, J. (2018). A Qualitative Exploration of Perceptions of Algorithmic Fairness. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–14).

Yan, J. N., Gu, Z., Lin, H., & Rzeszotarski, J. M. (2020). Silva: Interactively Assessing Machine Learning Fairness Using Causality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–13).

Yin, R. K. (1994). Case study research: Design and methods, applied social research. *Methods Series*, *5*.

Zhang, X.-D. (2020). *A Matrix Algebra Approach to Artificial Intelligence*. Springer Singapore.

A Data Protection Law Analysis of the New German Electronic Patient Records (ePA) in the Light of the GDPR

Eine datenschutzrechtliche Betrachtung der neuen elektronischen Patientenakten in Deutschland nach Einführung der DSGVO

Samuel Hammer

FOM Hochschule für Oekonomie und Management

Abstract

A lengthy discussion about the digitization and modernization of the German healthcare system is followed by the obligation for health insurance companies to introduce electronic patient records (ePA). This regulation faces resistance from Germany's highest data protection authority (BfDI). On the basis of relevant commentary literature and considering the arguments put forward by the stakeholders this study examines, whether the criticism of the authority is justified and whether a violation of the GDPR could lie in the implementation of the ePA. As a result of the study, no such violation can be determined. Especially the conditions for the effectiveness of consent to data processing are given. The introduction of the German ePA will take place in two stages, with the second stage including improvements regarding data protection. Thus, the result of the work can also be applied ,a maiore ad minus' to the second stage which is planned for 2022. It remains unclear whether the data protection authority (BfDI) will take further legal measures. This study affects also other research topics, such as the "right to data processing" or the role of German data protection authorities in legislative processes.

Zusammenfassung

Einer langwierigen Diskussion um die Digitalisierung und Modernisierung des deutschen Gesundheitssystems folgt 2021 die Verpflichtung der gesetzlichen Krankenkassen zur Einführung der elektronischen Patientenakte (ePA). Diese Regelung stößt bei Deutschlands oberster Datenschutzbehörde (BfDI) auf Widerstand. Anhand einschlägiger Kommentarliteratur und unter Berücksichtigung der durch die Interessenvertreter vorgebrachten Argumente untersucht diese Arbeit, ob die Kritik der Behörde begründet ist und ob ein Verstoß gegen die DSGVO mit der Einführung der ePA vorliegen könnte. Im Ergebnis lässt sich ein solcher Verstoß nicht feststellen. Insbesondere liegen die grundsätzlichen Anforderungen an die Wirksamkeit einer Einwilligung in die Datenverarbeitung vor. Die Einführung der ePA erfolgt in zwei Ausbaustufen, wobei die zweite Stufe datenschutzrechtlich relevante Verbesserungen beinhaltet. Daher gilt das Ergebnis der Arbeit „a maiore ad minus" auch für die 2022 geplante zweite Ausbaustufe. Offen bleibt, ob die Datenschutzbehörde (BfDI) weitere rechtliche Maßnahmen ergreifen wird. Die Untersuchung berührt weitere Forschungsthemen, wie beispielsweise das „Recht auf Datenverarbeitung" oder die Rolle der Datenschutzbehörden in Gesetzgebungsverfahren.

Keywords: Elektronische Patientenakte; ePA; DSGVO; Datenschutz; Telematikinfrastruktur.

1. Einleitung

1.1. Problemstellung, Erkenntnisinteresse und Forschungsfrage

Unter dem Begriff der Telematikinfrastruktur (TI) werden verschiedene Projekte und Reformen in Gesundheitswe-

Ich bedanke mich bei Herrn Professor Dr. Marcus Helfrich für die mir eingeräumte Freiheit bei der Themenwahl, sowie für die unkomplizierte Kommunikation und die konstruktive Kritik während der Betreuung dieser Arbeit.

sen und -wirtschaft vornehmlich zur Digitalisierung der Prozesse und Daten zusammengefasst. Die TI soll das deutsche Gesundheitswesen vernetzen und einen sicheren Austausch von Gesundheitsdaten ermöglichen.[1] Mit über 70 Millionen Versicherten[2], etwa 100 Krankenkassen[3] und 180.000 Vertragsärzten und Psychotherapeuten[4] ist die Vernetzung des deutschen Gesundheitswesens eines der größten Digitalisierungsvorhaben in Europa. Auf die Vernetzung des Gesundheitswesens durch die Telematikinfrastruktur baut die elektronische Patientenakte (ePA) auf, welche die Krankenkassen ihren Versicherten nach geltender Sozialgesetzgebung seit dem 1. Januar 2021 anbieten müssen, § 342 Abs. 1 SGB V.[5] Die ePA soll die Digitalisierung und Vereinfachung der bisher in Papierform ablaufenden Arbeitsschritte ermöglichen. Medizinische Informationen lassen sich besser sammeln und strukturieren.[6] Davon sollen Ärzte, Apotheker, Therapeuten, anderes medizinisches Fachpersonal und vor allem der Patient selbst profitieren.

Die hohe Aktualität des Themas wird deutlich durch die Tatsache, dass das seit Jahren geplante Konzept der Telematikinfrastruktur mit dem darauf aufbauenden Element der elektronischen Patientenakte immer noch nicht umgesetzt ist und dabei ständig neue Hindernisse, wie beispielsweise Fragen zum Datenschutz oder der Informationssicherheit, auftreten.[7] Bei den Beteiligten, insbesondere bei den Leistungserbringern und den Versicherten, ist eine Unsicherheit hinsichtlich der Datenverarbeitung und deren Auswirkungen auf die datenschutzrechtliche Behandlung vorhanden.[8] Eine nachhaltige Verbesserung der Prozesse kann im Gesundheitssystem hingegen nur eintreten, wenn die – für die Versicherten freiwilligen, für die Leistungserbringer und Kassen verpflichtenden – Maßnahmen von allen Beteiligten akzeptiert werden.[9] Einen wichtigen Beitrag dafür leisten Klarheit, Transparenz und Rechtssicherheit in datenschutzrechtlichen Fragestellungen.

Mit Einführung der Datenschutzgrundverordnung (DSGVO)[10] sind möglicherweise neue datenschutzrechtliche Anforderungen an die ePA hinzugekommen, die sich auch auf ihre prozessuale oder technische Ausgestaltung auswirken könnten. Die Unklarheiten hinsichtlich der Berücksichtigung von europäischen Datenschutzvorschriften

führten zu erheblicher Kritik, die den Gesetzgeber auf den Plan rief. Mit einer weiteren Novelle des SGB V, dem „Patientendatenschutzgesetz (PDSG)", will der Gesetzgeber Abhilfe schaffen.[11]

Beispielsweise ist unklar, ob es die Möglichkeit einer differenzierten Einwilligung geben muss, i.e. die Möglichkeit zu entscheiden, welche Beteiligten am Gesundheitssystem auf welche Daten zugreifen dürfen.[12] Nach derzeitiger Lage hätte ein Physiotherapeut Zugriff auf Daten, die eigentlich nur für einen Urologen bestimmt und relevant sind. Weiterhin soll durch das PDSG festgelegt werden, wer in der Telematikinfrastruktur, respektive im Rahmen der elektronischen Patientenakte für die verschiedenen Komponenten Verantwortlicher im Sinne von Artikel 4 Nr. 7 DSGVO sein wird.[13] Ferner wurde der Umgang mit Daten bezüglich Erbkrankheiten oder anderen genetischen Informationen kritisiert. Hier sind bei der Datenverarbeitung möglicherweise nicht nur die tatsächlich Erkrankten betroffen, sondern auch nahe Verwandte. Diese müssten möglicherweise auch über die Verarbeitung der Daten informiert werden. Die Bundesregierung sieht in diesem Fall eine Ausnahme von der Informationspflicht, die in dieser Arbeit zu prüfen wäre. Veraltete Sicherheitsstandards und nicht regelkonforme Authentifizierungsverfahren stehen ebenfalls in der Kritik.[14]

Auf die zuvor genannten Aspekte wird die vorliegende Arbeit eingehen, wobei das Hauptaugenmerk auf der datenschutzrechtlichen Beurteilung der Verarbeitungsvorgänge im Rahmen der elektronischen Patientenakte nach europäischem Recht liegen wird. Insbesondere die durch den Bundesbeauftragen für Datenschutz und Informationsfreiheit vorgebrachten Bedenken, welche in einer Warnung nach Art. 58 Abs. 2 lit. a DSGVO an die Krankenkassen kulminierten, sollen näher betrachtet werden.[15] Es soll festgestellt werden, *ob diesen Bedenken Rechnung zu tragen ist oder ob die Regelungen zur elektronische Patientenakte in ihrer derzeitigen Ausgestaltung mit der DSGVO vereinbar sind.*

1.2. Forschungsstand

Nach einigen Rückschlägen wurden die Bemühungen zur Schaffung einer Telematikinfrastruktur mit der elektronischen Patientenakte unter der Regie des Bundesgesundheitsministers Jens Spahn wieder deutlich vorangetrieben.[16] Mit der Verabschiedung des „Digitale-Versorgung-Gesetz"[17] erfolgte Ende 2019 eine Novelle des SGB V, mit der die Strukturen des Gesundheitssystems der Dynamik der digitalen Transformation und der Geschwindigkeit von Innovationsprozessen angepasst werden sollten.[18] Unter anderem

---

[1]Vgl. Koch & Henke, 2016, S. 309.

[2]Vgl. Bundesministerium für Gesundheit (BMG), 2021, o. S.

[3]Vgl. GKV-Spitzenverband, Gesetzliche Krankenkassen, 2021, o. S.

[4]Vgl. Kassenärztliche Bundesvereinigung (KBV), 2020, S.3.

[5]Zum Verhältnis der Telematikinfrastruktur und elektronischer Patientenakte s. unten 2.2.7.

[6]Vgl. Gematik, 2021b, o. S.

[7]Den rechtlichen Ausgangspunkt setzte das „E-Health-Gesetz" von 2015, Gesetz für sichere digitale Kommunikation und Anwendungen im Gesundheitswesen sowie zur Änderung weiterer Gesetze vom 21.12.2015, BGBl. I S. 2408.

[8]Vgl. für die Versicherten Klöckner & Olk, 2021; für die Ärzteschaft Bundesärztekammer, 2020a, S. 4.

[9]Vgl. auch Bundesärztekammer, Stellungnahme der Bundesärztekammer zum PDSG, 2020, S. 4.

[10]Verordnung 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016, ABl. L 119 vom 4. Mai 2016, S. 1-88.

[11]Vgl. Bundesministerium für Gesundheit, Gesetzesentwurf PDSG, 2020, S.2.

[12]Vgl. BT-Drs. 19/16228, S. 3.

[13]Vgl. a.a.O., S. 4.

[14]Vgl. a.a.O., S. 9 f.

[15]Vgl. BfDI, 2020,?.

[16]Dazu kritisch netzpolitik.org, Jens Spahn hat es eilig, netzpolitik.org (2020), o. S.

[17]Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation vom 9. Dezember 2019, BGBl. I S. 2562.

[18]Vgl. Jorzig und Sarangi (2020, S. 41).

sollten durch dieses Gesetz mehr Leistungserbringer an die Telematikinfrastruktur angeschlossen werden.[19] Mit dem „Terminservice- und Versorgungsgesetz"[20] wurden die Krankenkassen dazu verpflichtet, den Versicherten eine ePA ab spätestens 1. Januar 2021 zur Verfügung zu stellen. Mit diesen Novellen wurde versäumt, den mittlerweile durch die DSGVO geänderten Anforderungen an den Datenschutz ausreichend Rechnung zu tragen, so dass es nötig geworden war, ein weiteres Gesetz auf den Weg zu bringen: das bereits erwähnte „Patientendatenschutzgesetz".[21] Diese Entwicklungen sind zum Zeitpunkt der Anfertigung dieser Arbeit sehr aktuell. Dementsprechend fällt der Bestand an Literatur zur konkreten Problemstellung relativ gering aus. In einem vom *health innovation hub*[22] beauftragten Gutachten wird detailliert auf die Bedenken des BfDI eingegangen.[23] Der Verfasser der vorliegenden Arbeit versucht auf die vorgebrachten Gegenargumente einzugehen, eigene Überlegungen anzustellen und eine eigene Bewertung vorzunehmen.

1.3. Forschungsmethode

Die vorliegende Arbeit wurde auf Basis einer Literaturrecherche angefertigt.[24] Zu Beginn wurde der Suchraum auf europäisches Datenschutzrecht eingegrenzt, das den Bereich des deutschen Gesundheitswesens, der die elektronischen Patientenakte beinhaltet, betreffen könnte. Der Schwerpunkt der Arbeit liegt in der Ausarbeitung der Vereinbarkeit der elektronischen Patientenakte betreffenden Regelungen, insbesondere solche des SGB V, mit der Datenschutz-Grundverordnung. Vor der Beantwortung der Forschungsfrage waren Grundlagen der elektronischen Patientenakte und der Telematikinfrastruktur zu klären, um den Rahmen der Prüfung europäischen Datenschutzrechts zu identifizieren.

Während der Ausformulierung der zuvor genannten Grundlagen, wurden parallel dazu iterative Recherchezyklen zu den Kernthemen durchgeführt, die dann zum Teil wieder neue Recherchefelder eröffneten.[25]

Die Literatursuche wurde teilweise in öffentlichen juristischen Bibliotheken mittels Bibliothekssuchmaschine und Suchmaschinen im Internet durchgeführt. Die Nutzung der öffentlichen Bibliotheken war durch die Pandemie-Situation 2020/2021 jedoch stark eingeschränkt. Daher wurde diese Arbeit überwiegend anhand im Internet elektronisch verfügbarer Dokumente angefertigt, wobei die Recherche-Plattform Beck-Online und öffentlich verfügbare Stellungnahmen der Beteiligten zur PDSG-Gesetzgebung eine zentrale Rolle einnahmen.

In die Literaturauswahl wurden neben Sachbeiträgen aus Fachzeitschriften und Kommentarliteratur zu den Vorschriften, aktuelle Gesetzgebung, Gesetzesänderungen und Spezifikationen der Gesellschaft für Telematik (gematik) mit einbezogen. Die Auswahl der Literatur erfolgte anhand der Relevanz entsprechender Literatur für das Kernthema und ihrer Aktualität.

1.4. Aufbau der Arbeit

Die Kapitel 2.1 – 2.3 bilden die theoretische Grundlage dieser Arbeit. Hier soll dargelegt werden, in welchem Umfeld die datenschutzrechtliche Betrachtung einzuordnen ist. Kapitel 2.4 behandelt die allgemeinen Voraussetzungen für die Anwendung der DSGVO. Kapitel 2.5 liefert einen Überblick über datenschutzrechtliche Kritikpunkte, die im öffentlichen Diskurs auszumachen sind. Die Kapitel 2.6 – 2.10 beschäftigen sich detailliert mit dem Spannungsverhältnis zwischen DSGVO und nationaler Sozialgesetzgebung, wobei der Schwerpunkt der Arbeit auf der Prüfung der Einwilligung nach Art. 6 Abs. 1 Satz 1 lit. a bzw. Art. 9 Abs. 2 lit. a DSGVO liegt (Kapitel 2.7.1).

2. Hauptteil

2.1. Begriff der elektronischen Patientenakte

In der Vergangenheit gab es bereits schon einzelne Vorstöße, um die digitale Kommunikation zwischen Patienten, Leistungsträger und Krankenkassen zu verbessern. Beispielsweise veröffentlichte die Techniker Krankenkasse bereits 2018 eine Anwendung mit dem Namen „TK-Safe".[26] Allianz und DAK brachten kurze Zeit später die Anwendung „Vivy" auf den Markt, die unter anderem Versicherten der DAK, IKK Classic, Bahn-BKK ermöglichte, Dokumente elektronisch zu speichern.[27] Solche Anwendungen werden *elektronische Gesundheitsakte* genannt.[28] Die bis dahin veröffentlichten Anwendungen haben gemein, dass sie keine umfassende Vernetzung zwischen den Akteuren im Gesundheitswesen anbieten konnten.[29] Dafür fehlten einheitliche Standards, die dann später mit dem Terminservice- und Versorgungsgesetz herbeigeführt werden sollten. Es sollte eine *elektronische Patientenakte* entwickelt werden, die die Krankenkassen ihren Versicherten zur Verfügung stellen können.[30] Die Komponenten für die ePA müssen durch gematik als zentrale Stelle geprüft und zugelassen werden. Nach derzeitigem Stand gibt es 103 Anbieter eines ePA-Aktensystems, drei Hersteller von ePA-Aktensystemen und 13 verschiedene Produkte für das Frontend, die bei der gematik zugelassen sind.[31] Daher gibt

---

[19]Vgl. ebd.

[20]Gesetz für schnellere Termine und bessere Versorgung vom 6. Mai 2019, BGBl. I S. 646.

[21]Gesetz zum Schutz elektronischer Patientendaten in der Telematikinfrastruktur vom 14. Oktober, BGBl. 2115.

[22]Das health innovation hub ist eine Einrichtung des Bundesgesundheitsministeriums, vgl. https://hih-2025.de/about/.

[23]Vgl. Böllhoff, Cornelius, et al. (2020).

[24]S. dazu. vom v. Brocke, Jan, Simons, Alexander, et. al. (2009).

[25]Vgl. zur Organisation der Literatursichtung und überprüfung, a.a.O., S. 9.

---

[26]aerztezeitung (2018, o. S).

[27]aerztezeitung.de (2018, o. S).

[28]Vgl. Wortlaut der §§68 und 351 SGB V.

[29]Es fehlt im Wesentlichen die Anbindung an die TI. Vgl. dazu unten 2.2.7.

[30]§§291a Abs. 5c Satz 4, 291b Abs. 1a Satz 1 SGB V in der vom 11.05.2019 geltenden Fassung.

[31]Gematik (2021c, o. S).

es nicht die *eine* elektronische Patientenakte, sondern verschiedene Produkte, die sich gleichwohl an denselben Standards messen lassen müssen.[32]

2.2. Aufbau und Funktionsweise der elektronischen Patientenakten

Das deutsche Gesundheitswesen setzt sich im Wesentlichen aus drei Arten von Akteuren zusammen: den Leistungsempfängern, den Leistungserbringern und den Leistungsträgern.[33] An diesem Grundgerüst muss sich auch die Architektur der ePA orientieren. Um Aufbau und Funktionsweise der elektronischen Patientenakten zu verstehen, ist es erforderlich, die Beteiligten, deren Interessen und systembedingte Komponenten zu identifizieren. Die Rolle der zuletzt genannten Beteiligten und von den Beteiligten abgeleitete Anforderungen, sowie weitere für die ePA wichtige Elemente werden im Folgenden behandelt.

2.2.1. Leistungserbringer

Zu den Leistungserbringern im Sinne des Sozialgesetzbuches, §§ 69 ff SGB V, zählen unter anderem Vertragsärzte, Vertragszahnärzte, Apotheken, Vertragspsychotherapeuten, Krankenhäuser, Erbringer von Heil- und Hilfsmittelleistungen, wie beispielsweise Physiotherapeuten, Sprechtherapeuten oder Ergotherapeuten.

Die Ärzteschaft unterstützt das grundsätzliche Vorhaben der Digitalisierung des Gesundheitswesens, sieht aber die Gefahr einer Abwälzung von Aufgaben auf die Leistungsträger, die nur unmittelbar mit der medizinischen Versorgung des Patienten verbunden seien, wie beispielsweise datenschutzrechtliche Auskunfts- oder Beratungspflichten gegenüber den Versicherten.[34]

Die ePA ist so konzipiert, dass die Leistungserbringer entsprechend ihrer Rollen mit unterschiedlichen Verarbeitungsberechtigungen auf die Patientendaten zugreifen können.[35]

2.2.2. Leistungsempfänger

Leistungsempfänger sind die Empfänger der Gesundheitsleistungen, also Patienten der Ärzte in Krankenhäusern und Arztpraxen, Therapeuten oder anderen Leistungserbringern. Von der elektronischen Patientenakte im Sinne der Sozialgesetzgebung können zunächst nur Versicherte der gesetzlichen Krankenkassen profitieren.[36]

Ein wesentliches Interesse der Leistungsempfänger dürfte eine optimale Behandlung durch die Leistungserbringer sein. Einen wichtigen Beitrag dazu leisten ein schneller Informationsaustausch und die sinnvolle Zusammenarbeit der Leistungserbringer, damit jede notwendige Expertise schnell

und effizient an den richtigen Adressaten gelangt.[37] Auch eine bessere Übersicht über Dokumente und Daten und deren Strukturierung „an einem Ort" stellen ein bedeutsames Interesse der Leistungsempfänger dar. Jedoch dürfen diese Vorteile nicht zu Lasten anderer Rechtsgüter gehen, wie beispielsweise das Recht auf informationelle Selbstbestimmung als Ausprägung des allgemeinen Persönlichkeitsrechts, aus dem letztlich der Datenschutz erwächst.[38] Eine Auswahl an Kritikpunkten der Datenschützer werden in dieser Arbeit umrissen.[39] Auf die Bedenken des Bundesbeauftragten für Datenschutz und Informationsfreiheit wird detailliert eingegangen.[40]

Normen außerhalb der DSGVO, die auf den Schutz informationeller Selbstbestimmung abzielen, wie beispielsweise die Prüfung einer zulässigen Offenbarung gemäß § 203 StGB, werden im Rahmen dieser Arbeit nur behandelt, sofern dies im Rahmen der DSGVO nötig ist.[41]

2.2.3. gematik

Die gematik GmbH wurde von den Spitzenorganisationen des deutschen Gesundheitswesens nach gesetzlichem Auftrag gegründet, um den Aufbau der Telematikinfrastruktur als Basis für eine digitale und sichere Vernetzung im Gesundheitswesen voranzutreiben.[42] Die Geschäftsanteile sind in § 310 Abs. 2 SGB V geregelt. Aufgrund eingeschränkter Funktionsfähigkeit des Gremiums durch Patt- und Blockadesituationen, wurden Anfang 2019 die Mehrheitsverhältnisse und die Voraussetzungen für die Beschlussfähigkeit durch das Terminservice- und Versorgungsgesetz geändert.[43] Mithin hat die Bundesrepublik Deutschland mittlerweile mit 51% die Kontrolle über das Organ. Die Gesellschafter können gemäß § 310 Abs. 3 SGB V den Beitritt weiterer Spitzenorganisationen auf Bundesebene und des Verbandes der Privaten Krankenversicherung auf deren Wunsch beschließen. Die Geschäftsanteile sind innerhalb der Gruppen und Kostenträger entsprechend anzupassen. Demnach sind der Spitzenverband Bund der Krankenkassen (GKV) mit 22,05%, der Verband der Privaten Krankenversicherung (PKV) mit 2,45% und die übrigen in § 306 Abs. 1 Satz 1 SGB V genannten Organisationen der Leistungserbringer[44] mit insgesamt 24,5% beteiligt (Abbildung 1). Die gematik ist für die Einführung, Pflege, Weiterentwicklung der Telematikinfrastruktur und für datenschutzrechtlich relevante Festlegungen in Abstimmung mit dem Bundesamt für Sicherheit in der Informa-

---

[32]Im Verlauf der Arbeit wird der Einfachheit halber dennoch der Singular verwendet.

[33]Vgl. Jäschke und Hacks (2016, S. 7).

[34]Vgl. Bundesärztekammer (2020b, S. 1 und 7).

[35]Vgl. unten Abbildung 8.

[36]Zur Rolle der privaten Krankenkassen s. unten 2.3.

[37]Vgl. Stöferle (2016), Kooperations- und Kommunikationspartner aus Anwendersicht, S. 135.

[38]Vgl. BVerfGE 65, 1, Volkszählung.

[39]S. unten 2.5.

[40]S. unten 2.6 f.

[41]Vgl. ausführlich dazu Stöferle, Kooperations- und Kommunikationspartner aus Anwendersicht, S. 141.

[42]Vgl. Gesellschaft für Telematik (gematik) (2021e, o. S).

[43]Vgl. Bundesministerium für Gesundheit (BMG) (2019, o. S).

[44]Dies sind im Einzelnen: Die Kassenärztliche Bundesvereinigung (7,35%), die Kassenzahnärztliche Bundesvereinigung (2,45%), die Bundesärztekammer (2,45%), die Bundeszahnärztekammer (2,45%), die Deutsche Krankenhausgesellschaft (5,88%) sowie der Deutsche Apothekenverband (3,92%).

tionstechnik und mit der oder dem Bundesbeauftragten für den Datenschutz durch gesetzlichen Auftrag verantwortlich, § 311 SGB V. Die gematik ist weiterhin für die Zulassung der für die ePA benötigten Komponenten (§§ 341 Abs. 3, 325 SGB V) und für die Zulassung der von den Betreibern entwickelten Anwendungen und Services (§§ 311 Abs. 1 Nr. 4 und Nr. 5, 324, 325 SGB V) zuständig.

2.2.4. Hersteller ePA-Aktensysteme

Die Betreiber sind die Beteiligten, die einen wesentlichen Teil der Infrastruktur und des Backends für den Betrieb der digitalen Services bereitstellen. Die gematik bezeichnet diese Betreiber als Hersteller für „ePA-Aktensysteme".[45] Der Markt wird unter drei Unternehmen oder Unternehmenskonsortien aufgeteilt: *IBM* wird Services für die ePA im Backend für die Techniker Krankenkasse, die Barmer, die Knappschaft, die Hanseatische Krankenkasse und Viactiv BKK anbieten. Der zweite Anbieter ist ein Zusammenschluss des Unternehmens *Bitmarck* und dem österreichischen Unternehmen *Research Industrial Systems Engineering (RISE)*.[46] Die Unternehmensverbindung wird die Services für 87 Krankenkassen unterschiedlicher Größe, darunter Betriebskrankenkassen (BKK), DAK und IKK, anbieten.[47] Das Unternehmen *x-tention Informationstechnologie GmbH* wird zusammen mit internationalen Partnern die Services für die Allgemeinen Ortskrankenkassen (AOKen) bereitstellen.[48] Die Aktensysteme sind die Serverseite der Fachanwendung ePA und besteht aus folgenden Komponenten (Abbildung 2)[49]:

- „Authentisierung von Nutzern": Wird von anderen Komponenten verwendet, um die Authentifizierung von Versicherten und deren berechtigten Vertretern zu erstellen.

- Zugangsgateway des ePA-Aktensystems: Ermöglicht den Zugang des Versicherten auf das jeweilige Aktensystem über das Internet und dient dem Netzabschluss in Richtung Internet.

- Autorisierung und Schlüsselverwaltung: Stellt für den authentifizierten Nutzer eines Aktenkontos bei gegebener Autorisierung das jeweilige Schlüsselmaterial bereit.

- Dokumentenverwaltung: Dient dem sicheren Speichern und Auffinden von Dokumenten des Versicherten in seiner Akte. Diese Funktion können Versicherte selbst oder von ihm benannte Vertreter, Leistungserbringer und Kostenträger nutzen.

2.2.5. Leistungsträger als Anbieter der ePA

Die im Kontext der elektronischen Patientenakten relevanten Leistungsträger sind die gesetzlichen Krankenkassen, § 12 i.V.m. § 18 ff. SGB I. Die Krankenkassen sind die Kunden der oben beschriebenen Betreiber. Für Vertrieb und Teile der Frontendgestaltung der Smartphone Apps sind die Krankenkassen verantwortlich.[50] Die Anbieter der ePA sind damit in der Regel die gesetzlichen Krankenkassen.[51] Die Gestaltungsmöglichkeiten des Designs (customizing) hängen vom Angebot der Betreiber ab. Das Frontend kann beispielsweise mit unterschiedlichen Farben, Logos und Verweisen auf die Webseiten der jeweiligen Krankenkasse gestaltet werden.[52] Auch der Produktname wird von den Krankenkassen bestimmt. So wird die App bei der Barmer „eCare", bei der Techniker Krankenkasse „TK-Safe" und bei der AOK „Mein Leben" heißen.[53]

Der GKV-Spitzenverband unterstützt die Einführung und Weiterentwicklung der ePA, sieht aber eine Hürde in der vom BfDI geforderten Information des Versicherten über die im ersten Jahr fehlende Möglichkeit einer selektiven Rechtevergabe auf Dokumentenbasis vor jeder Speicherung eines Dokuments.[54] Unnötig sei auch das Festhalten einer Sanktionierung der Krankenkassen, wenn keine ePA angeboten werde.[55] Ablehnend steht der Spitzenverband auch der Neuaufnahme einer öffnenden Regelung zur Bereitstellung der ePA durch andere Anbieter als Krankenkassen gegenüber.[56] Dies führe zu erheblichen Rechtsunsicherheiten bei allen Beteiligten. Der GKV-Spitzenverband lehnte auch die Einrichtung einer Ombudsstelle bei der gematik für Betroffene im Sinne datenschutzrechtlicher Normen ab[57] und steht der Einrichtung von sogenannten Terminals kritisch gegenüber.[58]

2.2.6. ePA Frontend des Versicherten (ePA-FdV)

Das ePA Frontend des Versicherten bezeichnet die Client-Software auf einem mobilen Endgerät des Versicherten, mit dem dieser auf die ePA zugreifen kann (Abbildung 3).[59] Die Frontends wurden für bestimmte Krankenkassengruppen von verschiedene Herstellern entwickelt.[60] Beispielsweise wurde die „AOK Mein Leben" App von Ernst & Young entwickelt. Andere Frontends wurden von IBM, RISE, der Hanseatischen Krankenkasse und der Techniker Krankenkasse entwickelt.

2.2.7. ePA als Fachanwendung in der Telematikinfrastruktur

Die Telematikinfrastruktur (TI) ist die bevorzugte Informations-, Kommunikations- und Sicherheitsinfrastruktur des deutschen Gesundheitswesens mit allen technischen

---

[45]Vgl. Gesellschaft für Telematik (gematik) (2021a, o. S).
[46]Vgl. aerzteblatt.de Elektronische Patientenakte (2020, o. S).
[47]Vgl. ebd.
[48]Vgl. ebd.
[49]Vgl. Gematik (2021a, o. S).

[50]Vgl. aerzteblatt.de Elektronische Patientenakte (2020, o. S).
[51]Vgl. Gematik (2021c, o. S.): „Anbieter-ePA-Aktensystem"; nichtamtliche überschrift des §342 Abs. 1 SGB V.
[52]Vgl. ebd.
[53]Vgl. aerzteblatt.de Elektronische Patientenakte (2020, o. S).
[54]Vgl. GKV-Spitzenverband (2020b, S. 7); siehe dazu unten 2.7.1.3.
[55]Vgl. GKV-Spitzenverband (2020b, S. 7).
[56]Vgl. ebd.
[57]Vgl. a.a.O., S. 9.; §307 Abs. 5 Satz 2 und 3 SGB V.
[58]S. dazu unten 2.5.4.
[59]Gesellschaft für Telematik (gematik) (2021b, o. S).
[60]Vgl. Gematik (2021c, o. S).

Abbildung 1: Gesellschafteranteile der gematik

Quelle: *gematik*, Gesellschafteranteile, 2021
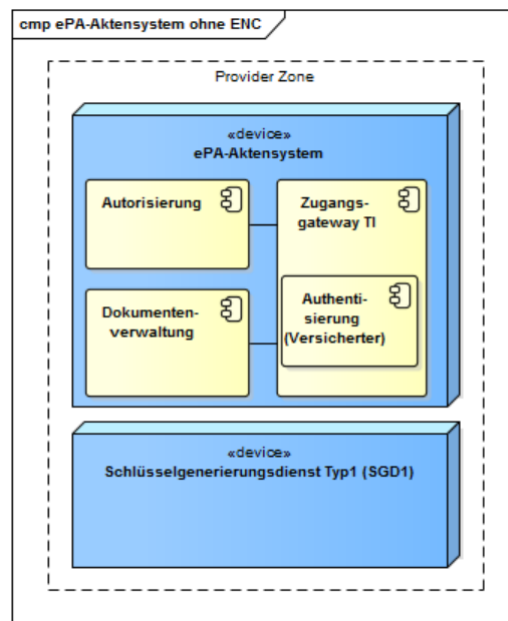
Abbildung 2: High Level Diagramm – ePA Aktensystem

Quelle: Gesellschaft für Telematik (gematik) (2021g, S.9).

und organisatorischen Anteilen.[61] Die TI unterstützt die Anwendungen der Versicherten gemäß § 306 Abs. 1 SGB V. Die ePA ist eine Fachanwendung der Telematikinfrastruktur gem. § 306 Abs. 1, § 327 SGB V und nutzt Teile der Telematikinfrastruktur, wie beispielsweise die Verbindung zu den „Konnektoren" der Leistungserbringer oder den „KTR-Consumern",

die den Kostenträgern den Zugang zur TI ermöglichen (unten Abbildung 5).[62]

Die obligatorischen Anforderungen an die Fachanwendung ePA ergeben sich aus § 341 Abs. 2 i.V.m. § 342 Abs. 2 SGB V. Enthalten sind unter anderem die „Einstellung" von

---

[61]Vgl. Gesellschaft für Telematik (gematik) (2021i, o. S).

[62]Vgl. Gesellschaft für Telematik (gematik) (2021f, o. S).

Medikationsplänen, Notfalldaten, Daten zu Befunden, Diagnosen, Therapiemaßnahmen, und Daten der Impfdokumentation.

2.2.8. Angebot nicht vorgeschriebener Services

Das ePA-FdV kann über die gesetzlichen Vorschriften hinaus zusätzliche Funktionalitäten enthalten. Diese werden dann allerdings nicht der Fachanwendung ePA zugeordnet und unterliegen somit nicht der Regelungshoheit der gematik (Abbildung 4).[63] Die Techniker Krankenkasse bietet beispielsweise ein persönliches Arztverzeichnis, Impf- und Vorsorgeempfehlungen, Übersicht zur Arbeitsunfähigkeit und Informationen über Arztbesuche und Krankenhausaufenthalte an.[64] Die Barmer will Gesundheitsinformationen, Prävention und saisonale Ratschläge anbieten.[65] Einige Krankenkassen möchten sogenannte Gesundheitshistorien, die auf Abrechnungsdaten basieren, bereitstellen.[66] Diese Zusatzfunktionen müssen für den Nutzer von den Funktionalitäten der Fachanwendung ePA jedoch unterschieden werden können.[67] Weiterhin muss das Design des Frontends sicherstellen, dass der Nutzer dem Verarbeiten der ePA-Daten in zusätzlichen Funktionalitäten bezüglich Umfang, Art und Dauer der Verarbeitung vor dem Zugriff der Zusatzfunktionen auf die ePA-Daten zustimmen muss und diese Zustimmung widerrufbar ist.[68] Darüber hinaus enthalten die Spezifikationen des ePA-FdV weitere Pflichten, die die *Sicherheit* betreffen. Die Zusatzfunktionen werden vom Produktgutachter in seinem Produktgutachten hinsichtlich ihrer Sicherheit mitberücksichtigt.[69] Der Begriff Sicherheit meint im Rahmen der Produkt- und Sicherheitsgutachten die Informationssicherheit, die den „Datenschutz" implizit mit einschließt.[70] Nach Ansicht des Verfassers wird der Datenschutz zwar nicht durch die Informationssicherheit impliziert,[71] sofern aber im Ergebnis den datenschutzrechtlichen Vorgaben entsprochen wird, ist diese Diskussion hinfällig.

2.2.9. Übersicht der technischen Komponenten der ePA

Schließlich ergibt sich ein komplexes Bild des Geflechts zwischen den Beteiligten unter Einbeziehung der verschiedenen technischen Komponenten (Abbildung 5).

2.3. Ausbauphase der ePA

Die Krankenkassen sind aus § 342 Abs. 1 SGB V verpflichtet, spätestens ab dem 1. Januar 2021 jedem Versicherten eine ePA mit den dort in Abs. 2 genannten Anforderungen zur Verfügung zu stellen. Mit der Bereitstellung zum 1. Januar 2021 begann eine „Testphase", die für die Dauer des

ersten Quartals des Jahres geplant war und in der die Anbindung einiger ausgewählter Leistungserbringer erprobt werden sollte.[72] Im zweiten Quartal soll nach erfolgreicher Testphase mit dem Rollout der Konnektoren begonnen werden, um eine flächendeckende Anbindung der Leistungserbringer bis zum Ende des Jahres fertigzustellen (Abbildung 6). Die Leistungserbringer sind dementsprechend aus § 341 Abs. 6 SGB V dazu verpflichtet, bis zum 30. Juni 2021 nachzuweisen, dass sie über die für den Zugriff auf die ePA erforderlichen Komponenten und Dienste verfügen. Aus Sicht des Verfassers ist es zumindest fragwürdig, dass eine Anwendung für über 70 Millionen Versicherte zur Verfügung gestellt wird, deren Anbindung an dezentrale Komponenten sich noch in der Testphase befindet.

Bis 2022 sind weitere Funktionalitäten in der ePA obligatorisch aufzunehmen, § 342 Abs. 2 Nr. 2 SGB V. Hinzukommen wird unter anderem ein feingranulares Rechtemanagement und das E-Rezept.[73] 2023 wird die ePA mit weiteren Modulen aufgerüstet, wie beispielsweise Daten pflegerischer Versorgung oder der Arbeitsunfähigkeitsbescheinigung, § 342 Abs. 1 Nr. 3 SGB V. Spätestens dann sollen Daten der ePA auch zu Forschungszwecken durch den Versicherten zur Verfügung gestellt werden können, § 342 Abs. 1 Nr. 4 SGB V.

Seit April 2020 ist auch wieder der Verband der Privaten Krankenversicherung (PKV) Gesellschafter der gematik.[74] Die privaten Krankenversicherungen wollen die ePA erst 2022 in ihrer erweiterten Form anbieten.[75] Darüber hinaus möchten sich einige private Versicherer mit weiteren digitalen Angeboten in der App hervorheben.[76] Diese Zusatzfunktionen müssten hinsichtlich Informationssicherheit und Datenschutz ebenfalls durch von der gematik beauftragte unabhängige Gutachter geprüft werden.

2.4. Allgemeines zur Anwendbarkeit der DSGVO im Kontext der ePA

Die Verordnung 2016/679 (DSGVO) hat kraft EU-Gesetzgebung allgemeine Geltung; sie ist in allen Teilen verbindlich und gilt unmittelbar in jedem Mitgliedsstaat der europäischen Union, Art. 288 UAbs. 2 AEUV.[77] Weiterhin sind bei der Auslegung einer Vorschrift des Unionsrechts neben ihrem Wortlaut und ihrem verfolgten Ziel auch der Kontext und das gesamte Unionsrecht zu berücksichtigen. Die Entstehungsgeschichte einer Vorschrift des Unionsrechts sowie deren Erwägungsgründe können ebenfalls Hinweise für ihre Auslegung geben.[78]

---

[63]Vgl. Gesellschaft für Telematik (gematik) (2021h, S. 11).
[64]Vgl. aerzteblatt.de Elektronische Patientenakte (2020, o. S.).
[65]Vgl. ebd.
[66]Vgl. ebd.
[67]Vgl. Gesellschaft für Telematik (gematik) (2021h, S. 24).
[68]Vgl. Gesellschaft für Telematik (gematik) (2021h, S. 25).
[69]Vgl. Gesellschaft für Telematik (gematik) (2020b, S. 2).
[70]Vgl. Gesellschaft für Telematik (gematik) (2020c, S. 6).
[71]Vgl. Hof (2020, S. 477 Rn. 2).

[72]Vgl. Gesellschaft für Telematik (gematik) (2021c, o. S.).
[73]Daten elektronischer Verordnungen nach §360 Abs. 1 SGB V.
[74]Vgl. Gesellschaft für Telematik (gematik) (2020a, o. S.). Der PKV war bereits Gesellschafter trat aber 2012 wieder aus der Organisation aus, vgl. Olk, Private Krankenversicherung will wieder an gesetzlicher Digitalisierung teilhaben, handelsblatt.com, 2020, o. S.
[75]PKV Verband der Privaten Krankenversicherung (2020, o. S.)
[76]Vgl. aerztezeitung (2020, o. S.): Signal Iduna möchte Mehrwertdienste anbieten, die „deutlich über die gesetzlichen Funktionen hinausgehen".
[77]Vgl. nur Ehmann/Selmayr/Selmayr/Ehmann, DSGVO, Einführung, A. Einl. Bemerkungen, Rn. 1.
[78]Vgl. EuGH NVwZ 2019, 143 Rn. 47 m.w.N.

Abbildung 3: High Level Design – ePA-Frontend des Versicherten

Quelle: Gesellschaft für Telematik (gematik) (2021h, S. 13).
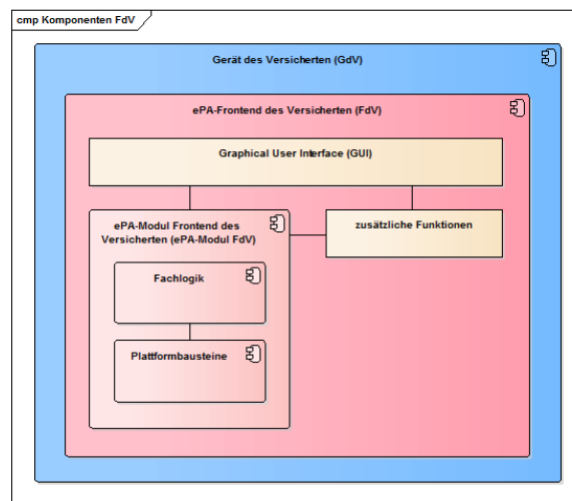
Abbildung 4: Komponenten ePA-Frontend des Versicherten

Quelle: Gesellschaft für Telematik (gematik) (2021h, S. 16).

2.4.1. Sachlicher Anwendungsbereich

Der sachliche Anwendungsbereich der DSGVO ist in Art. 2 Abs. 1 DSGVO geregelt. Die Daten müssen personenbezogen sein, Art. 4 Nr. 1 DSGVO.[79] Die betroffenen Versicherten sind identifizierbare natürliche Personen.

Es muss eine Verarbeitung der Daten vorliegen, Art. 1 Abs. 1, Art. 2 Abs. 1, Art. 4 Nr. 2 DSGVO. Erfasst sind unter anderem automatisierte oder teilautomatisierte Verarbeitungen personenbezogener Daten. Automatisierte Verarbeitungen sind solche, welche mit automatisierten Mitteln bewerkstelligt werden.[80] Darunter sind die meisten rechnergestützten Verarbeitungen personenbezogener Daten zu verstehen.[81]

Im Rahmen der elektronischen Patientenakte werden unterschiedlichste Patientendaten auf den Servern der Betreiber erhoben, gespeichert und verfügbar gemacht. Insbesondere das Ermöglichen des Zugriffs Dritter auf die personenbezogenen Daten ist ein zentrales Element der ePA.[82] Verarbeitet werden beispielsweise Stammdaten und Daten, die wegen ihres Gesundheitsbezugs durch Art. 9 Abs. 1 DSGVO besonderen Schutz genießen. Ein Anwendungsausschluss der Verordnung nach Art. 2 Abs. 2 und 3 DSGVO kommt bei den regelmäßigen Verarbeitungsvorgängen im Bereich der deutschen elektronischen Patientenakte nicht zum Tragen.

---

[82]Vgl. weiter unten Abbildung 8.

---

[79]S. auch ErwG 1 DSGVO.
[80]Vgl. Ehmann/Selmayr/Zerdick, DSGVO, Art. 2 Rn.3.
[81]Vgl. ebd.
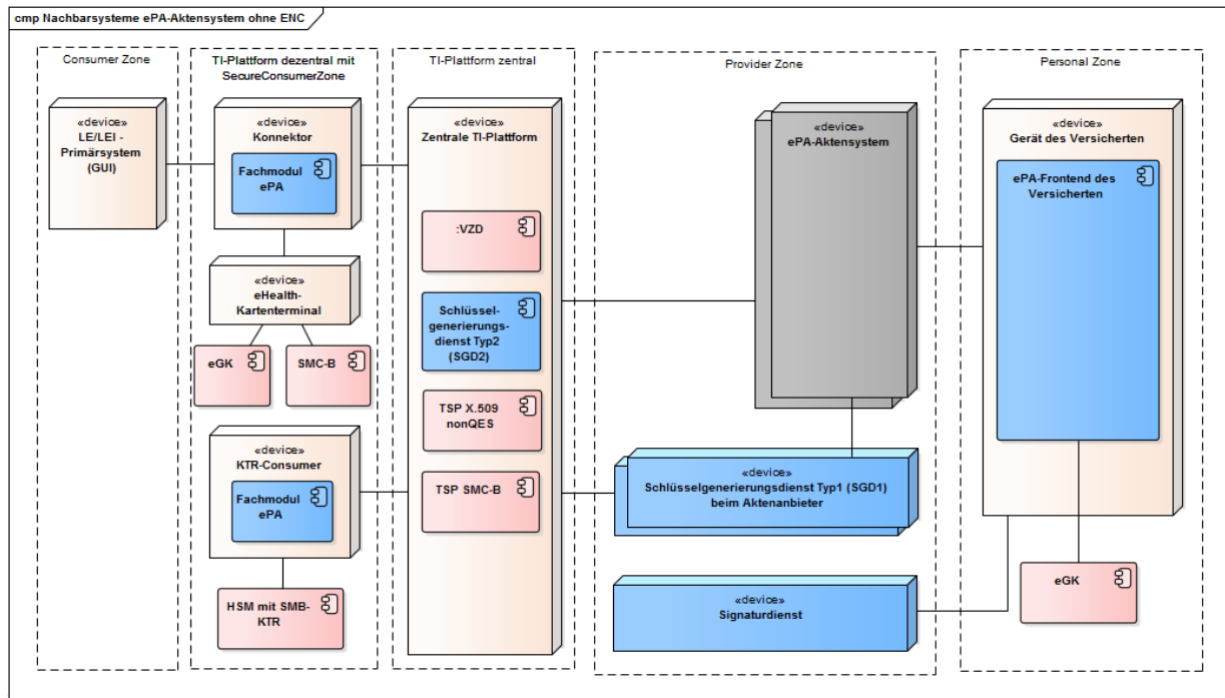
Abbildung 5: Gesamtübersicht der ePA

Quelle: Gesellschaft für Telematik (gematik) (2021g, S. 11).
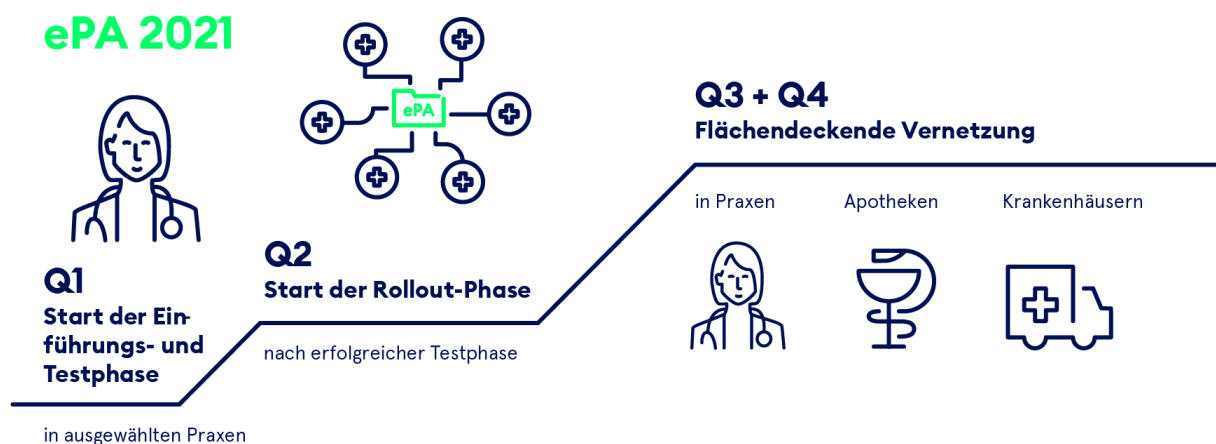
Abbildung 6: Ausbauphase ePA 2021

Quelle: Gesellschaft für Telematik (gematik) (2021c).

2.4.2. Räumlicher Anwendungsbereich

Der räumliche Anwendungsbereich ist in Art. 3 DSGVO geregelt. Die Verarbeitung personenbezogener Daten im Rahmen der deutschen elektronischen Patientenakte fallen in den räumlichen Bereich der DSGVO, da unabhängig vom Ort der Niederlassung der Verantwortlichen, gem. Art. 3 Abs. 1 DSGVO sich ein räumlicher Zusammenhang aus Art. 3 Abs. 2 lit. a DSGVO ergibt. Denn die betroffenen Personen, nämlich die deutschen Patienten, befinden sich in der Europäischen Union und die Datenverarbeitung erfolgt im Kontext des Angebots von Dienstleistungen im Gesundheitsbereich, wie beispielsweise Heilbehandlungen durch Ärzte und andere Leistungserbringer bzw. die Bereitstellung der ePA selbst.

2.4.3. Verantwortlicher

Den Verantwortlichen nach Art. 4 Nr. 7 DSGVO treffen verschiedene Pflichten, die sich aus der Verordnung ergeben, beispielsweise die Umsetzung technischer und organisatorischer Maßnahmen, um sicherzustellen, dass die Datenverar-

beitung mit der Verordnung konform abläuft, Art. 24 Abs. 1 DSGVO. Er ist darüber hinaus auch Adressat direkter Ansprüche einer betroffenen Person, unter anderem des Auskunftsrechts nach Art. 15 DSGVO oder des Schadensersatzes nach Art. 82 DSGVO.[83] Gemäß Art. 4 Nr. 7 1. HS DSGVO ist Verantwortlicher eine natürliche oder juristische Person, Behörde, Einrichtung oder andere Stelle, die allein oder gemeinsam mit anderen über die Zwecke und Mittel der Verarbeitung von personenbezogenen Daten entscheidet. Gemäß Halbsatz 2 des Art. 4 Nr. 7 DSGVO kann der Verantwortliche oder die Kriterien seiner Benennung nach dem Recht eines Mitgliedsstaats festgelegt werden, sofern auch die Zwecke und Mittel der Verarbeitung durch das Recht des Mitgliedsstaats vorgegeben sind. Während zuvor noch unklar war, wer Verantwortlicher im Rahmen der elektronischen Patientenakte sei[84], schafft das PDSG nun Abhilfe. Mit Einfügen des 11. Kapitels in das fünfte Buch des Sozialgesetzbuchs, liegt nun eine solche nationalstaatliche Regelung vor. Nach § 307 SGB V ergibt sich sinngemäß die Verantwortlichkeit aus der konkreten Zuständigkeit des Beteiligten und der Entscheidungskompetenz in den für ihn überblickbaren und beherrschbaren Strukturen.[85] Der für die Verarbeitung der Daten zum Zweck der *Nutzung* der elektronischen Patientenakte Verantwortliche ist gem. §§ 307 Abs. 4, 341 Abs. 4 SGB V der jeweilige Anbieter. Für die Nutzung von Komponenten der dezentralen Infrastruktur (vgl. oben Kapitel 2.2.1) sind die jeweiligen Nutzer dieser Komponenten verantwortlich, sofern sie „über die Mittel der Datenverarbeitung mit entscheiden", § 307 Abs. 1 SGB V. Diese Verantwortlichkeit soll sich im Wesentlichen auf die bestimmungsgemäße Nutzung der Komponenten, deren ordnungsgemäßen Anschluss und die Durchführung der Wartung beziehen.[86] Die Regelung in § 307 Abs. 5 Satz 1 SGB V soll eine durchgehende datenschutzrechtliche Verantwortlichkeit sicherstellen, in dem die gematik als Verantwortliche benannt wird, sofern keine speziellere Verantwortlichkeit vorliegt und soweit sie im Rahmen ihrer Aufgaben über die Mittel der Datenverarbeitung bestimmt.[87] Damit wird der in Art. 26 Abs. 1 Satz 3 DSGVO zum Ausdruck kommende Gedanke aufgegriffen.

Das bedeutet beispielsweise, dass ein Arzt für die Patientendaten in der Praxis verantwortlich ist und der TI-Konnektor vor Missbrauch geschützt wird. Die Verantwortung für die bestimmungsgemäße Funktionsweise des Konnektors liegt hingegen bei dessen Hersteller. Die Datenübertragung und Verarbeitung in der Telematikinfrastruktur fallen in den Verantwortungsbereich der gematik. Die Verantwortung für den Datenschutz im ePA-Aktensystem (s.o. Kapitel 2.2.4) obliegt der jeweiligen Krankenkasse.

Auch wenn die koordinierende Stelle der gematik gem. § 307 Abs. 5 Satz 2 SGB V als Auffangbecken für datenschutzrechtliche Anliegen dient, wird es durch die Aufteilung

der Verantwortlichkeit an so viele einzelne Beteiligte, deren Ineinanderwirken kaum vollständig getrennt werden kann, möglicherweise im konkreten Fall zu weiteren Diskussionen kommen.

2.5. Aktuelle Fragestellungen im Gesundheitsdatenschutz

Ein potenziell großer Teil der in der ePA gespeicherten Daten bezieht sich auf den Gesundheitszustand, die Gesundheitsgefährdungen von betroffenen Personen oder lässt zumindest Schlüsse auf eine dieser beiden Eigenschaften zu. Gesundheitsdaten betreffen den höchstpersönlichen Bereich eines Menschen. Mit Gesundheitsdaten können Menschen diffamiert, unter Druck gesetzt oder direkt gefährdet werden. Manchmal möchte eine Person auch gar nicht wissen, wie es um den eigenen Gesundheitszustand bestellt ist. Gesundheitsdaten können für das Individuum ein Leben lang relevant sein.

Der individuelle Schutz von Leben und Gesundheit ist gesetzlich garantiert.[88] Der besondere Schutz von Gesundheit und Leben kommt im europäischen Datenschutzrecht durch das Verbotsprinzip nach Art. 9 Abs.1 DSGVO zum Ausdruck. Danach ist die Verarbeitung von Gesundheitsdaten generell untersagt. Für die Verarbeitung ist das Vorliegen eines in Art. 9 Abs. 2 DSGVO genannten Erlaubnistatbestands erforderlich. Der Katalog in Abs. 2 ist zwar abschließend[89], enthält aber zahl- und weitreichende Ausnahmetatbestände, die zu einer weitgehenden Egalisierung des Verarbeitungsverbots führen.[90]

Im Folgenden wird eine Auswahl aktueller Fragen mit datenschutzrechtlichem Bezug im Rahmen der ePA ausgeführt[91].

2.5.1. Erbkrankheiten

Nur auf den ersten Blick problematisch erscheint die Verarbeitung von Daten über Krankheiten, die durch Vererbung von Vorfahren auf ihre Nachkommen übertragen werden können. Zwar enthalten die Daten Informationen sowohl über den direkt Betroffenen, als auch über dessen Nachkommen, i.e. die Möglichkeit des Vorhandenseins der Krankheit, – insofern könnte den Verarbeiter solcher Daten eine Informationspflicht aus Art. 14 DSGVO treffen – allerdings greifen in solchen Fällen die in Abs. 5 lit. c und lit. d DSGVO formulierten Ausnahmen einer einschränkenden nationalen Regelung[92]: Gemäß § 11 GenDG darf das Ergebnis einer genetischen Untersuchung grundsätzlich „nur der betroffenen Person und nur durch [. . .] die Ärztin oder den Arzt [. . .] mitgeteilt werden." Auch kommt eine Anwendung des § 203 StGB in Betracht. Ob eine Befugnis zur Weitergabe von Geheimnissen vorliegt, ist im Einzelfall zu prüfen. Schließlich

---

[83]Vgl. nur Ehmann/Selmayr/Nemitz, DSGVO, Art. 82 Rn. 4.

[84]Vgl. BT-Drs. 19/16228, S. 4.

[85]Vgl. BR-Drs. 164/20, Gesetzentwurf der Bundesregierung, S. 108.

[86]Vgl. BR-Drs. 164/20, Gesetzentwurf der Bundesregierung, S. 108 f.

[87]Vgl. a.a.O., S. 109.

[88]Art. 2, Art. 3, Art. 35 GRCh; Art. 2 Abs. 2 S. 1 GG.

[89] Vgl. Sydow/Kampert, DSGVO, Art. 9. Rn. 5 f; Paal/Pauly/Frenzel, DSGVO, Art. 9 Rn. 18 f.

[90]Vgl. Gola/Schulz, DSGVO, Art. 9 Rn. 1 f.

[91] Vgl. vor allem BT-Drs. 19/6628 und die Stellungnahmen des BfDI zum PDSG.

[92]Vgl. Antwort der Bunderegierung auf eine kleine Anfrage, BT-Drs. 19/6628, S. 4.

dürfte in diesem Bereich das Recht auf Nichtwissen relevant sein. Zur informationellen Selbstbestimmung gehört nämlich auch, dass der Betroffene darüber entscheiden kann, ob er sich mit dem Wissen um seine genetischen Besonderheiten belasten möchte.[93]

2.5.2. Aspekte der Datensicherheit

Kriterien der Datensicherheit[94] sind in Art. 32 DSGVO benannt. Sowohl Verantwortlicher als auch Auftragsverarbeiter sind dazu angehalten geeignete technische und organisatorische Maßnahmen zu treffen, um eine sichere Verarbeitung zu gewährleisten.[95] Die Datensicherheit ist außerdem als Grundsatz in Art. 5 Abs. 1 lit. f DSGVO benannt.

Im Zusammenhang mit Gesundheitsakten wurde in der Vergangenheit immer wieder Kritik hinsichtlich ihrer Datensicherheit geübt. Beispielsweise wurden in der Anwendung „Vivy" diverse Sicherheitsmängel identifiziert[96], hier einige Beispiele:

- Dokumente wurden mit einer 5-stelligen Session ID aus Kleinbuchstaben veröffentlicht

- Wurde eine Session gefunden, konnten Metadaten, wie Name, Versichertennummer, Adresse, Alter, Geschlecht, Bild, Arzt und Spezialisierung des Arztes ausgelesen werden

- Um das Dokument einzusehen musste eine 4-stellige PIN eingegeben werden (10000 Versuche mit der Brute Force Methode)

- In die App konnte HTML-Code, zum Beispiel Verlinkungen auf Phishing Seiten, eingeschleust werden

- Die 2-Faktor-Authentifizierung konnte mit Brute Force geknackt werden

Solche Mängel sind eine eklatante Gefahr für die Nutzer der Anwendung und eine Katastrophe für den Anbieter.

Auch bei der ePA bzw. der TI wurde hinsichtlich der Datensicherheit an einigen Punkten Kritik geäußert, z.B.:

- Der Einsatz des veralteten Verschlüsselungsprotokolls TLS 1.1 im Versichertenstammdatenmanagement. Mit dem Update der Konnektoren wurde die Unterstützung von TLS 1.1 entfernt. In den Spezifikationen der TI sind Vorgaben zur Verwendung von TLS 1.2 und TLS 1.3 festgeschrieben.[97]

- Alternative Authentifizierungsverfahren[98]: Der BfDI fordert beispielsweise, dass die Authentifizierungsverfahren der Sicherheitsstufe „hoch" im Sinne der BSI-Richtlinie TR-03147 genügen.[99] Der Bezug zur DSGVO und die Prüfung der Konformität der Authentifizierungsverfahren erfolgt unten in Kapitel 2.9.

- Schwachstelle Mensch: Im Dezember 2020 wurde bekannt, dass Konnektoren in Arztpraxen öffentlich über das Internet erreichbar waren.[100] Die sei darauf zurückzuführen, dass diese Konnektoren fehlerhaft angeschlossen wurden. Die Schwachstelle wurde dem Hersteller mitgeteilt. Eine Richtlinie der Kassenärztlichen Bundesvereinigung soll Sicherheitsstandards in Praxen verbindlich festlegen, § 75b Abs. 1 SGB V.[101] Des Weiteren wird der Einsatz von zertifizierten Fachleuten vorgeschrieben, § 75b Abs. 5 SGB V.[102]

- Gutachter der TU Graz fanden Schwachstellen im sogenannten VAU-Protokoll.[103] Nach Aussage der gematik führen diese Schwachstellen ohne weitere Folgen zu keinem praktisch durchführbaren Angriff.[104] Die Schwachstellen seien mit einer neuen Version umgehend ausgebessert worden.[105]

Bei Betrachtung dieser Probleme ist zu bedenken ist, dass ein System niemals vollständig sicher sein kann. Je komplexer, desto anfälliger ist es für Schwachstellen. Wichtig ist jedenfalls, dass die vorgeschriebenen Standards eingehalten und kontrolliert werden. Immerhin lässt die Aussage des IT-Experten Christoph Saatjohann hoffen: Auf die verschlüsselte Datenbank im Internet und die entsprechenden Server sei ein sehr hoher Wert gelegt worden.[106] Er sei vorsichtig optimistisch; die bekannten Probleme seien angegangen worden.

2.5.3. Datenschutz-Folgenabschätzung

Fraglich ist, ob vor der Verarbeitung von Daten besonderer Kategorien, wie den Gesundheitsdaten, grundsätzlich die Durchführung einer Datenschutz-Folgenabschätzung nach Art. 35 DSGVO angezeigt ist.[107] Der Maßstab des Art. 35

---

[93]Vgl. dazu Kühling/Buchner/Weichert, Art. 9 Rn. 34.

[94]Im Kontext des Datenschutzes ist oft von Datensicherheit die Rede, auch wenn weitgehende überschneidungen mit der Informationssicherheit bestehen.

[95]Vgl. Hof, Datenschutz mittels IT-Sicherheit, S. 477 Rn. 2.

[96]Vgl. Tschirsich, Martin, All Your Gesundheitsdaten Are Belong To Us, 2018, o. S.

[97]Vgl. BT-Drs. 19/16228, S. 10.

[98]i.e. Authentifizierungsverfahren ohne elektronische Gesundheitskarte. Vgl. BfDI (2020a, o. S).

[99]Bundesamt für Sicherheit in der Informationstechnik (BSI) (2018).

[100]Vgl. tagesschau.de (2020, o. S).

[101]S. auch KBV, Richtlinie nach §75b SGB V Anforderungen, 2020.

[102]S. auch KBV (2020).

[103]i.e. Vertrauenswürdige Ausführungsumgebung. Vgl. Slany (2020, S. 4).

[104]Vgl. Gesellschaft für Telematik (gematik) (2020d, S. 2).

[105]Vgl. ebd.

[106]Vgl. sueddeutsche.de (2020, o. S).

[107]Vgl. die Ansicht der Freien ärzteschaft, Gelten die Datenschutzgesetze nicht für Gesundheitsminister Spahn?, 2020, o. S. Die Freie ärzteschaft steht der Digitalisierung des Gesundheitswesens grundsätzlich kritisch gegenüber. Weiterhin wurde a.a.O. auch behauptet, der BfDI hätte die Durchführung einer Datenschutz-Folgenabschätzung eingefordert. Tatsächlich hat der BfDI in einer ersten Stellungnahme zum Entwurf des PDSG (vgl. BfDI, 2020b, S. 2) lediglich dazu angeregt von der Gestaltungsmöglichkeit in Art. 35 Abs. 10 DSGVO Gebrauch zu machen und eine Datenschutz-Folgenabschätzung verbindlich vorzusehen. In der Stellungnahme hinsichtlich der öffentlichen Anhörung einen Monat später ist dieser Vorschlag nicht mehr enthalten (vgl. BfDI, 2020a).

Abs. 1 DSGVO ist ein „voraussichtlich hohes Risiko für die Rechte und Freiheiten natürlicher Personen".[108] Die Beurteilung ob eines solches Risiko vorliegt, ergibt sich aus einer Synthese der in Abs. 1, 3, 4, 5 genannten Kriterien und auch aus den Erwägungsgründen 71, 75, 76, 89, 91.[109] Es genügt nicht, dass eines der beschriebenen Risiken oder Schadensszenarien möglich erscheint, es ist vielmehr erforderlich, dass das konkrete Risiko *im Einzelfall* über die allgemeinen Gefahren hinausgeht. Eine pauschale Verpflichtung zu einer Datenschutz-Folgenabschätzung aufgrund nur eines dieser Merkmale ist durch den EU-Gesetzgeber nicht vorgesehen. Allerdings könnte sich eine Verpflichtung aus Art. 35 Abs. 3 lit b DSGVO ergeben, nach der eine Datenschutz-Folgenabschätzung gemäß Abs. 1 erforderlich ist, wenn eine *umfangreiche* Verarbeitung besonderer Kategorien von personenbezogenen Daten gemäß Art. 9 Abs. 1 DSGVO erfolgt. Im Parlamentsentwurf der DSGVO wurde in einem ähnlichen Zusammenhang davon ausgegangen, dass bei einem Richtwert von 5.000 betroffenen Personen innerhalb von zwölf aufeinanderfolgenden Monaten ein „konkretes Risiko" erreicht sei.[110] Da die Verordnung in ihrer jetzigen Fassung aber keinen konkreten Richt- oder Schwellwert angibt, ist davon auszugehen, dass ein „voraussichtlich hohes Risiko" auch mit einem deutlich niedrigeren Wert vorliegen kann. Die Aufteilung der Verantwortlichkeiten im Rahmen der ePA erschwert eine Beurteilung der Sachlage. Konkretisiert werden kann die Norm durch eine Positivliste der Aufsichtsbehörde, Art. 35 Abs. 4 DSGVO. Der BfDI hat eine solche Liste veröffentlicht. Allerdings ergeben sich daraus keine besonderen Erkenntnisse: Am Ende des knappen Dokuments wird auf eine grundsätzliche Einzelfallprüfung verwiesen.[111] Für die Krankenkassen und Ärzte wären die Listen der entsprechenden LfDI zu berücksichtigen, deren Prüfung den Rahmen dieser Arbeit sprengen würde.

Schließlich ist davon auszugehen, dass stets im Einzelfall zu betrachten ist, welcher Verantwortliche eine Datenschutz-Folgenabschätzung durchführen muss. Auf der Hand liegt, dass eine einzelne Arztpraxis eher weniger dazu verpflichtet sein wird als eine Krankenkasse oder die gematik mit der gesamten TI im Rücken.[112] Im Angesicht der wenig greifbaren Norm und, zumindest auf deutscher Bundesebene, auch wenig greifbaren Konkretisierung, bleibt abzuwarten, inwiefern hier die Judikative Abhilfe schaffen muss. Gegen die Befugnisse einer Aufsichtsbehörde aus Art. 58 DSGVO, kann sich der Verantwortliche gem. Art. 78 Abs. 1 DSGVO gerichtlich wehren. Die Pflichten aus Art. 35 DSGVO betreffen übrigens nicht die Rechtmäßigkeit des einzelnen Verarbeitungsvorgangs.[113]

2.5.4. Ausschließlicher Zugang über Mobile Devices in 2021

Der Zugang zur elektronischen Patientenakte erfolgt in der ersten Ausbaustufe der ePA bis 1. Januar 2022 ausschließlich über Apps für iOS oder Android („Frontend-Nutzer"). In der zweiten Ausbaustufe ab 1. Januar 2022 sollen die Krankenkassen verpflichtet sein, auch „Frontend-Nichtnutzern" mit anderen Mitteln den Zugriff auf die ePA ermöglichen, § 338 Abs. 1 SGB V.[114] Ursprünglich war hierfür eine flächendeckende Bereitstellung von entsprechenden Terminals vorgesehen.[115] Diese Regelung wurde aber später wieder verworfen.[116] Nun soll von der gematik nur noch evaluiert werden, ob ein Bedarf für eine flächendeckende Bereitstellung solcher Einrichtungen besteht, § 338 Abs. 2 SGB V.

Der BfDI sieht in der mangelnden Ausstattung für Frontend-Nichtnutzer eine Ungleichbehandlung im Grundrecht auf informationelle Selbstbestimmung.[117] Er bemängelt einen Wertungswiderspruch zu § 336 SGB V der den Versicherten garantiert, auf die Daten ihrer ePA zugreifen zu können.[118]

Ob ein solcher Wertungswiderspruch in der Sozialgesetzgebung besteht, muss hier nicht geprüft werden, da der Wertungsmaßstab dieser Arbeit sich nur aus der DSGVO ergibt.[119]

2.5.5. Berechtigungsmanagement

Die Einführung der ePA erfolgt wie bereits dargestellt in mehreren Ausbaustufen mit unterschiedlichen Möglichkeiten in der Berechtigungsverwaltung für Dokumente. Ein wesentliche Kritikpunkt des BfDI ist das Ausbleiben eines differenzierten Berechtigungsmanagements, das den Zugriff der Leistungserbringer und anderer berechtigter Personen regelt.[120] Der allgemeine Zugriff auf Daten des Versicherten ist in § 339 SGB V und der Zugriff durch Leistungserbringer auf Daten der elektronischen Patientenakte ist in § 352 SGB V geregelt. Voraussetzungen an die Einwilligung des Versicherten bezüglich des Zugriffs anderer sind in § 342 Abs. 2 Nr. 1 lit. c, sowie Nr. 2 lit. b und speziell in § 353 SGB V festgelegt. Die ePA wird unter diesem Aspekt in zwei Schritten eingeführt. Mit dem Start der ePA ab 1. Januar 2021 gibt es noch kein ausdifferenziertes Rechtemanagement, während spätestens ab 1. Januar 2022 die Versicherten eine Einwilligung auf den Zugriff sowohl auf spezifische Dokumente und Datensätze als auch auf Gruppen von Dokumenten und Datensätzen erteilen können sollen. Dies ergibt sich aus den Formulierungen in § 342 Abs. 2 Nr. 1 lit. c und Abs. 2 Nr. 2 lit. b SGB V.

---

[108]Art. 35 Abs. 1 DSGVO „aufgrund".

[109]Vgl. Ehmann/Selmayr/Baumgartner, DSGVO, Art. 35 Rn. 24; Sydow/Schwendemann, DSGVO, Art. 35 Rn. 10.

[110]Vgl. ErwGe 63 und 75, Art, 25 Abs. 2 lit. b, Art. 32a Abs. 2 lit. a und Art. 35 Abs. 1 lit. b des Beschlusses des Europäischen Parlaments vom 12.3.2014, COM(2012)0011 – C7 – 0025/2012 – 2012/0011 (COD).

[111]Vgl. BfDI (2019, S. 2).

[112]Vgl. ähnlich auch Ehmann/Selmayr/Baumgartner, DSGVO, Art. 35 Rn. 39.

[113]Vgl. nur Ehmann/Selmayr/Baumgartner, DSGVO, Art. 35 Rn.78.

[114]Vgl. BR-Drs. 164/20, S. 119.

[115]Vgl. BR-Drs. 164/20, S. 33.

[116]Vgl. Verbraucherzentrale (2020, o. S).

[117]Vgl. BfDI (2020a, o. S).

[118]Vgl. BfDI (2020b, S. 3).

[119]Vgl. unten 2.7.1.1 zur Frage, ob diese potentielle Diskriminierungslage einen Einfluss auf die Einwilligung i.S.v. Art. 4 Nr. 11 DSGVO hat.

[120]Vgl. zuletzt BfDI (2020b, S. 3).

Damit die Berechtigungserteilung einer Einwilligung im Sinne der DSGVO entspricht, müssen die Voraussetzungen nach Art. 4 Nr. 11 DSGVO gegeben sein. Die technische Erteilung der Zugriffsberechtigung ist von der Einwilligung gemäß Art. 4 Nr. 11 DSGVO zu trennen, so dass eine Einwilligung widerrufen werden kann, auch wenn technisch der Zugriff noch möglich wäre. Die Grundlage der Verarbeitung würde dann dennoch wegfallen.

Die folgenden Abbildungen geben einen Überblick über die verschiedenen Möglichkeiten des Zugriffs durch die Leistungserbringer.

Die DSGVO-Konformität des Berechtigungsmanagement wird unten in Kapitel 2.7 f. näher betrachtet.

2.5.6. Erkenntnis für die vorliegende Arbeit

Im Rahmen der ePA gibt es eine Vielzahl von Ansatzpunkten zur Prüfung datenschutzrechtlicher Vorgaben. Auch eine abstrahierte – vom Einzelfall losgelöste – Prüfung der DSGVO-Konformität müsste, sofern sie den Anspruch an ihre Vollständigkeit genügen sollte, im Grunde für jede Art von Verarbeitungsvorgang gesondert erfolgen. Angenommen jeder unten abgebildete Anwendungsfall repräsentierte einen Verarbeitungsvorgang, so wird deutlich, dass die Prüfung der einzelnen Verarbeitungsvorgänge im Rahmen dieser Arbeit nicht möglich ist (Abbildung 9). Eine detaillierte Prüfung soll im folgenden Kapitel abschließend für die zuletzt vorgebrachten Kritikpunkte des BfDI hinsichtlich der ePA erfolgen, deren Relevanz derzeit am höchsten sein dürfte.

2.6. Spannungsverhältnis zwischen DSGVO und SGB V

Seit Oktober 2020 gilt das PDSG zur näheren Ausgestaltung der ePA und verpflichtet die gesetzlichen Krankenkassen zum Angebot der ePA ab 1. Januar 2021. Der BfDI kritisierte schon frühzeitig einige Aspekte des PDSG und in seinen Stellungnahmen zum PDSG wurden diese erneut formuliert.[121] Am 16. November 2020 übermittelte der BfDI eine Warnung an die gesetzlichen Kassen in seinem Zuständigkeitsbereich.[122] Aus Sicht des BfDI verstießen die Krankenkassen gegen die DSGVO, sofern sie die ePA nur nach den Vorgaben des PDSG umsetzten. Kritisiert werden zwei Eigenschaften der ePA: zum einen das Fehlen eines feingranularen Rechtemanagements[123], zum anderen die Verwendung sogenannter alternativen Authentifizierungsverfahren[124]. Der Aufsicht des BfDI unterliegen derzeit 63 gesetzliche Krankenkassen mit ca. 44 Millionen Versicherten.[125] Diese Krankenkassen müssten Zusatzfunktionen anbieten, die ein feingranulares Rechtemanagement ermöglichen und sie müssten von den alternativen Authentifizierungsverfahren absehen. Der BfDI behielt sich vor weitere Maßnahmen nach Art. 58 Abs. 2 DSGVO zu prüfen.[126]

In der Antwort auf eine kleine Anfrage ging die Bundesregierung auf die Auffassung des BfDI ein.[127] Die Bundesregierung teilt die Bedenken des BfDI nicht. Das PDSG sei von den Verfassungsressorts umfassend auf die Vereinbarkeit mit übergeordnetem Recht geprüft worden. Der BfDI sei fortlaufend in die fachliche Diskussion mit eingebunden und hätte maßgeblich an der Erarbeitung der Regelungen im Gesetzesentwurf zum PDSG mitgewirkt.

Das Bundesamt für Soziale Sicherung (BAS) formulierte seine Ansichten in einer Stellungnahme zur Warnung des BfDI und stärkte den Krankenkassen den Rücken.[128] Die Vereinbarkeit des PDSG mit übergeordnetem Recht sei geprüft worden. Die Nutzung der ePA sei freiwillig und die Verarbeitung geschehe aufgrund einer wirksamen Einwilligung (Art. 4 Nr. 11 DSGVO). Darüber hinaus seien die Informationspflichten der DSGVO auch in der Sozialgesetzgebung berücksichtigt worden (§ 343 Abs. 1 Nr. 3 SGB V). Das Diskriminierungsverbot in § 335 SGB V garantiere eine echte Wahlfreiheit. Auch allgemeine Grundsätze der DSGVO seien nicht verletzt worden (Art. 5, 25 und 32 DSGVO).

Zutreffend ist, dass die geforderten Maßnahmen des BfDI bis zum Einführungstermin der ePA nicht umsetzbar sind, da die Spezifikationen bei der gematik noch nicht definiert waren und erst eine Zulassung der gematik erfolgen müsste.[129] Das BAS müsste als Ausführungsbehörde die Sanktionen ohne Ermessen vollziehen, sofern dies der GKV-Spitzenverband in einem entsprechenden bestandskräftigen Bescheid feststellt, § 342 Abs. 5 S. 2 SGB V.[130]

Damit wurden die Krankenkassen vom Gesetzgeber und oberster Datenschutzaufsichtsbehörde in ein Dilemma bugsiert, das diese nicht auflösen können.

Im Dezember 2020 wurde ein Gutachten der Kanzlei Redeker Sellner Dahs veröffentlicht, das durch das *health innovation hub* des Bundesgesundheitsministeriums beauftragt wurde. Die Verfasser des Gutachtens gelangen zum Ergebnis, dass die Umsetzung der Sozialgesetzgebung nicht gegen das europäische Datenschutzrecht verstößt.[131]

Aufgrund des oben beschriebenen Konfliktpotentials der Normen und der dargelegten allgemeinen Anwendbarkeit der DSGVO (s.o. Kapitel 2.4), soll im Folgenden eine ausführliche Prüfung der Hauptkritikpunkte des BfDI – i.e. das Fehlen eines ausdifferenzierten Berechtigungsmanagements und die Sicherheit der alternativen Authentifizierungsverfahren – nach Maßgabe der DSGVO erfolgen.

2.7. Berechtigungsmanagement ab 1. Januar 2021, § 342 Abs. 2 Nr. 1 SGB V

Die allgemeine sachliche und räumliche Anwendbarkeit und das grundsätzliche Vorhandensein von Verarbeitungsvor-

---

[121] BfDI (2020b); BfDI (2020a); BfDI (2020a, o. S.).
[122] BfDI (2020b).
[123] Vgl. a.a.O., S. 1 ff.
[124] Vgl. BfDI (2020a, o. S.); BfDI (2020a, S. 7).
[125] BfDI (2021, o. S.); BfDI (2020a, o. S).
[126] BfDI (2020b, S. 3).

[127] Vgl. BT-Drs. 19/23243, S 2 f.
[128] Vgl. Bundesamt für Soziale Sicherung, Stellungnahme zur elektronischen Patientenakte, 2020.
[129] Vgl. Bundesamt für Soziale Sicherung, Stellungnahme zur elektronischen Patientenakte, 2020, S. 2 f.
[130] Vgl. a.a.O., S. 3.
[131] Vgl. Böllhoff, Cornelius, et al. (2020, S. 4 f.)

Berufsgruppe	Mögliche Verwendung medizinischer Informationen über Versicherte	Mögliche Verwendung der von Versicherten zur Verfügung gestellten Gesundheitsinformationen
Ärztinnen/Ärzte, Zahnärztinnen/Zahnärzte, Psychotherapeutinnen und -therapeuten, Krankenhäuser, Rehakliniken	Verarbeitung	Verarbeitung
Apothekerinnen/Apotheker	Auslesen, Speicherung und Verwendung sowie Verarbeitung des elektronischen Medikationsplans	Auslesen, Speicherung und Verwendung
Gesundheits- und Krankenpflegerinnen und -pfleger, Gesundheits- und Kinderkrankenpflegerinnen und -pfleger, Altenpflegerinnen und -pfleger, Pflegefachfrauen und Pflegefachmänner sowie deren Helferinnen und Helfer	Auslesen, Speicherung und Verwendung	Auslesen, Speicherung und Verwendung
Hebammen/Entbindungspfleger	Auslesen, Speicherung und Verwendung	Auslesen, Speicherung und Verwendung
Physiotherapeutinnen/Physiotherapeuten	Auslesen, Speicherung und Verwendung sowie Verarbeitung von Daten zu Befunden, Diagnosen, durchgeführten und geplanten Therapiemaßnahmen oder Früherkennungsuntersuchungen, von Behandlungsberichten und sonstigen untersuchungs- und behandlungsbezogenen medizinischen Informationen, die sich aus der physiotherapeutischen Behandlung ergeben	Auslesen, Speicherung und Verwendung
Arbeitsmedizinerinnen/Arbeitsmediziner, Betriebsärztinnen/Betriebsärzte	Auslesen, Speicherung und Verwendung	Auslesen, Speicherung und Verwendung
Öffentlicher Gesundheitsdienst	Verarbeitung, soweit für Aufgabenerfüllung erforderlich	Verarbeitung, soweit für Aufgabenerfüllung erforderlich

Abbildung 7: Zugriff der Leistungserbringer auf die ePA ab Januar 2021

Quelle: GKV-Spitzenverband (2020a, S. 15).

Kontrolle	Ausbaustufe	Berechtigung für	auf	wie
undifferenziert	1. Januar 2021	zugriffsberechtigte Leistungserbringer	Dokumente von Leistungserbringern UND ODER Dokumente vom Versicherten	App/eGK
differenziert	1. Januar 2022	zugriffsberechtigte Leistungserbringer	Vertraulichkeitsstufe ▪ normal UND ODER vertraulich UND ODER streng vertraulich	App/eGK/Vertreter
			Mittlere Granularität ▪ eine oder mehrere Kategorien einer Vertraulichkeitsstufe	App/eGK/Vertreter
			Feine Granularität ▪ einzelne Dokumente	App/eGK/Vertreter

Abbildung 8: Zusammenfassung Berechtigungsmanagement

Quelle: In Anlehnung an GKV-Spitzenverband (2020a, S. 15 ff).

gängen wurden oben dargelegt (oben 2.4). Im weiteren Verlauf der Prüfung sei stets von der Verarbeitung von Gesundheitsdaten, als Daten einer besonderen Kategorie im Sinne des Art. 9 Abs. 1 DSGVO, ausgegangen. Insofern wird die Einwilligung bezüglich der Registrierung bzw. Authentifizierung hier nicht berücksichtigt, weil dabei ausschließlich Stammdaten betroffen sind, die in der Regel keine Daten besonderer Kategorien enthalten.[132] Insofern sei auf Verarbeitungsvorgänge im Rahmen der Nutzung der ePA, i.e. die Verarbeitung von gesundheitsbezogenen Dokumenten und Datensätzen, abgestellt.

_____

[132] Zur Frage, ob die personenstandsrechtliche Geschlechtsangabe ein sensibles Datum sein kann: Ehmann/Selmayr/Schiff, Art. 9 Rn. 30 dort Fn. 80.
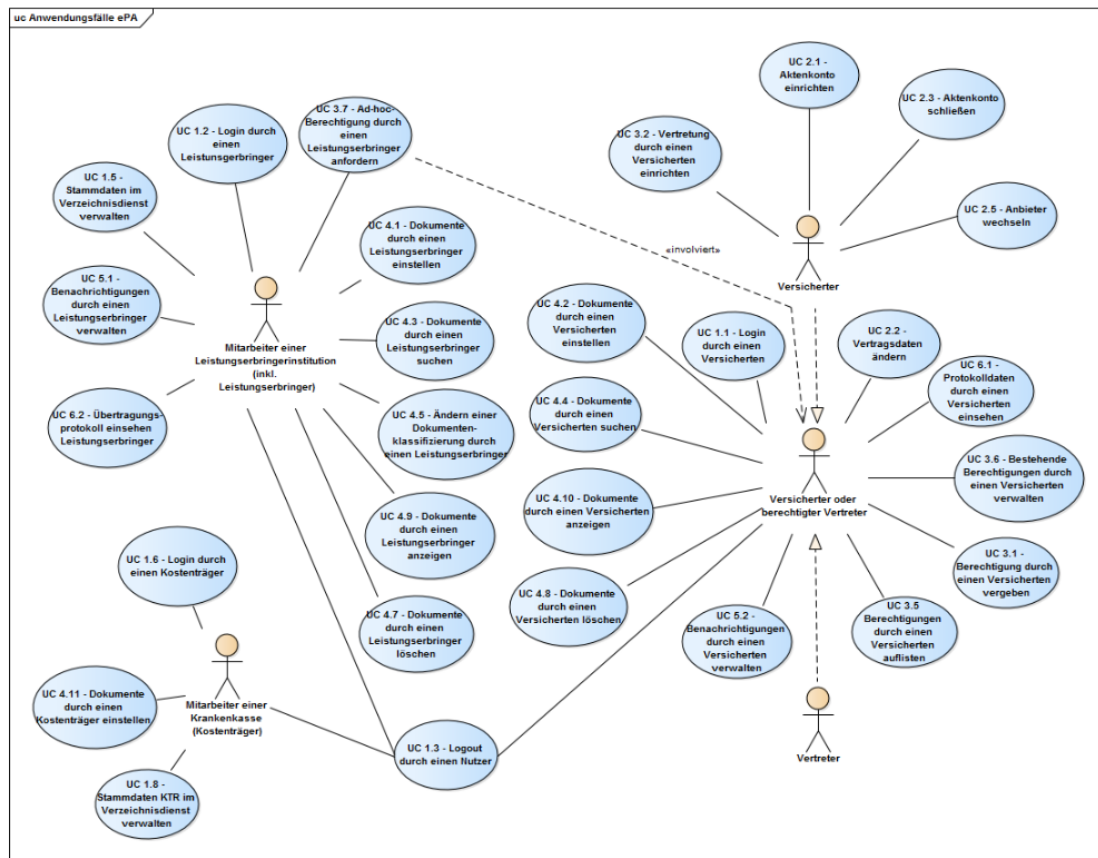
Abbildung 9: Anwendungsfälle der ePA

Quelle: Gesellschaft für Telematik (gematik) (2019, S. 37)
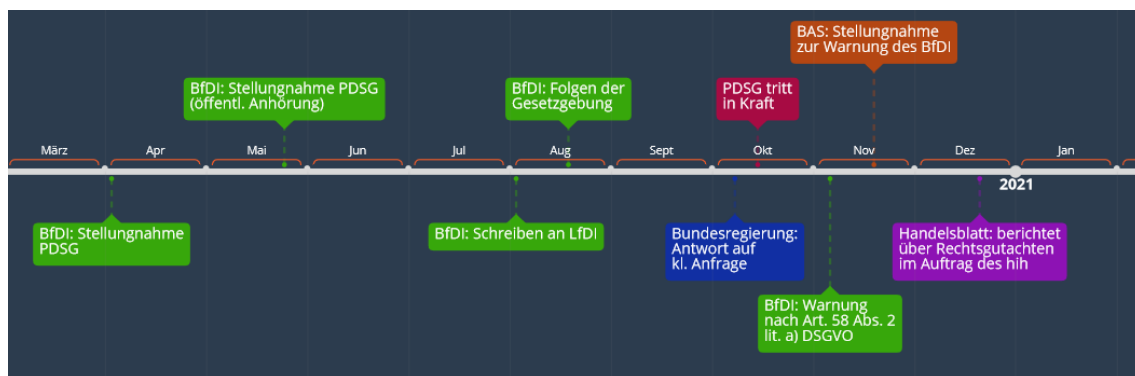
Abbildung 10: Zeitstrahl – BfDI vs ePA

Quelle: Eigene Darstellung

2.7.1. Rechtmäßigkeit der Datenverarbeitung aufgrund einer Einwilligung

Die Rechtmäßigkeit der Datenverarbeitung ergibt sich aus den Bestimmungen des Art. 6 Abs. 1 Satz 1 bzw. Art. 9 Abs.2 DSGVO.

Die Regelungen der Sozialgesetzgebung, insbesondere §§ 339 Abs. 1, 342 Abs. 2 Nr. 1 lit. c, 353 SGB V, verlangen für die Verarbeitung der in der ePA gespeicherten Gesundheitsdaten eine Einwilligung im Sinne des Art. 6 Abs. 1 lit. a bzw. Art. 9 Abs. 2 lit. a DSGVO.[133] Insofern besteht keine eigene Rechtsgrundlage in der Sozialgesetzgebung, die Verarbeitungsvorgänge in der ePA legitimieren könnten. Mithin scheidet eine

---

[133]Vgl. Böllhoff, Cornelius, et al. (2020, S. 23 f.)

Verarbeitung aufgrund von Art. 9 Abs. 2 und lit. h DSGVO aus.

Eine Verarbeitung aufgrund von Art. 9 Abs. 2 lit. c DSGVO kommt gegebenenfalls in Ausnahmefällen in Betracht. Beispielsweise, wenn in der ePA überlebenswichtige Informationen abgerufen werden können und die betroffene Person außer Stande ist, eine Einwilligung nach Art. 4. Nr. 11, Art. 7, Art. 9 Abs.2 lit. a DSGVO abzugeben.[134] Allerdings ist das nicht der Regelfall und taugt daher nicht als Gegenstand einer vom Einzelfall losgelösten Prüfung.

Eine Verarbeitung nach Art. 9 Abs. 2 lit. g und lit. i DSGVO scheidet in Ermangelung entsprechender Gesetzgebung aus.[135]

Fraglich ist aber, ob eine Datenverarbeitung auch gem. Art. 6 Abs. 1 lit. b DSGVO nicht schon aufgrund eines Nutzungsvertrag rechtmäßig ist. Dafür muss die Verarbeitung für die Vertragserfüllung erforderlich sein.[136] Hier könnte zwischen den Backend-seitigen und Frontend-seitigen Verarbeitungsvorgängen zu unterschieden werden. Wenn beispielsweise einem Leistungserbringer eine Berechtigung erteilt wird (nicht erforderlich), ist es zwingend notwendig, dass entsprechende Parameter im ePA-Aktensystem im Verantwortungsbereich der Krankenkasse gespeichert werden (erforderlich). Diese Speicherung der Einstellung können auch – zumindest mittelbare – Gesundheitsdaten enthalten, und zwar Metadaten wie die Art der Freigabe, der zugriffsberechtigter Arzt und dessen Fachrichtung. Die Erfüllung eines Vertrags ist allerdings kein Ausnahmetatbestand des Art. 9 Abs. 2 DSGVO. Insofern ist Art. 6 Abs. 1 lit. b DSGVO nicht auf Gesundheitsdaten anwendbar.[137] Grundlage einer Datenverarbeitung von Gesundheitsdaten in der ePA kann daher ausschließlich eine *Einwilligung gemäß Art. 6 Abs. 1 lit. a bzw. Art. 9 Abs. 2 lit. a, Art. 7 und Art. 4 Nr. 11 DSGVO* sein. Diese Einwilligung kann gleichwohl für den Zweck eingeholt werden, um Funktionalitäten bereitstellen zu können, die der Durchführung eines Nutzungsvertrages dienen.[138] Die Einwilligung bezieht sich im oben angeführten Beispiel daher auf das Auslesen der Daten durch den Leistungserbringer, und auch auf die Speicherung der Einstellung, dass diese Daten gelesen werden dürfen.

Im Sinne des Art. 4 Nr. 11 DSGVO ist eine Einwilligung jede freiwillig für den bestimmten Fall, in informierter Weise und unmissverständlich abgegebene Willensbekundung in Form einer Erklärung oder einer sonstigen eindeutigen bestätigenden Handlung, mit der die betroffene Person zu verstehen gibt, dass sie mit der Verarbeitung der sie betreffenden personenbezogenen Daten einverstanden ist, Art. 4 Nr. 11 DSGVO. Weitere Anforderungen an die Einwilligung werden in Art. 7 DSGVO aufgeführt.

Freiwilligkeit, Art. 4 Nr. 11, Art. 7 Abs. 4 DSGVO

Freiwilligkeit liegt vor, wenn die betroffene Person über eine echte Wahl und die Kontrolle bei der Abgabe der Einwilligung verfügt.[139] Die Einwilligung ist nichtig, sofern die betroffene Person keine echte Wahl hat, sich zur Einwilligung gedrängt fühlt oder negative Auswirkungen erdulden muss, wenn sie nicht einwilligt.[140] Dies ist beispielsweise der Fall, wenn die betroffene Person die Einwilligung nicht verweigern oder zurückziehen kann, ohne Nachteile zu erleiden.[141] Unangemessener Druck oder Einflussnahme auf die betroffene Person, die diese von der Ausübung ihres freien Willens abhalten, können demnach zur Unwirksamkeit der Einwilligung führen.[142]

Letztlich muss die Einwilligung zwar im konkreten Fall betrachtet werden, eine systematische Beeinflussung in der Willensbildung ist bei der ePA aber nicht erkennbar. Insbesondere die für den Vertrieb zuständigen Krankenkassen treten bezüglich der ePAen eher verhalten auf. Insbesondere wird auf intensive Bewerbung verzichtet.[143] Zudem könnte das Diskriminierungsverbot gem. § 335 SGB V zur Wahrung der Freiwilligkeit, auch im konkreten Fall, beitragen. Dem Versicherten dürfen keine Nachteile entstehen, wenn er die Nutzung der ePA oder den Zugriff auf Daten in der ePA ablehnt oder nachträglich widerruft. Demnach soll es sich bei der ePA um ein zusätzliches Angebot an den Versicherten handeln. Die Leistungen der Leistungserbringer sollen mit oder ohne Zugriff auf die ePA in der gleichen Güte erbracht werden.[144] Fraglich ist aber, ob diese gesetzliche Kodifizierung eine tatsächliche Ungleichgewichtslage zu verhindern vermag. Maßgeblich ist die tatsächliche Gestaltung der ePA und damit verbundene objektive feststellbare Vorteile und deren Wirkung auf die Zielgruppe. Dabei müssen Nutzen und Kosten gegeneinander abgewogen werden. Es wird kaum zu bestreiten sein, dass die Nutzung der ePA über mobile Endgeräte Vorteile mit sich bringt, beispielsweise den schnellen und unkomplizierten Zugriff und eine bes-

---

[134]Vgl. Forgó/Helfrich/Schneider/Arning/Born, Betrieblicher Datenschutz, Teil X. Kapitel 2 Rn. 47.

[135]Vgl. a.a.O., Rn. 48.

[136]Ehmann/Selmayr/Heberlein, DSGVO, Art. 6 Rn. 13; Paal/Pauly/Frenzel, DSGVO, Art. 6 Rn. 14; Kühling/Buchner/Buchner/Petri, DSGVO, Art. 6 Rn. 38 ff.

[137]Vgl. Forgó/Helfrich/Schneider/Arning/Born, Betrieblicher Datenschutz, Teil X. Kapitel 2 Rn. 52 dort Fn. 39; Engler, ZD 2018, 55, 58. Ungenau: DAK (2021, o. S.), Punkt 6.: „Die Rechtsgrundlage für die Datenverarbeitung im Rahmen der Nutzung der ePA ist Art. 6 Abs. 1 lit. b DSGVO, da diese Datenverarbeitungsvorgänge für die ordnungsgemäße Erfüllung des Nutzungsvertrags mit dem Nutzer über die ePA erforderlich sind."

[138]Vgl. so z.B: Die Techniker, Datenschutzerklärung für die elektronische Patientenakte und „TK-Safe", 2021, o. S.

[139]Vgl. ErwG 42 DSGVO; Artikel-29-Datenschutzgruppe, Leitlinien in Bezug auf die Einwilligung, S.6; Paal/Pauly/Ernst, DSGVO, Art. 4 Rn. 69; Sydow/Ingold, DSGVO, Art. 7 Rn. 29.

[140]Vgl. Böllhoff, Cornelius, et al. (2020, S. 26).

[141]ErwG 42 DSGVO, BeckOK Datenschutzrecht/Stemmer, 33 Art. 7 Rn. 38.

[142]Vgl. Gola/Schulz, DSGVO, Art. 7 Rn. 21.

[143]Vgl. aerzteblatt.de Elektronische Patientenakte (2020). Dies entspricht auch dem Eindruck des Verfassers: Auf den Internetauftritten der größten gesetzlichen Krankenkassen (AOKen, Techniker, BARMER, DAK) wird auf der Hauptseite nicht direkt auf die Verfügbarkeit einer elektronischen Patientenakte verwiesen. Auf der Internetpräsenz der KKH findet man überhaupt keine Verlinkung zu den App Distributionen von Google und Apple (Stand: Mai 2021).

[144]Vgl. Böllhoff, Cornelius, et al. (2020, S. 27).

sere Strukturierung und Sammlung von gesundheitsrelevanten Dokumenten (vgl. oben Kapitel 2.2.2). Ohne diese Vorteile wäre die Entwicklung und Bereitstellung der ePA allerdings auch nicht wirtschaftlich und überhaupt notwendig. Die Möglichkeit der Nutzung dieser Vorteile begründet keine Diskriminierungslage. Es wird insbesondere kein unbilliger emotionaler oder wirtschaftlicher Druck ausgeübt.

Insofern besteht auch keine Zwangslage bei Frontend-Nichtnutzern. Es ist kaum vorstellbar, dass Versicherte, die kein mobiles Endgerät nutzen können oder wollen, einem derartigen Druck unterliegen, dass sie sich zur Verwendung eines mobilen Endgeräts gezwungen fühlen. Mithin ist eine potentielle Ungleichbehandlung allenfalls innerhalb der Sozialgesetzgebung gegebenenfalls unter der Berücksichtigung von grundrechtlichen Garantien zu prüfen.[145]

Auch ein latentes Ungleichgewicht in der Arzt-Patienten-Beziehung, das je nach Schwere einer Erkrankung unterschiedlich sein dürfte und sich auch bei der Zugriffserteilung in der ePA auswirken könnte, schließt nicht von vornherein die Freiwilligkeit einer Einwilligung aus.[146] Eine von der Literatur teilweise geforderte Kompensation des Ungleichgewichts durch weitere Voraussetzungen dürfte bei der ePA erfüllt sein[147]:

- Umfassende Aufklärung des Patienten darüber, zu welchem spezifischen Zweck seine Daten verarbeitet werden sollen

Art. 13 Abs. 1 lit. c DSGVO und entsprechende Regelungen in der Sozialgesetzgebung beinhalten diese Anforderung (vgl. unten Kapitel 2.7.1.3).

- Behandlungsnachteile dürfen nicht an die Entscheidung des Patienten geknüpft werden

Hier greift das Diskriminierungsverbot nach § 335 SGB V.

- Die Ergebnisse der Datenverarbeitung können dem Patienten zu Gute kommen

- Die Datenverarbeitung selbst erfolgt zu nichtkommerziellen Zwecken

Auch das ist der Fall.

Unter dem Aspekt der Freiwilligkeit muss weiterhin das in Art. 7 Abs. 4 DSGVO und ErwG 43 S. 2 2. Alt. DSGVO zum Ausdruck kommende *Kopplungsverbot* geprüft werden.[148] Ein gewisser Druck könnte auf die betroffene Person einwirken, sofern keine Alternativen auf dem Markt verfügbar sind.[149]

Weil letztlich aber schwierig festzustellen ist, ob eine Leistung substituierbar ist, wird der Schutzzweck des Art. 7 Abs. 4 DSGVO eng auszulegen sein, nämlich insofern, dass geprüft werden muss, ob die Daten für den Vertragszweck erforderlich sind und die Verweigerung der Einwilligung nicht ohne Nachteil für die betroffene Person erfolgt.[150] Im Kern soll verhindert werden, dass Daten für vom Vertragsgegenstand „vollkommen losgelöste Zwecke" erhoben werden.[151] Wie oben festgestellt wird zum einen kein unbilliger emotionaler oder wirtschaftlicher Druck auf die betroffene Person ausgeübt. Zum anderen sind die Verarbeitungsvorgänge für die Erfüllung des Nutzungsvertrag des Versicherten mit der Krankenkasse unbedingt erforderlich, damit die Funktionalitäten vertragsgemäß bereitgestellt werden können.[152]

Auch eine Kopplung an das Grundverhältnis des Versicherten mit der Krankenkasse besteht nicht, da die Erteilung der Zugriffsmöglichkeit auf die Gesundheitsdaten der Versicherten an die Leistungserbringer demselben *Zweck* dient wie das Grundverhältnis.[153] Beides erfolgt zur Gewährleistung der Gesundheitsversorgung des Versicherten.[154] In § 352 SGB V ist zusätzlich festgeschrieben, dass der Zugriff auf Daten durch Personen in der ePA nur dann erfolgen darf, wenn er für die Versorgung des Versicherten erforderlich ist.

Ferner wird in der Literatur ein *horizontales Kopplungsverbot* genannt. Dieses soll verhindern, dass zu verschiedenen nebeneinander bestehenden Verarbeitungsvorgängen nicht gesondert eine Einwilligung des Betroffenen eingeholt wird, obwohl dies im Einzelfall angebracht wäre.[155] Wegen des fehlenden ausdifferenzierten Berechtigungsmanagements kann es im Kontext der ePA, beispielsweise durch das parallele Abfragen von Dokumenten durch verschiedene Leistungsträger zu nebeneinander bestehenden Verarbeitungsvorgängen kommen, die möglicherweise eine gesonderte Einwilligung des Betroffenen erfordern. Diese Frage soll jedoch nicht an dieser Stelle, sondern vielmehr nachfolgend im Rahmen des Merkmals des „bestimmten Falls" geprüft werden.

Bestimmtheit

Die Einwilligungserklärung muss gem. Art. 4 Nr. 11 DSGVO weiterhin „für den bestimmten Fall" abgegeben werden, damit der Betroffene die Reichweite seiner Erklärung überblicken kann und der Verantwortliche Grenzen für die Verarbeitung der personenbezogenen Daten vorfindet.[156] Die Bestimmtheit der Einwilligung ergibt sich de facto aus

[145]Vgl. dazu ausführlich, in ähnlichem Zusammenhang: Böllhoff, Cornelius, et al. (2020, S. 39).
[146]Vgl. Otto und Rüdlin (2017), 519, 521 m.w.N.
[147]Vgl. ebd. mit Verweis auf Lodzig, ZD-Aktuell 2012, 02952.
[148]Vgl. beispielsweise Ehmann/Selmayr/Heberlein, DSGVO, Art. 6 Rn. 7; Taeger/Gabel/Taeger, DSGVO, Art. 6 Rn. 29.
[149]Vgl. Teager/Gabel/Taeger, DSGVO, Art. 7 Rn. 85; Plath/Plath, DSGVO, Art. 7 Rn. 15; Albrecht, CR 2016, 88, 91; Paal/Pauly/Frenzel, DSGVO, Art. 7 Rn. 20; Gola/Gola, DSGVO Art. 4 Rn. 85.

[150]Vgl. Teager/Gabel/Teager, Art. 7 Rn. 93; Golland (2018), 130, 132; Laue/Kremer/Kremer, Das neue Datenschutzrecht in der betrieblichen Praxis, §2 Rn. 22 f.
[151]Vgl. Gola/Schulz, DSGVO, Art. 7 Rn. 24 f.
[152]Vgl. beispielsweise BKK, Nutzungsvertrag, o. S. Punkt 2: „Die ePA ermöglicht dem Nutzer die sichere Speicherung, übermittlung und Verwaltung ihrer Gesundheitsdaten (z.B., Befunde, Laborberichte, Arztbriefe, etc.)."
[153]Vgl. Böllhoff, Cornelius, et al. (2020, S. 32).
[154]Vgl. a.a.O., S. 33.
[155]Vgl. Gola/Gola, DSGVO, Art. 4 Rn. 85; ErwG 43 S. 2 2. Alt. DSGVO.
[156]Vgl. BeckOK Datenschutzrecht/Stemmer, Rn. 74; Forgó/Helfrich/Schneider/Arning/Born, Teil X. Kapitel 2 Rn.53.

dem Zusammenspiel der Informationen an den Betroffenen und der Einwilligungserklärung, womit sich folglich das Bestimmtheitserfordernis mit der Informationsplicht überschneidet (s. nachfolgendes Kapitel).[157] Umfasst vom Bestimmtheitsgrundsatz sind der Verantwortliche, die Art der Verarbeitung, die Verarbeitungszwecke und letztlich die zu verarbeitenden Daten selbst.[158]

Insbesondere hinsichtlich des zuletzt genannten Kriteriums, könnten sich im Rahmen der elektronischen Patientenakte Probleme ergeben. Denn in der ersten Ausbauphase der ePA ist kein ausdifferenziertes Berechtigungsmanagement vorgesehen. Die Dokumente lassen sich in dieser Ausbaustufe in zwei Kategorien einteilen (s.o. 2.5.5). Der Versicherte kann entscheiden, ob ein Leistungserbringer auf eine oder beide Kategorien zugreifen kann. Der Leistungserbringer kann technisch auf alle Dokumente dieser Kategorie zugreifen.

Der BfDI sieht in diesem Zusammenhang die Freigabe nach einem „Alles-oder-nichts-Prinzip".[159] An dieser Stelle muss bereits festgehalten werden, dass selbst in der ersten Ausbaustufe der ePA kein „Alles-oder-nichts-Prinzip" besteht, da eine Differenzierung nach den einzelnen Leistungserbringern und nach den Kategorien „Dokumente des Leistungserbringers" und „Dokumente des Versicherten" erfolgen kann. Außerdem kann eine zeitliche Befristung des Zugriffs erfolgen.[160]

In der DSGVO ist die Ausgestaltung der Einwilligung hinsichtlich ihrer Granularität nicht definiert. So kann die Einwilligung auch generell zweckbezogen erteilt werden, anstatt bezogen auf jeden einzelnen Verarbeitungsvorgang für denselben Zweck.[161] Dies kommt auch insbesondere in ErwG 32 S. 4 zum Ausdruck. Mithin sind an die Festlegung des Zweckbegriffs in der DSGVO keine klaren Anforderungen festgelegt. So können die Zwecke in ihrem Abstraktionsgrad variieren.

Auch die Zusammenfassung mehrerer Zwecke unter einer Einwilligung ist möglich, sofern diese bestimmt und abschließend bezeichnet sind und die Einwilligung insgesamt nicht gegen das Kopplungsverbot verstößt.[162] Ob dabei besondere Kriterien an die Spezifizierung des Zwecks anzulegen sind, kann letztlich dahinstehen, denn die Einwilligung in der ePA bezieht sich auf *einen* Zweck, nämlich der *gesundheitlichen Versorgung* des Versicherten.

Gegen eine restriktive Auslegung der Bestimmtheit spricht auch die Informationspflicht aus Art. 14 Abs. 1 lit. c DSGVO. Der Gesetzgeber geht davon aus, dass die betroffene Person nicht alle Daten konkret kennen muss, um wirksam einzu-

willigen.[163]

Ebenfalls ist in einem Arbeitspapier zu elektronischen Patientenakten der Artikel-29 Datenschutzgruppe kein datenschutzrechtliches Gebot zur Gewährleistung eines granularen Zugriffsberechtigungsmanagements zu erkennen.[164] Es wird vielmehr empfohlen, dass unterschiedliche Module nach verschiedenen Vertraulichkeitsstufen mit unterschiedlichen Zugangsvoraussetzungen eingerichtet werden sollen.[165]

Im Kontext der ePA ist weiterhin nicht ersichtlich, warum ein derart besonderer Sachverhalt vorliegen sollte, der die Voraussetzungen eines horizontalen Kopplungsverbots erfüllt, ErwG 43 S. 2 1. Alt. DSGVO. Dass die Daten einer besonderen Kategorie nach Art. 9 Abs. 1 DSGVO besonders schutzbedürftig sind reicht dafür allein nicht aus. Außerdem sollte auch keine vollständige Entmündigung der Person betrieben werden, indem ihr trotz maximaler Transparenz die Möglichkeit einer Einwilligung genommen wird.[166] Insofern sei hier teilweise auf die Erfüllung der Informationspflichten und des Transparenzgebots der DSGVO verwiesen.

Ein Minus in der Granularität der Einwilligung könnte durch ein Plus in Information und Transparenz ausgeglichen werden.

Informiertheit und Transparenzgebot

Die Einwilligung muss weiterhin dem Grundsatz der Informiertheit entsprechen, Art. 4 Nr. 11 DSGVO. Dafür muss die betroffene Person im Vorfeld abschätzen können, welche Auswirkungen die Erteilung einer Einwilligung für sie hat und sie muss die Umstände der Datenverarbeitung sowie die Tragweite ihrer Einwilligung eindeutig und klar erkennen können.[167] Die Person muss darüber in Kenntnis gesetzt werden, welche Arten von Daten zu welchem Zweck verarbeitet werden, wer die verantwortliche datenverarbeitende Stelle ist und wie diese zu erreichen ist, sowie an welche Dritten die Daten im Falle der Übermittlung weitergeben werden.[168]

Weiterhin muss die betroffene Person über die Widerrufbarkeit der Einwilligung und die Rechtmäßigkeit der bis zum Widerruf erfolgten Verarbeitung unterrichtet werden, Art. 7 Abs. 3 DSGVO.[169] In der Literatur ist umstritten, ob das Ausbleiben dieser Informationspflicht die Unwirksamkeit der Einwilligung nach sich zieht.[170]

---

[157]Vgl. BeckOK Datenschutzrecht/Stemmer, Art. 7 Rn. 75; Kugelmann, DuD 2016, 566, 568.

[158]Vgl. BeckOK Datenschutzrecht/Stemmer, Art. 7 Rn. 75.

[159]BfDI (2020b, S. 2).

[160]Diese ist standardmäßig auf 7 Tage festgelegt, kann vom Nutzer aber auch kürzer (ab 1 Tag) oder länger (bis 540 Tage) eingestellt werden. Vgl. dazu GKV-Spitzenverband (2020a), Kapitel 5.4, S. 16.

[161]Vgl. BeckOK Datenschutzrecht/Stemmer, Art. 7 Rn. 77.

[162]Vgl. ebd.; ErwG 43 DSGVO.

[163]Vgl. Böllhoff, Cornelius, et al. (2020, S. 29).

[164]Vgl. Artikel-29-Datenschutzgruppe (2007, S. 18).

[165]Vgl. ebd.

[166]Vgl. in ähnlichem Zusammenhang: Gola/Schulz, DSGVO, Art. 7 Rn. 35. In diese Richtung auch Schantz, Wolff, DSGVO, 2017, D II Rn. 517: Es hinge von der Komplexität der Einwilligung ab, ob eine differenzierte Einwilligung angebracht ist. Mit der Granularität der Einwilligung steigt auch ihre Komplexität. Man könnte befürchten, dass der Betroffene überfordert sein dürfte, wenn er eine Vielzahl zusätzlicher Entscheidungen treffen müsste.

[167]Vgl. Kühling/Buchner/Buchner/Kühling, DSGVO, Art. 4 Nr. 11 Rn. 8, Art. 7 Rn. 59.

[168]Vgl. a.a.O., Art. 7 Rn. 59.

[169]Vgl. Taeger/Gabel/Taeger, DSGVO, Art. 7 Rn. 57: Die Informationspflicht bezieht sich sowohl auf Satz 1 als auch auf Satz 2.

[170]Bejahend: Taeger/Gabel/Taeger, DSGVO, Art. 7 Rn. 5; verneinend: Eh-

Weitere Informationspflichten ergeben sich aus Art. 13 und 14 DSGVO und aus den allgemeinen Anforderungen in Art. 12 DSGVO.

Im Rahmen der ePA finden sich konkretisierende Vorschriften in § 343 SGB V und in § 353 Abs. 2 SGB V. § 343 SGB V regelt eine allgemeine Informationspflicht der Krankenkassen mit einer Auflistung zahlreicher Gesichtspunkte, die als nicht abschließend anzusehen sind.[171] Gemäß Absatz 2 der Vorschrift soll den Krankenkassen entsprechendes Informationsmaterial durch den Spitzenverband Bund der Krankenkassen im Einvernehmen mit dem BfDI zur Verfügung gestellt werden. Dieses Material findet man beispielsweise auf den Webseiten der Krankenkassen.[172] Gemäß § 353 Abs. 2 SGB V werden zusätzlich die Leistungserbringer verpflichtet, die Frontend-Nichtnutzer über die fehlende Granularität der Zugriffserteilung zu informieren.

Im Verhältnis zur DSGVO „extensivere Informationspflichten" sind mit Schaffung der genannten Vorschriften nicht von vornherein impliziert.[173] Beispielsweise ist nach europäischer Vorschrift über den Umstand zu informieren, dass es sich bei den zu verarbeitenden Daten um Daten einer besonderer Kategorie im Sinne des Art. 9 Abs. 1 DSGVO handelt, bereits weil die betroffene Person gem. Art. 9 Abs. 2 lit. a DSGVO „ausdrücklich" in die Verarbeitung einwilligen muss.[174] Der Regelkatalog in § 343 Abs. 1 SGB V enthält diesbezüglich beispielsweise keine ausdrückliche Regelung, wobei sich dieser Umstand aus dem Kontext der Norm ergeben sollte. Ob die Informationsübermittlung den Pflichten aus der DSGVO entspricht ist unabhängig von der Normierung im Sozialgesetzbuch für den konkreten Einzelfall zu prüfen.

Hinsichtlich der Präsentation von Informationen gibt es in der DSGVO keine konkreten Vorschriften. Neben den weitgehend unspezifischen Kriterien Art. 12 Abs. 1 S. 3 und 4 DSGVO (präzise, transparente, verständliche und leicht zugängliche Form; klare, einfache Sprache) ist Art. 12 Abs. 7 DSGVO zu nennen, der auf die Möglichkeit der Verwendung von unterstützenden Bildsymbolen verweist. Bislang hat sich keiner der verfügbaren Symbolsätze am Markt durchgesetzt.[175] Um das oben erwähnte Minus der Granularität der Einwilligung auszugleichen, hätte der Gesetzgeber von seinem Gestaltungsspielraum Gebrauch machen können und präzisere Anforderungen an die Präsentation der Informationen formulieren können. Eine spielerische, gestalterisch intelligente Auseinandersetzung, gegebenenfalls

in Verbindung mit einer Anleitung zur Verwendung der Anwendungen, hätte einen viel höheren Wirkungsgrad als die bloße Text-Information über eine Website, die nicht einmal direkt mit der Anwendung verknüpft ist.[176] Eine anschauliche Vermittlung der datenschutzrelevanten Informationen würde auch das Vertrauen der Versicherten in das Produkt ePA stärken.

Fraglich ist darüber hinaus, ob auch über die *Bedenken* des BfDI informiert werden muss.[177] Der BfDI plante die Krankenkassen anzuweisen ihren Versicherten einen „von ihm vorgegebenen Warntext" zukommen zu lassen.[178] Diese Forderung wurde den Krankenkassen bis jetzt nicht offiziell übermittelt.[179]

Unmissverständlichkeit und Form

Die Einwilligung muss gem. Art. 4 Nr. 11 DSGVO eine „unmissverständlich abgegebene Willensbeurkundung in Form einer Erklärung oder einer sonstigen eindeutigen bestätigenden Handlung" sein. Die größte Rechtssicherheit gewährleistet die ausdrückliche Einwilligungserklärung, wobei keine Schriftform erforderlich ist.[180] Es genügt demnach auch eine elektronisch zum Ausdruck gebrachte Einwilligung, beispielsweise durch Auswählen von Checkboxen und Bestätigung durch einen weiteren Klick auf einen Bestätigungs-Button. Der Verantwortliche muss gem. Art. 7 Abs. 1 DSGVO nachweisen können, dass eine dem Verarbeitungsvorgang entsprechende Einwilligung vorliegt. Diese Nachweispflicht erfordert, dass die Einwilligung in einer durch die Norm nicht näher spezifizierten Art dokumentiert werden kann.[181] Damit liegt in der Regel eine unmissverständliche Willensbekundung vor.[182] Sofern die Krankenkassen dieser Dokumentationspflicht, beispielsweise durch die Protokollierung der Einwilligungsparamater, nachkommen können, sind die Willensbekundungen als unmissverständlich zu betrachten.

Widerrufbarkeit

Nach Art. 7 Abs. 3 DSGVO muss die Einwilligung jederzeit widerrufbar sein. Die Widerrufbarkeit stellt sicher, dass der Betroffene nicht unwiederbringlich seines Selbstbestimmungsrechts beraubt wird, sondern ihm ein „Weg zu-

---

mann/Selmayr/Heckmann/Paschke, DSGVO, Art. 7 Rn. 89: Die ausdrückliche Unwirksamkeitsregelung in Absatz 2 spräche gegen eine entsprechende Auslegung für Absatz 3; Ernst, ZD 2017, 110, 112; Differenzierung, ob die Datenverarbeitung bereits begonnen hat oder nicht Simitis/Hornung/Spiecker/Klement, DSGVO, Art. 7 Rn. 95.

[171] Vgl. §343 Abs. 1 Satz 3 SGB V „insbesondere".

[172] So z.B. AOK https://www.aok.de/pk/fileadmin/user_upload/Universell/05-Content-PDF/RS_2020-852_Anlage_01_-_mit_AOK_Anpassung_ITSG_GKV-SV.pdf.

[173] So. z.B. Böllhoff, Cornelius, et al. (2020, S. 31).

[174] Vgl. Kühling/Buchner/Buchner/Kühling, DSGVO, Art. 7 Rn. 60a.; s. auch Böllhoff, Cornelius, et al. (2020, S. 31).

[175] Vgl. Kühling/Buchner/Bäcker, DSGVO, Art. 12 Rn. 20.

[176] So jedenfalls bei der AOK App „Mein Leben". Die Informationen nach §343 SGB V sind nicht direkt in der App enthalten. Vgl. zum Thema Medienbruch: Kühling/Buchner/Buchner/Kühling, DSGVO, Art. 7 Rn. 60a; Lüdemann, Pokrant, DuD 2019, 365, 368.

[177] z.B. Borchers, Detlef, Diagnose: digital!, c't 2021, 116, 118: „Zunächst werden sie [die Kassen] auf jeden Fall ihre Versicherten anschreiben und die Bedenken des obersten Datenschützers übermitteln.".

[178] Vgl. BfDI (2020a, S. 2); BfDI (2020b).

[179] Dementsprechend enthält die AOK App „Mein Leben" auch keinen Warntext.

[180] Vgl. BeckOK Datenschutzrecht/Stemmer, Art. 7 Rn. 80.

[181] Vgl. BeckOK Datenschutzrecht /Stemmer, Art. 7 Rn. 80, Rn. 88.

[182] Vgl. BeckOK Datenschutzrecht /Stemmer, Art. 7 Rn 80; Paal/Pauly/Frenzel, DSGVO, Art. 7 Rn. 6.

rück" offenbleibt.[183] Der Widerruf hinsichtlich der Nutzung der gesamten Anwendung ist in der Information des GKV-Spitzenverbands geregelt.[184] Die technische Rückabwicklung einzelner Zugriffsberechtigungen der Leistungserbringer ist für Frontend-Nichtnutzer derzeit über das erneute Setzen einer Zugriffsberechtigung mit der Mindestdauer (1 Tag) zu realisieren.[185] Das ist nicht elegant, aber zumutbar.

Zwischenergebnis

Damit sind die Voraussetzungen einer wirksamen Einwilligung nach Art. 6 Abs. 1 Satz 1 lit. a bzw. Art. 9 Abs. 2 lit. a DSGVO gegeben. Insbesondere die fehlende Granularität der Zugriffserteilung steht der Wirksamkeit der Einwilligung nicht entgegen.

2.7.2. Grundsätze des Datenschutzes gem. Art. 5 DSGVO

Fraglich ist, ob über die Rechtmäßigkeit der Verarbeitung hinaus, die in Art. 5 DSGVO benannten Grundsätze des Datenschutzes verletzt werden. Implikationen dieser Grundsätze wurden teilweise schon im Rahmen der Wirksamkeit der Einwilligung geprüft. Der BfDI stellt Verstöße gegen Art. 5 Abs. 1 lit. b, c und f DSGVO fest.[186]

Grundsätze der Datenminimierung und Erforderlichkeit

Gemäß Art. 5 Abs. 1 lit c DSGVO dürfen nur solche Daten verarbeitet werden, die auf den Zweck bezogen, der Datenverarbeitung *angemessen*, *erheblich*, und für den Zweck *erforderlich* sind.[187] Der Grundsatz der Erforderlichkeit ist damit eine Ausprägung des Grundsatzes der Datenminimierung.[188] Die Angemessenheit des Verhältnisses der Datenverarbeitung mit dem Zweck ist gegeben, wenn ihre Zuordnung zu dem Zweck nicht beanstandet werden kann.[189]

Es bestehen keine Anhaltspunkte, dass die Verarbeitung der Daten in der ePA dem Zweck der Gesundheitsversorgung nicht *angemessen* wäre.

Die *Zweckerheblichkeit* ist dann gegeben, wenn die Verarbeitung der Daten geeignet ist, ein legitimes Ziel zu erreichen.[190] Die Datenverarbeitung in der ePA ist geeignet einen bestimmten Grad der Gesundheitsversorgung zu erreichen und die Gesundheitsversorgung ist ein legitimes Ziel.

Die Prüfung, ob bei der ersten Ausbaustufe der ePA durch das grobgranulare Dokumentenmanagement nicht erforderliche Datenverarbeitungen vorgenommen werden könnten, muss nicht durchgeführt werden, weil die Verarbeitung nicht

erforderlicher Daten durch die Einwilligung legitimiert wäre.[191] Dafür spricht insbesondere, dass der *Erforderlichkeitsgrundsatz* in beinahe allen Rechtsgrundlagen der DSGVO genannt wird, nur nicht im Wortlaut der Rechtsgrundlage der Einwilligung in Art. 6 Abs. 1 lit. a oder Art. 9 Abs. 2 lit. DSGVO.[192] Die bewusste und informierte Entscheidung für die Nutzung von Diensten, die potentiell nicht dem Erforderlichkeitsgrundsatz entsprechen, muss gerade aufgrund der informationellen Selbstbestimmung möglich sein.[193] Diese Garantie beinhaltet auch, dass der einzelne mit seinen Daten verfahren kann, wie er möchte. Eine weitere Erforderlichkeitsprüfung würde den Sinn einer freiwilligen und informierten Einwilligung völlig konterkarieren. Der Umstand auf einen möglichen Verzicht der Erforderlichkeit von Datenverarbeitungen sollte in der Einwilligung aber entsprechend zum Ausdruck kommen.[194] Die Einwilligung erfolgt im Rahmen der ePA unter der Information, dass das Berechtigungsmanagement bis 1. Januar 2022 nicht auf Dokumentenebene steuerbar ist.[195]

Grundsatz der Integrität und Vertraulichkeit

Nach Art. 5 Abs. 1 lit. f DSGVO müssen Daten so verarbeitet werden, dass sie eine angemessene Sicherheit der Daten gewährleisten. Dafür sollen insbesondere geeignete technische und organisatorische Maßnahmen zum Einsatz kommen um vor etwaigen Risiken zu schützen. Etwas näher konkretisiert wird der Grundsatz in Art. 32 Abs. 1 und 2 DSGVO.[196] Die geeigneten technischen und organisatorischen Maßnahmen müssen ein dem Risiko angemessenes Schutzniveau gewährleisten. Die Maßnahmen sind vom Verantwortlichen unter Berücksichtigung des Stands der Technik, der Implementierungskosten und der Art, des Umfangs der Umstände, und der Zwecke der Verarbeitung, sowie der unterschiedlichen Eintrittswahrscheinlichkeit und Schwere des Risikos für die Rechte und Freiheiten natürlicher Personen zu treffen.

Art. 32 Abs. 1 und 2, sowie Art. 25 Abs. 1 DSGVO schreiben dem Verantwortlichen jedoch keine konkreten Maßnahmen zur Datensicherheit vor. Der Verantwortliche muss selbst einen abgestimmten Katalog technischer und organisatorischer Maßnahmen ermitteln und umsetzen, der in seiner Gesamtschau dazu geeignet ist, die in Art. 5 lit. f DSGVO benannten Risiken zu minimieren.[197]

In der ersten Ausbaustufe der ePA ist technisch nicht ausgeschlossen, dass Leistungserbringer auf Daten zugreifen

---

[183]Vgl. Teager/Gabel/Taeger, DSGVO, Art. 7 Rn. 69.

[184]Vgl. GKV-Spitzenverband (2020a, S. 8).

[185]Vgl. a.a.O., S. 14.

[186]Vgl. BfDI (2020b, S. 3).

[187]Vgl. Kühling/Buchner/Herbst, DSGVI, Art. 5 Rn. 57; Sydow/Reimer, DSGVO Art. 5 Rn. 32.

[188]BeckOK DatenschutzR/Schantz, Art. 5 Rn. 25.; Böllhoff, Cornelius, et al. (2020, S. 33).

[189]Vgl. Paal/Pauly/Frenzel. DSGVO, Art. 5 Rn. 35.

[190]Vgl. BeckOK DatenschutzR/Wolff, Art.5 Rn. 24.

[191]Vgl. BeckOK DatenschutzR/Wolff, Syst. A Prinzipien des Datenschutzrechts, Rn. 57.

[192]Vgl. Böllhoff, Cornelius, et al. (2020, S. 34).

[193]Vgl. ebd.; Sydow/Ingold, DSGVO, Art. 7 Rn. 10 DSGVO: Einwilligung als Realisierung der Autonomie als primärrechtlich garantierte Grundrechtsausübung.

[194]Vgl. BeckOK DatenschutzR/Wolff, Syst. A Prinzipien des Datenschutzrechts, Rn. 57.

[195]Vgl. GKV-Spitzenverband (2020a, S. 13 ff).

[196]Vgl. Kühling/Buchner/Herbst, DSGVO, Art. 5 Rn. 76.

[197]Vgl. Kühling/Buchner/Herbst, DSGVO, Art. 32 Rn. 5; Zur Vielfalt möglicher technischer und organisatorischer Maßnahmen Paal/Pauly/Martini, DSGVO, Art. 32 Rn. 29.

können, die sie nicht zwingend benötigen, um die Gesundheitsversorgung des Versicherten sicherzustellen.

Eine *unrechtmäßige* Verarbeitung nach Art. 5 lit. f DSGVO liegt aber wegen einer wirksam erteilten Einwilligung nicht vor.[198] Dem Schutz vor einer *unbefugten* Verarbeitung dient § 352 SGB V, der Leistungserbringer den Zugriff auf Daten verbietet, die zur Behandlung nicht erforderlich sind. Der Rückgriff auf das Berufsgeheimnis[199] vermag indes nicht zu überzeugen, da eine unbefugte Verarbeitung auch durch Berufsgeheimnisträger vorgenommen werden kann. Eine technische Umsetzung wäre zwar zu befürworten, aber am Ende kann auch hier diskutiert werden, inwiefern das Risiko einer unbefugten Verarbeitung von der Einwilligung des Versicherten gedeckt wäre.[200]

Grundsatz der Zweckbindung

Es liegt kein Verstoß gegen den Zweckbindungsgrundsatz nach Art. 5 Abs. 1 lit. b DSGVO vor, weil die Versicherten ihre freiwillige und informierte Einwilligung im Hinblick auf einen bestimmten Zweck abgeben (s.o. 2.7.1.2).[201] Eine Verarbeitung der Daten zu anderen Zwecken ist nicht ersichtlich.

2.7.3. Art. 25 DSGVO

Art. 25 DSGVO enthält ausschließlich Konkretisierungen der in Art. 5 DSGVO genannten datenschutzrechtlichen Grundsätze.[202] Wie zuvor ausgeführt ist aufgrund der wirksamen Einwilligung nicht von einem Verstoß gegen die Datenschutzgrundsätze auszugehen, daher können auch die Konkretisierungen des Art. 25 DSGVO nicht zu einem anderen Ergebnis führen.[203]

2.7.4. „Datensouveränität"

Der BfDI kritisiert mehrfach, dass Versicherte durch das defizitäre Berechtigungsmanagements in ihrer „Datensouveränität" beschnitten würden.[204] Richtig ist, dass eine Datensouveränität in der DSGVO weder datenschutzrechtlicher Grundsatz noch Wirksamkeitsvoraussetzung einer Einwilligung ist.[205] Allerdings muss festgehalten werden, dass der BfDI diesen Begriff nie ausdrücklich im Zusammenhang mit den Anforderungen der DSGVO, sondern vielmehr mit einem möglichen Wertungswiderspruch innerhalb der Sozialgesetzgebung nennt. Dies ist aber nicht Prüfgegenstand dieser Arbeit.

2.7.5. Zwischenergebnis

Das Berechtigungsmanagement der ePA in der ersten Ausbaustufe bis zum 1. Januar 2022 genügt den Anforderungen der DSGVO. Die freiwillige und informierte Einwilligung dient als Rechtsgrundlage für die Verarbeitung der Daten und damit verbundener potentieller Einschränkungen der Datenverarbeitungsgrundsätze.

2.8. Berechtigungsmanagement ab 1. Januar 2022, § 342 Abs. 2 Nr. 2 SGB V

Anhand des oben festgestellten Ergebnisses kann für die zweite Stufe des Berechtigungsmanagement, die noch eine feinere Differenzierung ermöglicht, mit einem Erst-Recht-Schluss festgestellt werden, dass hier ebenso kein Verstoß gegen europäisches Datenschutzrecht vorliegen wird.

2.9. Authentifizierungsanforderungen, § 336 Abs. 2 Nr. 2 SGB V

§ 336 Abs. 2 Nr. 2 SGB V regelt den Zugriff des Versicherten auf die ePA ohne Verwendung der elektronischen Gesundheitskarte (eGK) durch ein geeignetes technisches Verfahren, das einen hohen Sicherheitsstandard gewährleistet. Der BfDI kritisiert das aktuelle von der gematik zugelassene Zugriffsverfahren „Alternative Versicherten-Identität" (al.vi).[206] Er fordert eine Sicherheitseinstufung „hoch" entsprechend der eIDAS-VO.[207]

Im Rahmen der DSGVO kommt ein Verstoß gegen Art. 32 Abs. 1 und 2 DSGVO in Betracht. Wie oben in Kapitel 2.7.2.2 ausgeführt, schreibt die DSGVO dem Verantwortlichen jedoch keine konkreten Maßnahmen hinsichtlich der Ausgestaltung der Datensicherheit, sondern vielmehr Prüfungen vor, anhand derer er die notwendigen Maßnahmen selbst ermittelt.[208] Im Laufe der Verarbeitung besteht gleichwohl die Pflicht zu einer regelmäßigen Überprüfung und gegebenenfalls erforderlichen Anpassung der technischen und organisatorischen Maßnahmen.[209]

Diesen Anforderungen steht § 336 Abs. 2 Nr. 2 SGB V nicht entgegen. Denn dieser gibt lediglich vor, dass ein gewisser Sicherheitsstandard („hoch") nicht unterschritten werden darf.

2.10. Ergebnis

Im Ergebnis kann festgestellt werden, dass die ePA weder hinsichtlich des Berechtigungsmanagements der unterschiedlichen Ausbaustufen noch hinsichtlich der Sicherheit des Authentifizierungsverfahrens der Datenschutz-Grundverordnung nach zu beanstanden ist. *Insofern ist die elektronischen Patientenakte in ihrer derzeitigen Ausgestaltung mit der DSGVO vereinbar.*

---

[198]Vgl. Böllhoff, Cornelius, et al. (2020, S. 35).
[199]Vgl. ebd.
[200]Vgl. so dann wieder Böllhoff, Cornelius, et al. (2020, S. 37).
[201]Vgl. so auch Böllhoff, Cornelius, et al. (2020, S. 36).
[202]Vgl. Ehmann/Selmayr/Heberlein, DSGVO, Art. 5 Rn. 23; Kühling/Buchner/Herbst, DSGVO, Art. 5 Rn. 59; Teager/Gabel/Voigt, DSGVO, Art. 5 Rn. 29.; Gola/Pötters, DSGVO, Art. 5 Rn. 23.
[203]Vgl. so auch Böllhoff, Cornelius, et al. (2020, S. 37).
[204]Vgl. BfDI (2020a, S. 13.); BfDI (2020b, S. 3.); BfDI (2020a, S. 2).
[205]Vgl. so auch Böllhoff, Cornelius, et al. (2020, S. 24 f).

[206]Vgl. BfDI (2020a, S. 7 f); BfDI (2020a, o. S).
[207]Verordnung (EU) 910/2914.
[208]Vgl. so auch Böllhoff, Cornelius, et al. (2020, S. 50).
[209]Vgl. ebd.; Paal/Pauly/Martini, DSGVO, Art. 32 Rn. 56; Taeger/Gabel/Schultze-Melling, DSGVO, Art. 32 Rn. 13.

Versicherung	App	Berwertung (Distribution) (Anzahl Bewertungen)
AOKen	Mein Leben	2.5 (Apple) (206) 2.0 (Google) (323)
BARMER	eCare	3.6 (Apple) (48) 2.6 (Google) (92)
DAK	ePA	2.1 (Apple) (39) 2.0 (Google) (45)
IKK Classic	ePA	2.0 (Apple) (28) 1.6 (Google) (37)
KKH	ePA	2.1 (Apple) (25) N/A** (Google)
Knappschaft	Meine Gesundheit	3.6 (Apple) (54) 1.4 (Google) (155)
Mobil Krankenkasse	ePA	2.3 (Apple) (19) 1.7 (Google) (15)
SBK	Patientenakte	2.6 (Apple) (27) 2.6 (Google) (5)
Techniker Krankenkasse	TK-Safe*	4.8 (Apple) (197.492) 4.7 (Google) (52.961)

Stand 23.05.2021

Apple = App Store   Google = Google Play

*App ging aus bestehender Gesundheitsakte hervor

**App zum Zeitpunkt der Erhebung nicht im Store verfügbar

Abbildung 11: Bewertungen ePA FdV (Krankenkassen >1.000.000 Versicherte)

Quelle: Eigene Darstellung

3. Fazit und Ausblick

Die vorliegende Arbeit zeigt, dass es verschiedene Anknüpfungspunkte für eine Prüfung der ePA mit europäischem Datenschutzrecht gibt. So wurde beispielsweise festgestellt, dass die Rolle des Verantwortlichen gemäß Art. 4 Nr. 7 DSGVO durch § 307 SGB V zwar ausführlich konkretisiert wird, es aber möglich erscheint, dass es doch zu Diskussionen kommen wird, beispielsweise in Bereichen der Schnittstelle zwischen TI und ePA-Aktensystemen.

Datensicherheit ist ein unerlässlich wichtiges Thema im Bereich der Verarbeitung von Gesundheitsdaten, daher ist der Gesetzgeber angehalten, die Standards entsprechend hoch zu halten. Hinsichtlich der ePA kann mit vorsichtigem Optimismus von einer sicheren Anwendung innerhalb der TI ausgegangen werden. Schwachstellen sind den Beteiligten bewusst und werden beseitigt.

Die Datenschutz-Folgenabschätzung hat im Angesicht ihrer wenig konkreten Ausgestaltung durch Art. 35 DSGVO das Potential für weitere Diskussionen zu sorgen. Jedoch betreffen die Pflichten aus Art. 35 DSGVO nicht die Rechtmäßigkeit eines Verarbeitungsvorgangs.

Hauptkritikpunkt des Bundesbeauftragten für Datenschutz und Informationsfreiheit ist das aus seiner Sicht defizitäre Berechtigungsmanagement. Der Schwerpunkt der Arbeit richtet sich an dieser Kritik aus. Es wird festgestellt, dass die Rechtsgrundlage der Datenverarbeitungen in der ePA im Regelfall die Einwilligung nach Art. 6 Abs. 1 Satz 1 lit. a bzw. Art. 9 Abs. 2 lit a DSGVO ist. Im Detail werden die Anforderungen des Art.4 Nr. 11 DSGVO geprüft. Der Freiwil-ligkeit entgegen steht weder eine Ungleichbehandlung von Personen, die die ePA nutzen und solchen, die es nicht tun, noch ein Ungleichgewicht im Arzt-Patienten-Verhältnis. Auch eine unzulässige Kopplung der Datenverarbeitung an andere Zwecke liegt nicht vor. Das fehlende feingranulare Berechtigungsmanagement in der ersten Ausbaustufe spricht nicht gegen die Anforderungen an die Bestimmtheit der Einwilligung. Dem Berechtigungsmanagement liegt insbesondere kein „Alles-oder-nichts-Prinzip" zugrunde. Eine konkrete Anforderung an die Differenziertheit der Einwilligung lässt sich aus der DSGVO nicht ableiten. Ein Minus in der Granularität der Einwilligung sollte nach Ansicht des Verfassers durch ein Plus an Information und Transparenz ausgeglichen werden. Den Informationspflichten der DSGVO wird in der Sozialgesetzgebung und in der konkreten Ausgestaltung der Anwendungen hinreichend Rechnung getragen. Jedoch könnte eine anschaulichere Informationsvermittlung zu einer besseren Informiertheit und zu einem stärkeren Vertrauen des Versicherten in die Anwendung beitragen. Datenschutzgrundsätze nach Art. 5 DSGVO sind hinsichtlich des Erforderlichkeitsprinzips durch das undifferenzierte Rechtemanagement in der ersten Ausbaustufe zwar eingeschränkt; diese Einschränkung ist aber von der freiwilligen und informierten Einwilligung abgedeckt.

Die DSGVO schreibt keine konkreten Maßnahmen hinsichtlich der Datensicherheit vor. Die Ausgestaltung der Datensicherheit obliegt dem nationalen Gesetzgeber bzw. direkt dem Verantwortlichen durch Festlegung der technischen und organisatorischen Maßnahmen zur Risikovermeidung. Daher trifft die Kritik des BfDI hinsichtlich der alternativen Authen-

tifizierungsverfahren nicht zu. Eine differenzierte Auseinandersetzung mit den Sicherheitsstandards der ePA wäre ein interessantes Thema für eine weitere Ausarbeitung.

Wegen der eiligen Terminierung des Gesetzgebers trägt die ePA zu Beginn viele Softwarefehler in sich. Ein Blick in die Bewertungs- und Kommentarsektionen der Vertriebsplattformen von Apple und Google lässt vermuten, dass die Versicherten hier zu Beta-Testern gemacht werden (Abbildung 11). Vor allem scheinen Funktionalitäten und Dienste nicht zuverlässig verfügbar zu sein. Der Anzahl der Bewertungen im Verhältnis zu den Versicherten nach zu urteilen, stößt die ePA bei den Versicherten noch nicht auf große Resonanz.[210] Das mag auch daran liegen, dass die Krankenkassen die ePA derzeit nicht aktiv bewerben.

Auf der Leistungserbringerseite ist es für eine Bewertung der Testphase noch zu früh. Die Produkte der Konnektoren-Hersteller befinden sich derzeit noch in der Zulassungsprüfung.[211]

Fraglich ist indes das Vorgehen des BfDI. Er hat seine Kritik zwar schon frühzeitig geäußert, jedoch nie substantiiert dargelegt. Bezeichnend insofern ist auch der Verweis auf allgemeine Grundsätze (Art. 5, Art. 25, Art. 32 DSGVO) und politische Schlagworte („Datensouveränität"). Selbst die Warnung an die Krankenkassen enthält keine differenzierten Ausführungen zu den Kritikpunkten. Auf Nachfrage des Verfassers verwies die Behörde auf die öffentlich zugängliche Stellungnahme zum Entwurf des PDSG. Die Vorgehensweise des BfDI stieß angeblich auch in der eigenen Partei und in anderen Aufsichtsbehörden auf Kritik.[212] Ob der BfDI wie angekündigt[213] weitere Maßnahmen ergreifen wird, bleibt mit Spannung abzuwarten. Eine gerichtliche Klärung scheint dann nicht unwahrscheinlich.[214]

---

[210]Vgl. dazu auch golem.de (2021, o. S.); Klöckner, Jürgen, Olk, Julian (2021, o. S.)

[211]Vgl. Gesellschaft für Telematik (gematik) (2021d, o. S.)

[212]Vgl. Olk (2020a, o. S.)

[213]BfDI (2020a, S. 2); BfDI (2020c).

[214]Vgl. zur Befugnis des BAS als Rechtsaufsichtsbehörde der Krankenkassen: BAS, Stellungnahme zur elektronischen Patientenakte, S. 4.

Literatur

aerzteblatt.de Elektronische Patientenakte. (2020, Dezember). *Elektronische Patientenakte: Kassen warten mit Werbung ab.* Zugriff auf https://www.aerzteblatt.de/archiv/217189/Elektronische-Patientenakte-Kassen-warten-mit-Werbung-ab ([Zugriff 2021-05-15])

aerztezeitung. (2018). *TK-Versicherte erhalten E-Akte.* Zugriff am 2021-05-15 auf https://www.aerztezeitung.de/Politik/TK-Versicherte-erhalten-E-Akte-229937.html

aerztezeitung. (2020). *PKV bereitet sich auf ePA-Einstieg vor.* Zugriff am 2021-05-15 auf https://www.aerztezeitung.de/Wirtschaft/PKV-bereitet-sich-auf-ePA-Einstieg-vor-414315.html

aerztezeitung.de. (2018). *Neue Gesundheitsakte bringt GKV und PKV zusammen.* Zugriff auf https://www.aerztezeitung.de/Wirtschaft/Neue-Gesundheitsakte-bringt-GKV-und-PKV-zusammen-231694.html ([Zugriff 2021-05-15])

Artikel-29-Datenschutzgruppe . (2007). *Arbeitspapier zu elektronischen Patientenakten, 2007): Working Document on the processing of personal data relating to health in electronic health records (EHR), WP 131, Februar 2007.* Zugriff auf https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp131_en.pdf

BfDI. (2019). *Liste der Verarbeitungsvorgänge gem. Art. 35 Abs. 4 DSGVO, 2019): Liste von Verarbeitungsvorgängen gemäß Artikel 35 Absatz 4 DSGVO für Verarbeitungstätigkeiten öffentlicher Stellen des Bundes, Oktober 2019.* Zugriff auf https://www.bfdi.bund.de/SharedDocs/Downloads/DE/Datenschutz/Liste_VerarbeitungsvorgaengeArt35.pdf?

BfDI. (2020a). *BfDI zu Folgen der Gesetzgebung des PDSG, Pressemitteilung vom 19. August 2020.* Zugriff am 2021-05-21 auf https://www.bfdi.bund.de/DE/Infothek/Pressemitteilungen/2020/20_BfDI-zu-PDSG.html

BfDI. (2020b). *Bundespressekonferenz vom 19.08.2020.* Zugriff am 2021-05-23 auf https://www.youtube.com/watch?v=vsLM8EPH9NY

BfDI. (2020c). *Bundespressekonferenz vom 19.08.2020.* Zugriff auf https://www.youtube.com/watch?v=vsLM8EPH9NY

BfDI. (2020a). *Schreiben an die LfDI, 2020): Weiteres Vorgehen nach dem Beschluss des Deutschen Bundestages zum PDSG am 3. Juli 2020, August 2020.* Zugriff auf https://www.bfdi.bund.de/DE/Infothek/Transparenz/AccessforoneAccessforall/2021/2020_Anschreiben-LfDs-PDSG.pdf?

BfDI. (2020b). *Stellungnahme zum PDSG, 2020): Stellungnahme [...] zum Entwurf eines Gesetzes zum Schutz elektronischer Patientendaten in der Telematikinfrastruktur, April 2020.* Zugriff auf https://www.bfdi.bund.de/DE/Infothek/Transparenz/Stellungnahmen/2020/StgN_Patientendaten-Schutz-Gesetz.pdf?

BfDI. (2020). *Stellungnahme [...] zur öffentlichen Anhörung [...] zum Entwurf eines Geset-zes zum Schutz elektronischer Patientendaten in der Telematikinfrastruktur (Patienten-Datenschutz-Gesetz), Mai 2020.* Zugriff am 2021-05-15 auf https://www.bfdi.bund.de/DE/Infothek/Transparenz/Stellungnahmen/2020/StgN_Patienten-Datenschutz-Gesetz.pdf?

BfDI. (2020a). *Stellungnahme zur öffentlichen Anhörung zum PDSG, 2020): Stellungnahme [. . .] zur öffentlichen Anhörung [. . .] zum Entwurf eines Gesetzes zum Schutz elektronischer Patientendaten in der Telematikinfrastruktur (Patienten-Datenschutz-Gesetz), Mai 2020.* Zugriff auf https://www.bfdi.bund.de/DE/Infothek/Transparenz/Stellungnahmen/2020/StgN_Patienten-Datenschutz-Gesetz.pdf?

BfDI. (2020b). *Warnung an die gesetzlichen Krankenkassen, 2020): Warnung nach Artikel 58 Abs. 2 Buchst. a) DSGVO, November 2020.* Zugriff auf https://www.bfdi.bund.de/DE/Infothek/Transparenz/Rundschreiben/Allgemein/2020/Warnung-Krankenkassen-ePA.pdf?

BfDI. (2020). *Warnung nach Artikel 58 Abs. 2 Buchst. a) DSGVO, November 2020.* Zugriff am 2021-05-15 auf https://www.bfdi.bund.de/DE/Infothek/Transparenz/Rundschreiben/Allgemein/2020/Warnung-Krankenkassen-ePA.pdf?

BfDI. (2021). *Zuständigkeit für gesetzliche Krankenkassen, 2021.* Zugriff auf https://www.bfdi.bund.de/DE/Datenschutz/Themen/Gesundheit_Soziales/KrankenkassenArtikel/Krankenkassen-Zust%C3%A4ndigkeit-BfDI.html?nn=5217154

Böllhoff, Cornelius, et al. (2020). *Vereinbarkeit der ePA mit europäischem Datenschutzrecht, 2020): Vereinbarkeit der Regelungen zur elektronischen Patientenakte (ePA) nach dem Patienten-Datenschutz-Gesetz (PDSG) mit europäischem Datenschutzrecht – Rechtsgutachten im Auftrag des health innovation hub, November 2020.* Zugriff auf https://hih-2025.de/wp-content/uploads/2021/02/Redeker_Rechtsgutachten_ePA.pdf

Bundesamt für Sicherheit in der Informationstechnik (BSI). (2018). *Technische Richtlinie TR-03147, 2018): Technische Richtlinie TR-03147 Vertrauensniveaubewertung von Verfahren zur Identitätsprüfung natürlicher Personen, Dezember 2018.* Zugriff auf https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/TechnischeRichtlinien/TR03147/TR03147.pdf

Bundesärztekammer. (2020a). *Stellungnahme der Bundesärztekammer zum Referentenentwurf eines Gesetzes zum Schutz elektronischer Patientenakten in der Telematikinfrastruktur.* Zugriff am 2021-05-15 auf https://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/Stellungnahmen/SN_BAEK__PDSG_25022020.pdf

Bundesärztekammer. (2020b). *Stellungnahme zum PDSG, 2020): Stellungnahme der Bundesärztekammer zum Referentenentwurf eines Gesetzes zum Schutz elektronischer Patientenakten in der Telematikinfrastruktur, Februar 2020.* Zugriff auf https://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/Stellungnahmen/SN_BAEK__PDSG_25022020.pdf

Bundesministerium für Gesundheit (BMG). (2019). *TSVG, 2019): Terminservice- und Versorgungsgesetz (TSVG).* Zugriff auf https://www.bundesgesundheitsministerium.de/terminservice-und-versorgungsgesetz.html

Bundesministerium für Gesundheit (BMG). (2021, März). *Online-Ratgeber Krankenversicherung, 2021): Versicherte in der gesetzlichen Krankenversicherung.* Zugriff auf https://www.bundesgesundheitsministerium.de/gesetzlich-versicherte.html

DAK. (2021). *Datenschutzhinweise für die elektronische Patientenakte (ePA), 2021.* Zugriff auf https://www.dak.de/dak/unternehmen/datenschutzhinweise-fuer-die-elektronische-patientenakte-epa-2377742.html#/

Gematik. (2021a). *ePA-Aktensyteme.* Zugriff am 2021-05-15 auf https://fachportal.gematik.de/hersteller-anbieter/komponenten-dienste/epa-frontend-des-versicherten

Gematik. (2021b). *E-Patientenakte.* Zugriff am 2021-05-15 auf https://www.gematik.de/anwendungen/e-patientenakte/

Gematik. (2021c). *Zulassungsübersicht.* Zugriff am 2021-05-15 auf https://fachportal.gematik.de/zulassungs-bestaetigungsuebersichten

Gesellschaft für Telematik (gematik). (2019). *Systemspezifisches Konzept ePA, Oktober 2019.* Zugriff auf https://www.vesta-gematik.de/standard/formhandler/324/gemSysL_ePA_V1_3_0.pdf

Gesellschaft für Telematik (gematik). (2020a). *Gesellschafteranteile, 2020.* Zugriff auf https://www.gematik.de/news/news/welcome-pkv-wird-gesellschafter-der-gematik/

Gesellschaft für Telematik (gematik). (2020b). *Informationsblatt Sicherheits- und Produktgutachten, Mai 2020.* Zugriff auf https://fachportal.gematik.de/fileadmin/Fachportal/Downloadcenter/Informationen_fuer_Gutachter/gemInfo_Gutachter_ePA-FdV.pdf

Gesellschaft für Telematik (gematik). (2020c). *Richtlinie zur Prüfung der Sicherheitseignung, April 2020.* Zugriff auf https://fachportal.gematik.de/fileadmin/user_upload/fachportal/files/Service/Sicherheitsgutachten/gemRL_PruefSichEig_DS_V2.1.0.pdf

Gesellschaft für Telematik (gematik). (2020d). *Vorwort zum Gutachten der TU Graz, 2020): Vorwort zum Gutachten der TU Graz zur Sicherheitsanalyse der Kernkomponenten der elektronischen Patientenakte (ePA).* Zugriff auf https://www.gematik.de/fileadmin/user_upload/MediaUploads/Sicherheitsanalyse_TU_Graz

_zur_ePA_mit_Vorwort_der_gematik.pdf

Gesellschaft für Telematik (gematik). (2021a). *ePA-Aktensyteme, 2021.* Zugriff auf https://fachportal.gematik.de/hersteller-anbieter/komponenten-dienste/epa-aktensystem

Gesellschaft für Telematik (gematik). (2021b). *ePA-Frontend des Versicherten, 2021.* Zugriff auf https://fachportal.gematik.de/hersteller-anbieter/komponenten-dienste/epa-frontend-des-versicherten

Gesellschaft für Telematik (gematik). (2021c). *E-Patientenakte, 2021.* Zugriff auf https://www.gematik.de/anwendungen/e-patientenakte/

Gesellschaft für Telematik (gematik). (2021d). *Erste Bilanz der Testphase, 2021.* Zugriff auf https://www.gematik.de/news/news/elektronische-patientenakte-in-der-testphase-konnektorhersteller-ziehen-erste-bilanz

Gesellschaft für Telematik (gematik). (2021e). *Gesetzliche Grundlagen, 2021.* Zugriff auf https://www.gematik.de/ueber-uns/gesetzliche-grundlagen/

Gesellschaft für Telematik (gematik). (2021f). *KTR-Consumer, 2021.* Zugriff auf https://fachportal.gematik.de/hersteller-anbieter/komponenten-dienste/ktr-consumer

Gesellschaft für Telematik (gematik). (2021g). *Spezifikation ePA-Aktensystem, Februar 2021.* Zugriff auf https://fachportal.gematik.de/fachportal-import/files/gemSpec_Aktensystem_V1.7.0_Aend.pdf

Gesellschaft für Telematik (gematik). (2021h). *Spezifikation ePA-Frontend des Versicherten, Februar 2021.* Zugriff auf https://fachportal.gematik.de/fachportal-import/files/gemSpec_ePA_FdV_V1.8.0_Aend.pdf

Gesellschaft für Telematik (gematik). (2021i). *Telematikinfrastruktur (Glossar), 2021.* Zugriff auf https://www.gematik.de/glossar/begriffe/telematikinfrastruktur/397/

GKV-Spitzenverband. (2020a). *Informationen zur elektronischen Patientenakte): Informationen zur elektronischen Patientenakte (ePA) nach § 343 SGB V, November 2020.* Zugriff auf https://www.aok.de/pk/fileadmin/user_upload/Universell/05-Content-PDF/RS_2020-852_Anlage_01_-_mit_AOK_Anpassung_ITSG_GKV-SV.pdf

GKV-Spitzenverband. (2020b). *Stellungnahme zum PDSG, 2020): Stellungnahme des GKV-Spitzenverbandes vom 19.05.2020 zum Gesetzesentwurf [. . .] – PDSG vom 31.03.2020, Mai 2020.* Zugriff auf https://www.gkv-spitzenverband.de/media/dokumente/presse/p_stellungnahmen/200519_Stellungnahme_GKV-SV_PDSG_Gesetzentwurf.pdf

golem.de. (2021). *Die Patientenakte auf dem Smartphone bleibt ungeöffnet.* Zugriff auf https://www.golem.de/news/telemedizin-die-patientenakte-auf-dem-smartphone-bleibt-ungeoeffnet-2103-154751.html

Golland, A. (2018). Das Kopplungsverbot in der Datenschutz-Grundverordnung – Anwendungsbe-reich, ökonomische Auswirkungen auf Web 2.0-Dienste und Lösungsvorschlag. *Multimedia und Recht*, 130–135.

Hof, H.-J. (2020). Datenschutz mittels IT-Sicherheit. In M.-T. Tinnefeld, B. Buchner & et al. (Hrsg.), *Einführung in das Datenschutzrecht – Datenschutz und Informationssicherheit in europäischer Sicht, 7. Auflage.* de Gruyter.

Jäschke, T. & Hacks, S. (2016). Einführung Datenschutz im Gesundheitswesen. In T. Jäschke (Hrsg.), *Datenschutz im Gesundheitswesen: Grundlagen, Konzepte und Umsetzung* (S. 2–17). Medizinisch Wissenschaftliche Verlagsgesellschaft (MWV).

Jorzig, A. & Sarangi, F. (2020). Digitalisierung im Gesundheitswesen – Ein kompakter Streifzug durch Recht. *Technik und Ethik*.

Kassenärztliche Bundesvereinigung (KBV). (2020). *Statistische Informationen aus dem Bundesarztregister.* Zugriff am 2021-05-15 auf https://www.kbv.de/media/sp/2020-12-31_BAR_Statistik.pdf

KBV. (2020). *Richtlinie zur Zertifizierung nach § 75b Absatz 5 SGB V, Mai 2020.* Zugriff auf https://www.kbv.de/media/sp/RiLi___75b_Abs._5_SGB_V_Zertifizierung.pdf

Klöckner, J. & Olk, J. (2021). Geteiltes Bild zur elektronischen Patientenakte bei den Versicherten. *Handelsblatt*.

Klöckner, Jürgen, Olk, Julian. (2021). *Geteiltes Bild zur ePA, 2021): Geteiltes Bild zur elektronischen Patientenakte bei den Versicherten, Januar 2021.* Zugriff auf https://www.handelsblatt.com/inside/digital_health/umfrage-geteiltes-bild-zur-elektronischen-patientenakte-bei-den-versicherten/26774244.html

Koch, M. & Henke, M. (2016). Datenschutz in der Telematikinfrastruktur. In T. Jäschke (Hrsg.), *Datenschutz im Gesundheitswesen: Grundlagen, Konzepte und Umsetzung* (S. 309–322). Medizinisch Wissenschaftliche Verlagsgesellschaft (MWV).

netzpolitik.org. (2020). *Jens Spahn hat es eilig, Oktober 2020.* Zugriff auf https://netzpolitik.org/2020/jens-spahn-hat-es-eilig

Olk, J. (2020a). *Breite Front stellt sich hinter Jens Spahn.* Zugriff auf https://www.handelsblatt.com/inside/digital_health/streit-um-digitale-patientenakte-breite-front-stellt-sich-hinter-jens-spahn/26732890.html?ticket=ST-3044829-txdeZGvwodaQPYTHWmle-ap5

Olk, J. (2020b). *Private Krankenversicherung will wieder an gesetzlicher Digitalisierung teilhaben.* Zugriff auf https://www.handelsblatt.com/politik/deutschland/rueckkehr-zur-gematik-private-krankenversicherung-will-wieder-an-gesetzlicher-digitalisierung-teilhaben/25592704.html?ticket=ST-6959017-lRb22uKapYhKAnwchnMk-ap4

Otto, D. & Rüdlin, M. (2017). Standardisierung von Patienteneinwilligungen im Krankenhaus. *Zeitschrift für Datenschutz*, 519–524.

PKV Verband der Privaten Krankenversicherung. (2020). *PKV und eHealth.* Zugriff auf https://www.pkv.de/verband/presse/pressemitteilungen/gemeinsame-pressemitteilung-einheitliche-digitale-infrastruktur-fuer-das-gesundheitswesen/

Slany, W. (2020). *ePA Sicherheitsanalyse der TU Graz, 2020): Sicherheitsanalyse zur Sicherheit der kritischen Komponenten der elektronischen Patientenakte nach §291a SGB V – Fokus auf VAU und kryptographische Sicherheitsleistung SGD, 03.03.2020.* Zugriff auf https://www.gematik.de/fileadmin/user_upload/MediaUploads/Sicherheitsanalyse_TU_Graz_zur_ePA_mit_Vorwort_der_gematik.pdf

Stöferle, B. (2016). Kooperations- und Kommunikationspartner aus Anwendersicht. In T. Jäschke (Hrsg.), *Datenschutz im Gesundheitswesen: Grundlagen, Konzepte und Umsetzung* (S. 135–161). Medizinisch Wissenschaftliche Verlagsgesellschaft (MWV).

sueddeutsche.de. (2020). *Interview mit Christoph Saatjohann zur ePA, 2020): „Ich bin vorsichtig optimistisch".* Zugriff auf https://www.sueddeutsche.de/digital/elektronische-patientenakte-epa-gesundheit-it-sicherheit-1.5159783

tagesschau.de. (2020). *IT-Sicherheitslücken in Praxen.* Zugriff auf https://www.tagesschau.de/investigativ/br-recherche/sicherheit-telematik-101.html

v. Brocke, Jan, Simons, Alexander, et. al. (2009). *Reconstructing the Giant, 2009): Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process, 2009.* Zugriff auf https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1145&context=ecis2009

Verbraucherzentrale, v. (2020). *Digitalisierung im Gesundheitswesen muss allen offenstehen.* Zugriff auf https://www.vzbv.de/meldungen/digitalisierung-im-gesundheitswesen-muss-allen-offenstehen

The Effect of Micro-Entrepreneurship on Migration Plans of Young Adults in Rural Sub-Saharan Africa and the Mediating Role of Subjective and Economic Well-Being

Christian Stöhr

Technische Universität München

Abstract

Entrepreneurship can make a valuable contribution to global poverty reduction. However, poverty is often narrowly defined in terms of income and gross domestic product. Thus, a deep understanding of the impact of entrepreneurship on the development of the Global South cannot be obtained. To address this issue and uncover the transformative potential of entrepreneurship, this paper proposes a model in which both economic and subjective well-being mediate the relationship between micro-entrepreneurship and migration plans of young adults in rural sub-Saharan Africa. The results provide tentative evidence of a negative impact of micro-entrepreneurship on migration plans. Further, it is shown that subjective well-being explains part of the negative association between micro-entrepreneurship and migration plans. While those results hold for entrepreneurs in low-income countries and agricultural entrepreneurs, no evidence is found for entrepreneurs in middle/high-income countries and non-farm entrepreneurs. Contrary to expectations, the mediating role of economic well-being is not significant for either the entire group of micro-entrepreneurs or its subgroups. Finally, economic and subjective well-being as measures of poverty alleviation are found to be positively correlated.

Keywords: Entrepreneurship; Migration; Subjective well-being; Sub-saharan africa.

1. Introduction

Despite strong economic growth in many African countries in recent decades (Beegle, Christiaensen, Dabalen, & Gaddis, 2016), a large share of the young population sees emigration from their home country as the best option to break the cycle of unemployment and poverty (Milasi, 2020). Instead of migrating, however, people are increasingly trying to escape their lack of prospects through entrepreneurship. This is reflected in a large number of existing micro-enterprises in sub-Saharan Africa (SSA), as well as the high willingness to start an own business there (Kew, 2015). Although some entrepreneurs fail in the early stages of their business due to insufficient support, misplaced incentives, and lack of skills (GEM, 2013), entrepreneurship can offer a way out of unemployment (Kew, 2015). By providing both access to a regular income and increasing the perceived quality of life, entrepreneurship can drive poverty alleviation (Tobias, Mair, & Barbosa-Leiker, 2013) and ultimately even reduce young adults' migration plans (Cai, Esipova, Oppenheimer, & Feng, 2014; Chindarkar, 2014).

To better understand such transformative potential of entrepreneurship (Tobias et al., 2013), an integrative model is developed that takes a holistic view of microenterprise outcomes in rural SSA. Using linear and logistic regression, it is first estimated how micro-entrepreneurs differ from unemployed in terms of subjective well-being (SWB), economic well-being (EWB), and migration plans. Subsequent application of the quasi-experimental method Propensity Score Matching (PSM) controls for the influence of confounding variables and reduces selection bias. Structural Equation Modeling (SEM) is finally conducted to simultaneously determine the direct and indirect effects of entrepreneurship in explaining migration plans.

The thesis adds to the existing literature on outcomes of entrepreneurship by examining its impact in poverty-stricken regions of the Global South. The extensive analysis of Afrobarometer data counteracts the underrepresentation of quantitative empirical work in the research field. By including subjective well-being, the exclusively economic focus on assessing poverty based on income or gross domestic product (GDP) is countered. Moreover, the comprehensive investigation of the relationship between micro-entrepreneurship and migration plans in rural SSA contributes to an area of research that is still sparsely studied. Finally, by looking at

EWB, SWB, and migration plans in an interconnected way, this paper shows how poverty has been overcome from the perspective of the ordinary entrepreneur, thereby illustrating the transformative potential of entrepreneurship.

Essentially, the purpose of this thesis is to study the effect of micro-entrepreneurship on migration plans of rural sub-Saharan youth and to analyze the role of EWB and SWB in this relationship. To this end, the central research questions are:

1. *How does micro-entrepreneurship affect the migration plans of rural, young sub-Saharan Africa?*
2. *What function do subjective and economic well-being have in this relation?*

The thesis proceeds as follows. Section 2 reviews previous literature on the impact of (micro-)entrepreneurship in the context of socio-economic constraints. Based on theoretical and empirical findings, Section 3 establishes the theoretical model and formulates four hypotheses. Section 4 presents the data, variables, and research approach on which the empirical analysis in Section 5 builds. Section 6 discusses the results of the thesis in terms of relevance to theory and practice, as well as their limitations. The final section concludes from this research.

2. Literature Review

The thesis deals with the broad field of (micro-)entrepreneurship and its outcomes in the context of socio-economic constraints, specifically the rural region of SSA. The section below defines important terms and summarizes what is already known about the research field. Moreover, it is explained to what extent the study can contribute to the previous research.

2.1. Terminology of Rural Micro-Entrepreneurship

The Global Entrepreneurship Monitor (GEM) defines entrepreneurship as "any attempt at new business or new venture creation, such as self-employment, a new business organization, or the expansion of an existing business, by an individual, a team of individuals, or an established business" (Reynolds, Hay, & Camp, 1999, p.3). A similar classification is made by the Organization for Economic Co-operation and Development (OECD). Accordingly, entrepreneurs are "people who sense opportunities, innovate, take risks, and develop new goods and services. They drive business dynamics [...] and fuel overall economic growth" (OECD, 2000, p.3). Micro-entrepreneurs, as distinguished from entrepreneurs, are those who engage in entrepreneurial activities using the services of no more than five people (Jayachandran, 2020; Quagrainie, Adams, Kabalan, & Dankwa, 2021).

In rural SSA, these micro-enterprises are generally mostly owner-operated with no other employees (Gough & Langevang, 2016). If micro-entrepreneurship takes place in rural areas, it is also often referred to as rural entrepreneurship (Koyana

& Mason, 2017). The empirical unit of observation in the thesis is precisely this microentrepreneur or rural entrepreneur. Those people engage either in income-generating activities within agriculture, i.e. farming as a business instead of subsistence farming (Gough & Langevang, 2016), or outside agriculture as non-farm entrepreneurs (Alemu & Adesina, 2017; Nagler & Naudé, 2017). Non-agricultural activities comprise mainly handicraft production, trade in goods, and provision of services (Gough & Langevang, 2016). It should be noted that most micro-entrepreneurs in developing countries are part of the informal sector. This usually means that they are not formally registered with the government (Gough & Langevang, 2016; Jayachandran, 2020; Nagler & Naudé, 2014).

Entrepreneurship literature often distinguishes between the terms opportunity entrepreneurship (pull) and necessity entrepreneurship (push), among them for instance Reynolds, Camp, Bygrave, Autio, and Hay (2001), Z. Acs (2006) and Dencker, Bacq, Gruber, and Haas (2021). While in the first case one actively decides to start a venture to take advantage of a potential business opportunity, in the second one acts out of lack of alternatives (Z. Acs, 2006). Notwithstanding, such a distinction plays a minor role in answering the research questions. Indeed, strict divisions into necessity and opportunity or push and pull bear the risk of obscuring the dynamics of entrepreneurship (Langevang, Namatovu, & Dawa, 2012; Welter, Baker, Audretsch, & Gartner, 2017) and limiting its theoretical development (Dencker et al., 2021). Furthermore, previous literature confirms that both necessity and opportunity entrepreneurs operate in rural SSA (Gough & Langevang, 2016; Nagler & Naudé, 2017). For these reasons, this paper accepts heterogeneity and diversity in entrepreneurship instead of defining it strictly according to existing concepts (Dencker et al., 2021; Welter et al., 2017).

2.2. Outcomes of Entrepreneurship in the Global South

To date, there is little research on the impact of entrepreneurship in developing countries (Dencker et al., 2021; Kolk, Rivera-Santos, & Rufín, 2014; Vivarelli, 2013). Especially SSA has been poorly researched (George, Corbishley, Khayesi, Haas, & Tihanyi, 2016; Igwe & Icha-Ituma, 2020; Sheriff & Muffatto, 2015). Moreover, the focus of the entrepreneurship literature has so far been on influencing factors rather than outcomes (Dencker et al., 2021; Shepherd, Parida, & Wincent, 2021), and on institutions or countries rather than ordinary entrepreneurial participants, which it has largely ignored (Tobias et al., 2013).

In the last decade, however, the outcomes of (micro-) entrepreneurship in the context of poverty or resource constraints have attracted growing attention among researchers. The relevant literature on this, including studies in SSA, is presented in Table 1 to provide a quick overview. The impact of entrepreneurship has been studied mainly from a European and North American perspective, both representing the Global North (Bruton, Ahlstrom, & Obloj, 2008). Surprisingly, so little research exists on the poor in the developing world. First, low-income countries tend to have higher

levels of entrepreneurial activity than advanced economies (Z. J. Acs, Desai, & Hessels, 2008). Second, especially in developing countries, economic performance and employment situation depend on entrepreneurship, which is mostly informal (Bruton, Ireland, & Ketchen Jr, 2012).

Third, a deep understanding of the impact of entrepreneurship in developing countries or poverty-stricken regions can make an important contribution to theory and practice (Bruton et al., 2008; Naudé, 2010). Because context plays a critical role in studying the consequences of entrepreneurship (Wiklund, Nikolaev, Shir, Foo, & Bradley, 2019), researchers must also gain their insights within developing countries. Ultimately, this could impact the overall understanding of the theory. In fact, the boundary conditions in developing countries seem to influence the entrepreneurship theory of developed countries (Reid, Roumpi, & O'Leary-Kelly, 2015; West, Bamford, & Marsden, 2008).

2.3. Economic Outcomes

More than 650 million people are living in extreme poverty, meaning each of them has to get by on less than 1.90 USD per day (Roser & Ortiz-Ospina, 2019). Eradicating this poverty is therefore one of the greatest challenges facing humanity (Chliova et al., 2015). The UN has written out this challenge as goal number 1 in its Sustainable Development Goals (SDGs): "End Poverty in All Its Forms Everywhere" (UN, 2020, p.6). Research increasingly sees entrepreneurship as a means for sustainable and economic development (Adusei, 2016; Bruton et al., 2013; Johnson & Schaltegger, 2020; London & Hart, 2010; Shepherd & Patzelt, 2017). This is partly because at the core of entrepreneurship is opportunity and wealth creation (Drucker, 1998; Hitt, Ireland, Camp, & Sexton, 2001; Schumpeter, 1934; Venkataraman, 2004). Under these aspects, numerous scientists are further interested in poverty alleviation through entrepreneurship (Alvarez & Barney, 2014; Bradley et al., 2012; Chliova et al., 2015; Kimmitt et al., 2020; Shepherd et al., 2021; Sutter et al., 2019; Vermeire & Bruton, 2016).

The theoretical and practical insights on the economic impacts of entrepreneurship under socio-economic constraints are still incomplete and controversial (Sutter et al., 2019). This is particularly true for SSA (Vermeire & Bruton, 2016) and its rural population (Tobias et al., 2013). Tobias et al. (2013) address this research gap with a quantitative examination of poverty alleviation from the perspective of coffee farmers in rural Rwanda. They report a significant increase in personal wealth as a result of entrepreneurial activity. Some qualitative studies indicate that at least some of the rural population sees entrepreneurship as a way to improve their livelihoods (Adewumi, 2020; Kolawole & Ajila, 2015; Okeke & Nwankwo, 2017). However, other studies are doubtful about the contribution of entrepreneurship to income growth and rural development (Kew, 2015; Nagler & Naudé, 2014, 2017). This is because most youth enterprises are undifferentiated, low-growth businesses that essentially offer employment only to the owner (Kew, 2015) and primarily serve for risk diversification (Nagler & Naudé, 2014). Beyond that, it

was found that rural non-farm enterprises mostly leave the market due to lack of profitability, finance, or idiosyncratic shocks (Nagler & Naudé, 2017).

2.4. Non-Economic Outcomes
2.4.1. Outcomes on Subjective Well-Being

The widely used concept of SWB depicts the presence of positive affect, the absence of negative affect, and the experience of happiness (Diener, 1984). The SWB builds on two different well-being accounts. First, there is cognitive-evaluative well-being. It describes general life satisfaction, i.e. what people think about their living (Diener, Emmons, Larsen, & Griffin, 1985). Second, there is hedonic-experienced well-being, which expresses how people experience their lives. Hedonic well-being is measured with positive affect, which involves feelings of happiness or joy, and negative affect, which reflects stress, depression, anger, or sadness (Diener, 1984). With the eudaimonic approach, another stream of research has emerged. It suggests that psychological functioning is a critical component of a fulfilling life (Nikolaev, Boudreaux, & Wood, 2020). Specifically, it outlines a dynamic process in which the fulfillment of psychological needs and desires should enable positive feelings and life satisfaction (Ryan & Deci, 2017; Seligman, 2012). The hedonic approach and the eudaimonic approach of SWB are not covered in the thesis since the available data do not allow them to be accurately represented. Instead, the emphasis lies on the evaluative approach because it can be well replicated with available survey data (Abdallah, Thompson, & Marks, 2008) and there is already sufficient academic evidence of its link to entrepreneurship (Nikolaev et al., 2020) and migration (Chindarkar, 2014). Usually, evaluative well-being is captured with life satisfaction scales by retrospectively assessing life or individual life domains. According to the evaluative concept, people can best judge their own lives and well-being by assigning relative weights to various hedonic experiences over time (Kahneman, Diener, & Schwarz, 1999; Kahneman & Krueger, 2006).

The exploration of well-being in entrepreneurship literature is still in its nascent stage (Wiklund et al., 2019). In the last decade, however, well-being as a significant cause or effect of entrepreneurship has received increasing attention from management, entrepreneurship, and economics scholars (Andersson, 2008; Binder & Coad, 2013; Nikolaev et al., 2020; Shepherd & Patzelt, 2017; Shir, 2015; Shir, Nikolaev, & Wincent, 2019; Stephan, 2018; Uy, Foo, & Song, 2013; Wiklund et al., 2019; Wiklund, Yu, Tucker, & Marino, 2017). There is already some knowledge on this in the context of socio-economic constraints. For example, Nikolova (2019) studies the relationship between entrepreneurship and the mental health of necessity entrepreneurs. According to her article, not only opportunity entrepreneurs but also necessity entrepreneurs benefit health-wise from their activity. Abreu et al. (2019) find that self-employed people from deprived neighborhoods have higher life satisfaction than self-employed people from wealthier neighborhoods.

Table 1: Research on Outcomes of (Micro-)Entrepreneurship under Socio-economic Constraints

Authors (Year)	Journal/Issuer	Arcticle Summary
Bradley, McMullen, Artz, and Simiyu (2012)	J. Manag. Stud.	combine prior research on poverty alleviation, which focuses on the allocative view of opportunity with the creation and discovery views of opportunity
Binder and Coad (2013)	Small Bus. Econ.	apply a matching approach to self-employment and life satisfaction
Bruton, Ketchen Jr, and Ireland (2013)	JBV	review extant knowledge about entrepreneurship as a solution to poverty and discuss future research
Tobias et al. (2013)	JBV	examine poverty alleviation and conflict resolution in rural Ruanda's entrepreneurial coffee sector
Alvarez and Barney (2014)	ETP	explore the extent to which people in poverty are capable of forming and exploiting different entrepreneurial opportunities
Nagler and Naudé (2014)	IZA	provide a comparative empirical analysis of non-farm entrepreneurship in rural Africa in terms of patterns and determinants
Chliova, Brinckmann, and Rosenbusch (2015)	JBV	conduct a meta-analysis towards the impact of microcredit on outcomes of client entrepreneurs
Kolawole and Ajila (2015)	WJEMSD	present a practical approach for enhancing rural entrepreneurship development as a driver of transformation and development
Vermeire and Bruton (2016)	Afr.J.Bus.Manage.	provide a review and agenda for entrepreneurial opportunities and poverty in SSA
Nagler and Naudé (2017)	Food Policy	report on the prevalence, patterns and performance of non-farm enterprises SSA
Okeke and Nwankwo (2017)	APSDPR	determine the perceptions of rural entrepreneurs on the role of entrepreneurship in rural economic development
Sutter, Bruton, and Chen (2019)	JBV	analyze entrepreneurship literature to poverty alleviation over the period from 1990 to 2017
Abreu, Oner, Brouwer, and van Leeuwen (2019)	JBV	investigate how entrepreneurial well-being varies across urban-rural continuum and wealthy-deprived neighbourhoods
Nikolova (2019)	JBV	provides causal evidence of physical and mental health consequences of self-employment
Bhuiyan and Ivlevs (2019)	JBV	investigate subjective well-being of microcredit-funded entrepreneurs in rural Bangladesh
Adewumi (2020)	IJBMS	assesses the challenges and outcomes of rural entrepreneurship in Nigeria
Kimmitt, Munoz, and Newbery (2020)	JBV	revisit the relationship between poverty and entrepreneurship under a eudaimonic perspective
Giambra and McKenzie (2021)	World Bank	examine the relation between self-employment or entrepreneurship and migration
Shepherd et al. (2021)	ETP	explore entrepreur's beliefs about how entrepreneurship can alleviate poverty

Note: The full titles of abbreviated journals or issuers are given in the List of Abbreviations.

Likewise, there exists evidence on entrepreneurship and well-being in a context similar to that of the present work. Tobias et al. (2013) find an improvement in quality of life through entrepreneurship for specialty coffee farmers in rural Rwanda. Bhuiyan and Ivlevs (2019) examine the extent to which microcredit-supported entrepreneurship affects the SWB of entrepreneurs in rural Bangladesh. They note higher levels of worry and depression, which indirectly reduces life

satisfaction. It should be emphasized, though, that the findings of Bhuiyan and Ivlevs (2019) cannot be generalized to micro-entrepreneurs. Instead, they are due to microcredit borrowing by entrepreneurs. Overall, the evaluation of research in the Global South shows that the evidence is not yet sufficient to gain a comprehensive understanding of outcomes on SWB.

2.4.2. Outcomes on Migration Plans/Intentions

Before the entrepreneurial impacts on migration aspirations are addressed, there is a conceptual classification and an introduction to the main causes of migration. First of all, actual migration has to be differentiated from migration plans or intentions. In short, migration intentions or plans describe the step between migration preferences and actual migration (Carling & Schewel, 2018). Migration aspirations do not necessarily lead to actual migration, partly due to missing ability to migrate (Carling & Schewel, 2018), over-reporting (Etling, Backeberg, & Tholen, 2020; Migali & Scipioni, 2019) and the dynamics of migration considerations (De Jong, Root, Gardner, Fawcett, & Abad, 1985). In fact, about 20% of the world's population considered emigration in the recent past, but less than 1% actually migrated (Abel, 2018). Hence, as with Ozaltin, Shakir, and Loizides (2020), the thesis does not aim to test the robustness of migration aspirations in predicting actual migration. Section 4 explains why it is additionally necessary to distinguish between plans and intentions to answer the research questions.

Although it has been demonstrated that people choose to emigrate primarily at a young age, there is still little evidence on the factors that drive the migration behavior of young people worldwide (Milasi, 2020). A comprehensive, empirical study of young people's migration intentions and plans was conducted by Milasi (2020) using data from 139 countries over the period 2010-2016. He finds that unemployment, involuntary part-time work, and post-secondary school completion increase both migration intentions and the likelihood of turning these intentions into concrete plans.

Most of the literature distinguishes between push factors, which induce people to leave their place of residence, and pull factors, which attract people to a new place of residence (Ozaltin et al., 2020; Salameh, 2019). From migration and sociology research, the crucial push and pull factors are economic conditions (Neumayer, 2005; Schmeidl, 1997), violence (Moore & Shellman, 2007; Steele, 2009; Weiner, 1996), as well as cultural networks (Davenport, Moore, & Poe, 2003; Schmeidl, 1997; Wood, 2008). Since entrepreneurship serves as an employment opportunity, it is seen primarily as a pull factor (Dako-Gyeke, 2016; Ozaltin et al., 2020).

Besides migration causes, a significant part of previous research on migration and entrepreneurship deals with the so-called refugee or immigrant entrepreneurship (Desai, Naudé, & Stel, 2021; Kachkar, 2019; Kerr & Kerr, 2020; Wauters & Lambrecht, 2008). This research field examines entrepreneurs after their migration. Howbeit, those who do not migrate, although previously intending to do so, will

not appear in such studies. Hence, they are subject to a selection bias (Ozaltin et al., 2020). The direct impact of entrepreneurship on migration plans has been insufficiently studied in the context of the Global South. While there is research that delves into the relationship between rural entrepreneurship and rural-urban migration in SSA (Adewumi, 2020; Bello-Bravo, 2015; Okeke & Nwankwo, 2017), migration across borders is only marginally addressed. Giambra and McKenzie (2021) provide descriptive and causal insights into the relationship between self-employment and migration. Nonetheless, their focus is not on rural SSA but on a variety of countries around the world, and the theoretical relationship between the two variables remains unclear. Moreover, there are doubts about equating self-employment with entrepreneurship, since essentially a distinction is made between the two (Parker, 2004; Shir et al., 2019).

Even though the direct effects of entrepreneurship on migration plans remain largely unexplored, there is some evidence of the indirect effects of entrepreneurship on migration via economic and non-economic factors. Studies looking at the influence of EWB or SWB on migration plans or intentions include those by Tartakovsky and Schwartz (2001) in the Urals in Russia, Chindarkar (2014) in 18 Latin American countries, Cai et al. (2014) across 154 countries worldwide, Salameh (2019) in Arab Spring countries, Bhoojedhur and Isbell (2019) in Malawi, and Ozaltin et al. (2020) in the Baghdad region of Iraq.

2.5. Transformative Entrepreneuring and Reform Perspective

According to Tobias et al. (2013), entrepreneurship can unleash potential for transformation in regions characterized by poverty and conflict. The authors present a model that follows the change-oriented idea of entrepreneurship to remove economic and social constraints through entrepreneurial processes (Mair, Battilana, & Cardenas, 2012; Rindova, Barry, & Ketchen Jr, 2009). The model by Tobias et al. (2013) consists of intentional and unintentional mechanisms with the participation of various entrepreneurial players. Critically depending on region-specific characteristics, these mechanisms can then generate both social and economic value, in turn extending the concept of transformative entrepreneurship by Venkataraman (2004). In the study of Tobias et al. (2013), the generated value consists of increased perceived quality of life and the reduction of prejudice against other ethnic groups. Ultimately, this can lead to sustainable, economic development (Tobias et al., 2013).

The thesis adopts the idea of transformative entrepreneuring by examining the mechanisms through which entrepreneurship can change the lives of ordinary entrepreneurs in poverty-stricken rural SSA. The transformative effect is obtained by looking at the intertwining of poverty alleviation with migration plans, comparable to the intertwining of poverty alleviation with conflict resolution found in Tobias et al. (2013). This work thus explores a new form of the transformative potential of entrepreneurship and responds to the request of Tobias et al. (2013) for new research on

its various components. An important difference, however, is that the longitudinal study by Tobias et al. (2013) examines entrepreneuring, and this views entrepreneurship as a process rather than a state. The thesis, in contrast, deals with entrepreneurship as a state. Therefore, it is not possible to examine how the transformative potential unfolds over time. Only an assessment can be made by comparing entrepreneurs with a control group, the unemployed. This is the reason, why a static model is formed instead of a process model.

Sutter et al. (2019) review the literature on entrepreneurship and poverty alleviation from three perspectives: remediation, reform, and revolution. The thesis examines entrepreneurship from the reform perspective. This is in part because it assumes that entrepreneurship can overcome poverty by enabling social change (Sutter et al., 2019). In the assumption of the thesis, this change is achieved by increased SWB and EWB and reduced migration plans. By contrast, the reform perspective is characterized by the fact that one is primarily interested in the lived experiences and views of those living in poverty. Consequently, one takes a more subjective view than in the remediation or revolution perspective (Sutter et al., 2019). In this work, the interest is in just those, namely the perceptions of young rural entrepreneurs.

2.6. Contribution to the Previous Literature

The thesis contributes to the existing literature on the impacts of entrepreneurship in poverty-stricken regions (see Table 1) by examining the outcomes of entrepreneurship for young adults in rural SSA. In doing so, it responds to calls from academia for more context-specific entrepreneurship research (Wiklund et al., 2019), particularly in developing countries (Bruton et al., 2008). In this way, it also reflects the need for more empirical and conceptual management research on the African continent (George et al., 2016).

To date, most entrepreneurship research on poverty has been qualitative. This is largely due to the difficulty of obtaining reliable, quantitative data in impoverished settings (Bruton et al., 2013). Now that a solid foundation of theoretical evidence is available, calls for more quantitative work on entrepreneurship and poverty are getting louder (Bruton et al., 2013; Tobias et al., 2013). This paper answers this call with a comprehensive empirical analysis of 31 countries using data from the pan-African research institute Afrobarometer.

The focus of entrepreneurship research has so far been on firm-level outcomes being growth and performance. In recent years, other outcomes like well-being have received increasing attention (Wiklund et al., 2019). The thesis ties in with the idea of Johannisson (2011) of liberating entrepreneurship research from an exclusively economic position and instead regard it as a practice of everyday life. A combination of economic indicators such as individual wealth with non-economic indicators such as perceived quality of life allows a better understanding of social and economic growth in developing countries (Tobias et al., 2013).

This approach is taken up by combining EWB and SWB as outcomes of entrepreneurial action.

Research on the relationship between entrepreneurship and migration plans or intentions has been sparse. In particular, the literature review revealed that the direct impact of entrepreneurship on migration plans in the context of rural SSA has been insufficiently studied. In this regard, this work aims to provide new insights. Furthermore, SWB is not only a desirable outcome but may itself generate other outcomes just as future life events (Diener, 2012; Wiklund et al., 2019). Nevertheless, research to date on the drivers of international migration has concentrated on objective factors such as income (Cai et al., 2014). Since not only EWB but also SWB may have an impact on migration decisions (Cai et al., 2014), an investigation of both constructs in a still under-researched context can offer new insights. Further, while the relationship between income and migration and life satisfaction and migration has previously been studied separately, there is a lack of studies connecting the three concepts (Chindarkar, 2014). In this regard, the thesis also adds to the existing literature. Additionally, one picks up the thoughts of Ozaltin et al. (2020) and circumvents the selection bias that the literature around refugee entrepreneurship often faces. This is made possible by examining the plans of entrepreneurs before a potential migration rather than surveying them after the actual migration.

Last but not least, the thesis is a response to Bruton et al. (2013), who argue that entrepreneurship is a long-term solution to poverty when scholars understand how to help people in poverty start their businesses. Thus, it takes a different perspective than the one that has been widely used. The common one is based on the fact that people in poverty are seen as a large potential market for the sale of a variety of goods and services (Bruton et al., 2013). The view of the thesis, on the contrary, is whether entrepreneurs can break the cycle of poverty on their own, and consequently is consistent with the reform perspective (Sutter et al., 2019). So, the focus shifts to impacts on the micro-entrepreneur himself rather than on effects at the country or regional level. This is in response to the claims of Tobias et al. (2013), that previous literature is primarily limited to the latter two, neglecting local and community levels. For example, entrepreneurship to overcome poverty has often been viewed from the perspective of governmental organizations, government agencies, or individual heroic entrepreneurs (Rindova et al., 2009; Tobias et al., 2013). Now, comparable to coffee farmers in rural Rwanda, young entrepreneurs in rural SSA form the empirical unit of study. For these same people, entrepreneurship might have the effect of unleashing a transformative potential (Tobias et al., 2013). In addition to the benefits for EWB and SWB, this could even lead to a reduction in migration considerations. With this, a new type of transformation may be integrated into the theory of transformative entrepreneuring. Ultimately, it becomes easier to understand whether poverty has been eliminated from the perspective of an ordinary entrepreneur.

3. Theoretical Model and Hypotheses Development

The purpose of this chapter is to develop a testable theoretical model with corresponding hypotheses based on the literature presented earlier. The model will then be used to explore the relationship between micro-entrepreneurship and migration plans of young people in rural SSA.

3.1. Entrepreneurship of Young Adults in Rural SSA as Research Context

SSA generally provides a wide range of opportunities for entrepreneurship research (George et al., 2016). Economically, it is one of the fastest-growing regions in the world (Igwe & Icha-Ituma, 2020). That aside, the population is expected to increase by about one billion people between 2019 and 2050, which is half of the global population growth during that period (UN, 2019). With the recent reduction in the fertility rate, the young, working-age population is growing faster than other age groups. In 2015, 62% of the population was already under 25 years old (Kew, 2015). From 2010 to 2035, about 450 million people are expected to join the labor force (George et al., 2016). This demographic dividend offers the opportunity for accelerated economic growth (UN, 2019).

The flip side of this demographic, however, is an increasing threat of being affected by youth unemployment (Bommes, Fassman, & Sievers, 2014). The likelihood of becoming unemployed is already three times higher for young Africans than for adults (Kew, 2015). The opportunities in the formal sector are decreasing and the proportion engaging in informal entrepreneurship, in turn, is rising (Gough & Langevang, 2016). The labor shortage and widespread extreme poverty (Beegle et al., 2016) leave many with no choice but to flee their home countries (Milasi, 2020). In global refugee flows, this is noticeable in that a large proportion of migrants are coming from sub-Saharan countries (UNHCR, 2021). According to the Pew Research Center, 8 out of 10 of the fastest-growing international migrant populations since 2010 have come from SSA. More than that, the number of emigrants from SSA has increased by at least 50% between 2010 and 2017 (Connor, 2018).

On top of the strong growth of the young population, there are other reasons to focus on this age group. Partly because the entrepreneurial propensity is stronger among young Africans than older Africans. In Uganda, for example, it is 1.8 times as high (Kew, 2015). Even if the entrepreneurial tendency does not automatically mean that the respondents then actually start a business, it demonstrates a generally positive attitude towards entrepreneurship as a career choice. This openness is also demonstrated by the fact that young Africans are less concerned about the failure of their business than older ones (Kew, 2015).

In terms of migration, young people are the ones most likely to leave the home country (IOM, 2020; Migali & Scipioni, 2019; Milasi, 2020). Global estimates show that between 2010 and 2015, the net migration of young migrants (15-29 years) was five times higher than that of older people (30-64 years). This is partly due to the high mobility of young people (Milasi, 2020). Another point in favor of focusing on the young population is that tackling unemployment has a stronger impact on young people's intentions to migrate than it does for older people (Milasi, 2020).

In the context of this work, it is necessary to distinguish between rural and urban regions. The share of entrepreneurship in rural areas is similar to that in urban areas, however, opportunities and constraints for youth entrepreneurship differ (Gough & Langevang, 2016). Rural areas are more affected by extreme poverty than urban areas. They are often characterized by low incomes, low productivity, lack of infrastructure, high unemployment, and high population density (Adewumi, 2020; Gough & Langevang, 2016; Kew, 2015). In addition, farmland is becoming scarcer due to population growth, mineral exploitation, and commercial farming (Gough & Langevang, 2016). Entrepreneurship offers the potential to change such deplorable conditions especially in rural settings (Tobias et al., 2013). For all these reasons, the context of the thesis is set on the young, rural population of SSA.

3.2. Theoretical Model

The scientific evidence discussed previously, the idea of the transformative potential of entrepreneurship in SSA (Tobias et al., 2013), and the view from the reform perspective (Sutter et al., 2019) provide the essential components of the logic underlying the link between micro-entrepreneurship and migration plans. According to the thesis' assumption, this link is mediated by economic and subjective well-being. In this way, EWB and SWB form the construct of poverty alleviation, comparable to Tobias et al. (2013). The individual relationships are illustrated in Figure 1.

As the model shows, the basis is the direct relationship between microentrepreneurship and migration plans. Rural micro-entrepreneurs are included in the thesis as such if they live in the rural region and are engaged in full-time self-employment that generates an income. Section 4 will give a detailed description of the variables.

3.3. Hypotheses Development

3.3.1. Micro-Entrepreneurship and Migration Plans

Throughout this section, it is discussed how the migration plans of young entrepreneurs and young unemployed people differ in SSA's rural areas. Unemployment affects many young people in SSA and is seen as a major driver of migration. It is relevant to the thesis in that a reduction of unemployment through entrepreneurship could mitigate this key driver of migration. The following empirical findings first explain the role of unemployment as a reason for emigration. Unemployment does not have to mean that people are simply inactive. Instead, it might also mean housework, irregular piecework and, illegal activities such as prostitution or robbery (Gough & Langevang, 2016). According to Bhoojedhur and Isbell (2019), turning to the migration plans and
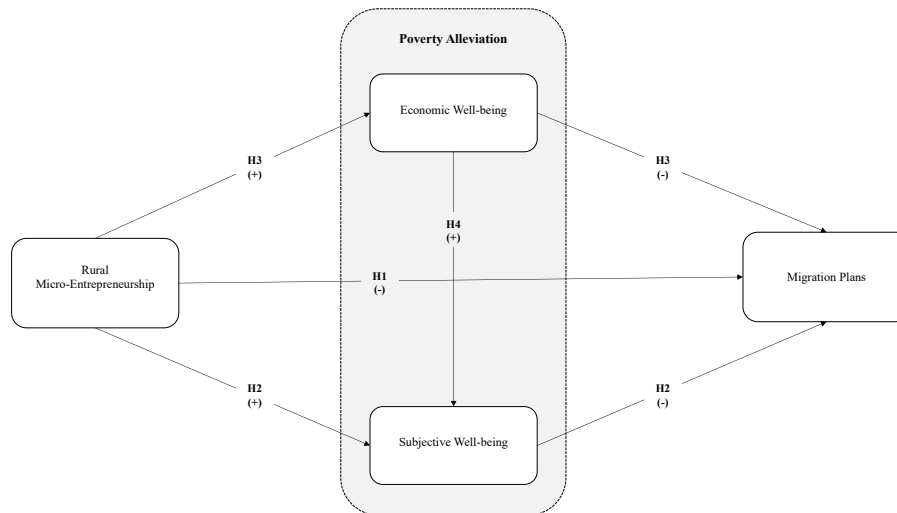
Figure 1: Hypothesized Model Predicting Migration Plans of Rural Micro-Entrepreneurs

Note: own illustration, following the mediation model of Nikolaev et al. (2020).

intentions of Malawians, 40% of respondents cite the search for work as a motive for considering emigration. Similar findings were obtained in the Arab Spring countries. Bommes et al. (2014) argue that the limited labor opportunities result mainly from the strong population growth combined with slow economic growth. Migali and Scipioni (2019) find that across different country groups, the unemployed are more likely to want to migrate than the people in employment, i.e., employees and the self-employed. The migration-inducing effect of unemployment has also been observed specifically among the young population (Milasi, 2020).

Regarding unemployment as a cause of migration, the thesis differentiates between intended unemployment and unintended unemployment. People outside the employment phase, i.e., with intended unemployment, appear to have a lower desire to emigrate than individuals in the employment phase (Migali & Scipioni, 2019). This is particularly true for young people, most of whom have not yet reached the labor force phase. Milasi (2020) reasons that young people away from the labor market are usually students and have fewer concrete plans about permanent emigration at this stage of their lives than later in their working lives.

In addition to unemployment, uncertainty about employment status is also seen as a driver of migration intentions (Bhoojedhur & Isbell, 2019; Milasi, 2020). Those in uncertain employment conditions, such as involuntary part-time employment, are particularly likely to migrate. Low pay and poor working conditions play a major role in this (Milasi, 2020). The fraction of those who strongly consider emigration is 43% for part-time employment, but only 23% each for full-time employment or voluntary unemployment (Bhoojedhur & Isbell, 2019). Full-time entrepreneurship without dependence on the existing employer could reduce such uncertainty about employment status (Nikolova, 2019).

The question is whether entrepreneurial activity per se

can reduce migration plans. Studies examining the relationship between self-employment and migration confirm that the self-employed in developing and emerging economies are less likely to migrate than the inactive or wage earners. This is the case regardless of age, gender, education, and income (Giambra & McKenzie, 2021; Milasi, 2020). Using randomized experiments to estimate causal effects, Giambra and McKenzie (2021) also find that programs that lead to more self-employment and entrepreneurship reduce actual migration. Although these effects are rather modest and take several years to materialize, there is at least no evidence that self-employment-enhancing policies would increase migration (Giambra & McKenzie, 2021).

There are different attempts to explain this negative relationship. For example, self-employment may increase employment opportunities in the home country and thus the opportunity costs of migration abroad. This argument is also usually invoked by policy-makers, who advocate more jobs in the home country to prevent emigration (Giambra & McKenzie, 2021). Moreover, self-employment may entail a lock-in effect, which means one has already invested money and other resources in the own business and this capital could be lost in the event of migration. In addition to physical capital, this also affects intangible capital such as customer or supplier relationships or the value of the brand. Such capital would be difficult to monetize in the event of emigration (Giambra & McKenzie, 2021). Furthermore, there is a risk that the skills learned through self-employment are less transferable to other places than the skills learned through wage labor. This in turn may reduce career prospects abroad (Giambra & McKenzie, 2021).

According to the studies, it is still unclear whether self-employment per se is the cause of weaker migration plans or whether it is more a problem of self-selection. After all, self-employment is often the last resort for people in poverty,

and the poor simply do not have the means to migrate (Giambra & McKenzie, 2021; Milasi, 2020). In addition, some factors might make self-employed people even more likely to migrate, such as increased risk-taking and the ability to discover new business opportunities (Giambra & McKenzie, 2021; Naudé, Siegel, & Marchand, 2017; Williamson, 1965). Overall, the research findings do not yet allow for theoretical clarity on the actual effect of self-employment or entrepreneurship on migration, thus showing that further research is needed (Giambra & McKenzie, 2021). However, when the unintentionally unemployed and entrepreneurs are contrasted, evidence to date suggests that the latter are likely to have a lower propensity to migrate.

Based on the evidence presented, it is hypothesized that rural young microentrepreneurs have fewer migration plans than rural young unemployed. Last, it should be emphasized that the thesis is more concerned with identifying microentrepreneurship as a meaningful indicator of migration plans but less concerned with obtaining a concrete causal effect.

> **Hypothesis 1:** *Micro-entrepreneurs experience lower levels of migration plans than those who are unemployed.*

3.3.2. Subjective Well-Being as a Pathway to Migration Plans

The following section presents findings comparing entrepreneurs and the unemployed in terms of SWB and its influence on migration plans. Essentially, work has a positive impact on people's life satisfaction or happiness, whereas unemployment has a negative impact. This is true regardless of any selection effects, which would imply that people deliberately choose unemployment because of their dissatisfaction (Binder & Coad, 2013). In a recent study, Nikolova (2019) finds positive health effects of entrepreneurship, among both opportunity and necessity entrepreneurs. Regardless of whether they have employees or work alone and regardless of income, risk preferences, and personality traits, entrepreneurs benefit health-wise from their activities (Nikolova, 2019). Several factors must be considered to identify the underlying route cause. Among other things, one can escape the stigma of unemployment and gain full control over one's job. In addition, one can achieve more flexibility and experience an identity boost (Nikolova, 2019). Besides, the procedural utility theory (Benz & Frey, 2008) implies that health benefits can result solely from the process of entrepreneurial activity (Nikolova, 2019).

Apart from the health benefits, there are also indications of higher life satisfaction through entrepreneurship. Especially in underprivileged areas, where there are greater restrictions on wage employment, self-employment represents a viable alternative for the labor market. Abreu et al. (2019) find an increased life satisfaction in such deprived areas, as also present in rural SSA. On top of that, the entrepreneurial experience helps young people develop new skills, such as discovering new business opportunities, training innovative and critical thinking, and embracing teamwork and leadership skills. Finally, entrepreneurship gives young people the

opportunity to be involved in society (Kew, 2015). Such potentials indicate why life satisfaction may be higher among self-employed young people in SSA (Kew, 2015). Based on previous findings, it is hypothesized that young microentrepreneurs have higher SWB relative to the unemployed in rural SSA.

The question arises whether SWB increased by entrepreneurship in turn influences migration plans and consequently acts as a mediator. Tartakovsky and Schwartz (2001) find a significant link between SWB and various motivations for migration, but do not establish a direct link between SWB and migration itself. Salameh (2019) states that migration is influenced by social factors, including social and psychological aftereffects of economic factors. Further empirical evidence indeed shows that the desire for international migration is negatively associated with SWB and that this is especially true for individuals with high levels of education (Cai et al., 2014; Chindarkar, 2014). It should be noted that the studies are not looking for a causal relationship, but a correlation. Complementing this, greater dissatisfaction with living situations and a pessimistic view of future living conditions are associated with a stronger propensity to migrate (Migali & Scipioni, 2019).

Overall, research shows that entrepreneurial activity can increase SWB. Especially when compared to the unemployed, micro-entrepreneurs seem to have a higher SWB. And the latter seems to mediate, at least in part, the effect on migration plans. Based on the predominantly empirical findings explained above, the following second hypothesis is formulated.

> **Hypothesis 2:** *Micro-entrepreneurs experience higher levels of subjective well-being than those who are unemployed, which partially mediates the negative relationship between entrepreneurship and migration plans.*

3.3.3. Economic Well-Being as a Pathway to Migration Plans

This section discusses scientific evidence on how the EWB of micro-entrepreneurs differs from the unemployed and how this relates to migration plans. The review of youth businesses by Kew (2015) reveals that the economic impact of entrepreneurship varies across sub-Saharan countries. Accordingly, the proportion of entrepreneurs expressing a positive impact rather than a negative or no impact on their livelihood varies significantly. In Malawi, Angola, Uganda, Ghana, and Nigeria, about half of youth businesses have a positive impact on livelihoods. In contrast, in Botswana, Namibia, Zambia, and South Africa, the rate is only around 30% (Kew, 2015). Notwithstanding the heterogeneity, such figures are suggestive of the economic potential of entrepreneurship for young adults.

Qualitative studies reinforce the role of entrepreneurship as a solution to mass unemployment and underemployment (Kolawole & Ajila, 2015). They point to the potential for wealth and income, and hence overall living standards, among the rural population (Okeke & Nwankwo, 2017). A

survey in rural Nigeria shows for instance that entrepreneurship empowers residents to generate at least enough income for a proper meal (Adewumi, 2020). Quantitative evidence allows drawing further conclusions about the economic impact for the ordinary entrepreneur in SSA. Tobias et al. (2013) find that the income and personal wealth of specialty coffee farmers in Rwanda have increased significantly since they began to produce.

Since previous empirical insights suggest that entrepreneurs have a higher EWB than the unemployed, the question arises to what extent the EWB is related to migration plans. The following research assumes that higher EWB is associated with reduced migration plans. According to Bhoojedhur and Isbell (2019), 51% of Malawian respondents cite economic hardship or escaping poverty as a reason to consider migration. Respectively, those least affected by poverty would be least likely to migrate. That income levels play a crucial role in determining migration flows also emerges from the study by Kandemir (2012). It says that people in 25 of the 27 largest migration corridors worldwide migrate to countries with a higher Human Development Index (HDI). Only in two cases, the decision falls on countries with a lower HDI. Survey results from Ozaltin et al. (2020) in the conflict-ridden Baghdad region underscore the importance of income levels for migration intentions. Alongside this, evaluations of surveys in Arab Spring states reveal that, in addition to violence and persecution, fleeing economic hardship and the desire for a better living situation are major reasons for emigrating (Salameh, 2019). Overall, the great importance of materialism, wealth, financial well-being, or control over material resources for migration intentions is widely accepted by researchers (Chindarkar, 2014; Dako-Gyeke, 2016; Milasi, 2020; Tartakovsky & Schwartz, 2001).

In contrast to the above insights, however, some studies report a strengthening effect of higher EWB on migration. As such, people living in extreme poverty seem more likely to emigrate after a certain improvement in income (Adhikari & Gentilini, 2018; Clemens, 2020; UNDP, 2009). Especially budget constraints seem to discourage young people in low-income countries from turning their migration intentions into concrete plans (Milasi, 2020). Self-employment might be one reason for overcoming such budget restrictions (Giambra & McKenzie, 2021). Also, uncontrolled, improperly designed cash transfers to poor people can overcome such income constraints on migration (Adhikari & Gentilini, 2018). Logically, the relationship between EWB and migration plans could also run in the opposite direction once a certain income threshold is exceeded. This, in turn, suggests a non-linear relationship (Dao, Docquier, Parsons, & Peri, 2018; Friebel, Gallego, & Mendola, 2013). The above evidence shows that the relationship between EWB and migration efforts is still controversial in the literature. Therefore, it is worthwhile to examine the role of EWB in the context of the thesis. Altogether, the discoveries of previous literature outweigh the cognition that higher EWB through entrepreneurial activity is negatively correlated with migration plans. The following third hypothesis is put forward.

Hypothesis 3: *Micro-entrepreneurs experience higher levels of economic well-being than those who are unemployed, which partially mediates the negative relationship between entrepreneurship and migration plans.*

3.3.4. Economic Well-Being and Subjective Well-Being

Finally, the thesis addresses the question of whether the increased EWB resulting from entrepreneurial activity in turn positively influences SWB. According to Tobias et al. (2013), the linkage between economic and social value creation is initially triggered by the economic value created. Specifically, the increased personal wealth of specialty coffee farmers in rural Rwanda is associated with a positive change in the quality of life and thus higher SWB (Tobias et al., 2013). Since the study context of Tobias et al. (2013) is very similar to that of the underlying thesis, it is assumed that this connection also applies to young, rural entrepreneurs in other countries of SSA. Furthermore, it is demonstrated that poverty alleviation consists not only of economic factors but also of life satisfaction (Tobias et al., 2013).

Apart from this, Tobias et al. (2013) find that increased SWB, in turn, reduces outgroup prejudice. Specifically, only when higher income leads to higher SWB are entrepreneurs more likely to engage with people from other ethnic backgrounds. Income, therefore, does not unfold its effect alone, but in conjunction with SWB. In the context of the thesis, this would mean that only if higher EWB leads to higher SWB, entrepreneurs are less likely to emigrate. If so, the SWB would at least partially explain the impact of the EWB on migration plans.

Salameh (2019) also points to the close linkage of social factors with economic factors as causes of migration. Economic factors such as poverty, hunger, unemployment, and low living standards promote social and psychological constraints such as social exclusion and inequality, and alienation. For this reason, one should not look at EWB and SWB independently of each other, but rather take a closer look at possible linkages. Based on this knowledge, the final hypothesis is therefore that increased EWB due to micro-entrepreneurship leads to higher SWB.

Hypothesis 4: *Economic well-being of micro-entrepreneurs is expected to have a positive effect on their subjective well-being.*

4. Methodology

This chapter first presents the selected data set and the relevant characteristics of the sample. It then explains how the different variables are constructed and which challenges are involved. Finally, the procedure of the subsequent empirical analysis is introduced.

4.1. Data and Sample

The empirical analysis is conducted with data from Afrobarometer. This is a panAfrican, non-partisan research

project that has been running surveys since 1999. The cross-national surveys are intended to examine the social, economic, and political conditions in Africa. Furthermore, they should be representative of a country's population in a given year (Mattes, Dulani, & Gyimah-Boadi, 2016). By drawing primary samples randomly from the population, every adult citizen has an equal chance of being selected for an interview (Afrobarometer, 2021). The database provides comprehensive crosssectional data for more than 30 countries of SSA. This allows the region to be largely mapped as the context of the study. Besides, the data are publicly available and can be considered up-to-date due to their recent collection. Beyond that, the specific questions on migration preferences and the economic and social situation of the respondents are suitable for approximating the required variables. Taking all aspects into consideration, Afrobarometer is the preferred database over alternatives such as Global Entrepreneurship Monitor, World Bank, or Gallup World Poll.

Having chosen Afrobarometer as its database, the thesis draws on the seventh round of surveys conducted between 2016 and 2018 (Mattes, 2020). Only this survey round contains the specific migration module needed to study the dependent variables migration plans and intentions. The Afrobarometer data set is processed and analyzed using the statistical software RStudio. As a first step, unneeded variables are dropped. Then the countries Morocco and Tunisia are removed, as they do not belong to SSA due to their geographical classification (UNDP, 2013). Kenya is also excluded, as migration plans were not surveyed there. This leaves 41,825 observations from 31 of the total 46 SSA countries (UNDP, 2013). Thereupon, the dataset is restricted to the rural sampling units. Lastly, to reflect the young population, only respondents between the ages of 18 and 35 are included in the analysis. This classification is comparable to that of Kew (2015) and Milasi (2020). The proportion of young adults in the sample is 52.61%, accounting for about half of the rural working-age population. Overall, after limiting to the rural youth population and deleting missing values, the dataset consists of 10,106 individual-level observations.

A common measure of entrepreneurial engagement is the Total Early-Stage Entrepreneurial Activity (TEA). This metric indicates participation in new or nascent businesses (Gough & Langevang, 2016) and averages 26.6% for 18-64-year-olds (Amoros & Bosma, 2014) and around 20% for 18-34-year-olds in SSA (Kew, 2015). The TEA of 18-35-year-olds in the present rural sample is a relatively low 10.33%. The difference in proportions might be due to several reasons. First of all, this thesis considers only rural entrepreneurs. However, the TEA may be higher in urban areas than in rural areas (Amha, Woldehanna, Tamrat, & Gebremedhin, 2015). In addition, the identification of micro-entrepreneurs in the thesis might differ from that by Amoros and Bosma (2014) and Kew (2015). Possibly, they use a less restrictive definition of micro-entrepreneurs and thus obtain a higher TEA. Third, large heterogeneity in TEA rates exists across SSA countries (Kew, 2015; Nagler & Naudé, 2017). For example, the minimum TEA is 13% in South Africa, whereas the maximum is

56% in Uganda (Kew, 2015). Last but not least, the results of the above studies are limited to only eight (Amoros & Bosma, 2014) and nine (Kew, 2015) instead of 31 countries, which might lead to further discrepancies.

4.2. Variables

4.2.1. Rural Micro-Entrepreneur

The empirical analysis compares the micro-entrepreneurs as the treatment group with the unemployed as the control group. The two groups are mapped with a dummy variable. Rural micro-entrepreneur is given the value 1 if the respondent lives in the rural region (Gough & Langevang, 2016; Koyana & Mason, 2017), has a full-time job that generates an income (Nikolova, 2019) and is self-employed (Dale, 2015; Hébert & Link, 1989; Sohns & Diez, 2018).

Rural micro-entrepreneur receives the value 0 if the respondent is unemployed. To this end, the respondent states that he lives in the rural region and is involuntary without work, so is looking for a job (Kew, 2015; Klasen & Woolard, 2009). The Afrobarometer statements for the variable are listed in Table 12 in Annex A.1. The dummy mapping in the rural sample yields a total of 1,044 micro-entrepreneurs and 3,386 involuntarily unemployed. Thus, the N for the two groups in the empirical analysis is 4,430. Table 2 shows the country-specific partitioning of the sample groups.

4.2.2. Subjective Well-Being

As already mentioned in Section 2, the concept of evaluative well-being, i.e. life satisfaction and basic happiness, is utilized to approximate SWB. Hedonic well-being and eudaimonic well-being are likely to be imprecisely captured by the available data and are therefore not considered. The components of SWB including their weighting are presented in Table 3. The modeling of SWB follows a similar procedure in the study of Abdallah et al. (2008) on the estimation of global life satisfaction. The authors also use Afrobarometer data and follow the evaluative well-being approach to frame the SWB. Specifically, their SWB is based on statements about satisfaction with personal and national economic and political conditions. In addition, it considers assertions about fear of crime and violence, family life, and how often respondents have gone without food in the past year. As with Abdallah et al. (2008), the SWB index builds on the Hierarchy of Needs by Maslow (1943). The relevance and applicability of Abraham Maslow's model in the context of entrepreneurship in developing countries has already been confirmed in further studies (Dencker et al., 2021; Hagerty, 1999; Sirgy, 1986).

The variable SWB attempts to replicate the five levels of the Hierarchy of Needs, consisting of physiological, safety, love/belongingness, esteem, and self-actualization (Maslow, 1943). To replicate the pyramidal structure of the Hierarchy of Needs, the components are first scaled and then weighted accordingly. Nevertheless, the methodology is not without limitations. For the top level of self-actualization, as with Abdallah et al. (2008), no matching variable was found in the Afrobarometer dataset. Thus, the index is restricted to

Table 2: Number of Observations by Country for Different Sample Groups

#	Country	SSA Sample	Rural Youth	Rural Micro-Entrepr.	Rural Unemployed
				N (Number of Observations)	
1	Benin	1,200	338	26	149
2	Botswana	1,198	116	3	71
3	Burkina Faso	1,200	422	12	115
4	Cabo Verde	1,200	177	16	71
5	Cameroon	1,202	260	9	90
6	Côte d'Ivoire	1,200	248	25	59
7	Eswatini	1,200	448	13	203
8	Gabon	1,199	123	1	58
9	Gambia	1,200	214	27	67
10	Ghana	2,400	460	172	118
11	Guinea	1,194	323	5	44
12	Lesotho	1,200	219	2	117
13	Liberia	1,200	344	2	228
14	Madagascar	1,200	331	33	69
15	Malawi	1,200	512	24	95
16	Mali	1,200	191	16	27
17	Mauritius	1,200	192	20	32
18	Mozambique	2,392	656	6	277
19	Namibia	1,200	229	3	103
20	Niger	1,200	468	22	177
21	Nigeria	1,600	565	86	178
22	São Tomé & Príncipe	1,200	189	15	70
23	Senegal	1,200	256	23	77
24	Sierra Leone	1,200	320	40	149
25	South Africa	1,840	242	3	133
26	Sudan	1,200	389	13	116
27	Tanzania	2,400	672	308	93
28	Togo	1,200	333	13	127
29	Uganda	1,200	405	79	132
30	Zambia	1,200	271	15	69
31	Zimbabwe	1,200	193	12	72
		41,825	10,106	1,044	3,386

Note: based on Afrobarometer survey round 7 conducted within 2016-2018.

the other four levels. Further, the seventh round of the Afrobarometer survey does not include any questions on physical or mental health. So, the safety level is shaped solely by personal economic conditions, which slightly overweights them in the index. The remaining three levels of Maslow's Hierarchy of Needs are modeled identically to Abdallah et al. (2008).

Given the poor data availability on SWB in SSA, there is no better alternative in the context of the thesis. Apart from that, Abdallah et al. (2008) are convinced of the basic appropriateness of their methodology in measuring the subjective quality of life. Moreover, the sample of the thesis with a total of 31 countries represents SSA far more completely than the sample of Abdallah et al. (2008). Beyond that, the ranking of the countries in terms of their median SWB seems plausi-

ble. An overview of the country ranking in the sample can be found in Figure 5 in Annex A.2. Over the survey period from 2016 to 2018, less fragile countries with a higher HDI, such as Mauritius or São Tomé and Príncipe, perform significantly better than more fragile, less developed countries such as Guinea or Madagascar (TheGlobalEconomy, 2021; UNDP, 2016).

4.2.3. Economic Well-Being

The concept of lived poverty is applied to reflect EWB. Specifically, the EWB variable is based on an experiential measure developed by Afrobarometer, the Lived Poverty Index (LPI). The LPI measures the frequency with which people experience shortages of basic needs. It complements official poverty and development statistics because information on

the scarcity of basic needs is a core concept of poverty and this is often inadequately captured by other measures (Mattes, 2020). In its original form, the higher the value, the higher the experienced poverty. The scale of the original variable is therefore reversed. Consequently, higher values mean lower lived poverty and, hence, higher EWB. Apart from reports published by Afrobarometer (Dulani, Mattes, & Logan, 2013; Mattes, 2020; Mattes et al., 2016), other studies use the LPI in identical (Hodler, Srisuma, Vesperoni, & Zurlinden, 2020; Odhiambo, 2019) or modified form (Cheeseman, 2015) as a measure of personal wealth.

The LPI is considered to have an acceptable degree of reliability and validity with a Cronbach's alpha of 0.71 to 0.77 (Meyer & Keyser, 2016). Individual income is not asked in the Afrobarometer surveys because, among other reasons, there are doubts about the reliability of such data (Cheeseman, 2015). Due to limited data and the advantage over a purely asset-based wealth index (Odhiambo, 2019), the thesis, therefore, relies on the LPI as an approximation for the EWB. As with the SWB, there are significant differences between the countries. The country-specific EWB-ranking is highlighted in Figure 6 in Annex A.3. For example, while the young, rural population in Mauritius rarely experiences shortages in basic needs, they quite often suffer deprivation in Gabon and Togo (Mattes, 2020).

4.2.4. Migration Plans/Intentions

In the empirical analysis, both variables migration plans and migration intentions serve to depict emigration aspirations. Despite their essentially similar classification (Carling & Schewel, 2018), migration intentions and plans must be distinguished from each other. The crucial difference between the two terms is their concreteness (Carling & Schewel, 2018; Tjaden, Auer, & Laczko, 2019; Van Dalen & Henkens, 2008). Migration plans describe concrete planning already undertaken, possibly even preparation for emigration, and not just the mere idea or consideration of emigration described by migration intentions (Carling & Schewel, 2018). If they are more abstract, they tend to reflect dissatisfaction with the personal situation in the country. If they are more concrete, they express specific plans to emigrate (Tjaden et al., 2019).

The Afrobarometer items including the description of both dependent variables can be found in Table 3. Migration plans become more and more concrete on a fourpoint scale from "Considering migration not at all", through "You are not currently making any specific plans or preparations" and "You are planning to move in the next year or two but not yet making preparations" to "You are currently making preparations to move, like getting a visa" (Afrobarometer, 2020, p.49). Migration intentions are increasingly ranked on a four-point scale from "Not at all", through "A little bit" and "Somewhat" to "A lot" (Afrobarometer, 2020, p.48). Similar to the distinction in the thesis, (Milasi, 2020, p.5) uses the variables "Plan", whether the respondent has already made plans to emigrate within the next 12 months, and "Desire",

whether the respondent would like to emigrate permanently if possible.

The difference becomes clearer if one compares the number of similar answers to both questions. The number of answers with concrete emigration plans is significantly lower than the number of answers with only emigration considerations (Carling & Schewel, 2018). According to the Gallup World Poll, 32% of respondents in SSA had the desire to emigrate, but only 4% planned to do so in the next year. Only 1% have already made preparations (OECD, 2015). In other words, only a fraction of the people who wish to emigrate actually become active. Tjaden et al. (2019) strengthen this by claiming that the exact wording in the question about migration aspirations is crucial to interpreting the answer. The more abstract the question, the more likely people might show a desire to migrate.

The literature review revealed a research preference for migration intentions over migration plans. Based on those findings, more research should be done on the latter. Hence, the main model relies on migration plans. Migration intentions are then used to test the main model for robustness. This way, migration aspirations can be captured both in their non-binding and concrete form.

4.2.5. Challenges in Variable Measurement

The empirical analysis is restricted to Afrobarometer data, which were not specifically collected to test the hypotheses. This implies that the explanatory and mediating variables, meaning micro-entrepreneurship, SWB, and EWB, do not emerge directly from the survey. The definition of micro-entrepreneurship lacks information on entrepreneurial motivation, type of business, and the number of employees. Such information might have been useful, as it would enable more accurate classification of entrepreneurs and non-entrepreneurs. In the case of SWB, the dataset does not contain specific questions on mental and physical health. Thus, unlike the methodology of Abdallah et al. (2008), the safety level cannot be fully replicated along the lines of Maslow (1943). The low internal consistency of the SWB index with a Cronbach's alpha of 0.5 is therefore not surprising. For comprehensive measurement, one should also consider hedonic and eudaimonic aspects of well-being.

Similar challenges are faced in the construction of EWB. Carter (2011) points out that it is important to use multidimensional constructs to measure EWB, as previous studies have often been limited to static measures such as income. This means that other measures such as assets, savings, pensions, or consumption estimates should be taken into account in order to capture relative wealth over a longer period of time. The restriction to static measures, however, bears the risk that some economic outcomes for entrepreneurs are not included in the EWB (Carter, 2011). Whether the LPI is therefore suitable as an approximation for a multidimensional measure has been sparsely studied. Meyer and Keyser (2016), for instance, classify the LPI as a unidimensional construct controlling for different samples and countries. Therefore, other measures may be needed instead of LPI to capture

Table 3: Variable Description for Poverty Alleviation and Migration Plans/Intentions

Variable Name Need level	*Afrobarometer Item* & Description	Weighting
Economic Well-being (EWB)	*Lived Poverty Index (reversed)* "... an experiential measure that is based on a series of survey questions about how frequently people actually go without basic necessities during the course of a year." (Mattes, 2020, p.3)	
Subjective Well-being (SWB)		
Physiological Needs	*Access to food* "Over the past year, how often, if ever, have you or anyone in your family: Gone without enough food to eat?"	40%
Safety Needs	*Personal economic conditions* "In general, how would you describe: Your own present living conditions?"	30%
Love/Belongingness Needs	*Violence in the family* "[...] It can always be justified, sometimes be justified, or never be justified for... ... parents to use physical force to discipline their children? ... a man to beat his wife?" *Experience of crime* "Over the past year, how often, if ever, have you or anyone in your family... ... felt unsafe walking in your neighbourhood? ... feared crime in your own home?"	20%
Esteem Needs	*Satisfaction with democracy* "Overall, how satisfied are you with the way democracy works in [country]?" *Optimism* "Looking ahead, do you expect economic conditions in this country to be better or worse in 12 months' time?"	10%
Migration Plans	*Planned emigration to other country* "How much planning or preparation have you done in order to move to another country to live?"	
Migration Intentions	*Considered emigration to other country* "How much, if at all, have you considered moving to another country to live?"	

Note: All items are extracted from Afrobarometer survey round 7 conducted within 2016-2018 (Afrobarometer, 2020).

EWB in a multidimensional way.

4.2.6. Control Variables

To take other socio-economic conditions into account that may influence (micro-) entrepreneurship, SWB, EWB, and migration plans or intentions, several control variables are

included in the empirical analysis, in line with the existing evidence. The variable age squared is added to capture for non-linearity in the relationship of age and other correlating variables (Theodossiou, 1998). The literature on control variables in the research context is summarized in Table 4. For simplicity and clarity, literature on control variables relating to EWB is not included in the table. However, it is found that EWB or LPI are also influenced to some extent by factors such as gender, age, or education (Beegle et al., 2016; Kew, 2015; Mattes, 2020). EWB is therefore controlled just like the other variables. The individual variable descriptions for the control variables can be found in Table 12 in Annex A.1.

4.3. Research Design and Methods

The thesis performs a quantitative analysis using secondary data from Afrobarometer. For this purpose, both a cross-sectional research design and a quasi-experimental design are used. To answer the first research question and to gain initial insights for the second research question, the models on the relationships of the variables are estimated using Linear Regression (SWB, EWB) and Logistic Regression (Migration Plans, Migration Intentions).

The first research question is further answered by reducing bias due to potential confounders. A Randomized Controlled Trial (RCT) would be useful for this, as one could additionally infer a causal effect of entrepreneurship on migration plans (Hariton & Locascio, 2018). Because the restriction to observational data does not allow an RCT, it is a methodological challenge to obtain credible causal effects. First, unobserved factors such as risk tolerance or other personality traits may influence both the decision to engage in entrepreneurial activity and to emigrate (Nikolova, 2019). Second, individuals prone to migration may consciously choose to engage in entrepreneurial activity. That raises the problem of self-selection. Third, it can not be ruled out that stronger migration plans, in turn, affect entrepreneurial outcomes, implicating reverse causality (Nikolova, 2019).

Besides randomized experiments, such challenges of endogeneity in observational data can also be tackled by instrumental variable analysis or matching techniques (Nikolova, 2019). Finding adequate instrumental variables that correlate with entrepreneurship but do not influence the outcomes is usually difficult to realize (Nikolova, 2019). Accordingly, the application of a matching technique remains an option. Consequently, the quasi-experimental approach PSM, pioneered by Rosenbaum and Rubin (1983), is implemented with the *MatchIt* package in RStudio (**?**).

To answer the second research question, the theoretical model from Section 3 is tested using SEM analysis. The analysis is carried out with the *Lavaan* package in RStudio (Rosseel, 2012). SEM offers the advantage of estimating all parameters of a mediation model simultaneously. It is considered a powerful technique for mediation analysis and provides a robust language for expressing causal relationships (Gunzler, Morris, & Tu, 2016). Since the SWB and EWB indices were already formed before the SEM estimation, no la-

tent variables exist in the SEM model. Therefore, one can also speak of a path analysis that only uses observed variables (Grapentine, 2000). SEM is a confirmatory rather than exploratory technique. This is because a model is first specified, then estimated and analyzed, and eventually modified (**?**). Since the theoretical model was derived from previous scientific findings, its validity can now be tested by SEM. The model fit is finally assessed using various measures such as the Chi-square, p-value, CFI, RMSEA, NFI, and GFI (Hoyle, 2012). However, due to the sensitivity to sample size, the Chi-squared test plays a subordinate role (Hu & Bentler, 1999).

Overall, the entire empirical analysis, consisting of linear and logistic regression, PSM, and SEM, borrows from the methodology applied by Nikolaev et al. (2020). Like in the thesis, the researchers establish an integrative mediation model for entrepreneurship outcomes including SWB, explore the direct and indirect pathways therein, and rely on cross-sectional data.

5. Empirical Evidence

The chapter first reviews the main descriptive statistics including the most highly correlated variables. The first part of the results section presents the linear regression, logistic regression, and PSM estimates. In the second part, the SEM results are discussed and finally tested for robustness.

5.1. Descriptive Statistics

The summary statistics of the variables included in the empirical analysis can be found in Table 5.

To highlight the most strongly correlated variables in the sample, the relationships of the variables are visualized by a correlation plot in Figure 2. In this plot, the correlation coefficients are colored according to their Pearson's r. It is shown that rural micro-entrepreneurship positively correlates with both EWB and SWB. The relationship of micro-entrepreneurship with migration plans or migration intentions, however, is slightly negative. Migration plans and migration intentions have a strong positive correlation in the young, rural population. Aside from that, there is a relatively strong positive link between EWB and SWB. The correlations between rural micro-entrepreneurs and the remaining variables are relatively weak, though present to some extent, as indicated by the bottom row of the correlation plot. Lastly, the noticeably blue box in the plot signifies that the head of the household is usually male.

The exact values and significance levels are reported in the correlation matrix in Table 6. The matrix suggests, in line with previous literature, that the young, rural micro-entrepreneurs are typically older (Liang et al., 2018) and lower educated (Kew, 2015) compared to the unemployed. They also tend to be male (Fairlie & Robb, 2009), albeit with weaker significance ($p < .01$). Regarding migration plans and intentions, it is notable that stronger migration tendency

Table 4: Control Variables

| | Research on the Relationship of the Control Variable with: | | |
	Entrepreneurship	SWB	Migration Plans/Intentions
Age and Age Squared	Liang, Wang, and Lazear (2018)	Blanchflower and Oswald (2004)	Migali and Scipioni (2019); Ozaltin et al. (2020)
Gender	Fairlie and Robb (2009)	Stevenson and Wolfers (2009)	Ruyssen and Salomone (2018); Migali and Scipioni (2019)
Education	Brown, Farrell, and Harris (2011); Kew (2015)	Nikolaev and Rusakov (2016)	Migali and Scipioni (2019); Ozaltin et al. (2020)
Head of Household	Bhuiyan and Ivlevs (2019)	Bhuiyan and Ivlevs (2019)	De Jong and Steinmetz (2006); Friebel et al. (2013)
Household Size	Bhuiyan and Ivlevs (2019)	Bhuiyan and Ivlevs (2019)	Friebel et al. (2013); Manchin and Orazbayev (2018)
Remittances Dependency	Naudé et al. (2017)	Ivlevs, Nikolova, and Graham (2019)	Van Dalen, Groenewold, and Fokkema (2005)
Migration Experience	Black and Castaldo (2009)	Ivlevs et al. (2019)	Manchin and Orazbayev (2018); Migali and Scipioni (2019)
Perceived Safety	Fatoki and Chindoga (2011)	Wills-Herrera, Orozco, Forero-Pineda, Pardo, and Andonova (2011)	Weiner (1996); Moore and Shellman (2007); Steele (2009)
Country Income Group	Z. J. Acs et al. (2008)	Diener, Diener, and Diener (2009)	Milasi (2020)

Note: The variable *Country Income Group* distinguishes between low-income and middle/high-income countries according to GNI per capita in USD 2016-2018 (WorldBank, 2021).

Table 5: Summary Statistics

Variable	N	Mean	SD	Min	Max
Rural Micro-Entrepreneur	4,430	0.24	0.42	0	1
Economic Well-Being (LPI Reversed)	4,430	-1.34	0.89	-4	0
Subjective Well-Being	4,430	0.00	0.83	-3.52	1.77
Migration Plans	4,430	1.62	0.84	1	4
Migration Intentions	4,430	1.93	1.23	1	4
Age	4,430	26.08	5.09	18	35
Gender (Female=2)	4,430	1.52	0.50	1	2
Education Level	4,430	2.46	0.86	1	4
Head of Household (Yes=1)	4,430	0.31	0.46	0	1
Household Size (Number of Adults)	4,430	3.95	2.75	1	20
Remittances Dependency	4,430	1.32	0.74	1	4
Migration Experience (Yes=2)	4,430	1.25	0.43	1	2
Perceived Personal Safety	4,430	4.25	1.05	2	6
Country Income Group (Middle/High-income=1)	4,430	0.43	0.49	0	1

Note: based on Afrobarometer survey round 7 conducted within 2016-2018.

is expressed by younger, better educated (Giambra & McKenzie, 2021) and male respondents (Milasi, 2020) who consider themselves as head of household (De Jong & Steinmetz, 2006). Those willing to emigrate typically receive regular remittances from relatives or friends abroad (Van Dalen et al., 2005). They tend to have spent a long time outside the country (Migali & Scipioni, 2019) or have family members with migration experience (Manchin & Orazbayev, 2018). In addition, the correlations suggest that the perception of higher personal security is associated with a lower willingness to emigrate. This in turn underlines the role of violence and crime as drivers of migration (Moore & Shellman, 2007;
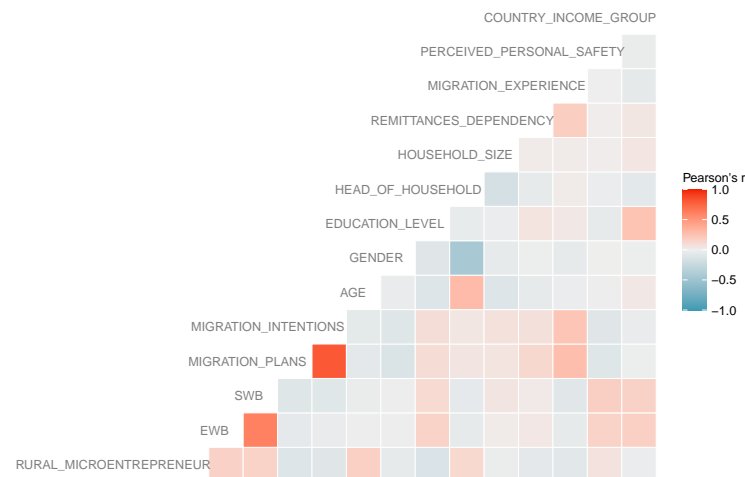
Figure 2: Correlation Plot

Note: own illustration based on Afrobarometer survey round 7 conducted within 2016-2018.

Steele, 2009; Weiner, 1996).

5.2. Main Results

5.2.1. Linear Regression for SWB and EWB

The empirical analysis begins with linear regression in the form of OLS estimations of the relationship between micro-entrepreneurship and SWB (Model 1) and microentrepreneurship and EWB (Model 2). The linear regression models predict the values for the dependent variables EWB and SWB using the independent variable rural microentrepreneur plus the control variables. To run OLS, some assumptions must be met. The first is the linearity assumption, which states there must be a linear dependence between the predictors and the outcome variables (Poole & O'Farrell, 1971). Since the scale level of EWB and SWB is continuous and the residual plot, as well as the QQ-Plot, do not indicate a fitted pattern, linearity is assumed to be given (Cohen, Cohen, West, & Aiken, 2013).

The second assumption to be checked involves a normally distributed error term. As the sample size is sufficiently large, the central limit theorem ensures that the error term approaches a normal distribution (Cohen et al., 2013). Hence, the assumption is met. Subsequently, it is tested for homoscedasticity by checking the residual plot. In other words, the variance of the error term has to remain constant. If heteroscedasticity is present instead, the OLS estimates would still be valid but the standard errors would not (Hayes & Cai, 2007). While the assumption of homoscedasticity is satisfied in Model 1, it might be violated in Model 2, since in the latter the variance seems to follow a pattern. To address this, heteroscedasticity-robust standard errors are estimated using the HC3 method that weights each squared OLS residual (Hayes & Cai, 2007). HC3 is seen by Cribari-Neto, Ferrari, and Oliveira (2005) as the most reliable estimator. The robust standard errors take into account a potential correlation

between the variation in the outcome variable with the explanatory variables (Hayes & Cai, 2007). Finally, it is controlled for multicollinearity. If two or more variables move together, it may result in estimation instability and structural misspecification (Farrar & Glauber, 1967). By obtaining variance inflation factors (VIF) significantly smaller than four, multicollinearity can be ruled out (O'Brien, 2007). The results of the OLS regressions are presented in Table 7.

It is found that the surveyed micro-entrepreneurs have both higher SWB (ß= .344, p < .01) and EWB (ß= .406, p < .01) than unemployed respondents. Both coefficients are statistically significant. The correlations are robust to the inclusion of the control variables age, age squared, gender, education level, head of household, household size, remittances dependency, migration experience, perceived personal safety as well as country income group. The models each explain about 10-11% of the variance in the measurement of SWB and EWB. According to Cohen (1992), these results imply a small to medium effect size. This is not surprising as there might be other factors influencing SWB and EWB that have not been measured. Overall, the obtained positive significant correlations of entrepreneurship with both EWB and SWB provide initial evidence for hypotheses 2 and 3.

5.2.2. Logistic Regression for Migration Plans/Intentions

The next part of the empirical analysis examines the direct relationship between micro-entrepreneurship and migration plans (Model 3) and intentions (Model 4). The variables migration plans and migration intentions are originally ordinal and categorically scaled. This allows for interpretation of the rank of the values, but not the distance between the values. Howbeit, linear regression requires both. For this to happen, continuous dependent variables would be needed (Wu & Leung, 2017). Since this assumption is violated, binary logistic regression is performed for Model 3 and Model 4.

Table 6: Correlation Matrix

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1 Rural Micro-Entrepren.	1												
2 Economic Well-Being	0.169*	1											
3 Subjective Well-Being	0.157*	0.610*	1										
4 Migration Plans	-0.102*	-0.055*	-0.093*	1									
5 Migration Intentions	-0.090*	-0.033	-0.087*	0.803*	1								
6 Age	0.175*	-0.009	-0.025	-0.065*	-0.061*	1							
7 Gender	-0.046	0.004	0.004	-0.128*	-0.096*	-0.030	1						
8 Education Level	-0.123*	0.159*	0.115*	0.098*	0.097*	-0.102*	-0.091*	1					
9 Head of Household	0.123*	-0.047	-0.057*	0.055*	0.048	0.291*	-0.435*	-0.038	1				
10 Household Size	-0.017	0.021	0.052	0.060*	0.076*	-0.103*	-0.045	-0.022	-0.158*	1			
11 Remittances Dependency	-0.064*	0.043	0.029	0.127*	0.082*	-0.047	-0.014	0.058*	-0.043	0.023	1		
12 Migration Experience	-0.077*	-0.046	-0.084*	0.275*	0.237*	-0.023	-0.050	0.040	0.021	0.024	0.187*	1	
13 Perceived Personal Safety	0.068*	0.159*	0.181*	-0.096*	-0.091*	-0.008	0.002	-0.043	-0.020	0.012	0.011	0.000	1
14 Country Income Group	-0.052	0.178*	0.141*	-0.028	-0.044	0.045	-0.009	0.275*	-0.057*	-0.006	0.084*	-0.042	-0.019

Note: Correlation based on 4,430 observations. *Significance at the <.001 level.

For this purpose, the coefficients are calculated with Maximum likelihood (ML) instead of OLS. The main difference to OLS regression is that probabilities for the categories of the dependent variable are estimated (Hosmer & Lemeshow, 2013). To begin with, the two outcome variables are transformed into binary variables. For migration plans, this means that the variable is given the value 1 if migration plans or preparations have already been undertaken, and 0 otherwise. In the case of migration intentions, the variable is given the value 1 if migration considerations are expressed, and 0 otherwise. Before the logistic regression is performed, the main assumptions are checked.

After the transformation, the binarity of the dependent variable is given. Second, the observations are independent of each other. This means that the observations do not come from repeated measurements of the same person or are in any way related to each other (Peng, Lee, & Ingersoll, 2002). Third, there is no multicollinearity among the explanatory variables. Fourth, there is presumably linearity between the independent variables and the log odds (Peng et al., 2002). In addition to the met assumptions, there is a relatively large sample size, which is advantageous for the regression (Peng et al., 2002). Hence, the logistic regression is performed using heteroscedasticityrobust standard errors and keeping the control variables constant.

The results are summarized in Table 7. The estimation of the coefficient for rural micro-entrepreneurship returns a negative significant value (ß = -.346, p < .01). The micro-entrepreneurs thus seem to be less likely to have migration plans than the unemployed.

For further interpretation, the coefficients are exponentiated to receive odds ratios. The odds ratio represents the ratio of the probability of an event (migration plans=1) occurring in the presence of predictor x (rural micro-entrepreneur=1) compared to the probability of the event occurring in the absence of this predictor (rural micro-entrepreneur=0) (Peng et al., 2002). In terms of the study, this means that for the micro-entrepreneur, the odds of having migration plans are 0.708 times as high as the odds for an unemployed person. Similar results are obtained when Model 4 is estimated (ß = -.414, p < .01; odds ratio = .661).

Model fit is assessed by various metrics. The Akaike Information Criterion (AIC) can be used to compare the model fit between different models. As a lower AIC indicates a better model fit, Model 3 has an advantage over Model 4. The Area under the ROC Curve (AUC) indicates the probability that the model correctly classifies pairs of observations. The AUC is 0.71 for Model 3 and 0.69 for Model 4. These values testify to an acceptable discrimination level (Hosmer & Lemeshow, 2013). Finally, McFadden's pseudo R2 as a measure of the overall effect size is 0.09 for Model 3 and 0.08 for Model 4. Again, the values are considered acceptable (McFadden, 1978). The results of the logistic regression give initial support for hypothesis 1, which will be strengthened in the next step.

5.2.3. Propensity Score Matching

Now, PSM is employed to estimate the average marginal effect of microentrepreneurship on migration plans, taking into account confounding by the included covariates (Austin, 2011). PSM is an essential component of Propensity Score Analysis which was pioneered by Rosenbaum and Rubin (1983). It is considered one of the most widely accepted causal inference methods for observational studies (Nikolaev et al., 2020). As the core of PSM, matching attempts to balance the covariates. This means that the distributions of the covariates in the two groups are about the same as they would be in a successful randomized experiment (Austin, 2011). After matching, a treatment effect is estimated. Based on the matched sample, this treatment effect reveals the difference in the value of the outcome (migration plans) between treatment (micro-entrepreneurs) and control group (unemployed) (Nikolaev et al., 2020). The PSM follows the five steps recommended by Olmos and Govindasamy (2015):

Preliminary Analysis.
By estimating the standardized/normalized difference (Imbens & Wooldridge, 2009) and performing an omnibus test (Hansen & Bowers, 2008), it is first checked whether the two groups of micro-entrepreneurs and unemployed are unbalanced. The standardized differences of the covariates age, age squared, education level, and head of household are greater than 0.25, which suggests an imbalance for these specific variables. Plus, the omnibus test yields a statistically significant chi-square, verifying that at least one variable in the model causes an imbalance. Thus, matching seems appropriate (Olmos & Govindasamy, 2015). *Estimation of the Propensity Scores.* The propensity score reveals the probability of belonging to the treatment group of micro-entrepreneurs under the condition of the observed covariances age and age squared, gender, education level, head of household, household size, remittances dependency, migration experience, perceived personal safety, and country income group. The propensity score is estimated with logistic regression. The Back-to-Back Histogram before matching in Figure 7 in Annex B.1 visualizes the imbalance in the number of observations at similar propensity scores between the treatment and control groups (Olmos & Govindasamy, 2015). *Propensity Score Matching.* The nearest neighbor matching method is used to match the propensity scores. Participants from the control group are matched with participants from the treatment group based on the smallest distance to each other. The procedure is carried out 1:1 without replacement. In other words, one participant in the control group is always matched with exactly one participant in the treatment group (Olmos & Govindasamy, 2015). Figure 8 in Annex B.1 reveals that the unmatched individuals in the control group were evenly distributed across the propensity-score-continuum. The matching yields a reasonable balance, which is confirmed visually by the Back-to-Back Histogram after matching in Figure 7 in Annex B.1. As can be

Table 7: Linear and Logistic Regression Results

| Variables | Linear Regression | | Logistic Regression | | | |
| | (1) SWB | (2) EWB | (3) Migration Plans | | (4) Migration Intentions | |
	ß(SE)	ß(SE)	ß(SE)	Odds Ratio	ß(SE)	Odds Ratio
Rural Micro-Entrepreneur	0.344*** (0.028)	0.406*** (0.029)	-0.346*** (0.112)	0.708	-0.414*** (0.081)	0.661
Age	-0.017 (0.026)	0.000 (0.028)	-0.094 (0.094)	0.910	0.042 (0.071)	1.043
Age Squared	0.000 (0.000)	0.000 (0.001)	0.001 (0.002)	1.001	-0.001 (0.001)	0.999
Gender						
Female	0.000 (0.028)	0.014 (0.029)	-0.562*** (0.094)	0.570	-0.321*** (0.074)	0.725
Education Level						
Primary	0.035 (0.039)	0.090** (0.040)	-0.034 (0.135)	0.967	0.271*** (0.103)	1.311
Secondary	0.169*** (0.038)	0.267*** (0.039)	-0.025 (0.129)	0.975	0.533*** (0.100)	1.704
Post-Secondary	0.329*** (0.051)	0.454*** (0.055)	0.375** (0.174)	1.455	0.756*** (0.140)	2.129
Head of Household	-0.081*** (0.031)	-0.075** (0.032)	0.202* (0.105)	1.224	0.209** (0.083)	1.233
Household Size	0.013*** (0.004)	0.006 (0.005)	0.038*** (0.015)	1.039	0.045*** (0.012)	1.046
Remittances Dependency						
A Little Bit	-0.001 (0.042)	0.050 (0.045)	0.260* (0.135)	1.297	0.107 (0.111)	1.113
Somewhat	0.150*** (0.047)	0.095* (0.051)	0.559*** (0.145)	1.748	0.403*** (0.128)	1.496
A Lot	0.025 (0.070)	0.119 (0.075)	0.640*** (0.208)	1.896	0.172 (0.186)	1.187
Migration Experience	-0.155*** (0.029)	-0.080*** (0.030)	1.126*** (0.088)	3.085	1.087*** (0.076)	2.967
Perceived Personal Safety	0.139*** (0.012)	0.130*** (0.013)	-0.174*** (0.040)	0.840	-0.175*** (0.031)	0.840
Country Income Group	0.197*** (0.025)	0.259*** (0.027)	-0.053 (0.090)	0.948	-0.297*** (0.070)	0.743
Observations	4,430	4,430	4,430		4,430	
R² adjusted	0.100	0.109	x		x	
McFadden's Pseudo R2	x	x	0.087		0.080	
AIC	x	x	3,663		5,563	
AUC	x	x	0.708		0.685	

Note: Models (1) and (2) are estimated with OLS regressions. Models (3) and (4) are estimated with Log regressions. Robust standard erros (HC3) are shown in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

seen, the two groups are significantly more similar in terms of their propensity scores than they were before matching. Statistically, the good balance is verified by low standardized differences and a non-significant chi-square (Olmos & Govindasamy, 2015). After matching, all standardized mean differences for the covariates are below 0.1. These results provide sufficient evidence for a significant reduction of the previous selection bias (Austin, 2009). Table 8 summarizes the balancing diagnostics for the matching estimates.

Outcomes Analysis.
To estimate the treatment effect and its standard error, a logistic regression model is fitted with the binary variable migration plans as an outcome and micro-entrepreneurship

with covariates as additive predictors. As matching is done without replacement, clustered standard errors are applied instead of regular robust standard errors to obtain valid estimations (Abadie & Spiess, 2021). Additionally, full matching weights are included in the estimation. Table 9 contains the Average Treatment Effect on the Treated (ATT) for the corresponding migration plans. The ATT has the advantage of being explicitly restricted to the treatment group sample, hence to the micro-entrepreneurs. This helps to examine the effect of the treatment applied (Ho, Imai, King, & Stuart, 2007), namely the pursuit of entrepreneurial activity, and is thereby in line with the aim of the thesis. The estimated effect is -0.253 (p < .10, odds ratio = .777). Compared to the simple logistic regression in Model 3, the effect is weaker and only slightly significant. Nonetheless, the direction of the effect has not changed. Therefore, micro-entrepreneurship is assumed to reduce migration plans. The same applies to migration intentions.

Sensitivity Analysis.
Finally, it is tested how sensitive the results are to hidden bias. For this purpose, the Wilcoxon signed-rank test according to Rosenbaum (2002) is applied. It checks how large the bias must become for the results to change their status from significant to non-significant. The results suggest that even a relatively small unobserved difference in a covariate would change the statistical inference. Thus, it cannot be ruled out that crucial variables are missing. That is why the results of the PSM should be viewed with caution (Olmos & Govindasamy, 2015). Nevertheless, the results of the linear and logistic regression are underpinned and hypothesis 1 is further substantiated.

5.2.4. Main SEM Model

SEM is conducted to test hypotheses 2, 3, and 4. Therefore, the exogenous variable is used to compare the group of micro-entrepreneurs with the unemployed. The SEM examines both the direct path from micro-entrepreneurship to migration plans and the indirect paths via poverty alleviation, consisting of SWB and EWB. This allows us to deduce to what extent SWB and EWB mediate the relationship between micro-entrepreneurship and migration plans. Since migration plans are ordinally scaled, Diagonally Weighted Least Squares (DWLS) are implemented to estimate the parameters (Rosseel, 2012). The estimated model is illustrated in Figure 3.

Overall, the results of the SEM analysis bolster hypotheses 2 and 4, but not hypothesis 3. Table 10 summarizes the direct, indirect, and total effects of microentrepreneurship on migration plans. A direct, negative, and statistically significant effect of micro-entrepreneurship on migration plans is found (ß= -.282, p < .01). Besides, a positive, also statistically significant effect of micro-entrepreneurship on SWB (ß= .110, p < .01) is identified, which is then associated with weaker migration plans (ß = -.124, p < .01).

The combined indirect effect of micro-entrepreneurship via SWB on migration plans is therefore a reduction of -0.014 (p < .01), supporting hypothesis 2. On the other hand, there is an effect running from micro-entrepreneurship to EWB (ß = .352, p < .01). EWB in turn is positively associated with migration plans, though not statistically significant (ß = .028, p > 0.1). The combined indirect effect via EWB on migration plans is 0.010 (p > 0.1) and thus cannot provide evidence for hypothesis 3. In the end, one finds a positive direct effect of EWB on SWB (ß = .566, p < .01), which strengthens hypothesis 4. This positive effect of EWB on SWB indirectly leads to a reduction of migration plans by -0.025 (p < .01). Hence, the total indirect effect on migration plans across the paths of SWB and EWB amounts to -0.029 (p < .01). The total effect on migration plans ultimately reflects the sum of the direct (-.282) and indirect effects (-.029) and is -0.310 (p < .01).

The results in Table 10 indicate that the direct effect of micro-entrepreneurship on migration plans accounts for about 91% of the total effect[1]. The total indirect effect via poverty alleviation makes up the remaining 9% of the total effect[2]. It follows that the negative effect of micro-entrepreneurship on migration plans works mainly direct and is possibly influenced by omitted variables. Moreover, a slight mediating role of SWB in the relationship between entrepreneurship and migration plans can be observed, which is also called partial mediation. However, no mediating role of EWB is found in this relation. Instead, EWB seems to influence migration plans only indirectly through the path of SWB.

Overall, the main model provides evidence for hypotheses 2 and 4, but not for hypothesis 3. Hypothesis 1 is enhanced by the positive path of micro-entrepreneurship on migration plans. Since the SEM model has no degrees of freedom, the model is saturated or just-identified (Hoyle, 2012). That implies that the hypotheses about the paths in the model can be tested, but not the adequacy of the model itself (?).

5.2.5. Adapted SEM Model

Because the path from EWB to migration plans is not statistically significant and thereby hypothesis 3 is not supported, a modified model is proposed. The nonsignificant path via EWB is deleted and the analysis is performed again. Figure 4 presents the adapted model without the direct association of EWB with migration plans.

The only way EWB can now be associated with migration plans is indirectly through SWB. Table 11 shows the indirect and direct effects of the SEM analysis based on the adapted model.

The differences to the main model are only marginal and all effects are significant (p < 0.01). Since the model now has one degree of freedom, it becomes an overidentified model (Hoyle, 2012). At this point, the adequacy of the model

---

[1]derived by dividing direct effect (-0.282) by total effect (-0.310).
[2]derived by dividing total indirect effect (-0.029) by total effect (-0.310).

Table 8: Balance Diagnostics for Matching Estimates

	Sample	Treated	Control	SMD	% Bias-Reduction
Age	Unmatched	27.68	25.59	0.42	
	Matched	27.68	27.52	0.03	92.1
Age Squared	Unmatched	791.15	679.95	0.41	
	Matched	791.15	781.31	0.04	91.1
Gender	Unmatched	1.48	1.53	-0.11	
	Matched	1.48	1.47	0.00	96.4
Education Level	Unmatched	2.27	2.52	-0.31	
	Matched	2.27	2.27	0.00	100.0
Head of Household	Unmatched	0.42	0.28	0.27	
	Matched	0.42	0.42	-0.01	97.1
Household Size	Unmatched	3.87	3.98	-0.04	
	Matched	3.87	3.85	0.01	83.2
Remittances Dependency	Unmatched	1.24	1.35	-0.17	
	Matched	1.24	1.21	0.04	76.0
Migration Experience	Unmatched	1.19	1.27	-0.20	
	Matched	1.19	1.19	0.00	98.8
Perceived Personal Safety	Unmatched	4.37	4.21	0.17	
	Matched	4.37	4.37	0.00	98.9
Country Income Group	Unmatched	0.38	0.44	-0.12	
	Matched	0.38	0.35	0.06	50.6

Note: SMD=Standardized Mean Difference. Treated and control columns report the means for each group.

Table 9: PSM Estimates

Dependent Variable	Matching Method	ATT (SE)	Odds Ratio
Migration Plans	1:1 nearest neighbor matching without replacement	-0.253* (0.137)	0.777
Migration Intentions	1:1 nearest neighbor matching without replacement	-0.400*** (0.097)	0.670

Note: Estimates report the average treatment effect on the treated (ATT) with cluster-robust standard errors including all covariates. Sample size: N(Matched)=2,088; N(Treatment)=1,044; N(Control)=1,044. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

can be tested. The global goodness-of-fit measures are Chi-square=0.618, p-value=0.432, CFI=1.000, RMSEA=0.000, NFI=0.999, GFI=1.000. Being aware of the positive influence of having a large sample, these measures suggest a very good fit (Sivo, Fan, Witta, & Willse, 2006). The relevance of the model as an alternative to the main model is discussed in Section 6.

5.3. Robustness Checks for SEM

A number of checks are carried out to verify the robustness of the SEM results. The tables and figures are gathered in Annex B.2.

Replication with Bootstrap Standard Errors.
If non-normal data are present, which cannot be ruled out in this setting, ordinary non-parametric bootstrapping is useful (Hoyle, 2012; Rosseel, 2012). In this case, random samples, called replications, are generated from the available sample to obtain corresponding standard errors. Based on those standard errors, bootstrap-based confidence intervals are created (Hoyle, 2012). Consequently, the proposed main model is computed using 5,000 replications (Hayes & Preacher, 2013) to receive the new standard errors and confidence intervals. The effects are listed in Table 13 in Annex B.2.1 and underpin the previous findings.

Including Control Variables.
Next, the control variables are added to the main model. The SEM is structured in such a way that the control variables have a connection both to the exogenous (micro-
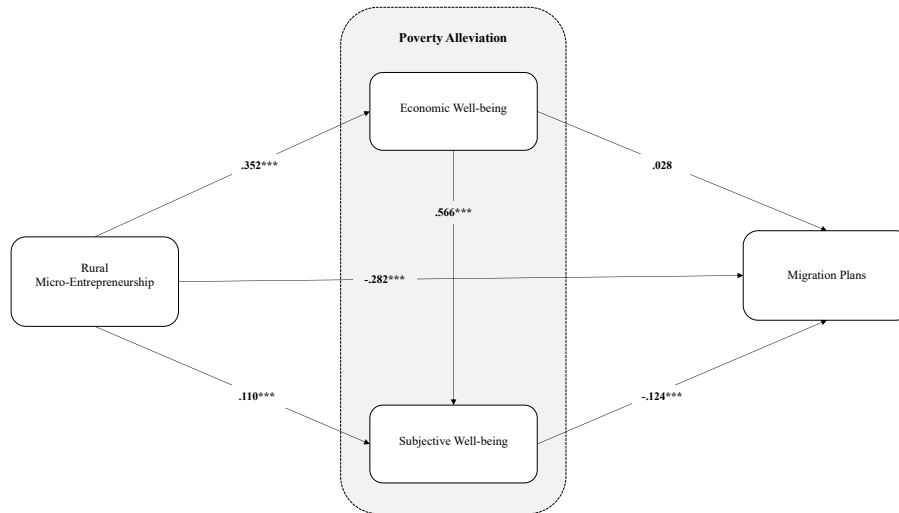
Figure 3: Main SEM Model

Note: own illustration based on Nikolaev et al. (2020).
Estimator: DWLS. N=4,430. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 10: Direct and Indirect Effects of Micro-Entrepreneurship on Migration Plans

						95% Conf. Int.	
	ß	std.est	SE	z	p>z	LLCI	ULCI
Direct effects							
ME → MP	-0.282***	-0.119	0.043	-6.529	.000	-0.366	-0.197
ME → SWB	0.110***	0.056	0.026	4.167	.000	0.058	0.162
ME → EWB	0.352***	0.169	0.033	10.553	.000	0.286	0.417
SWB → MP	-0.124***	-0.103	0.025	-4.898	.000	-0.174	-0.075
EWB → MP	0.028	0.025	0.024	1.179	.238	-0.019	0.075
EWB → SWB	0.566***	0.601	0.012	47.318	.000	0.543	0.590
Indirect effects							
ME → SWB → MP	-0.014***	-0.006	0.004	-3.102	.002	-0.022	-0.005
ME → EWB → MP	0.010	0.004	0.008	1.174	.240	-0.007	0.026
ME → EWB → SWB → MP	-0.025***	-0.01	0.006	-4.479	.000	-0.036	-0.014
Total indirect effect							
ME → MP	-0.029***	-0.012	0.008	-3.509	.000	-0.045	-0.013
Total effect (dir.+indir.)							
ME → MP	-0.310***	-0.131	0.042	-7.306	.000	-0.394	-0.227

Note: ME=Micro-Entrepreneur. MP=Migration Plans. std.est.=standardized estimate. LLCI=Lower Limit Confidence Interval. ULCI=Upper Limit Confidence Interval. Results are based on SEM model from Figure 3. Estimation Method: DWLS. ***p < 0.01, **p < 0.05, *p < 0.1.

entrepreneur) and endogenous variables (SWB, EWB, migration plans). The variable country income group is omitted because it is controlled separately in another check. The results in Table 14 in Annex B.2.2 confirm the findings of the main model. The significance and direction of the effects remain the same. It is noteworthy that the direct effects of micro-entrepreneurship on SWB and EWB are larger than before. While the other direct and indirect effects, including the total effect on migration plans, are smaller. However, due to their small magnitude, the differences to the main model seem negligible.
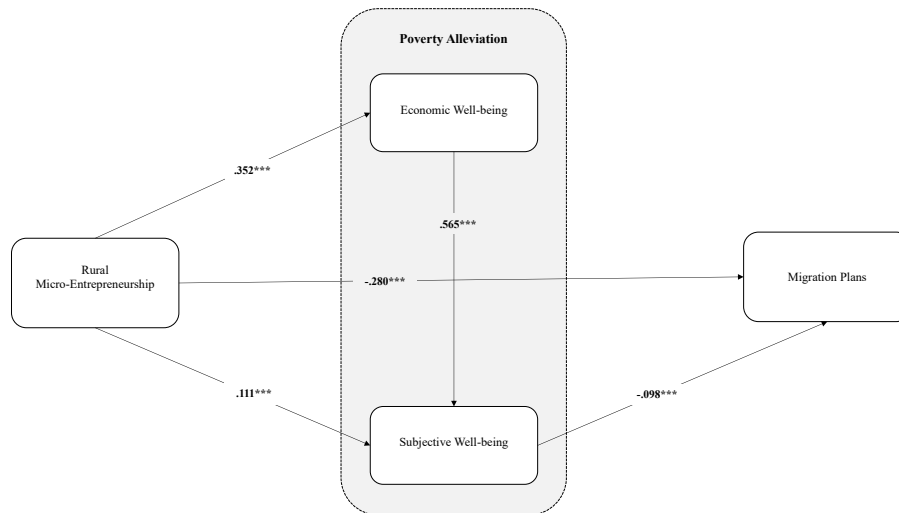
Figure 4: Adapted SEM Model

Note: own illustration based on Nikolaev et al. (2020).
Estimator: DWLS. N=4,430. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 11: Adapted Model: Direct and Indirect Effects of Micro-Entrepreneurship

	ß	std.est	SE	z	p>z	95% Conf. Int. LLCI	ULCI
Direct effects							
ME → MP	-0.280***	-0.118	0.043	-6.500	.000	-0.364	-0.196
ME → SWB	0.110***	0.056	0.026	4.179	.000	0.059	0.163
ME → EWB	0.352***	0.169	0.033	10.553	.000	0.286	0.417
SWB → MP	-0.098***	-0.081	0.022	-4.470	.000	-0.141	-0.055
EWB → SWB	0.565***	0.599	0.012	46.898	.000	0.542	0.589
Indirect effects							
ME → SWB → MP	-0.011***	-0.005	0.004	-3.069	.002	-0.018	-0.004
ME → EWB → SWB → MP	-0.020***	-0.008	0.005	-4.106	.000	-0.029	-0.010
Total indirect effect							
ME → MP	-0.030***	-0.013	0.007	-4.074	.000	-0.045	-0.016
Total effect (dir.+indir.)							
ME → MP	-0.310***	-0.131	0.042	-7.306	.000	-0.394	-0.227

Note: ME=Micro-Entrepreneur. MP=Migration Plans. std.est.=standardized estimate. Results are based on SEM model from Figure 4. Estimation Method: DWLS. ***p < 0.01, **p < 0.05, *p < 0.1.

Using Migration Intentions as Dependent Variable.
The next check aims to control for a difference in SEM results when the dependent variable is migration intentions instead of migration plans. As already explained in Section 4, the two outcome variables differ concerning the concreteness and prevalence of migration aspirations. Such differences could influence the measured correlations and should therefore be checked. Annex B.2.3 shows the model in Figure 9 and the effects in Table 15. The results are similar to the

ones of the main model, even though a small positive effect is found for the direct path from EWB to migration intentions (ß = .057, p < .05). This, in turn, may account for the significant indirect effect of entrepreneurship on migration intentions via EWB (ß = .020, p < .05). In other words, with higher EWB, the migration intentions of micro-entrepreneurs might even be stronger. This is in contrast to both hypothesis 3, according to which one would expect a negative mediating role, and the main model, which finds no mediating role.

Apart from this, the other hypotheses are strengthened.

Subset to Low-income Countries & Middle/High-income Countries.

To control for country-specific differences such as income grouping, the main model is estimated separately for low-income countries and middle/high-income countries. After subsetting, the two country income groups each account for about half of the sample size. The findings are listed in Table 16 and 17 in Annex B.2.4. It can be seen that when limiting to low-income countries, the effects differ only marginally from the main model and the same hypotheses are supported. Notably, the direct effect of entrepreneurship on migration plans (ß = -.452, p < .01) and the total effect (ß = -.488, p < .01) are even stronger for the low-income subset compared to main model. However, the effect of entrepreneurship on EWB is slightly smaller. After limiting to middle/high-income countries, the deviations from the main model are stronger. The negative, direct effect of entrepreneurship on migration plans is significantly weaker and no longer significant (ß = -.051, p = .468). As such, hypothesis 1 is not confirmed for middle/high-income countries. The indirect effects likewise shrink and are no longer significant. Hypotheses 2 and 3 would consequently not apply either. Solely hypothesis 4 is still valid.

Distinction between Non-farm and Agricultural Entrepreneurs.
The entrepreneurs in the sample are predominantly active in the agriculture sector, followed by the trade, retail, and craft sectors (Kew, 2015; Okeke & Nwankwo, 2017). However, the main model does not distinguish between agricultural and nonfarm entrepreneurs. It may be useful to consider them explicitly, such as Nagler and Naudé (2014, 2017) who focused on outcomes of non-farm entrepreneurs. For simplification, the sector "Agriculture/Farming/Fishing/Forestry" (Afrobarometer, 2020, p.61) within the sample is referred to as agricultural sector. Removing this sector from the sample leads to a significant reduction in the number of micro-entrepreneurs from N=1,044 to N=469. Anyways, the share of non-farm entrepreneurs remains substantial (Nagler & Naudé, 2017). The SEM is calculated again and the results are summarized in Table 18 in Annex B.2.5. Overall, limiting to non-farm entrepreneurs has a strong impact on the results. It is noted that the negative, direct effect of entrepreneurship on migration plans is no longer significant (ß = -0.050, p = .401). Besides, the direct effect of entrepreneurship on SWB shrinks and loses its significance. Hypotheses 1, 2, and 3 are no longer sustained in this way. There is only a marginally significant effect measured across the path from EWB to SWB on migration plans (ß = -0.026, p < .01). This might be the reason for the small, indirect effect on migration plans (ß = -0.032, p < .01). Hypothesis 4 still holds. As the last step, the sample is limited to the agricultural sector. The number of micro-entrepreneurs is now N=575. The effects in Table 19 reconfirm the results of the main model, strengthening hypotheses 1, 2, and 4. Compared to non-agricultural entrepreneurs, migration plans of agricultural entrepreneurs

are reduced more, while EWB and SWB are increased more. Remarkably, the direct (ß = -0.481, p < .01) and total effects (ß = -0.527, p < .01) on migration plans are even stronger than in the main model.

6. Discussion

Drawing on the idea of transformative entrepreneuring in entrenched poverty zones (Tobias et al., 2013), the thesis presents a mediation model in which subjective and economic well-being explain partially the relationship between engagement in micro-entrepreneurship and migration plans. The model is tested using a sample of 10,106 young adults from rural areas of 31 sub-Saharan countries by distinguishing between micro-entrepreneurs and the unemployed. It can be noted that the group of micro-entrepreneurs expresses migration plans to a lesser extent than the group of unemployed. In addition, a positive relationship is detected between engagement in micro-entrepreneurship and both subjective and economic well-being. Moreover, EWB seems to increase SWB, at least there is sufficient evidence for a positive association. While SWB is negatively related to migration plans, no significant results are found for the relationship between EWB and migration plans. Different robustness checks largely confirm the results. This applies, in particular, to the alternative examination of migration intentions and the restriction to low-income countries and agricultural businesses. However, for the subgroups of middle/highincome countries and non-farm entrepreneurs, micro-entrepreneurship does not seem to contribute to reducing migration plans.

6.1. Interpretation of Results

The first part of the empirical analysis examined the direct relationships between micro-entrepreneurship and its potential outcomes EWB, SWB, and migration plans. Linear regression was applied to compare the EWB and SWB for microentrepreneurs and the unemployed. Considering various control variables, microentrepreneurship was found to be positively related to SWB. The same was identified for the relation of micro-entrepreneurship and EWB. In this way, young microentrepreneurs seem to benefit from their job not only economically, but also in terms of their evaluative well-being, comparable to the Rwandan coffee farmers in the article of Tobias et al. (2013).

Logistic regression was then conducted to contrast the migration plans of the two groups. The estimates revealed that micro-entrepreneurs have a lower level of migration plans than the unemployed. Utilizing PSM, tentative conclusions were made about the effect of entrepreneurship on migration plans. The bias-reduced results suggested, albeit with low significance, the negative effect that entrepreneurial activity seems to have on migration plans. On the one hand, this complements Milasi (2020) and Giambra and McKenzie (2021), whose results point to a negative relationship between self-employment and migration aspirations. On the other hand, the weak significance might reveal confounding

through other variables. This first part of the analysis was additionally implemented with migration intentions as the dependent variable. Only marginal differences were noted, strengthening the previous results.

In the second part of the analysis, SEM was used to test the integrative model of micro-entrepreneurship, EWB, SWB, and migration plans. It was found that SWB partially mediates the negative relationship between entrepreneurship and migration plans. This is in line with Chindarkar (2014) and Cai et al. (2014), who claim that higher SWB associates with lower migration aspirations. In contrast, the path via EWB was found to be not significant. Specifically, EWB would affect migration plans only indirectly via the path of SWB. For this reason, an alternative model was proposed that does not include a link from EWB to migration plans.

Having found similar results and a convincing model fit, the question is whether the adapted model could even be superior to the main model. For one thing, it was affirmed that SWB can explain variations in international migration preferences better than established objective measures like income (Cai et al., 2014). For another, the empirical evidence such as by Kandemir (2012), Salameh (2019) or Ozaltin et al. (2020), which suggested a correlation between personal wealth and migration aspirations, is too extensive to fully abandon it in the integrative model. Plus, the stronger mediating role of SWB might arise because it captures both evaluative factors (i.e. life satisfaction) and economic factors (i.e. living conditions) (Cai et al., 2014). EWB instead only covers economic factors. For these reasons, the adapted model does not seem to be the superior one, and the link between EWB and migration plans should not be ignored. Anyway, the unclear direction of this relationship highlights the need for further investigation. The use of migration intentions as a dependent variable suggested that a positive effect of EWB on migration aspirations cannot be ruled out either, especially in less developed countries (Adhikari & Gentilini, 2018; Dao et al., 2018). In such cases, entrepreneurship could serve to overcome the budget constraints that young adults face when they emigrate (Giambra & McKenzie, 2021). More broadly, the unclear link between EWB and migration considerations may result from the non-linear relationship attributed to the two variables (Friebel et al., 2013).

The ensuing differentiation between country income groups signaled that countryspecific differences appear to have a considerable impact on the outcomes of rural entrepreneurship (Nagler & Naudé, 2014). Restricting the sample to lowincome countries initially revealed a weaker positive relationship between microentrepreneurs and EWB. It is therefore assumed that in poorer countries the income of entrepreneurs differs less strongly from unemployed people than in more developed countries. Besides, migration plans for entrepreneurs in low-income countries were even lower than in the main model. In contrast, for middle/high-income countries the effect of entrepreneurship on migration plans was insignificant. This is consistent with Milasi (2020) who argued that the negative effect of young people's self-employment on migration intentions is stronger in develop-

ing countries than in emerging countries. It also underscores the argument of Clemens (2020) that emigration increases as low-income countries evolve into middle-income countries, and decreases again as they continue to evolve into high-income countries.

The non-farm sector in SSA's rural areas has grown steadily in recent decades, attracting mostly younger people, while the agricultural sector appears more and more unpopular among youth (Gough & Langevang, 2016). Sohns and Diez (2018) and Giambra and McKenzie (2021) suggested that the causal effects of farm work on migration may be different from those of non-farm self-employment. Furthermore, it was alleged that agricultural activities can only be classified as entrepreneurial to a limited extent since one is not involved in the creation of new businesses and they are often carried out on a subsistence basis (Gough & Langevang, 2016).

From these points of view, the question arose to what extent the non-farm and agricultural businesses would differ in their outcomes. Restricting to the non-farm sector lead to effects that were significantly weaker and non-significant compared to the main model. Restricting to the agricultural sector, on the other hand, confirmed the results of the main model with even lower migration plans and higher EWB and SWB for the group of entrepreneurs. This would mean that the agricultural sector would be preferable to the non-farm sector in terms of the transformative potential proposed in the thesis. It cannot be denied that agriculture plays an essential role as a source of livelihood for rural youth (AFI, 2014; Gough & Langevang, 2016; Kolawole & Ajila, 2015). Nevertheless, the data in this work are not sufficient to fully justify why agricultural entrepreneurship yields better returns than non-farm entrepreneurship. The lower migration plans for the agriculture sector will even raise the question of increased self-selection bias. For example, the portability of the business for a farmer in the case of emigration might be significantly lower than for someone who only sells clothes or cooks food (Giambra & McKenzie, 2021). This could already explain why farmers want to emigrate less frequently.

In the main model, a separation between the two sectors was intentionally not undertaken. Agriculture provides a resilient livelihood strategy when combined with other economic activities. Therefore, many young people engage in farming as a supplementary activity (Gough & Langevang, 2016). Often, in the agricultural off-season, the non-farm business is targeted and vice versa. This underscores the high degree of connectedness between the rural non-farm sector and agriculture (Gough & Langevang, 2016). Further, both non-farm and agriculture activities can be considered entrepreneurial according to the thesis if they serve to generate an income (Gough & Langevang, 2016). This blending makes it difficult to distinguish between non-farm and agricultural entrepreneurship, especially if only limited data is available.

Overall, only the positive effect from EWB to SWB remained significant for all robustness checks. Those results are in line with Tobias et al. (2013), thereby implying that

micro-entrepreneurs' increased personal wealth might improve their satisfaction with life. It should be noted that the empirical methodology is only sufficient to confirm a correlation, but not causality.

6.2. Implications for Theory and Practice

6.2.1. Implications for Theory.

The theoretical contribution of this work lies mainly in studying the outcomes of micro-entrepreneurship for young people in rural SSA. It describes a context that is particularly characterized by socio-economic constraints and has been little researched to date. Hereby, the thesis responds to the demand for more context-specific entrepreneurship research (Wiklund et al., 2019) in developing countries (Bruton et al., 2008) and specifically the African continent (George et al., 2016).

By investigating Afrobarometer data, the underrepresentation of quantitative studies in entrepreneurship research is countered (Bruton et al., 2013; Tobias et al., 2013). With the consideration of a total of 31 countries, the thesis is thereby able to map SSA comprehensively. In addition to looking at SSA as an entire geographic region, the sample is divided into low-income countries and middle/high-income countries by GNI per Capita. Withal, the quantitative analysis is extended by a distinction between agricultural and non-agricultural sectors and migration intentions rather than plans.

The proposed model enables a holistic view of the impact of microentrepreneurship. This view is not limited to personal wealth but also takes into account non-economic factors such as SWB and migration plans. In this sense, one embraces the idea of freeing entrepreneurship research from an exclusively economic focus or heroic achievements (Johannisson, 2011; Rindova et al., 2009; Tobias et al., 2013) and tackles an overly narrow definition of poverty.

As a core of the work, the thesis delivers new insights into the so-far poorly studied relationship between entrepreneurship and migration plans under conditions of poverty. First, the direct relationship between the two variables is to be deciphered. Then one empirically demonstrates the extent to which SWB and EWB mediate this relationship. Understanding such linkages can be important for both practice and theory (Diener, 2012). As requested by Chindarkar (2014), economic, non-economic factors and migration plans are linked to better understand the influencing factors of migration plans.

To date, post-migration studies help to better define the potential and the challenges for refugees who engage in entrepreneurship after leaving the country (Kachkar, 2019). One key limitation exists, though. The studies only screen entrepreneurs who have factually emigrated. Those who decided against emigration despite migration plans are left out. With this selection bias, it is difficult to fully grasp the impact of entrepreneurship on migration plans. Asking people before they might emigrate, and therefore ex-ante, reduces selection bias and also allows understanding people's decision-making phase (Ozaltin et al., 2020). In this way, those who

have considered emigration, or perhaps even had a strong tendency to do so but ultimately decided against it, will also be heard. By extension, focusing on migration plans rather than actual migration flows helps reduce biases that persist due to reliance on data from receiving countries (Ozaltin et al., 2020).

The thesis responds to the call for more research on the role of entrepreneurship in poverty reduction, for instance by Sutter et al. (2019) and Vermeire and Bruton (2016). Entrepreneurship can be a long-term solution to poverty if science understands how to help people in poverty start their business (Bruton et al., 2013). So, the view of research should move away from seeing people in poverty as a potential market for goods and instead focus on how people can overcome poverty through their efforts. To this end, the impact from the perspective of ordinary entrepreneurs is examined (Bruton et al., 2013; Tobias et al., 2013). Like the specialty coffee farmers in rural Rwanda (Tobias et al., 2013), the young rural entrepreneurs of SSA form the empirical unit of study. As such, one moves to the micro-level rather than analyzing entrepreneurship at the mesoor macro-level (Johnson & Schaltegger, 2020). Once poverty is alleviated at the micro-level, development at the macro-level is spurred in the next step to ultimately achieve sustainable development (Kolawole & Ajila, 2015).

By and large, entrepreneurship in developing countries is often disparaged as having little impact, when in fact it can lead to growth (Welter et al., 2017). Therefore, failing to understand the potential of various entrepreneurial opportunities can be detrimental to the success of poverty reduction efforts (Alvarez & Barney, 2014). This paper shows that micro-entrepreneurship in particular, which is widespread in just such developing countries, can unleash transformative spirit (Tobias et al., 2013) by reducing migration plans. With the awareness that entrepreneurship cannot be equated with entrepreneuring, the thesis can serve as a support to the theorybuilding of Tobias et al. (2013). At last, after the detailed exploration of the SDG of poverty reduction, research may have been inspired to explore the achievement of other UN goals in SSA, for example, Goal 4: Quality Education, Goal 5: Gender Equality, or Goal 10: Reduce Inequalities (UN, 2020).

6.2.2. Implications for Practice.

Wage and salary employment is limited in most developing countries (Cho, 2015). Usually, the only alternative to unemployment is self-employment. The thesis demonstrates that entrepreneurship as a form of selfemployment can help to overcome poverty by having a positive impact on EWB and SWB. This is in response to Sutter et al. (2019)'s advice that intensive research can provide critical insights for entrepreneurs, policymakers, and organizations on the extent to which poverty can be overcome through entrepreneurship. It also helps to assess whether the entrepreneurial potential of young people in rural areas should be nurtured and whether it is worthwhile to improve their financial access (Cho, 2015). Especially when it comes to achieving the UN

goals of poverty eradication and sustainable economic development by 2030, new insights as those given in this work can be valuable for the various stakeholders.

The results are also interesting in light of the political and social relevance of migration, to name for example the 2015 refugee crisis, drownings in the Mediterranean, and right-wing populism (Sekeris & Vasilakis, 2016). The effects of growing migration are felt by both the countries of origin and the countries of destination. Truly, beyond positive influences such as filling labor shortages in destination countries (Rapoza, 2017) and providing financial support to the family in the home country through remittances (FAO, IFAD, IOM, & WFP, 2018), increased migration brings several challenges. For instance, countries of origin lose the capable workforce, known as brain-drain (Capuano & Marfouk, 2013). Meanwhile, in destination countries, immigration is perceived by populist movements as a threat to domestic jobs, security, and culture (Brubaker, 2017). The results of the thesis suggest that rural entrepreneurship opens up opportunities to improve living conditions in the home countries. This could potentially reverse the brain-drain (Sriram & Mersha, 2006) and, at the same time, counteract populism in the recipient countries by reducing refugee flows.

The thesis can give policymakers a means of discouraging people from emigrating by offering them prospects for the future in their country of origin. First, focusing on migration plans instead of actual migration can help make laws to manage, restrict, or prevent migration flows more effective. This is because one then identifies factors that turn desires into actions before a migration occurs (Ozaltin et al., 2020). Second, positive economic development in countries causing emigration (Salameh, 2019) rather than uncontrolled monetary payments from abroad may ultimately be an effective means of preventing migration. Of course, it is known that migration plans and intentions increase with higher levels of education among youth (Milasi, 2020). 46% of those who intend to migrate have at least a university degree and 63% have secondary education (Ozaltin et al., 2020). While raising educational attainment is important, it is not the only policy challenge. Job opportunities, a positive economic outlook, and upward mobility - all signs of positive economic development - are also crucial factors in reducing migration and retaining the young, willing, and partly well-educated population in the country (Milasi, 2020). Promoting entrepreneurship through technological support, access to capital and training in job-specific and soft skills can increase the likelihood of success for young people (Kluve et al., 2017) and thus corresponds very well to such crucial factors. Ultimately, the results of the paper highlight that such entrepreneurship support could be valuable not only for well-educated young people in cities (Milasi, 2020) but also for less well-educated young people in rural areas.

On a final note, the assumption behind policy efforts to promote self-employment to curb emigration is that there is a causal effect through which engaging in self-employment lowers individuals' migration intentions (Giambra & McKenzie, 2021; Kluve et al., 2017). While the thesis cannot pro-

vide concrete causal effects for policy, it confirms that entrepreneurial activities are negatively related to migration plans, especially in low-income countries and for agricultural entrepreneurs. Furthermore, both income and life satisfaction tend to be higher for entrepreneurs than for the unemployed. Even though causal derivations were not feasible, those results could be explained with existing logical considerations. The most obvious assumption is that a new job opportunity increases the income one earns at home and thus increases the opportunity cost of migration (Giambra & McKenzie, 2021). Moreover, the lower migration considerations could be attributed to the lock-in effect that entrepreneurship entails. This is demonstrated by the fact that the investments made in the company cannot be fully reversed and would lead to a loss of the invested capital in the event of a migration (Koelle, 2019). Last, the portability of acquired skills might be stronger through self-employment than through wage employment, which again would drive migration (Giambra & McKenzie, 2021).

6.3. Limitations and Future Research

In the course of the study, several limitations emerged which point to important directions for future research.

First, to keep the dynamics of entrepreneurship (Langevang et al., 2012; Welter et al., 2017) and avoid limiting its theoretical development (Dencker et al., 2021), it was not distinguished between entrepreneurship by necessity and entrepreneurship by opportunity. Thus, it is not possible to determine the entrepreneur's motivation for running a business. Instead, the data only allow defining the micro-entrepreneur as a self-employed person with a job that generates an income. At this point, the question arises to what extent the measured effects on SWB, EWB, and migration plans would differ between different entrepreneur types. According to Binder and Coad (2013), a lack of distinction between entrepreneurial motivations may affect the robustness of the results. Then again, a distinct consideration could help to better specify and conceptualize the impact of entrepreneurship in developing countries (Chliova et al., 2015). Nagler and Naudé (2017), for example, find that necessity businesses in SSA are less productive than opportunity businesses. Such findings are of high importance for rural development because more productive businesses are better able to grow, survive, and create new jobs, which is fundamental to rural development (Nagler & Naudé, 2017). In future studies, one could build on the suggested model and control for entrepreneurial motivation. Depending on how the results differ, new insights for theory and practice may be gained. If only opportunity entrepreneurs were shown to have higher subjective and economic well-being and lower migration plans, they would likely become the focus of policymakers.

Furthermore, the empirical analysis is performed with cross-sectional data. Such do not allow to study the proposed relationships in their full dynamic (Nikolaev et al., 2020). For example, a relationship between entrepreneurship and well-being measured at a particular point in time does not

necessarily exist in the medium and long term (Abreu et al., 2019). This dynamic might also occur in the context of entrepreneurship and migration plans. In the same way, the entrepreneurs in the sample are likely to find themselves at different stages of their business lives. Indeed, entrepreneurs often have to go through various business paths to increase their income (Vermeire & Bruton, 2016). It is therefore possible that the proposed relationships vary at different stages of the entrepreneurial journey. Additionally, there is evidence that entrepreneurs do not always seek revenue for themselves, but also their future family generation (Shepherd et al., 2021). To fully understand the benefits of entrepreneurship for individuals, a longer time horizon is needed. Future research could adopt the model presented in the thesis and re-examine the effects of entrepreneurship using longitudinal data. This would complement the growing longitudinal studies of the relationship between entrepreneurship and well-being (Abreu et al., 2019; Binder & Coad, 2013, 2016; Shir, 2015).

The restriction to cross-sectional data ultimately makes it challenging to measure causal effects. The PSM method is insufficient for this purpose because it only allows adjustment for observed confounders, not for unobserved confounders (Rubin, 1997). A further limitation is reflected in the fact that well-being and migration plans are only compared between the treatment group of entrepreneurs and the control group of unemployed. However, the data do not reveal what the employment status was before starting with entrepreneurship. Direct effects of a transition from unemployment to entrepreneurship cannot be examined in this way (Binder & Coad, 2013). With data on this, the effect of entrepreneurship could be isolated to a greater extent and causality assumptions would be easier to achieve.

In addition to the complicated measurability of the variables, which was discussed in Section 4, the restriction to Afrobarometer data leads to further empirical challenges. Since a study of individual countries would drastically reduce the sample size, robust results would be difficult to obtain (Kline, 2015). Thereby, the empirical studies are not carried out for individual countries, but for SSA as a whole. This lack of distinction does not allow accounting for the country-specific heterogeneity of entrepreneurship in SSA (Nagler & Naudé, 2014). Moreover, other control variables may have been forgotten, that may play a role in the studied relations. For example, it would make sense to control for the effects of climate change, such as intensification of droughts and floods, on microentrepreneurs' migration plans (Naudé, 2008). Additionally, one could control for violent conflicts before and within the survey period, such as the Boko Haram insurgency in Nigeria (Mukhtar, Rose, Choy, & Bibi-Farouk, 2018) or the Northern Mali conflict (Hoogeveen, Rossi, & Sansone, 2019). It makes sense because the influence of coercion causes people to emigrate regardless of their abilities and desires (Ozaltin et al., 2020). On these grounds, future research should collect appropriate primary data or rely on other databases to accurately obtain the desired information. This would enable to specifically answer the research

questions, acquire a larger sample, and ultimately increase statistical power.

A final limitation is the risk of self-selection bias distorting the results. The charge is regularly made that the effects of entrepreneurship reflect people's self-selection into entrepreneurship rather than the actual consequences of entrepreneurial activity (Bhuiyan & Ivlevs, 2019). Consistent with Nikolaev et al. (2020) and Tobias et al. (2013), logical and empirical evidence was provided that entrepreneurship could enable young adults to enhance their well-being. However, especially in terms of the impact on migration plans, entrepreneurs may be more inclined to migrate than employees or the unemployed because of their personality (Giambra & McKenzie, 2021). Accordingly, the entrepreneurial activity would not be the cause of weaker migration plans, but rather, personal characteristics that lead to becoming an entrepreneur in the first place. Such limitations of values-driven job fit (Nikolaev et al., 2020) were attempted to be mitigated through PSM. Focusing on migration considerations rather than actual migration further lessened the self-selection (Ozaltin et al., 2020). Still, this might not be sufficient to completely rule out selection bias.

7. Conclusion

This paper develops a comprehensive model for better understanding the impact of micro-entrepreneurship on young adults' migration plans in rural sub-Saharan Africa. To this end, the study examines the role that both subjective and economic well-being take in this relation. In addition, several hypotheses were tested and confirmed based on survey data collected by Afrobarometer in 31 countries of SSA. Logistic regression and Propensity Score Matching provide evidence of a negative association, possibly even effect of micro-entrepreneurship on migration plans. Structural Equation Modeling subsequently reveals a mediating role of subjective wellbeing in this relationship. Specifically, this holds for the subgroups of low-income countries and agricultural entrepreneurs, but not for middle/high-income countries and non-farm entrepreneurs. Contrary to expectations, economic well-being was not found to play a mediating role between entrepreneurship and migration plans. It may only indirectly influence migration plans via subjective well-being. Finally, a positive correlation is observed between economic and subjective well-being, both of which are argued to be essential components of poverty reduction, with the former well-being presumably conditioning the latter.

The study contributes to the existing entrepreneurship literature by portraying the importance of entrepreneurship for young people's migration plans in the context of socioeconomic constraints. By considering evaluative well-being, poverty is defined in more than just economic terms, and examining migration plans from the perspective of the ordinary entrepreneur allows us to conclude about the transformative potential of entrepreneurship for those in abject poverty. However, methodological and data limitations cannot ensure complete robustness and generalizability. Future

studies should test the proposed model using specifically collected and longitudinal survey data to verify the results and derive causal effects. Since entrepreneurial characteristics are considered an important driver of rural development, governments should adopt appropriate policies and provide more incentives for the establishment of an entrepreneurial culture among citizens. It is hoped that the thesis will inspire policymakers and development organizations in their efforts to help unemployed young people become entrepreneurs and ultimately escape poverty.

References

Abadie, A., & Spiess, J. (2021). Robust post-matching inference. *Journal of the American Statistical Association*. (Advance online publication, doi: 10.1080/01621459.2020.1840383)

Abdallah, S., Thompson, S., & Marks, N. (2008). Estimating worldwide life satisfaction. *Ecological Economics*, *65*(1), 35–47.

Abel, G. J. (2018). Estimates of global bilateral migration flows by gender between 1960 and 20151. *International Migration Review*, *52*(3), 809–852.

Abreu, M., Oner, O., Brouwer, A., & van Leeuwen, E. (2019). Well-being effects of self-employment: A spatial inquiry. *Journal of Business Venturing*, *34*(4), 589–607.

Acs, Z. (2006). How is entrepreneurship good for economic growth? *Innovations: Technology, Governance, Globalization*, *1*(1), 97–107.

Acs, Z. J., Desai, S., & Hessels, J. (2008). Entrepreneurship, economic development and institutions. *Small Business Economics*, *31*(3), 219–234.

Adewumi, S. A. (2020). Challenges and prospects of rural entrepreneurship: A discourse analysis of selected local government areas of Osun state, Nigeria. *International Journal of Business and Management Studies*, *12*(2), 544–560.

Adhikari, S., & Gentilini, U. (2018). Should I stay or should I go: do cash transfers affect migration? *World Bank Policy Research Working Paper*.

Adusei, M. (2016). Does entrepreneurship promote economic growth in Africa? *African Development Review*, *28*(2), 201–214.

AFI. (2014). Small and Growing: Entrepreneurship in African Agriculture. *Agriculture For Impact, A Montpellier Panel Report, June 2014*.

Afrobarometer. (2020). *Data Codebook for a Round 7 Afrobarometer Survey.* Retrieved 2021-07-02, from https://afrobarometer.org/sites/default/files/r7_merged.codebook_final_14dec20.pdf

Afrobarometer. (2021). *Sampling principles and weighting.* Retrieved 2021-07-01, from https://afrobarometer.org/surveys-and-methods/sampling-principles

Alemu, A. E., & Adesina, J. O. (2017). In search of rural entrepreneurship: Non-farm household enterprises (NFEs) as instruments of rural transformation in Ethiopia. *African Development Review*, *29*(2), 259–271.

Alvarez, S. A., & Barney, J. B. (2014). Entrepreneurial opportunities and poverty alleviation. *Entrepreneurship Theory and Practice*, *38*(1), 159–184.

Amha, W., Woldehanna, T., Tamrat, E., & Gebremedhin, A. (2015). The characteristics and determinants of entrepreneurship in Ethiopia. *Ethiopian Journal of Economics*, *24*(1), 131–165.

Amoros, J. E., & Bosma, N. (2014). *Global entrepreneurship monitor 2013 global report: Fifteen years of assessing entrepreneurship across the globe* (Tech. Rep.). Global Entrepreneurship Research Association (GERA) and Universidad del Desarrollo and Utrecht University. Retrieved from https://www.gemconsortium.org/report/gem-2013-global-report (accessed: 2021-07-02)

Andersson, P. (2008). Happiness and health: Well-being among the self-employed. *The Journal of Socio-Economics*, *37*(1), 213–236.

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*, *28*(25), 3083–3107.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, *46*(3), 399–424.

Beegle, K., Christiaensen, L., Dabalen, A., & Gaddis, I. (2016). *Poverty in a Rising Africa*. World Bank Publications.

Bello-Bravo, J. (2015). Rural-urban migration: A path for empowering women through entrepreneurial activities in West Africa. *Journal of Global Entrepreneurship Research0*, *5*(1), 1–9.

Benz, M., & Frey, B. S. (2008). Being independent is a great thing: Subjective evaluations of self-employment and hierarchy. *Economica*, *75*(298), 362–383.

Bhoojedhur, S., & Isbell, T. (2019). *Almost half of Malawians consider emigration; most-educated are most likely to look overseas* (Tech. Rep.). Afrobarometer Dispatch No. 281. Retrieved from https://media.africaportal.org/documents/ab _r7_dispatchno281_malawi_emigration.pdf (accessed: 2021-07-02)

Bhuiyan, M. F., & Ivlevs, A. (2019). Micro-entrepreneurship and subjective well-being: Evidence from rural Bangladesh. *Journal of Business Venturing*, *34*(4), 625–645.

Binder, M., & Coad, A. (2013). Life satisfaction and self-employment: A matching approach. *Small Business Economics*, *40*(4), 1009–1033.

Binder, M., & Coad, A. (2016). How satisfied are the self-employed? a life domain view. *Journal of Happiness Studies*, *17*(4), 1409–1433.

Black, R., & Castaldo, A. (2009). Return migration and entrepreneurship in Ghana and Côte d'Ivoire: the role of capital transfers. *Tijdschrift voor Economische en Sociale Geografie*, *100*(1), 44–58.

Blanchflower, D. G., & Oswald, A. J. (2004). Well-being over time in Britain and the USA. *Journal of Public Economics*, *88*(7-8), 1359–1386.

Bommes, M., Fassman, H., & Sievers, W. (2014). *Migration from the Middle East and North Africa to Europe: Past Developments, Current Status and Future Potentials*. Amsterdam University Press.

Bradley, S. W., McMullen, J. S., Artz, K., & Simiyu, E. M. (2012). Capital is not enough: Innovation in developing economies. *Journal of Management Studies*, *49*(4), 684–717.

Brown, S., Farrell, L., & Harris, M. N. (2011). Modeling the incidence of self-employment: individual and employment type heterogeneity. *Contemporary Economic Policy*, *29*(4), 605–619.

Brubaker, R. (2017). Why populism? *Theory and Society*, *46*(5), 357–385.

Bruton, G. D., Ahlstrom, D., & Obloj, K. (2008). Entrepreneurship in emerging economies: Where are we today and where should the research go in the future. *Entrepreneurship Theory and Practice*, *32*(1), 1–14.

Bruton, G. D., Ireland, R. D., & Ketchen Jr, D. J. (2012). Toward a research agenda on the informal economy. *Academy of Management Perspectives*, *26*(3), 1–11.

Bruton, G. D., Ketchen Jr, D. J., & Ireland, R. D. (2013). Entrepreneurship as a solution to poverty. *Journal of Business Venturing*, *28*(6), 683–689.

Cai, R., Esipova, N., Oppenheimer, M., & Feng, S. (2014). International migration desires related to subjective well-being. *IZA Journal of Migration*, *3*(1), 1–20.

Capuano, S., & Marfouk, A. (2013). African brain drain and its impact on source countries: What do we know and what do we need to know? *Journal of Comparative Policy Analysis: Research and Practice*, *15*(4), 297–314.

Carling, J., & Schewel, K. (2018). Revisiting aspiration and ability in international migration. *Journal of Ethnic and Migration Studies*, *44*(6), 945–963.

Carter, S. (2011). The rewards of entrepreneurship: Exploring the incomes, wealth, and economic well–being of entrepreneurial households. *Entrepreneurship Theory and Practice*, *35*(1), 39–55.

Cheeseman, N. (2015). "No bourgeoisie, no democracy"? The political attitudes of the Kenyan middle class. *Journal of International Development*, *27*(5), 647–664.

Chindarkar, N. (2014). Is subjective well-being of concern to potential migrants from Latin America? *Social Indicators Research*, *115*(1), 159–182.

Chliova, M., Brinckmann, J., & Rosenbusch, N. (2015). Is microcredit a blessing for the poor? a meta-analysis examining development outcomes and contextual considerations. *Journal of Business Venturing*, *30*(3), 467–487.

Cho, Y. (2015). Entrepreneurship for the poor in developing countries. *IZA World of Labor*(167).

Clemens, M. A. (2020). The emigration life cycle: How development shapes emigration from poor countries. *IZA Discussion Paper No. 13614*.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Taylor & Francis.

Connor, P. (2018). *International migration from sub-Saharan Africa has grown dramatically since 2010*. Retrieved 2021-07-05, from https://www.pewresearch.org/fact-tank/2018/02/28/international-migration-from-sub-saharan-africa-has-grown-dramatically-since-2010/

Cribari-Neto, F., Ferrari, S. L. P., & Oliveira, W. A. S. C. (2005). Numerical evaluation of tests based on different heteroskedasticity-consistent covariance matrix estimators. *Journal of Statistical Computation and*

Simulation, *75*(8), 611–628.

Dako-Gyeke, M. (2016). Exploring the migration intentions of Ghanaian youth: A qualitative study. *Journal of International Migration and Integration*, *17*(3), 723–744.

Dale, A. (2015). Self-Employment and entrepreneurship: notes on two problematic concepts. In *Deciphering the enterprise culture (routledge revivals)* (pp. 55–72). Routledge.

Dao, T. H., Docquier, F., Parsons, C., & Peri, G. (2018). Migration and development: Dissecting the anatomy of the mobility transition. *Journal of Development Economics*, *132*, 88–101.

Davenport, C., Moore, W., & Poe, S. (2003). Sometimes you just have to leave: Domestic threats and forced migration, 1964-1989. *International Interactions*, *29*(1), 27–55.

De Jong, G. F., Root, B. D., Gardner, R. W., Fawcett, J. T., & Abad, R. G. (1985). Migration intentions and behavior: Decision making in a rural Philippine province. *Population and Environment*, *8*(1-2), 41–62.

De Jong, G. F., & Steinmetz, M. (2006). Migration intentions in South Africa and elsewhere. In *Migration in South and Southern Africa: Dynamics and Determinants* (pp. 249–267). Human Sciences Research Council South Africa.

Dencker, J. C., Bacq, S., Gruber, M., & Haas, M. (2021). Reconceptualizing necessity entrepreneurship: A contextualized framework of entrepreneurial processes under the condition of basic needs. *Academy of Management Review*, *46*(1), 60–79.

Desai, S., Naudé, W., & Stel, N. (2021). Refugee entrepreneurship: context and directions for future research. *Small Business Economics*, *56*, 933–945.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, *95*(3), 542–575.

Diener, E. (2012). New findings and future directions for subjective well-being research. *American Psychologist*, *67*(8), 590–597.

Diener, E., Diener, M., & Diener, C. (2009). Factors predicting the subjective well-being of nations. In *Culture and well-being: The collected works of ed diener* (pp. 43–70). Springer Netherlands.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71–75.

Drucker, P. F. (1998). The discipline of innovation. *Harvard Business Review*, *76*(6), 149–157.

Dulani, B., Mattes, R. B., & Logan, C. (2013). *After a decade of growth in Africa, little change in poverty at the grassroots* (Tech. Rep.). Afrobarometer Policy Brief No. 1. Retrieved from https://afrobarometer.org/publications/pp1-after-decade-growth-africa-little-change-poverty-grassroots (accessed: 2021-07-10)

Etling, A., Backeberg, L., & Tholen, J. (2020). The political dimension of young people's migration intentions: evidence from the Arab Mediterranean region. *Journal of Ethnic and Migration Studies*, *46*(7), 1388–1404.

Fairlie, R. W., & Robb, A. M. (2009). Gender differences in business performance: evidence from the Characteristics of Business Owners survey. *Small Business Economics*, *33*(4), 375–395.

FAO, IFAD, IOM, & WFP. (2018). *The linkages between migration, agriculture, food security and rural development* (Tech. Rep.). FAO Rome, Licence: CC BY-NC-SA 3.0 IGO. Retrieved from http://www.fao.org/3/ca0922en/ca0922en.pdf (accessed: 2021-07-10)

Farrar, D. E., & Glauber, R. R. (1967). Multicollinearity in regression analysis: the problem revisited. *The Review of Economics and Statistics*, *49*(1), 92–107.

Fatoki, O., & Chindoga, L. (2011). An investigation into the obstacles to youth entrepreneurship in South Africa. *International Business Research*, *4*(2), 161–169.

Friebel, G., Gallego, J. M., & Mendola, M. (2013). Xenophobic attacks, migration intentions, and networks: evidence from the South of Africa. *Journal of Population Economics*, *26*(2), 555–591.

GEM. (2013). *Supporting Africa's young entrepreneurs: an investment in job creation and future prosperity for all: Nigeria* (Tech. Rep.). Global Entrepreneurship Monitor Nigeria. Retrieved from https://www.gemconsortium.org/economy-profiles/nigeria (accessed: 2021-07-10)

George, G., Corbishley, C., Khayesi, J. N. O., Haas, M. R., & Tihanyi, L. (2016). Bringing Africa in: Promising directions for management research. *Academy of Management Journal*, *59*(2), 377–393.

Giambra, S., & McKenzie, D. (2021). Self-employment and Migration. *World Development*, *141*. (105362)

Gough, K. V., & Langevang, T. (2016). *Young Entrepreneurs in sub-Saharan Africa*. Routledge.

Grapentine, T. (2000). Path analysis vs. structural equation modeling. *Marketing Research*, *12*(3), 12–20.

Gunzler, D., Morris, N., & Tu, X. M. (2016). Causal mediation analysis using structure equation models. In *Statistical causal inferences and their applications in public health research* (pp. 295–314). Springer International Publishing.

Hagerty, M. R. (1999). Testing Maslow's hierarchy of needs: National quality-of-life across time. *Social Indicators Research*, *46*(3), 249–271.

Hansen, B. B., & Bowers, J. (2008). Covariate balance in simple, stratified and clustered comparative studies. *Statistical Science*, *23*(2), 219–236.

Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials–the gold standard for effectiveness research. *BJOG: An International Journal of Obstetrics and Gynaecology*, *125*(13), 1716–1716.

Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods*, *39*(4), 709–722.

Hayes, A. F., & Preacher, K. J. (2013). Conditional process modeling: Using structural equation modeling to examine contingent causal processes. In *Structural Equation Modeling: A Second Course (2nd ed.)* (pp. 219–266). IAP Information Age Publishing.

Hébert, R. F., & Link, A. N. (1989). In search of the meaning of entrepreneurship. *Small Business Economics*, *1*(1), 39–49.

Hitt, M. A., Ireland, R. D., Camp, S. M., & Sexton, D. L. (2001). Strategic entrepreneurship: Entrepreneurial strategies for wealth creation. *Strategic Management Journal*, *22*(6-7), 479–491.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, *15*(3), 199–236.

Hodler, R., Srisuma, S., Vesperoni, A., & Zurlinden, N. (2020). Measuring ethnic stratification and its effect on trust in Africa. *Journal of Development Economics*, *146*.

Hoogeveen, J. G., Rossi, M., & Sansone, D. (2019). Leaving, Staying or Coming Back? Migration Decisions During the Northern Mali Conflict. *The Journal of Development Studies*, *55*(10), 2089–2105.

Hosmer, D. W., & Lemeshow, S. (2013). *Applied Logistic Regression (Second Edition)*. John Wiley & Sons, Inc.

Hoyle, R. H. (2012). *Handbook of Structural Equation Modeling*. Guilford Press.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Igwe, P. A., & Icha-Ituma, A. (2020). A review of ten years of African entrepreneurship research. In *Research handbook on entrepreneurship in emerging economies* (pp. 325–353). Edward Elgar Publishing.

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*(1), 5–86.

IOM. (2020). *International Dialogue on Migration No. 29: Youth and Migration* (Tech. Rep.). International Organization for Migration. Retrieved from https://publications.iom.int/books/international-dialogue-migration-no-29-youth-and-migration (accessed: 2021-07-02)

Ivlevs, A., Nikolova, M., & Graham, C. (2019). Emigration, remittances, and the subjective well-being of those staying behind. *Journal of Population Economics*, *32*(1), 113–151.

Jayachandran, S. (2020). Microentrepreneurship in developing countries. *National Bureau of Economic Research*.

Johannisson, B. (2011). Towards a practice theory of entrepreneuring. *Small Business Economics*, *36*(2), 135–150.

Johnson, M. P, & Schaltegger, S. (2020). Entrepreneurship for sustainable development: A review and multilevel causal mechanism framework. *Entrepreneurship Theory and Practice*, *44*(6), 1141–1173.

Kachkar, O. A. (2019). Refugee entrepreneurship: empirical quantitative evidence on microenterprises in refugee camps in Turkey. *Journal of Immigrant & Refugee Studies*, 17(3), 333–352.

Kahneman, D., Diener, E., & Schwarz, N. (1999). *Well-being: Foundations of Hedonic Psychology*. Russell Sage Foundation.

Kahneman, D., & Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, 20(1), 3–24.

Kandemir, O. (2012). Human development and international migration. *Procedia-Social and Behavioral Sciences*, 62, 446–451.

Kerr, S. P., & Kerr, W. (2020). Immigrant entrepreneurship in America: Evidence from the survey of business owners 2007 & 2012. *Research Policy*, 49(3).

Kew, J. (2015). *Africa's Young Entrepreneurs: Unlocking the Potential for a Brighter Future*. International Development Research Centre (IDRC) = Centre de recherches pour le développement international (CRDI).

Kimmitt, J., Munoz, P., & Newbery, R. (2020). Poverty and the varieties of entrepreneurship in the pursuit of prosperity. *Journal of Business Venturing*, 35(4).

Klasen, S., & Woolard, I. (2009). Surviving unemployment without state support: Unemployment and household formation in South Africa. *Journal of African Economies*, 18(1), 1–51.

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling (Fourth Edition)*. Guilford Press.

Kluve, J., Puerto, S., Robalino, D., Romero, J. M., Rother, F., Stoeterau, J., … Witte, M. (2017). Interventions to improve the labour market outcomes of youth: A systematic review of training, entrepreneurship promotion, employment services and subsidized employment interventions. *Campbell Systematic Reviews*, 13(1), 1–288.

Koelle, M. (2019). *Microenterprises and the Lure of Wage Work: Theory and Evidence from Mexican Export Manufacturing* (Tech. Rep.). Centre for the Study of African Economies, University of Oxford.

Kolawole, O. D., & Ajila, K. (2015). Driving local community transformation through participatory rural entrepreneurship development. *World Journal of Entrepreneurship, Management and Sustainable Development*, 11(2), 131–139.

Kolk, A., Rivera-Santos, M., & Rufín, C. (2014). Reviewing a decade of research on the "base/bottom of the pyramid" (BOP) concept. *Business & Society*, 53(3), 338–377.

Koyana, S., & Mason, R. B. (2017). Rural entrepreneurship and transformation: the role of learnerships. *International Journal of Entrepreneurial Behavior & Research*, 23(5), 734-751.

Langevang, T., Namatovu, R., & Dawa, S. (2012). Beyond necessity and opportunity entrepreneurship: motivations and aspirations of young entrepreneurs in Uganda. *International Development Planning Review*, 34(4), 339-459.

Liang, J., Wang, H., & Lazear, E. P. (2018). Demographics and entrepreneurship. *Journal of Political Economy*, 126(S1), S140–S196.

London, T., & Hart, S. L. (2010). *Next Generation Business Strategies for the Base of the Pyramid: New Approaches for Building Mutual Value*. FT Press.

Mair, J., Battilana, J., & Cardenas, J. (2012). Organizing for society: A typology of social entrepreneuring models. *Journal of Business Ethics*, 111(3), 353–373.

Manchin, M., & Orazbayev, S. (2018). Social networks and the intention to migrate. *World Development*, 109, 360–374.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50(4), 370–396.

Mattes, R. (2020). *Lived poverty on the rise: Decade of living-standard gains ends in Africa* (Tech. Rep.). Afrobarometer Policy Paper No. 62. Retrieved from https://afrobarometer.org/publications/pp62-lived-poverty-rise-decade-living-standard-gains-ends-africa (accessed: 2021-07-02)

Mattes, R., Dulani, B., & Gyimah-Boadi, E. (2016). *Africa's growth dividend? lived poverty drops across much of the continent* (Tech. Rep.). Afrobarometer Policy Paper No. 29. Retrieved from https://afrobarometer.org/publications/pp29-africas-growth-dividend-lived-poverty-drops-across-the-continent (accessed: 2021-07-02)

McFadden, D. (1978). Quantitative Methods for Analysing Travel Behaviour of Individuals: Some Recent Developments. In *Behavioural travel modelling* (pp. 279–318). Croom Helm.

Meyer, D. F., & Keyser, E. (2016). Validation and testing of the Lived Poverty Index Scale (LPI) in a poor South African community. *Social Indicators Research*, 129(1), 147–159.

Migali, S., & Scipioni, M. (2019). Who's about to leave? a global survey of aspirations and intentions to migrate. *International Migration*, 57(5), 181–200.

Milasi, S. (2020). 7What Drives Youth's Intention to Migrate Abroad? Evidence from International Survey Data. *IZA Journal of Development and Migration*, 11(1).

Moore, W. H., & Shellman, S. M. (2007). Whither will they go? a global study of refugees' destinations, 1965–1995. *International Studies Quarterly*, 51(4), 811–834.

Mukhtar, S., Rose, R. A. C., Choy, L. K., & Bibi-Farouk, A. U. I. (2018). Boko Haram and the Geopolitics of Forced Migration in Nigeria. *Journal of International Studies*, 14, 51–63.

Nagler, P., & Naudé, W. (2014). Non-farm entrepreneurship in rural Africa: Patterns and determinants. *IZA Discussion Paper*.

Nagler, P., & Naudé, W. (2017). Non-farm entrepreneurship in rural sub-Saharan Africa: New empirical evidence. *Food Policy*, 67, 175–191.

Naudé, W. (2008). Conflict, disasters and no jobs: Reasons for international migration from Sub-Saharan Africa. *WIDER Research Paper*.

Naudé, W. (2010). Entrepreneurship, developing countries, and development economics: new approaches and insights. *Small Business Economics*, 34(1), 1–12.

Naudé, W., Siegel, M., & Marchand, K. (2017). Migration, entrepreneurship and development: critical questions. *IZA Journal of Migration*, 6(1), 1–16.

Neumayer, E. (2005). Bogus refugees? The determinants of asylum migration to Western Europe. *International Studies Quarterly*, 49(3), 389–409.

Nikolaev, B., Boudreaux, C. J., & Wood, M. (2020). Entrepreneurship and subjective well-being: The mediating role of psychological functioning. *Entrepreneurship Theory and Practice*, 44(3), 557–586.

Nikolaev, B., & Rusakov, P. (2016). Education and happiness: an alternative hypothesis. *Applied Economics Letters*, 23(12), 827–830.

Nikolova, M. (2019). Switching to self-employment can be good for your health. *Journal of Business Venturing*, 34(4), 664–691.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673–690.

Odhiambo, F. O. (2019). Assessing the predictors of lived poverty in Kenya: A secondary analysis of the Afrobarometer survey 2016. *Journal of Asian and African Studies*, 54(3), 452–464.

OECD. (2000). *Small and medium-sized enterprises: local strength, global reach* (Tech. Rep.). Public Affairs Division, OECD Policy Brief. Retrieved from https://www.oecd.org/cfe/leed/1918307.pdf (accessed: 2021-07-10)

OECD. (2015). *Connecting with Emigrants: A Global Profile of Diasporas 2015*. OECD Publishing, Paris. Retrieved from https://books.google.de/books?id=UZRmnQAACAAJ

Okeke, C., & Nwankwo, F. (2017). Rural entrepreneurship and rural development in Nigeria. *Africa's Public Service Delivery and Performance Review*, 5(1).

Olmos, A., & Govindasamy, P. (2015). Propensity scores: a practical introduction using R. *Journal of MultiDisciplinary Evaluation*, 11(25), 68–88.

Ozaltin, D., Shakir, F., & Loizides, N. (2020). Why do people flee? Revisiting forced migration in post-Saddam Baghdad. *Journal of International Migration and Integration*, 21(2), 587–610.

Parker, S. C. (2004). *The economics of Self-employment and Entrepreneurship*. Cambridge University Press.

Peng, C.-Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research*, 96(1), 3–14.

Poole, M. A., & O'Farrell, P. N. (1971). The assumptions of the linear regression model. *Transactions of the Institute of British Geographers*(52), 145–158.

Quagrainie, F. A., Adams, S., Kabalan, A. A. M., & Dankwa, A. D. (2021). Micro-entrepreneurship, sustainable development goal one and cultural expectations of Ghanaian women. *Journal of Entrepreneurship*

in Emerging Economies, *13*(1), 86–106.

Rapoza, K. (2017). *Here's Why Europe Really Needs More Immigrants.* Retrieved 2021-07-19, from https://www.forbes.com/sites/kenrapoza/2017/08/15/heres-why-europe-really-needs-more-immigrants/?sh=9fda5fb49171

Reid, M., Roumpi, D., & O'Leary-Kelly, A. M. (2015). Spirited women: The role of spirituality in the work lives of female entrepreneurs in Ghana. *Africa Journal of Management*, *1*(3), 264–283.

Reynolds, P. D., Camp, M. S., Bygrave, W. D., Autio, E., & Hay, M. (2001). 2001 Executive Report. *Global Entrepreneurship Monitor*. Retrieved from https://www.gemconsortium.org/report/gem-2001-global-report (accessed: 2021-07-02)

Reynolds, P. D., Hay, M., & Camp, M. S. (1999). *Global Entrepreneurship Monitor: 1999 Executive Report* (Tech. Rep.). Global Entrepreneurship Monitor. Retrieved from https://www.gemconsortium.org/report/gem-1999-global-report (accessed: 2021-07-02)

Rindova, V., Barry, D., & Ketchen Jr, D. J. (2009). Entrepreneuring as emancipation. *Academy of Management Review*, *34*(3), 477–491.

Rosenbaum, P. R. (2002). Overt bias in observational studies. In *Observational studies* (pp. 71–104). Springer.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Roser, M., & Ortiz-Ospina, E. (2019). *Global Extreme Poverty*. Retrieved 2021-05-02, from https://ourworldindata.org/extreme-poverty

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, *48*(2), 1–36.

Rubin, D. B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine*, *127*(8, Pt.2), 757–763.

Ruyssen, I., & Salomone, S. (2018). Female migration: A way out of discrimination? *Journal of Development Economics*, *130*, 224–241.

Ryan, R. M., & Deci, E. L. (2017). *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. Guilford Publications.

Salameh, M. T. B. (2019). Migration from the Arab Spring Countries to Europe: Causes and Consequences. In *Smart technologies and innovation for a sustainable future* (pp. 243–254). Springer International Publishing.

Schmeidl, S. (1997). Exploring the causes of forced migration: A pooled time-series analysis, 1971-1990. *Social Science Quarterly*, *78*(2), 284–308.

Schumpeter, J. A. (1934). *The Theory of Economic Development*. Cambridge, MA: Harvard University Press.

Sekeris, P., & Vasilakis, C. (2016). The Mediterranean refugees crisis and extreme right parties: Evidence from Greece. *MPRA Paper 72222, University Library of Munich, Germany*.

Seligman, M. E. P. (2012). *Flourish: A Visionary New Understanding of Happiness and Well-being*. Simon and Schuster.

Shepherd, D. A., Parida, V., & Wincent, J. (2021). Entrepreneurship and poverty alleviation: the importance of health and children's education for slum entrepreneurs. *Entrepreneurship Theory and Practice*, *45*(2), 350–385.

Shepherd, D. A., & Patzelt, H. (2017). *Trailblazing in Entrepreneurship: Creating New Paths for Understanding the Field*. Springer Nature.

Sheriff, M., & Muffatto, M. (2015). The present state of entrepreneurship ecosystems in selected countries in Africa. *African Journal of Economic and Management Studies*, *6*(1), 17–54.

Shir, N. (2015). Entrepreneurial Well-Being: The Payoff Structure of Business Creation. *Stockholm School of Economics*.

Shir, N., Nikolaev, B. N., & Wincent, J. (2019). Entrepreneurship and well-being: The role of psychological autonomy, competence, and relatedness. *Journal of Business Venturing*, *34*(5). (105875)

Sirgy, M. J. (1986). A Quality-of-Life Theory Derived from Maslow's Developmental Perspective: 'Quality' Is Related to Progressive Satisfaction of a Hierarchy of Needs, Lower Order and Higher. *American Journal of Economics and Sociology*, *45*(3), 329–342.

Sivo, S. A., Fan, X., Witta, E. L., & Willse, J. T. (2006). The search for" optimal" cutoff properties: Fit index criteria in structural equation

modeling. *The Journal of Experimental Education*, *74*(3), 267–288.

Sohns, F., & Diez, J. R. (2018). Explaining micro entrepreneurship in rural Vietnam–a multilevel analysis. *Small Business Economics*, *50*(1), 219–237.

Sriram, V., & Mersha, T. (2006). Facilitating entrepreneurship in Sub-Saharan Africa: what governments can do. *Journal for International Business and Entrepreneurship Development*, *3*(1-2), 136–151.

Steele, A. (2009). Seeking safety: Avoiding displacement and choosing destinations in civil wars. *Journal of Peace Research*, *46*(3), 419–429.

Stephan, U. (2018). Entrepreneurs' mental health and well-being: A review and research agenda. *Academy of Management Perspectives*, *32*(3), 290–322.

Stevenson, B., & Wolfers, J. (2009). The paradox of declining female happiness. *American Economic Journal: Economic Policy*, *1*(2), 190–225.

Sutter, C., Bruton, G. D., & Chen, J. (2019). Entrepreneurship as a solution to extreme poverty: A review and future research directions. *Journal of Business Venturing*, *34*(1), 197–214.

Tartakovsky, E., & Schwartz, S. H. (2001). Motivation for emigration, values, wellbeing, and identification among young Russian Jews. *International Journal of Psychology*, *36*(2), 88–99.

TheGlobalEconomy. (2021). *Compare countries with annual data from official sources: Fragile state index*. Retrieved 2021-07-02, from https://www.theglobaleconomy.com/compare-countries/

Theodossiou, I. (1998). The effects of low-pay and unemployment on psychological well-being: a logistic regression approach. *Journal of Health Economics*, *17*(1), 85–104.

Tjaden, J., Auer, D., & Laczko, F. (2019). Linking migration intentions with flows: Evidence and potential use. *International Migration*, *57*(1), 36–57.

Tobias, J. M., Mair, J., & Barbosa-Leiker, C. (2013). Toward a theory of transformative entrepreneuring: Poverty reduction and conflict resolution in Rwanda's entrepreneurial coffee sector. *Journal of Business Venturing*, *28*(6), 728–742.

UN. (2019). World population prospects 2019: Highlights. *United Nations, Department for Economic and Social Affairs, Population Division*. Retrieved from https://population.un.org/wpp/Publications/Files/WPP2019_Highlights.pdf (accessed: 2021-07-19)

UN. (2020). *The Sustainable Development Goals Report 2020* (Tech. Rep.). United Nations, United Nations Statistics Division (UNSD). Retrieved from https://unstats.un.org/sdgs/report/2020/The-Sustainable-Development-Goals-Report-2020.pdf (accessed: 2021-07-19)

UNDP. (2009). *Bericht Über die menschliche Entwicklung 2009: Barrieren Überwinden: Migration und menschliche Entwicklung* (Tech. Rep.). United Nations, United Nations Development Programme (UNDP). Retrieved from http://hdr.undp.org/sites/default/files/hdr2009-german.pdf (accessed: 2021-07-19)

UNDP. (2013). *UNDP in Sub-Saharan Africa: Supporting a region on the move* (Tech. Rep.). United Nations, United Nations Development Programme (UNDP). Retrieved from https://www.africa.undp.org/content/rba/en/home/library/outreach-material/undp-africa-brochure.html (accessed: 2021-07-19)

UNDP. (2016). *Africa Human Development Report 2016 Accelerating Gender Equality and Women's Empowerment in Africa* (Tech. Rep.). United Nations, United Nations Development Programme (UNDP). Retrieved from https://www.undp.org/sites/g/files/zskgke326/files/publications/AfHDR_2016_lowres_EN.pdf (accessed: 2021-07-19)

UNHCR. (2021). *Global Appeal 2021 Update* (Tech. Rep.). United Nations, United Nations High Commissioner for Refugees (UNHCR). Retrieved from https://reporting.unhcr.org/sites/default/files/ga2021/pdf/Global_Appeal_2021_full_lowres.pdf (accessed: 2021-07-19)

Uy, M. A., Foo, M.-D., & Song, Z. (2013). Joint effects of prior start-up experience and coping strategies on entrepreneurs' psychological well-being. *Journal of Business Venturing*, *28*(5), 583–597.

Van Dalen, H. P., Groenewold, G., & Fokkema, T. (2005). The effect of remittances on emigration intentions in Egypt, Morocco, and Turkey. *Population Studies*, *59*(3), 375–392.

Van Dalen, H. P., & Henkens, K. (2008). *Emigration intentions: Mere words*

or true plans? explaining international migration intentions and behavior (Tech. Rep.). Discussion Paper, Tilburg University, Center for Economic Research.

Venkataraman, S. (2004). Regional transformation through technological entrepreneurship. *Journal of Business Venturing*, *19*(1), 153–167.

Vermeire, J. A. L., & Bruton, G. D. (2016). Entrepreneurial opportunities and poverty in sub-Saharan Africa: a review & agenda for the future. *Africa Journal of Management*, *2*(3), 258–280.

Vivarelli, M. (2013). Is entrepreneurship necessarily good? microeconomic evidence from developed and developing countries. *Industrial and Corporate Change*, *22*(6), 1453–1495.

Wauters, B., & Lambrecht, J. (2008). Barriers to refugee entrepreneurship in Belgium: Towards an explanatory model. *Journal of Ethnic and Migration Studies*, *34*(6), 895–915.

Weiner, M. (1996). Bad neighbors, bad neighborhoods: An inquiry into the causes of refugee flows. *International Security*, *21*(1), 5–42.

Welter, F., Baker, T., Audretsch, D. B., & Gartner, W. B. (2017). Everyday entrepreneurship–a call for entrepreneurship research to embrace entrepreneurial diversity. *Entrepreneurship Theory and Practice*, *41*, 311–321.

West, G. P., Bamford, C. E., & Marsden, J. W. (2008). Contrasting entrepreneurial economic development in emerging Latin American economies: Applications and extensions of resource-based theory. *Entrepreneurship Theory and Practice*, *32*(1), 15–36.

Wiklund, J., Nikolaev, B., Shir, N., Foo, M.-D., & Bradley, S. (2019). Entrepreneurship and well-being: Past, present, and future. *Journal of Business Venturing*, *34*(4), 579–588.

Wiklund, J., Yu, W., Tucker, R., & Marino, L. D. (2017). ADHD, impulsivity and entrepreneurship. *Journal of Business Venturing*, *32*(6), 627–656.

Williamson, J. G. (1965). Regional inequality and the process of national development: a description of the patterns. *Economic Development and Cultural Change*, *13*(4, Pt.2), 1–84.

Wills-Herrera, E., Orozco, L. E., Forero-Pineda, C., Pardo, O., & Andonova, V. (2011). The relationship between perceptions of insecurity, social capital and subjective well-being: Empirical evidences from areas of rural conflict in Colombia. *The Journal of Socio-Economics*, *40*(1), 88–96.

Wood, E. J. (2008). The social processes of civil war: The wartime transformation of social networks. *Annual Review of Political Science*, *11*, 539–561.

WorldBank. (2021). *World Bank Country and Lending Groups: historical classification by income in XLSX format*. Retrieved 2021-07-01, from https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

Wu, H., & Leung, S.-O. (2017). Can Likert scales be treated as interval scales?—A Simulation study. *Journal of Social Service Research*, *43*(4), 527–532.

Battle of Realities – a Qualitative-Empirical Analysis of the Use of Augmented and Virtual Reality in Sales

Battle of Realities – eine qualitativ-empirische Analyse über den Einsatz von Augmented Reality und Virtual Reality im Vertrieb

Daniel Schachtsieck

Ruhr-Universität Bochum

Abstract

Many companies are unaware of the possible uses of augmented reality and virtual reality and the resulting potential. Current research also shows large gaps in this area. The aim of this work is to develop a decision-making aid for their use in sales. In addition, a general understanding of the application of both technologies in sales is to be created and specific potentials are to be revealed by analyzing them against the background of selected criteria in different sales areas. For this purpose, a qualitative empiricism is carried out in the form of expert interviews. Finally, implications for practice are derived in an assignment matrix with application examples. This paper shows that companies are already successfully using the technologies in different sales areas and that salespeople are aware of their added value. The use of the technologies will increase, as will the user acceptance of the customers, due to the progressing generational change. By combining theoretical basics with expert knowledge from practice, the work expands the level of knowledge about the use of augmented reality and virtual reality in sales.

Zusammenfassung

Zahlreiche Unternehmen kennen die möglichen Einsatzmöglichkeiten von Augmented Reality und Virtual Reality und die daraus resultierenden Potenziale nicht. Auch die aktuelle Forschung weist hier große Lücken auf. Das Ziel dieser Arbeit ist es, eine Entscheidungshilfe für deren Einsatz im Vertrieb zu erarbeiten. Zusätzlich soll ein allgemeines Anwendungsverständnis beider Technologien im Vertrieb geschaffen und konkrete Potenziale aufgedeckt werden, indem diese vor dem Hintergrund ausgewählter Kriterien in unterschiedlichen Vertriebsbereichen analysiert werden. Dafür wird eine qualitative Empirie in Form von Experteninterviews durchgeführt. Abschließend werden in einer Zuordnungsmatrix mit Anwendungsbeispielen Implikationen für die Praxis abgeleitet. Die vorliegende Arbeit zeigt, dass Unternehmen die Technologien bereits erfolgreich in unterschiedlichen Vertriebsbereichen einsetzen und sich die Vertriebler über deren Mehrwert bewusst sind. Der Einsatz der Technologien wird ebenso wie die Benutzerakzeptanz der Kunden durch den fortschreitenden Generationswechsel zunehmen. Durch die Zusammensetzung theoretischer Grundlagen mit dem Expertenwissen aus der Praxis erweitert die Arbeit den Wissensstand über den Einsatz von Augmented Reality und Virtual Reality im Vertrieb.

Keywords: Augmented Reality; Virtual Reality; Sales; Digitalisierung.

1. Einleitung in die Thematik

„Unsere Prognosen zeigen, dass VR, AR [...] in Deutschland zum Milliardenmarkt werden kann. [...] Jetzt gilt es, Konsumenten zu begeistern und für einen „Must have"- Effekt zu sorgen, um aus einer faszinierenden, innovativen Technologie auch einen kommerziellen Erfolg zu machen". Mit dieser Einschätzung zeigt Klaus Böhm, Leiter der Media

Practice bei Deloitte, das Potenzial, das in der Anwendung von Augmented Reality (AR) und Virtual Reality (VR) steckt. Weitere Zahlen, welche die Relevanz des Themas verdeutlichen, wurden in einer Studie von Deloitte im Jahr 2020 prognostiziert. Demnach werden im Jahr 2024 in Deutschland mit VR-Hardware und -Inhalten 530 Millionen Euro umgesetzt. Insgesamt wird die durchschnittliche jährliche Wachstumsrate (CAGR) zwischen 2019 und 2024 bei jeweils

30 % liegen (vgl. für den gesamten Absatz Deloitte, 2020).

Zudem gab die Wirtschaftsprüfungsgesellschaft KPMG 2016 in einer Studie über die Potenziale von AR und VR in Unternehmen an, dass Marketing und Vertrieb mit 52% das größte Anwendungsgebiet sind (vgl. Deutsch et al., 2016, S. 8). Weltweit sollen die Umsätze der AR und VR-Anwendungen laut Goldman Sachs bis 2025 auf 80 Milliarden US-Dollar steigen (vgl. Bellini et al., 2016, S. 4).

Im Jahr 2015 haben Roland Berger und Google Deutschland gemeinsam eine Befragung unter 2745 Vertriebsverantwortlichen im B2B-Umfeld durchgeführt, wie umfassend die Digitalisierung in dem befragten Unternehmen betrieben wird. 60% der Befragten geben an, dass die Digitalisierung des Vertriebs künftig in ihrer Branche ausschlaggebend für den Geschäftserfolg sein wird (vgl. für den gesamten Absatz Lässig et al., 2015).

1.1. Problemstellung, Forschungsstand und Ziele der Arbeit

Derzeit herrscht in Bezug auf die Technologien die Unkenntnis der Unternehmen über mögliche Einsatzmöglichkeiten und über daraus resultierende Potenziale. Während viele Studien die Auswirkungen der digitalen Transformation auf Marketingstrategien und -aktivitäten untersuchten, wurden die Auswirkungen digitaler Technologien auf Verkaufsprozesse und folglich auf die Aktivitäten des Vertriebs vernachlässigt (vgl. Guenzi & Habel, 2020, S. 58). Dennoch erkennen die meisten Unternehmen an, wie wichtig es ist, ihre Vertriebsorganisation digital zu verändern (vgl. Guenzi & Habel, 2020, S. 81). Darüber hinaus ist das Zukunftspotenzial und die Innovationskraft von AR und VR durch Studien unverkennbar belegt (vgl. Hilken, de Ruyter, Chylinski, Mahr & Keeling, 2017, S. 898). Jedoch stellt eine entsprechende Umsetzung die Unternehmen vor Herausforderungen (vgl. Guenzi & Habel, 2020, S. 81).

Das durch die Corona Pandemie bedingte Problem des abnehmenden Kundenkontakts fordert von den Unternehmen, auf alternative Kommunikationswege zu wechseln, wodurch die Entwicklung von Technologien wie AR und VR gefördert werden. Unternehmen, die in diese Technologien bereits investiert haben, profitieren nun in Form von Effizienzsteigerungen im Vertrieb sowie einer Verbesserung des Kundenerlebnisses (vgl. Kober, 2020, S. 51).

In der Literatur wird sich primär mit den technologischen Aspekten oder Einzelfallanwendungen, wie beispielsweise (bspw.) die Medizin, beschäftigt. Konkrete Analysen über mögliche Vorteile und Anwendungspotenziale im Vertrieb sind derzeit kaum erkennbar. Lediglich rudimentäre Ausführungen finden sich in einzelnen Journals wieder (vgl. Anhang i). Somit ergibt sich eine erhebliche Forschungslücke im Hinblick auf den Einsatz von AR und VR im Vertrieb, welche bereits in der Literatur identifiziert wurde (vgl. Wedel, Bigné & Zhang, 2020, S. 2). Diese Bachelorarbeit setzt an dieser Forschungslücke an und leistet eine umfassende Analyse beider Technologien.

Dabei liegt der Schwerpunkt auf den Einsatz im B2B-Bereich. Der Vertrieb liegt im Fokus der Analyse, da dieser

einen essenziellen Anteil an der wirschaftlichen Entwicklung eines Unternehmens hat (vgl. Winkelmann, 2013, S. 287 f.). Der Zuwachs an Komplexität der Angebote und die zunehmende Vielschichtigkeit führen dazu, dass die Bedeutung des Vertriebs weiter zunimmt (vgl. Maas, 2021, S. 17). Die resultierende zentrale Forschungsfrage lautet, welche Technologie (AR und/oder VR) sich vor dem Hintergrund ausgewählter Kriterien für den Einsatz in welchen Vertriebsbereichen eignet. Das Ziel dabei ist es, aus den neu gewonnenen Erkenntnissen eine bisher in der Literatur nicht erkennbare Entscheidungshilfe für Unternehmen zu erarbeiten, welche eine praktische Anwendung dieser Technologien ermöglicht. Zusätzlich soll ein allgemeines Anwendungsverständnis dieser Technologien im Vertrieb geschaffen und konkrete Potenziale aufgedeckt werden.

1.2. Aufbau der Untersuchung

Um die oben genannte Forschungsfrage fundiert und umfangreich beantworten zu können, wird im zweiten Kapitel ein allgemeines Verständnis der Thematik im Rahmen der theoretischen Grundlagen geschaffen, um einen einheitlichen Wissensstand für den weiteren Verlauf der Arbeit sicherzustellen. Literaturbasiert werden die verschiedenen Vertriebsbereiche voneinander abgegrenzt und AR und VR definiert. Im Anschluss werden im dritten Kapitel wissenschaftliche Theorien in ihren Grundzügen erläutert, um daraus Kriterien abzuleiten, die die Basis der Analyse von AR und VR bilden.

Im vierten Kapitel wird der methodische Rahmen der Experteninterviews erläutert. Die Methodik der qualitativen Empirie bietet einen tiefen Einblick über den aktuellen Einsatz der Technologien in der Praxis.

Im fünften Kapitel werden die beiden Technologien AR und VR basierend auf den theoretischen Grundlagen, den wissenschaftlichen Theorien und den neuen Erkenntnissen aus der qualitativen Empirie analysiert. Dabei bilden der Nutzen, die Integration, die Benutzerakzeptanz, die Kunden- und die Produktzufriedenheit den Kriterienkatalog und somit die Basis der Analyse. Innerhalb der Analyse werden die Ergebnisse der Experteninterviews im Lichte der Kriterien dargestellt.

Am Ende der Bachelorarbeit werden die Ergebnisse diskutiert, die bedeutsamsten Erkenntnisse zusammengefasst und Implikationen für die Praxis beschrieben. Dabei wird auf die Generalisierbarkeit der Ergebnisse eingegangen und Empfehlungen für den Einsatz der Technologien in den jeweiligen Vertriebsbereichen anhand einer Abbildung dargestellt. Zuletzt wird untersucht, wo die Probleme dieser Arbeit lagen, um darauf aufbauend Implikationen für die zukünftige Forschung ableiten zu können.

2. Theoretische Grundlagen zum Vertrieb und zu den Technologien

In diesem Kapitel werden die Grundlagen der Bachelorarbeit präsentiert, indem alle relevanten Begriffe definiert

sowie wichtige Informationen zum tieferen Verständnis des Sachverhaltes erklärt und miteinander in Zusammenhang gebracht werden.

2.1. Der Vertrieb und die verschiedenen Vertriebsbereiche

Zunächst wird der Vertrieb definiert, bevor die unterschiedlichen Vertriebsbereiche herausgestellt werden, um die Anwendungen von AR und VR auf die jeweiligen Bereiche beziehen zu können.

Im Rahmen dieser Bachelorarbeit wird unter dem Begriff Vertrieb jene organisatorische Einheit im Unternehmen verstanden, die für die Koordination und Umsetzung der Verkaufsziele, -strategien und -maßnahmen verantwortlich ist (vgl. Steiner, 2016, S. 18). Vor dem Hintergrund der sich überschneidenden Aufgabengebiete von Marketing und Vertrieb wird im Rahmen dieser Arbeit der Zeithorizont als Abgrenzungskriterium ausgewählt (vgl. Homburg, Jense & Krohmer, 2008, S. 148). Während der Zeithorizont aufgrund der Anreize für Vertriebsmitarbeiter und ihre Ergebnisorientierung tendenziell kurzfristig ausgelegt ist, ist das Marketing dagegen hinsichtlich der Wettbewerbsvorteile eher langfristig ausgerichtet (vgl. Rouziès et al., 2005, S. 115).

Der Vertrieb wird in die Teilbereiche Vertriebssystempolitik, Vertriebskanalpolitik, Verkaufspolitik und Vertriebslogistik eingeteilt. Die Vertriebssystempolitik konzentriert sich auf die effiziente Gestaltung der Verkaufsform, der Vertriebsorganisation und der Vertriebssteuerung. Die Vertriebskanalpolitik hingegen bestimmt die Art und die Anzahl der Vertriebspartner, um das Produkt bestmöglich vom Hersteller zum Endkunden zu bekommen. Dafür müssen die jeweiligen Vertriebswege- und Partner koordiniert und aufeinander abgestimmt sein. Die erfolgreiche Integration von neuen Technologien in bestehende Strukturen spielt hier bspw. eine entscheidende Rolle. Relevante Einflussfaktoren für die Wahl des Vertriebskanals sind unter anderem die Vertriebskosten, der gewünschte Distributionsgrad, Wachstumsmöglichkeiten, Lieferantentreue sowie die Beeinflussbarkeit und die Kontrolle der Absatzermittler (vgl. Hofbauer & Hellwig, 2016, S. 91; Busch, Fuchs & Unger, 2008, S. 333). Der dritte Bereich, die Verkaufspolitik, umfasst die Gewinnung, Pflege und Bindung des Kunden. Die Vertriebslogistik beinhaltet sämtliche Entscheidungen über die physischen Bewegungen wie Lieferservice, Auftragsabwicklung, Lagerhaltung und Transport (vgl. für den gesamten Absatz Steiner, 2016, S. 19-20).

2.2. Die Einordnung von Augmented Reality und Virtual Reality in das Virtualitätskontinuum

In einer Arbeit aus 1994 entwickelten Milgram und Kishino das Virtualitätskontinuum, anhand dessen man die unterschiedlichen Technologien mit visuellen Displays mit gemischten Realitäten einordnen und voneinander abgrenzen kann. In dieser Arbeit werden sechs Klassen definiert und es wird versucht, diese sinnvoll voneinander zu unterscheiden. AR und VR stehen sich in diesem Virtualitätskontinuum als umgekehrten Fall entgegen. An dem einen Ende des Virtualitätskontinuums steht die reale beziehungsweise (bzw.)

physische Umgebung, am gegenüberliegenden Extremum die virtuelle Umgebung. Um im weiteren Verlauf der Arbeit die besonderen Merkmale von AR und VR zu verstehen, ist es wichtig, die physische Realität zu erläutern, um sie von der virtuellen Realität abgrenzen zu können. Daher wird die physische Realität als die Umgebung, die wir natürlich wahrnehmen, ohne zusätzliche Informationen, die von Kommunikationstechnologien bereitgestellt werden, definiert (vgl. Steffen, Gaskin, Meservy, Jenkins & Wolman, 2019, S. 686). Um die Technologien AR und VR nun in dem Virtualitätskontinuum einzuordnen, werden beide Technologien definiert. Augmented Reality bezeichnet jeden Fall, in dem eine reale Umgebung durch virtuelle Objekte erweitert wird. Somit wird die Umgebung um ein virtuelles Objekt in Echtzeit ergänzt (vgl. für den gesamten Absatz Milgram & Kishino, 1994, S. 2-3).

Demgegenüber betrachten wir nun VR. VR wird definiert als eine computergenerierte Simulation einer Situation, die den Benutzer einbezieht, der sie über einen oder mehrere Sinne wahrnimmt und auf einer Art und Weise mit ihr interagiert, die real zu sein scheint (vgl. LaValle, 2017, S. 1; Sherman & Craig, 2002, zitiert nach Wedel et al., 2020, S. 1). Es findet ein komplettes Eintauchen in eine virtuelle Umgebung statt. Dabei handelt es sich um eine Echtzeit-Interaktion (vgl. Burdea & Coiffet, 2003, S. 2). Grundsätzlich wird zwischen zwei Arten von VR unterschieden. Zum einen nicht immersive und zum anderen die immersive VR, oder auch die sogenannte (sog.) low immersion und high immersion VR (vgl. Lee & Wong, 2014, S. 50). Die erste Form, die low immersion VR, wird definiert als einen computergenerierten dreidimensionalen virtuellen Raum, der mit audiovisuellen Standardgeräten wie einem Desktop-Computer mit einem zweidimensionalen Monitor erlebt wird. High immersion VR hingegen ist ein computergenerierter virtueller 360 Grad-Raum, der mit Hilfe eines am Kopf montierten Geräts als räumlich realistisch wahrgenommen werden kann. Der zentrale Unterschied der beiden Formen des VR besteht in dem Grad des Eintauchens in die virtuelle Umgebung. Low immersion VR mögen expansiv und interaktiv sein, erscheinen jedoch für die Sinne nicht realistisch. High immersion VR muss nicht unbedingt soziale Interaktionen oder eine expansive Umgebung beinhalten, impliziert jedoch ein räumliches Bewusstsein, das die Sinne der Benutzer in ein Gefühl physischer Präsenz versetzt. Der Grad des Eintauchens hängt primär von der Ausrüstung ab, die zum Erleben des VR-Inhalts verwendet wird. Benutzer erleben low immersion VR, indem sie einen zweidimensionalen Standardmonitor anzeigen und eine Tastatur oder eine Maus oder beides für die Interaktion verwenden. Für high immersion VR ist ein am Kopf montiertes Gerät bzw. ein Headset notwendig. Das Headset trennt Benutzer physisch von der realen Welt und ermöglicht so ein tieferes Eintauchen in den virtuellen Raum (vgl. für den gesamten Absatz Kaplan-Rakowski & Gruber, 2019, S. 1-2).

Letztlich lässt sich der Unterschied zwischen AR und VR wie folgt zusammenfassen: Während man beim Einsatz von AR eine virtuelle Komponente in der realen Welt ergänzt, kann man mit VR eine komplett neue, virtuelle Welt erschaf-

fen, in die der Anwender vollständig eintaucht. Bezogen auf das Virtualitätskontinuum lässt sich festhalten, dass AR deutlich näher als VR an dem Ende der realen Umgebung steht und VR auf der anderen Seite der virtuellen Umgebung.

3. Die Erstellung des Kriterienkatalogs

Auf Basis wissenschaftlicher Theorien werden Kriterien herausgearbeitet, die im weiteren Verlauf der Arbeit die Grundlage für die Analyse von AR und VR bilden. Dafür werden im Folgenden die Theorien aufgezeigt und zusammenfassend erläutert, bevor daraus der Kriterienkatalog abgeleitet wird.

3.1. Zusatznutzen für Unternehmen

Um eine theoretische Grundlage für die Analyse des Nutzens von AR und VR zu schaffen, wird ein Blick auf die Nutzenlehre von Vershof geworfen (vgl. Vershofen, 1959, S. 81 ff.). Aus dieser Arbeit resultierten eine Vielzahl neuer Nutzentheorien. Die Gemeinsamkeit aller Theorien bildet die Idee, dass der Grundnutzen um Zusatznutzenkomponenten zu ergänzen ist, die von funktionalen Aspekten losgelöst ein zusätzliches Nutzenversprechen generieren. Ein Beispiel für eine solche Zusatznutzenkomponente sei auf magisch-mysteriöser Natur ausgelöste Emotionen (vgl. Gutsche, Herrmann, Huber, Kressmann & Algesheimer, 2005, S. 639). Die Nutzentheoretiker sehen in dieser Erkenntnis die Chance für Unternehmen, sich durch diese neuartigen Nutzenversprechen einen Vorzug gegenüber Produkten von Wettbewerbern zu verschaffen (vgl. Wiswede, 2012, S. 28 ff.). Mit Hilfe der Theorieentwicklung kann nun die Bedeutung des Grundnutzens und die der Zusatznutzenarten erfasst werden. Nach Sweeney und Soutar bildet einer dieser Zusatznutzenarten die Dimension des Markennutzens, also das Produkterlebnis, das sich aus den Empfindungen ableiten lässt (vgl. Sweeney & Soutar, 2001, S. 211). Darauf aufbauend soll das Image eines Unternehmens die Markeneinstellung stark beeinflussen (vgl. Suh & Yi, 2006, S. 153). In der Vershofenschen Terminologie wäre der Markennutzen ein Zusatznutzen, bei dem es um imaginär-symbolische Nutzenstiftungen geht (vgl. Gutsche et al., 2005, S. 640).

Nach der empirischen Studie von Gutsche und Partner scheint der Markennutzen der Treiber der Wiederkaufsabsicht zu sein. Damit stünden Unternehmen vor der Herausforderung, den Wandel zu erfassen und zu verstehen und das Leistungsversprechen entsprechend zu modifizieren (vgl. Gutsche et al., 2005, S. 653). AR und VR könnten daher diesen Markennutzen verstärken und so einen Zusatznutzen schaffen, der bspw. auf die von den Technologien ausgelösten Emotionen basiert. Zudem könnten die Technologien die Wiederkaufsabsicht der Konsumenten verstärken.

3.2. Benutzerakzeptanz von Informationstechnologien

Fetscherin und Lattemann beschäftigten sich 2008 in Ihrer Arbeit mit der Bewertung der Benutzerakzeptanz von virtuellen Welten durch die Anwendung des Technology Acceptance Model (TAM-Theorie). Im Rahmen dieser Arbeit wurden aus einer Online-Umfrage mit 249 Teilnehmern Hypothesen zur Benutzerakzeptanz und Akzeptanz virtueller Welten empirisch getestet. Im Allgemeinen besagt die TAM-Theorie, dass die Behavioral Intention to Use, also die Verwendungsabsicht, ein direkter Prädiktor für die Akzeptanz und Einführung neuer Technologien ist, was auch für virtuelle Welten sinnvoll ist. Von allen getesteten Faktoren scheint die wichtigste Determinante für die Einführung der virtuellen Welt der wahrgenommene Wert von Kommunikation, Zusammenarbeit und Kommunikationskanälen zu sein. Die Auswirkungen dieser Komponenten auf die Benutzerakzeptanz waren signifikant größer als die der nächstwichtigsten Determinante, der wahrgenommenen Nützlichkeit der virtuellen Welten. Weitere Einflussfaktoren, die für die Benutzerakzeptanz der virtuellen Welt relevant sind, sind soziale Normen, die Einstellung zur Technologie und sozialdemografische Faktoren. Geschäftsmodelle für virtuelle Welten sollen den Benutzern ein klares und attraktives Wertversprechen bieten. Insbesondere müssen sie sich mit der Notwendigkeit von Community-Building-Elementen wie Kommunikation, Zusammenarbeit und Kooperationsfunktionen befassen, um virtuelle Welten als neuen Kanal für die Interaktion mit Kunden voll ausnutzen zu können. Die Studie von Fetscherin und Lattemann bezog sich auf die virtuelle Welt Second Life und gibt zu verstehen, dass weitere Forschung zu weiteren Modellen durchgeführt werden sollte. Für diese Arbeit sind für die Benutzerakzeptanz der Kunden für die Anwendung von AR und VR insbesondere die Faktoren Bewusstsein über die Verwendungsabsicht, der wahrgenommene Wert sowie die grundsätzliche Einstellung der Kunden der Technologien gegenüber ausschlaggebend (vgl. für den gesamten Absatz Fetscherin & Lattemann, 2008, S. 231-240).

3.3. Theorie zur Kundenzufriedenheit über die Diskrepanz zwischen der vorherigen Erwartung und der tatsächlichen Leistung

Die Kundenzufriedenheit wird definiert als eine bewertende Zusammenfassung der (direkten) Konsumerfahrung, basierend auf der Diskrepanz zwischen der vorherigen Erwartung und der tatsächlichen Leistung, die nach dem Konsum wahrgenommen wird. Tse und Wilton erweitern die Literatur zur Kundenzufriedenheit, indem sie theoretisch und empirisch den Effekt der wahrgenommenen Leistung anhand eines zuerst von Churchill und Surprenant (1982) vorgeschlagenen Modells untersuchten und erweiterten. In dieser Arbeit wurde festgestellt, dass die wahrgenommene Leistung mehr Einfluss hat als andere Modellelemente, wie bspw. die Erwartung. Trotzdem haben die anderen Elemente signifikante, direkte Einflüsse. Die Studie zeigt, dass der Zusammenhang zwischen Erwartung und Zufriedenheit viel geringer ist und sich die Erwartung vor oder nach dem Produktversuch verschiebt. Diesbezüglich kann man analysieren, inwiefern AR und VR die wahrgenommene Leistung beeinflussen und die mögliche Diskrepanz zwischen der vorherigen Erwartung und der tatsächlichen Leistung verringern kann (vgl. für den gesamten Absatz Tse & Wilton, 1988, S. 204, S. 210).

3.4. Theorie zur Produktzufriedenheit über zusätzliche Affektquellen

Die Zufriedenheit ist nicht nur ein kognitives Phänomen. Sie umfasst vielmehr auch ein Element der Beeinflussung oder des Gefühls, indem sich die Verbraucher im Zusammenhang mit der Zufriedenheit subjektiv gut und im Zusammenhang mit der Unzufriedenheit subjektiv schlecht fühlen. Daher wird die Zufriedenheit in Kunden- und Produktzufriedenheit unterteilt. Die Produktzufriedenheit beinhaltet im Vergleich zur Kundenzufriedenheit weitere Aspekte, welche insbesondere hinsichtlich der Anwendung von AR und VR von Bedeutung sein könnten. Westbrook nennt in seiner Arbeit eine Reihe unterschiedlicher Affektquellen, die die Produktzufriedenheit der Verbraucher beeinflussen können. So nennt er als Erstes den Optimismus und Pessimismus, als Zweites vorübergehende affektive Einflüsse wie die Stimmung und als Drittes konsumentenbezogene affektive Einflüsse wie die Einstellung zum Unternehmen, Markt und Konsum. Die vierte und letzte Affektquelle ist das vorübergehend günstige oder ungünstige Gefühl des Verbrauchers, das sich im Zusammenhang mit Elementen des Konsumbereichs, wie zum Beispiel mit Produktangeboten oder mit Marketingpraktiken entwickelt und durch vorherige Erfahrungen des Verbrauchers hervorgerufen wird. Die Studie kam hinsichtlich der vier Affektquellen auf folgende Ergebnisse. Die Produktzufriedenheit hängt direkt von der Realisierung der Erwartungen sowie von der allgemeinen Lebenszufriedenheit ab und variiert umgekehrt mit der allgemeinen Unzufriedenheit der Verbraucher. Hingegen verändert sich die Produktzufriedenheit nicht direkt mit dem Optimismus oder Pessimismus oder direkt mit der Gunst der Stimmung. Im Hinblick auf den Einsatz von AR und VR könnten die vier Affektquellen, insbesondere die Beeinflussung der Gefühle, die durch den Einsatz solche Technologien hervorgerufen werden sowie der Optimismus oder der Pessimismus der Verbraucher, hinsichtlich der Anwendung der Technologien, von großer Bedeutung sein. Daher werden im weiteren Verlauf der Arbeit die Kunden- und die Produktzufriedenheit unter Berücksichtigung der vier Affektquellen getrennt voneinander analysiert (vgl. für den gesamten Absatz Westbrook, 1980, S. 49-52).

3.5. Ausarbeitung des Kriterienkatalogs

Neben den erläuterten Kriterien des Nutzens, der Benutzerakzeptanz und der Kunden- und Produktzufriedenheit wird noch die Integration der Technologie in bestehende Technologien eingeführt. Der Nutzen soll sich insbesondere auf die Unternehmerseite beziehen und in Verbindung mit der Produktzufriedenheit die (wahrgenommene) Qualitätserhöhung der Produkte und die damit einhergehende Abgrenzung zu Wettbewerbern untersuchen. Die Integration soll sich näher mit den Erfahrungen der Unternehmen bei der Einführung der Technologien und mit den Wechselwirkungen zwischen bestehenden Strukturen und neuen Technologien beschäftigen. Die Integration steht in einem unmittelbaren Zusammenhang mit der Benutzerakzeptanz, da auch dort die Kooperation der unterschiedlichen Kommunikations- und Vertriebskanäle relevant ist. Somit bilden der Nutzen, die Integration, die Benutzerakzeptanz, die Kundenzufriedenheit und die Produktzufriedenheit den Kriterienkatalog und die Grundlage der Analyse.

4. Experteninterviews als qualitativ empirische Studie

Im folgenden Kapitel wird die Methodik der Experteninterviews erläutert.

4.1. Methodisches Vorgehen – die Experteninterviews

Zu Beginn wird das methodische Vorgehen der Experteninterviews näher beschrieben. Anschließend wird der Begriff Experte genau definiert sowie der Ablauf des Interviews und wichtige Rahmenbedingungen erklärt, bevor die Auswahl der Interviewpartner begründet vorgestellt wird. Zum Schluss werden die Gedanken und die Struktur hinter dem Leitfaden aufgezeigt und die Auswertungsmethode der Daten erläutert.

Experteninterviews zählen zu der Methodik der qualitativ-empirischen Sozialforschung. Dabei beschreibt die empirische Sozialforschung den Prozess, einen bestimmten Ausschnitt der sozialen Welt zu beobachten und zu untersuchen, um mit Hilfe der gemachten Beobachtungen die Weiterentwicklung von Theorien zu fördern (vgl. Gläser & Laudel, 2010, S. 24). Qualitativ-empirisch bedeutet in diesem Zusammenhang, dass im Zuge der Forschung offenere Erhebungsverfahren genutzt werden und sich auf eine kleinere Zahl untersuchter Fälle berufen wird, anstatt mit einer großen Menge standardisierter Daten zu operieren, wie es bei der quantitativen Empirie der Fall ist. Dabei steht das Ziel, in Auseinandersetzung mit dem empirischen Material und dem eigenen theoretischen Vorwissen etwas Neues zu erfahren und theoretische Einsichten zu erlangen, im Vordergrund (vgl. Universität Bielefeld, 2006).

Die für das Erhebungsinstrument ausgewählten Interviewpartner gelten als Experten. Ein Experte kann dabei wie folgt definiert werden: „Experte beschreibt die spezifische Rolle des Interviewpartners als Quelle von Spezialwissen über die zu erforschenden sozialen Sachverhalte. Experteninterviews sind eine Methode dieses Wissen zu erschließen" (Gläser & Laudel, 2010, S. 12). Dabei ist im Einzelnen zu entscheiden, wer als Experte behandelt werden soll und wem somit ein spezifischer Status als Interviewpartner zugewiesen wird (vgl. Helfferich & Baur, 2019, S. 671). Experteninterviews dienen dazu, exklusives Wissen hinsichtlich eines bestimmten Themas und schwer zugängliche Informationen für die weitere Analyse zu erschließen (vgl. Kaiser, 2014, S. 5).

Die Experteninterviews im Rahmen dieser Bachelorarbeit wurden entweder per Videoanruf oder per Telefon durchgeführt und hatten eine Durchschnittsdauer von 25 Minuten. In beiden Fällen wurde lediglich der Ton aufgenommen, um die Inhalte und Ergebnisse im Rahmen dieser Arbeit festzuhalten. Die Struktur der Interviews war dabei teilstrukturiert: Sie lassen sich nicht quantitativ auswerten und sind an einen Leitfaden gebunden (vgl. Bogner, Littig & Menz, 2005, S. 17).

Die Basis des Interviews bildet eine Liste vorbereiteter Fragen. Für das Interview ist es wichtig, einen Kommunikationsprozess zu planen, der alle für die Untersuchung relevanten Informationen einbringt und sich darüber hinaus an den kulturellen Kontext des Gesprächspartners orientiert (vgl. Gläser & Laudel, 2010, S. 114).

Bei der Durchführung des Interviews ist nicht zu vernachlässigen, dass das Erkunden des Erfahrungsbereiches des Befragten im Vordergrund steht (vgl. Atteslander, 2010, S. 147). Der Interviewer hat das Gespräch so zu steuern, dass die erwarteten Informationen auch tatsächlich generiert werden können. Dies wird durch offene Fragen gewährleistet, damit die Experten nicht beeinflusst werden und ihrer Rolle als „Lieferant von Informationen" gerecht werden können (vgl. Kaiser, 2014, S. 2). Die offenen Fragen ermöglichen einen flexiblen und variablen Gesprächsverlauf, wobei der Interviewer nicht die Kontrolle des Gesprächs verlieren darf, um mit Hilfe der gewonnenen Informationen im Anschluss die Forschungsfragen beantworten zu können (vgl. Kaiser, 2014, S. 5).

4.2. Die Auswahl der Interviewpartner

Die Auswahl der Interviewpartner verlief anhand bestimmter Kriterien, die die Gesprächspartner erfüllen mussten, um als Experte klassifiziert werden zu können. Dabei war das erste Kriterium die Tätigkeit im B2B-Vertrieb. Wie eben definiert müssen die Experten dabei in einem bestimmten spezifischen Gebiet, dem Vertrieb, über einen überdurchschnittlichen Wissenstand verfügen. Dies ist erfüllt, da sämtliche Interviewpartner entweder im Vertrieb arbeiten oder täglich mit den Verantwortlichkeiten im Vertrieb in Berührung kommen. Darüber hinaus war es wichtig, Gesprächspartner aus verschiedenen Unternehmen aus unterschiedlichen Branchen zu finden, um eine Vielzahl an verschiedenen Perspektiven und Eindrücken zu erlangen und somit den B2B-Vertrieb flächendeckend repräsentieren zu können. Die Namen der Interviewpartner sowie die der Unternehmen werden anonym gehalten und anhand der folgenden Tabelle mit den Bezeichnungen A1 bis A6 hinsichtlich der Position, des Geschlechts und der Branche zugeordnet. In den Transkriptionen im Anhang wurden die Namen durch die jeweilige Bezeichnung und die Namen der Unternehmen mit den Abkürzungen U1 bis U6, orientiert an der folgenden Tabelle, anonymisiert (vgl. Anhang b bis g).

4.3. Der Aufbau des Leitfadens

Im Folgenden werden die Gedanken und die Struktur erläutert, die hinter dem Aufbau des Leitfadens stecken. Die Erstellung des Leitfadens folgt dem Prinzip, so offen wie möglich und so strukturierend wie nötig zu sein (vgl. Helfferich & Baur, 2019, S. 670). Dabei bildet nach der Einleitung die erste offen gestellte Frage des Hauptteils die Grundlage des Leitfadens, indem diese nach der bereits vorhandenen Anwendung bzw. alternativ nach dem geplanten Einsatz von AR und VR fragt.

Die Bearbeitung der wissenschaftlichen Arbeit und die Analyse verschiedener Quellen trugen zu der Erarbeitung

weiterer thematischer Aspekte und der Verfassung der relevanten Fragestellungen bei. Dabei wurden die Themenbereiche geclustert und hinsichtlich der Analyse anhand der Kriterien geordnet. Die Fragen wurden in die Themenblöcke Stärken und Schwächen, Ziele des Einsatzes, Implementierungshürden, Nutzen, Kunden- und Produktzufriedenheit sowie Akzeptanz der Kunden zugeordnet. Diese Bereiche wurden im nächsten Schritt in diese chronologisch sinnvolle Reihenfolge gebracht. Die Reihenfolge spiegelt den Prozess der Einführung der Technologien im Unternehmen und den damit verbundenen Gedankengang wider. Zum Schluss wurden die Interviewpartner nach ihrer Einschätzung hinsichtlich der Entwicklung der Technologien in fünf bis zehn Jahren gefragt. Die Fragen wurden neutral verfasst, um zu gewährleisten, dass die gewünschten Erkenntnisse und Auskünfte als Antwort zwar folgen können, ohne dabei jedoch eine bestimmte Antwort vorzugeben oder den Experten in eine thematische Richtung zu drängen. Anschließend wurde innerhalb der Themenblöcke eine Klassifizierung vorgenommen und zwischen übergeordneten, inhaltlichen Fragen und untergeordneten Nachfragen unterschieden. Letzteres wurde nur gestellt, wenn die Ergebnisse der ursprünglichen Fragen als nicht ausreichend erachtet wurden. Der Leitfaden für die Experteninterviews befindet sich im Anhang (vgl. Anhang a).

4.4. Die Inhaltsanalyse

Um die Experteninterviews für die Bachelorarbeit auswerten und die Informationen in die Arbeit integrieren zu können, wurden alle Interviews transkribiert. Diese Transkription erfolgte entsprechend den Regeln nach Dresing und Pehl. Diese beinhalten, dass Satzbrüche, Pausen und unverständliche Wörter gekennzeichnet, Wortschleifen angepasst und Stottern, Wortdoppelungen und Verständnissignale ausgelassen werden (vgl. Dresing & Pehl, 2018, S. 21 f.). Diese Regeln finden im Rahmen dieser Arbeit Anwendung, da der Fokus auf die inhaltlichen Aussagen der Interviewpartner liegt und Zögern, Wortfindungshemmungen oder Ähnliches irrelevant im Sinne der neuen Erkenntnisgewinnung sind. Diese Regeln wurden um die Transkriptionsregel nach Kruse ergänzt, die besagt, dass gleichzeitige Redebeiträge gekennzeichnet werden (vgl. Kruse, 2015, S. 354). Damit wird gewährleistet, dass die Chronologie des Interviews richtig wiedergegeben wird.

Die Experteninterviews wurden nach der qualitativen Inhaltsanalyse nach Mayring ausgewertet. Diese Methode stellt die Bewältigung großer Informationsmengen sicher und bleibt dabei qualitativ-interpretativ und kann latente Sinngehalte erfassen (vgl. Mayring & Fenzl, 2014, S. 543). Zudem folgt sie strengen Regeln und ist somit intersubjektiv überprüfbar (vgl. Mayring & Fenzl, 2014, S. 543). Dafür werden zunächst Textpassagen einzelnen Kategorien zugeordnet (vgl. Mayring & Fenzl, 2014, S. 544). Diese Kategorien stammen aus dem Kriterienkatalog aus dem dritten Kapitel und basieren auf den dort skizzierten Theorien. Hinzu kommen die Kategorien „Model free Findings, Chancen und Risiken der Technologien im Vertrieb" sowie die Zukunftsaussichten

Tabelle 1: Zuordnung der Interviewpartner (Quelle: Eigene Darstellung)

Bezeichnung	Position	Geschlecht	Branche und Unternehmensbezeichnung
A1	Vorstandsvorsitzender	Männlich	Holzwerkstoff- und Kunststofferzeugnissen (U1)
A2	Head of Global Marketing	Männlich	Hersteller von Elektro- und Druckluftwerkzeugen (U2)
A3	Stellvertretender Beiratsvorsitzender	Männlich	Hersteller von Schalungs- und Gerüstsystemen (U3)
A4	Ehemaliger Head of Group Division Sales Europe	Männlich	Sanitärtechnik (U4)
A5	Sales Excellence	Männlich	Pumpenhersteller (U5)
A6	Inhaber	Männlich	Medienunternehmen (U6)

der Interviewpartner. Die letzte Spalte beinhaltet Notizen über konkrete Anwendungsbeispiele der Unternehmen, die zur Beantwortung der Forschungsfrage herangezogen werden (vgl. Abbildung 1). Diese Kategorien, nach denen die Interviews ausgewertet werden, bilden die Gliederung der Analyse im fünften Kapitel, wo die Ergebnisse der Interviews aufbereitet werden, wohingegen die Einschätzungen der Interviewpartner hinsichtlich der zukünftigen Entwicklung im Rahmen der Diskussion im sechsten Kapitel erläutert werden. Im zweiten Schritt wird geprüft, ob die Textstellen mehreren Kriterien zugeordnet werden können (vgl. Mayring & Fenzl, 2014, S. 544). Zuletzt wird bei Kategorien mit zahlreichen Textstellen zwischen Gemeinsamkeiten und Unterschiede der Aussagen der Experten sowie hinsichtlich weiteren Unterkategorien, die ebenfalls in Kapitel 3 erläutert wurden, differenziert. Das Ziel der strukturierenden Inhaltsanalyse ist es, die Informationen unter vorher festgelegten Kriterien einzuschätzen und nachvollziehbar auszuwerten. Ein Beispiel der Auswertung bildet die Tabelle 2 und befindet sich im Anhang, wobei die Seitenzahlen sich auf die im Anhang der Arbeit befindenden Transkriptionen beziehen (vgl. Anhang h).

5. Die Analyse der Ergebnisse des Einsatzes von Augmented Reality und Virtual Reality anhand ausgewählter Kriterien

Dieses Kapitel bildet den Kern der Arbeit, der die theoretischen Grundlagen aus der Literatur sowie aufbauend auf den Theorien, die im dritten Kapitel erläutert wurden, nun mit den neu gewonnenen Erkenntnissen aus den Experteninterviews zusammenführt. Diese Zusammenführung findet in der Form einer Analyse anhand des herausgearbeiteten Kriterienkatalogs statt. Hierzu werden relevante und interessante Standpunkte der Interviews hinsichtlich der einzelnen Kriterien herausgearbeitet, die Meinungen der verschiedenen Gesprächspartner wiedergegeben, bei Bedarf verglichen und Differenzen sowie Gemeinsamkeiten aufgezeigt. Zuletzt werden, losgelöst von den in Kapitel 3 erläuterten Theorien, wei-

tere Erfahrungen und Einschätzungen der Experten hinsichtlich der Chancen und Risiken des Einsatzes der Technologien im Vertrieb dargestellt.

5.1. Nutzen

Alle Interviewpartner vertreten hinsichtlich der emotionalen Zusatznutzenstiftung, die unabhängig von funktionalen Vorteilen durch Emotionen ausgelöst wird, die gleiche Meinung. So bietet der Einsatz von AR und VR bspw. im Bauwesen die Möglichkeit, Fehler frühzeitig im Bauprozess zu erkennen, sodass die Projektdurchführung des Unternehmens für die Kunden zuverlässiger und sicherer wird. Die Technologien ermöglichen es den Unternehmen, dem Kunden ein vertrauenswürdiges und professionelles Bild von sich zu geben, indem Kunden das Unternehmen als innovativ ansehen. Der Experte des Unternehmens aus dem Sanitärbereich berichtet von den Auftritten auf medialen Messen, wo der Kunde wahrnimmt, wie sich das Unternehmen präsentiert. Seiner Meinung nach ist es wichtig, dass die Kunden nicht den Eindruck gewinnen, die Unternehmen wissen nicht, was sie in Bezug auf AR und VR tun. Insbesondere die Kunden, denen Innovationen wichtig sind und die bereits mit dem Einsatze solcher Technologien in Kontakt waren, dürfen nicht enttäuscht werden. Umgekehrt liegt in dieser Aussage die Chance, sich als Unternehmen zu präsentieren und einen nachhaltig positiven Eindruck bei dem Kunden zu hinterlassen. Die Argumentation des Experten des Herstellers von Schalungs- und Gerüstsystemen legt ebenfalls nahe, dass das Unternehmen durch den Einsatz von Holo Lens Brillen, also von high immersion VR, als kompetentester Partner wahrgenommen wird.

Bei der Frage, ob die Technologien einen Wettbewerbsvorteil generieren können, gab es Unterschiede in den Antworten der Experten. Während im Bauwesen Störstellen früh vorhergesagt und sich somit ein Vorteil gegenüber Wettbewerbern, die diese Technologien noch nicht einsetzen, geschaffen werden kann, tut sich der Experte A2 schwer, einen Wettbewerbsvorteil identifizieren zu können. Viele weitere Einflüsse, die sich direkt auf die Produkte beziehen, spielen

eine entscheidende Rolle. Es ist daher schwierig, den Wettbewerbsvorteil allein auf die Technologien zurückzuführen. Alle anderen Experten sind sich dennoch in dem Punkt einig, dass der Einsatz von AR und VR eine Möglichkeit darstellt, sich von Wettbewerbern zu differenzieren. Aufgrund der Tatsache, dass die Technologien noch nicht häufig branchenübergreifend Anwendung finden und laut der Experten des Pumpenherstellers und des Herstellers von Holzwerkstoff- und Kunststofferzeugnissen nicht Marktstandard sind, kann man sich gezielt von der Konkurrenz abgrenzen. Seiner Ansicht nach muss man den Einsatz der Technologien in Verbindung mit der Gesamtstrategie setzen, um sich bspw. als Premiumlieferant mit Innovationen von Wettbewerbern abzugrenzen. Dafür sind AR und VR eine geeignete Lösung. Der Experte A4 vertritt die gleiche Meinung und ergänzt, dass große Unternehmen weiter an dem Einsatz dieser Technologien arbeiten sollen, um sich ein Feld zu schaffen, sich aufgrund neuer Kompetenzen auf dem Markt zu behaupten. Dabei ist es auch von Bedeutung, wie die Technologien eingesetzt werden. Der Gesprächspartner des Medienunternehmens betont zudem, dass sie sich aufgrund der individuelleren Lösung, umgesetzt durch AR und VR, gegen zwei deutlich größere Unternehmen bei einer Auftragsausschreibung durchsetzen konnten.

Im Folgenden wird sich der Markennutzen näher angeschaut, der in den Nutzentheorien als bedeutsamer Treiber der Wiederkaufsabsicht gilt. Diesbezüglich ergaben die durchgeführten Experteninterviews ein klares Bild. So sind sich alle Experten einig, dass der Einsatz von AR und VR das Image des Unternehmens stärkt. Dabei unterstreicht der Experte A3, dass die Technologien zur Imageförderung richtig eingesetzt werden müssen. Diese Aussage wird vom Vertreter des Sanitärunternehmens bestärkt, der sagt, dass das Image bei falscher oder schlechter Anwendung solcher Technologien leiden kann. Kunden können bei falscher Anwendung den Eindruck gewinnen, dass das Unternehmen nicht auf dem neusten Stand der Entwicklungen ist und können das Vertrauen in die Produkte des Unternehmens verlieren. Zusätzlich stellt er fest, dass trotz dieses Risikos Unternehmen die Entwicklungen dieser Technologien mitgehen und fördern müssen, um ihr Image zu prägen. Und das ungeachtet der Annahme, dass der Einsatz der Technologien kein Geschäft generiert.

In einem letzten Schritt wird sich basierend auf den Nutzentheorien angeguckt, ob der emotionale Zusatznutzen, der Wettbewerbsvorteil sowie die Stiftung des Markennutzen durch die Technologien dazu führen, dass sich die Wiederkaufsabsicht der Kunden verändert. Nach den Meinungen der Experten ist vor allem die Imageförderung der in der Theorie in Kapitel 3 erläuterte Markennutzen, der durch den Einsatz der Technologien gefördert und laut den Theorien unmittelbar die Wiederkaufsabsicht fördern soll. Dessen ungeachtet vertreten die Experten jedoch unterschiedliche Meinungen, wenn es um eindeutige Rückschlüsse des Einsatzes der Technologien auf die Wiederkaufsabsicht geht. Der Experte A3 und der Experte A6 sind sich sicher, dass der Einsatz von AR und VR ausschlaggebend bei der Auftrags-

gewinnung sowie zahlreichen Folgeaufträgen ist. So betont Letzterer, dass nach Vorstellung der Technologien auf Messen oder im Rahmen von Präsentationen die Kunden begeistert waren und nach ersten Simulationen und Erfahrungen in der Regel immer weiter mit diesen Technologien arbeiten wollen. Demgegenüber stellen die Experten A2 sowie A4 fest, dass die Wiederkaufsabsicht sehr schwierig zu messen ist. Weitere Attribute, die sich lediglich auf die Produkteigenschaften beziehen, sind nach ihren Vermutungen ausschlaggebender.

Somit lässt sich die Analyse des Einsatzes von AR und VR im Hinblick auf den Nutzen wie folgt zusammenfassen. Die Experten erkennen affektive Zusatznutzen und die Möglichkeit, sich durch die Technologien vom Wettbewerb differenzieren zu können. Dabei ist jedoch die Qualität des Einsatzes entscheidend. Zudem lässt sich ein Markennutzen in Form der Imageverbesserung identifizieren. Diese Feststellungen führen jedoch nicht zu der eindeutigen Schlussfolgerung, dass der Einsatz von AR und VR zwangsläufig zu einer Verbesserung der Wiederkaufsabsicht der Kunden führt. Dies kann wiederum an der fehlenden Messbarkeit des Effekts liegen.

5.2. Integration

Wenn es um die Integration von AR und VR in bestehende Vertriebskanäle und deren Interaktion miteinander geht, vertreten die Experten mit Ausnahme des Vertreters des Herstellers von Schalungs- und Gerüstsystemen eine vergleichbare Meinung. Ihm zufolge gibt es im Bauwesen keine weiteren Probleme bei der Integration. Demgegenüber sagt der Experte A1, dass die Einführung solcher Technologien einige Voraussetzungen mit sich bringt und nennt als Beispiel die Notwendigkeit von qualitativ hochwertigem Fotomaterial für digitale Kataloge, die man in Verbindung mit AR verwenden kann, die digitale Kompetenzen der Mitarbeiter sowie die Infrastruktur, um solche Konzepte realisieren zu können. Der Vertriebsexperte aus dem Sanitärbereich fügt zusätzlich die technologischen Hürden hinzu, die aufeinander abzustimmen sind. Darüber hinaus sind die Machtverhältnisse in der Kunden-Lieferantenbeziehung von großer Bedeutung, damit klar ist, wer den Inhalt solcher Anwendungen bestimmen darf. Als Beispiel erläutert der Experte den dreistufigen Vertrieb, indem das Unternehmen Produkte an einen Zwischenhändler und nicht unmittelbar an den Endkunden verkauft. Der Gesprächspartner des Herstellers von Elektro- und Druckluftwerkzeugen verdeutlicht, dass bei Anwendung der Technologien die Interaktionsmöglichkeiten unterschiedlicher Vertriebskanäle besonders attraktiv sind. Als Beispiel nennt er einen QR-Code auf einem digitalen Katalog mit AR-Anwendungen, die den Kunden unmittelbar auf eine Internetseite führen, wo man den Kaufauftrag abschließen oder sich weitere Videos über das Produkt angucken kann. Der Experte A6, der bereits zahlreiche Einführungen der Technologien betreut hat, berichtet von einem Verlauf der Integration, der nie ganz reibungslos funktioniert. So ist die Größe des Unternehmens relevant und damit einhergehend die Frage, ob das Unternehmen in Zusammenarbeit mit einer Werbeagentur steht. Wenn man die Technologien bspw. in

die Unternehmensseite im Internet integrieren möchte und diese Internetseite von externen Werbeagenturen koordiniert wird, wird die Kommunikation und Absprache hinsichtlich der Einführung der neuen Technologien erschwert. Sollte dieser Zwischenschritt nicht erforderlich sein und kein externes Unternehmen mit in den Prozess eingebunden sein, berichtet der Experte von erfolgreichen Integrationen. Im anderen Fall ist ein weiterer Zwischenschritt in Form von Absprachen mit dem externen Unternehmen notwendig, der das Endergebnis jedoch nicht verändert.

Zusammenfassend ist somit die erfolgreiche Integration der Technologien in den Vertrieb eines Unternehmens mit gewissen Hürden wie den technischen Voraussetzungen und der inhaltlichen Absprache verbunden. Zeitgleich bietet sich Unternehmen die Chance, unterschiedliche Vertriebskanäle mit dem Einsatz der Technologien interagieren zu lassen.

5.3. Benutzerakzeptanz

Die Experten berichteten von ihren Erfahrungen und kommen hinsichtlich des Verständnisses für die Verwendungsabsicht der Kunden, wieso die Unternehmen die Technologien einsetzen und welchen Mehrwert dieser Einsatz generiert, zum gleichen Ergebnis. Die Kunden der Experten sind sich der Verwendungsabsicht des Einsatzes von AR und VR nicht vollumfänglich bewusst. Jeder Experte sprach den aktuellen Generationswechsel und den demografischen Wandel an. Kunden, die bereits lange im Geschäft sind, sind in der Regel diejenigen, die die Arbeit mit innovativen Technologien wie AR und VR ablehnen und den Nutzen nicht nachvollziehen können. Insbesondere Handwerker und Baustoffhändler sind bei solchen Themen sehr zurückhaltend. Demgegenüber stehen junge Berufseinsteiger, die diesen Technologien mit Offenheit und hoher Technikaffinität begegnen. Bspw. berichtete der Experte A3 von Ende 20-jährigen Bauleitern, die den Mehrwert der Technologien verstehen und sich für den Einsatz begeistern. Seiner Meinung nach muss eine Kundendifferenzierung erfolgen. Die allgemeine Schwäche der Akzeptanz sieht er jedoch nicht, da Kunden im Projekt frühzeitig über die Verwendungsabsicht aufgeklärt werden können. Der Gesprächspartner A2 erkennt eine derzeitige Erweiterung des Themas AR und VR und ein steigendes Bewusstsein für die Technologien in den Köpfen der Kunden. Sollten weitere große Unternehmen wie Ikea diese Technologien populärer machen, ist er sich sicher, dass das Verständnis der Kunden steigen und der Einsatz in vielen weiteren Branchen erfolgsversprechend sein wird. Dabei ist es von besonderer Bedeutung, die Vorteile für den Kunden transparent und eindeutig zu übermitteln. Als Beispiel nennt er die Promotion über soziale Netzwerke. Dieses Bewusstsein über den Nutzen des Einsatzes der Technologien ist bei den Kunden noch nicht im gewünschten Ausmaß vorhanden. Außerdem spricht er Kunden an, die große Fans der Produkte des Unternehmens sind und leicht für die Anwendung der Technologien zu begeistern sind. Somit bildet die grundsätzliche Einstellung der Kunden zum Unternehmen einen weiteren Einflussfaktor auf die Verwendungsabsicht. In die gleiche Richtung gehen die Aussagen des Experten

des Medienunternehmens. Seine Kunden sind zunächst aufgrund mangelnder Informationen hinsichtlich der Kosten der Technologien, der Datensicherheit und der benötigten Zeit sehr vorsichtig. Darüber hinaus ist es für Unternehmen schwierig, den Gedankensprung hinzukriegen, die gesehene Leistung der Technologien auf die eigene Anwendung im Unternehmen und den Nutzen zu beziehen. Insbesondere ist es im Marketing und im Vertrieb schwierig, Unternehmen von dem nachhaltigen Wert zu überzeugen. Dennoch sind Kunden spätestens nach ersten Demoanwendungen von den Technologien begeistert.

Somit ist laut den Experten und nach der TAM-Theorie die Verwendungsabsicht, die als direkter Prädiktor für die Akzeptanz und Einführung neuer Technologien gilt, einem Generationswechsel untergestellt und durchlebt einen Wandel. Ältere Kunden, die lange im Geschäft sind haben Probleme, den Mehrwert und den Nutzen zu verstehen, wohingegen jüngere Leute, die mehr und mehr auf den Arbeitsmarkt kommen, aufgrund ihrer Offenheit und hohen Technikaffinität ein größeres Verständnis für den Einsatz von AR und VR haben.

5.4. Kundenzufriedenheit

Die Kundenzufriedenheit umfasst die wahrgenommene Leistung der Kunden sowie die vorherigen Erwartungen der Kunden. Im Bauwesen berichtet der Experte hinsichtlich der wahrgenommenen Leistung von durchweg positiven Erfahrungen. Sobald die Kunden wahrnehmen, was die Technologien leisten, würden sie nichts anderes mehr wollen. Dies spiegelt sich in den Feedbacks wider, die das Unternehmen erhielt. Der Experte A6 berichtet von ähnlichen Ereignissen. Nach Kongressen und Veranstaltungen, wo das Unternehmen AR-Lösungen präsentiert hat, gab es positives Feedback der Unternehmen und Anfragen, solche technischen Lösungen für sie zu entwickeln. Zudem ergänzt der Experte, dass in einem nächsten Schritt nach Vorstellung eines Demokonzepts die Kunden mit „leuchtenden Augen und stellenweise offenem Mund" dastanden. Der Gesprächspartner A2 blieb in seinen Aussagen zurückhaltender. Sie entwickelten eine App, die mit Augmented Reality arbeitet. Dort gibt es bis jetzt nur wenig positives Feedback der Kunden. Zudem wünscht sich das Unternehmen, dass die Kunden die App häufiger und langfristiger verwenden. Sie vermuten, dass Kunden die App nach dem Kauf ihrer Produkte einmal ausprobieren und dann nicht mehr verwenden. Im Hinblick auf die vorherigen Erwartungen der Kunden kommt der Experte aufgrund der geringen Nutzung ihrer App auf den Entschluss, dass die Kunden den Einsatz der Technologien nicht immer benötigen und darüber hinaus nicht honorieren, dass so etwas benötigt wird. Die Kunden fragen nicht proaktiv nach dem Einsatz derartiger Innovationen. Auf der anderen Seite stehen die Meinungen der Experten aus dem Bauwesen sowie des Experten aus dem Medienunternehmen. Letzterer stellt fest, dass die Kunden zunächst nichts von dem Einsatz von AR und VR erwarten, weil die Kunden, wie im Kapitel der Benutzerakzeptanz erläutert, nichts über die Technologien wissen. Deswegen sind Kunden erstaunt, wenn sie erfahren, was mit

dem Einsatz der Technologien alles zu realisieren ist. Sie starten zunächst mit einer gewissen Anwendungsstruktur und erweitern den Einsatz Schritt für Schritt in verschiedenen Bereichen. Der Experte A3 berichtet, wie sich die Erwartungen seiner Kunden nach den ersten Berührungspunkten mit den Technologien verändert haben. Die Erwartungen sind gestiegen und die Kunden gewöhnen sich schnell an den professionellen Service, sodass der Druck auf das Unternehmen steigt. Dies bewertet der Experte jedoch nicht negativ, sondern betrachtet die steigende Erwartung als wünschenswerte Entwicklung.

Somit kristallisiert sich kein eindeutiges Ergebnis im Hinblick auf die Kundenzufriedenheit heraus. Während die einen Experten von durchweg positivem Feedback berichten, gibt es in einem anderen Unternehmen den Wunsch, Kunden würden bereits existierende Anwendungen der Technologien häufiger nutzen. Es lässt sich daher keine branchenübergreifende, allgemeingültige Aussage treffen.

5.5. Produktzufriedenheit

Die Produktzufriedenheit steht in engem Zusammenhang mit der grundsätzlichen Einstellung des Kunden zum Unternehmen sowie erlebten Erfahrungen mit den Technologien und ergänzt die Kundenzufriedenheit um affektive Quellen. Diesbezüglich gibt sich der Experte des Herstellers von Schalungs- und Gerüstsystemen aus dem Bauwesen positiv gestimmt. Seine Kunden sind den Technologien gegenüber optimistisch eingestellt. Gerade im Bereich komplexer Anwendungen hat er mit Kunden zu tun, die AR und VR bereits erlebt und entsprechend erwartungsvoll eingestellt sind. Nichtsdestotrotz differenziert er zwischen Kunden, die bereits Erfahrungen mit den Technologien haben und diejenigen Kunden, die bisher keine Berührungspunkte hatten und daher zunächst zurückhaltend sind. Jedoch hat er bis jetzt nicht erlebt, dass zurückhaltende Kunden die Anwendung der Technologien abgelehnt haben. Der Experte A2 beschreibt den Prozess, Kunden mit Hilfe der Technologien das Gefühl zu übermitteln, sich stärker mit dem Unternehmen identifizieren zu können und spricht in diesem Zusammenhang von „Fans". Zudem erlebt er eine große Bereitschaft seitens der Kunden, den Einsatz der Technologien anzunehmen und begründet dies mit offensichtlichen Verbesserungen der Leistungen und Produkten, die mit AR und VR im direkten Zusammenhang stehen.

Bezogen auf die grundsätzliche Einstellung des Kunden zum Unternehmen wägt der eben genannte Experte ab, in welchem Umfeld man sich bewegt und ob eine emotionale Bindung zum Produkt möglich ist. In seinem Unternehmen hält er dies für möglich, da die Produkte beim Kunden eine große Beliebtheit genießen. Er spricht von Kunden, die sich ihr Unternehmenslogo tätowieren lassen. Der Experte des Unternehmens aus dem Sanitärbereich verdeutlicht allerdings, dass dies nicht der Regelfall ist.

Zusammenfassend lässt sich für die Produktzufriedenheit festhalten, dass die Einstellung der Kunden zu der Anwendung der Technologien von Vorerfahrungen im Umgang mit AR und VR abhängt. Sind diese gegeben, berichten die Experten von Optimismus und Begeisterung. Zudem scheint es für Unternehmen, dessen Produkte bei den Kunden beliebt sind, die Möglichkeit zu geben, die emotionale Bindung zwischen Unternehmen und Kunden durch den Einsatz von AR und VR stärken zu können.

5.6. Model free Findings: Chancen und Risiken der Technologien im Vertrieb

Im Folgenden werden losgelöst von den vorgestellten Theorien weitere Ergebnisse der Experteninterviews ausgewertet, die sich in Bezug auf den Einsatz der Technologien im Vertrieb in Chancen und Risiken einteilen lassen.

Was die Chancen angeht, sind sich alle Experten sicher. Sie teilen die Einschätzung, dass der Mehrwert der Technologien schnell erkennbar ist. So kann sich der Kunde Produkte schnell grafisch vorstellen und in komplexen Anwendungsfällen und Projekten, insbesondere im B2B-Bereich, Unklarheiten aus dem Weg räumen. Unternehmen selbst haben die Möglichkeit, mit Hilfe von AR und VR Zeit und Kosten zu sparen. Außerdem können sich Unternehmen durch die Anwendung von AR und VR als Premiumlieferant gegenüber den Wettbewerbern differenzieren. Eine weitere gemeinsame Einschätzung der Experten bildet die an Tempo gewonnene Entwicklung in Bezug auf die Anwendung solcher Technologien. Jeder Experte identifiziert im Interview unabhängig voneinander die derzeitige Corona-Pandemie als Grund für die schnellere Entwicklung derartiger Innovationen. Der Gesprächspartner aus dem Medienunternehmen stellt zudem fest, dass insbesondere virtuelle Messen und Veranstaltungen im Zuge dessen zugenommen haben. Der Experte A4 geht in diesem Kontext noch einen Schritt weiter und betont, dass Unternehmen diesen Trend aktiv anstoßen und weiterentwickeln müssen. Die Technologien bieten in diesem Zusammenhang den Unternehmen die Chance, professionell und innovativ auf die derzeitigen, durch die Pandemie bedingten Einschränkungen wie den wegfallenden direkten Kundenkontakt zu reagieren.

Neben den einheitlichen Einschätzungen der Chancen ist besonders interessant, einen Blick auf die kontroversen Risikoeinschätzungen der Experten zu werfen. In einem Punkt ähneln sich zunächst alle Aussagen. Um die genannten Chancen realisieren zu können, ist die erfolgreiche Umsetzung entscheidend. Dabei betonen alle Experten, die Technologien „richtig" anzuwenden. Der Experte A3 fügt hinzu, dass der Einsatz von AR und VR gut durchdacht und strukturiert umgesetzt werden muss. Diesbezüglich grenzt der Experte aus dem Sanitärbereich die für die Umsetzung verantwortliche Person dahingehend ein, dass es sich hierbei um einen jüngeren Mitarbeiter handeln sollte. Mit dieser Aussage schlägt er eine erste Orientierungshilfe für das beobachtete Problem des Experten des Medienunternehmens vor, der feststellt, dass Unternehmen oft nicht wissen, wer für die Verwirklichung des Einsatzes der Technologien im Vertrieb zuständig ist. Der Hintergrund des Vorschlags einer jungen Leitung als verantwortliche Person ist der fortlaufende Generationswechsel, der laut den Experten ein weiteres Risiko mit sich

bringt. Um die Technologien „richtig" anzuwenden, muss, nach den Aussagen mehrerer Experten, die Kundenzielgruppe genau betrachtet werden. Dabei geht es nicht nur um das Alter, sondern auch um die IT-Kompetenzen des Kunden. Der Experte A1 stuft die Qualifikation bei Mitarbeitern des klassischen Handels als sehr gering ein und sieht darin die Schwierigkeit, die Anwendungen der Technologien maximal endverbrauchergerecht gestalten zu müssen, damit sie den gewollten Erfolg erzielen können. Zudem sieht er seine Branche grundsätzlich als rückständig an. Ähnliches bestärkt der Experte aus dem Bauwesen, der vor allem ein Risiko bei sich selbst im Unternehmen sieht, da die Technologien und die Einsatzmöglichkeiten nicht schnell genug weiterentwickelt werden. In die gleiche Richtung gehend argumentiert der Interviewpartner von dem Hersteller der Elektro- und Druckluftwerkzeugen, der nach einem „Triggerpunkt" für die App sucht, damit Anwender diese langfristiger und häufiger nutzen. Der Experte A4 grenzt die Bedeutung des Einsatzes der Technologien in seiner Meinung ein, indem er sagt, dass dieser nicht als Ersatz des persönlichen Kundenkontakts verstanden werden darf, sondern lediglich zur Unterstützung dient. Dem klassischen Außendienst schreibt er einen besonderen Stellenwert zu. Nichtsdestotrotz muss man die Technologien verwenden, ungeachtet der fehlenden Messbarkeit von Umsatzerfolgen.

Ein weiteres Risiko, das vor allem vom Experten A2 beschrieben wird, sind die hohen Startinvestitionen, die die Technologien mit sich bringen. Sein Unternehmen investiert derzeit viel Geld in AR und VR und tut sich schwer damit, zu beweisen, dass die Investitionen einen finanziellen Ertrag generieren. Im Gegenteil dazu vertritt der Experte des Medienunternehmens die Meinung, dass sich die Investitionen in die Technologien schnell amortisieren. Diese unterschiedlichen Aussagen sind die Folge des dargestellten Problems aller Experten, konkrete Umsatzerlöse nicht gezielt auf die Technologien zurückführen und messen zu können. Darüber hinaus verbirgt sich hinter dem Einsatz der Technologien ein weiteres Risiko. Die Chance, sich vom Wettbewerb zu differenzieren impliziert zeitgleich die Gefahr, bei schlechter Umsetzung einen negativen Eindruck bei den Kunden zu hinterlassen. Diesbezüglich erwähnt der Experte A4 einen Auftritt seines Unternehmens auf einer virtuellen Messe, wo viel Geld ausgegeben wurde, um dies zu vermeiden.

Im internationalen Vergleich beschreiben die Experten A2 und A6 alarmierende Entwicklungen für die Anwendung von AR und VR in Deutschland. Deutschland ist gegenüber asiatischen Ländern, europäischen Nachbarn und den USA weit hinterher. Erste Versuche, die Technologien anzuwenden, werden daher im Ausland durchgeführt. Deutschland ist neuen Innovationen gegenüber zu pessimistisch eingestellt und zurückgeblieben. So ist der „Rückstand so groß, dass wir den mit den [...] uns zur Verfügung stehenden Mitteln einfach nicht mehr aufholen können."

6. Zusammenfassung und Fazit

Im Zuge dieser Bachelorarbeit wurde der Einsatz von AR und VR im Vertrieb anhand konkreter Beispiele analysiert. Dabei wurden über Experteninterviews branchenübergreifende Meinungen und Know-how eingeholt. Im abschließenden Kapitel werden nun die Ergebnisse diskutiert, Implikationen für die Praxis und Forschung abgeleitet und Limitationen der Arbeit erläutert.

6.1. Diskussion und Zusammenfassung der Ergebnisse

In dieser Bachelorarbeit wurde bestätigt, dass Unternehmen bereits AR und VR erfolgreich in unterschiedlichen Branchen im Vertrieb einsetzen. Es lassen sich zentrale Aussagen, die von den Experten bestätigt wurden, zusammenfassen. Die wichtigste Erkenntnis ist, dass sich nach Angaben der Experten jeder Vertriebler bewusst ist, welchen Mehrwert die Anwendung der Technologien mit sich bringen kann. Im Zuge des derzeitigen Generationswechsels wird der Einsatz solcher Technologie mit steigendem Tempo zunehmen und die Benutzerakzeptanz der Kunden steigen. Die Arbeit erfasst zudem das Problem der schwierigen Messbarkeit des Einflusses von AR und VR auf die Umsätze. Zusätzlich ergeben sich Schwierigkeiten bei der quantitativen Beurteilung der Effekte, wie die mögliche Steigerung der Wiederkaufsabsicht, die durch den Einsatz der Technologien gefördert werden soll.

Im Hinblick auf die zentrale Forschungsfrage dieser Arbeit, welche Technologie (AR und/oder VR) sich vor dem Hintergrund ausgewählter Kriterien für den Einsatz in welchen Vertriebsbereichen eignet, kann festgestellt werden, dass die Experten in der Praxis derzeit kaum zwischen AR und VR unterscheiden und sich mit ihren Aussagen größtenteils auf die Anwendung beider Technologien beziehen. Den Grund dafür erkennt man in den Anwendungsbeispielen (vgl. Abbildung 1). Häufig kommen auf digitalen Messen, in digitalen Katalogen oder in den Apps sowohl AR als auch VR zum Einsatz. Da sich die Ergebnisse somit auf beide Technologien beziehen und häufig von Mischformen die Rede ist, wird eine Differenzierung im Sinne der Forschungsfrage erschwert. Außerdem zeigt sich, dass die Technologien vielseitig einsetzbar sind und nicht nur einem einzigen Vertriebsbereich zugeordnet werden müssen. Der Einsatz von high immersion VR kann bspw. in Form von VR-Brillen sowohl bei der Produktpräsentation und somit bei der Gewinnung von Kunden in der Verkaufspolitik als auch bei der Auftragsabwicklung in der Vertriebslogistik angewendet werden.

Dies spiegelt sich ebenfalls in den Zukunftsaussichten der Experten wider. Der Experte A6 ist sich, genauso wie jeder andere Experte, sicher, dass die Anwendung und die Einsatzmöglichkeiten von AR und VR in Zukunft rasant zunehmen und somit an Bedeutung gewinnen werden. Er betont, dass AR und VR kein kurzfristiger Hype sind und es in ein paar Jahren zu einer Symbiose kommen wird. Schwierig ist lediglich zu prognostizieren, wie schnell eine entsprechende umfassende Umsetzung geschehen wird.

Technologie		
VR	Mischform	AR

Vertriebsbereiche

	VR	Mischform	AR
Vertriebs-systempolitik		Virtuelle Messe	Digitale Kataloge
		QR-Codes, um in virtuelle Welten via Handy einzutauchen	
Vertriebs-kanalpolitik	VR-Room mit Avataren		
Verkaufs-politik	VR-Showrooms der Produkte	Extended Reality App in der Akquisitionsphase	App zur Produkt-vorstellung
	3D-Showrooms		"How to App"
	Einsatz von Holo-Lens Brillen zur Vorstellung des Projekts und Betreuung der Kunden	360° Modell zur Produktvorstellung	
Vertriebs-logistik			Einsatz von Googel-Glasses im Servicebereich

Abbildung 1: Die Zuordnungsmatrix der Anwendungsbeispiele von AR und VR in die jeweiligen Vertriebsbereiche (Quelle: Eigene Darstellung)

6.2. Implikationen für die Praxis

Diese Bachelorarbeit greift das eingangs beschriebene Problem der Unkenntnis der Unternehmen über mögliche Einsatzmöglichkeiten der Technologien und der daraus resultierende Potenziale auf und bietet einen Überblick über aktuelle Anwendungsbeispiele, die in der folgenden Abbildung dargestellt werden. Darauf aufbauend wird mit dieser Abbildung die zentrale Forschungsfrage anschaulich beantwortet.

Die Grafik zeigt auf der X-Achse, ob es sich um Augmented Reality, Virtual Reality oder zentral auf der Achse dargestellt um Mischformen beider Technologien handelt. Die Y-Achse zeigt, in welchen der in Kapitel 2 herausgearbeiteten Vertriebsbereichen die Anwendungsbeispiele eingesetzt werden. Dabei gibt es Beispiele, die in einem oder in mehreren Vertriebsbereichen vorkommen. Die abgebildeten Einsatzbeispiele stammen alle aus den Experteninterviews und werden so von den Unternehmen angewendet oder wurden als Pilotprojekt ausprobiert. Es ist zu erkennen, dass die Technologien entlang der Vertriebsbereiche vielseitig einsetzbar sind und somit in jedem Vertriebsbereich Unterstützung leisten können. Dennoch sind die meisten Anwendungen der Verkaufspolitik zuzuordnen. Dieses Ergebnis überrascht nicht, da die Technologien eine anschauliche Darstellung von Produkten ermöglichen und diese insbesondere bei der Gewinnung neuer Kunden ausschlaggebend ist. Der Vertriebskanalpolitik kann nur der VR-Room mit Avataren zugeordnet werden, der sowohl den unternehmensinternen Austausch als

auch den Austausch mit externen Vertriebspartnern auf ein neues Level hebt. Dieser Bereich scheint noch viel ungenutztes Potenzial zu haben, da viele Unternehmen die Technologien noch nicht in dem Umfang nutzen, dass die Kommunikation umgestellt werden kann.

Die Grafik schafft somit einen Überblick über die aktuelle Situation möglicher Einsatzfelder der Technologien und beantwortet die Forschungsfrage, welche Technologie in welchen Vertriebsbereichen eingesetzt wird. Dabei gelten die Ergebnisse unabhängig von der Branche für Unternehmen, die primär im B2B-Bereich tätig sind, sodass ein gewisses Maß an Generalisierbarkeit hergestellt wurde. Diesen Unternehmen ist zu empfehlen, mit ersten Anwendungen der Technologien in der Verkaufspolitik anzufangen, um ihre Produkte anschaulicher zu präsentieren und diese Anwendungen Schritt für Schritt in unterschiedlichen Vertriebsbereichen auszubauen.

6.3. Limitationen der Arbeit

Die Stärke dieser Bachelorarbeit liegt in der Zusammensetzung theoretischer Grundlagen mit dem Expertenwissen aus der Praxis, sodass der Wissensstand über den Einsatz von AR und VR im Vertrieb erweitert werden konnte. Dabei beziehen sich die Ergebnisse lediglich auf den B2B-Bereich. Somit konnten die Ziele, die zentrale Forschungsfrage beantworten und einen Überblick über den aktuellen Stand der Anwendungen zu schaffen, erreicht werden. Diesbezüglich muss man jedoch einschränkend festhalten, dass nicht wie

erwartet und in der Forschungsfrage formuliert, eindeutig zwischen Einsatzmöglichkeiten von AR und davon abgrenzend von VR analysiert werden konnten, da in der Praxis häufig Mischformen existieren. Der Rahmen dieser Arbeit reicht nicht aus, um diese häufig auftretende Mischform von AR und VR, die sog. Mixed Reality, näher zu analysieren.

Die aus Kapitel 5 gewonnen Erkenntnisse zeigen eindeutig, dass weiterhin Unsicherheiten hinsichtlich des Einsatzes von AR und VR existieren. Dafür ausschlaggebend ist insbesondere das geringe Wissen über die Technologien der Allgemeinheit. Zudem wurden neue Probleme offenbart, die durch die Bachelorarbeit nicht gelöst werden konnten. Mehrere Experten schätzen Deutschland international hinsichtlich der Anwendung der Technologien rückständig und langsam ein. Daneben konnte kein eindeutiges Zukunftsszenario ausgemacht werden, wie schnell sich der Einsatz von AR und VR weiterentwickeln wird.

6.4. Implikationen für die Forschung

Basierend auf der aktuellen wissenschaftlichen Literatur und dem externen Expertenwissen wird in dieser Arbeit ein erster Überblick über den Einsatz von AR und VR im Vertrieb gewonnen und basierend auf theoretisch hergeleiteten Kriterien umfassend analysiert. Somit dient diese Arbeit Unternehmen als Entscheidungshilfe für Anwendungsfelder der Technologien im Vertrieb. Die zukünftige Forschung sollte an diesem Punkt ansetzen und den Wissensstand über die Erkenntnisse aus den Erfahrungen der Anwendung von AR und VR in der Praxis erweitern, damit ein tieferes Verständnis über AR und VR gewährleistet wird und das Thema an Popularität gewinnt. Das ist die Grundvoraussetzung, um das volle Potenzial der Technologien entfalten zu können.

Darüber hinaus zeigt die Arbeit das Problem der fehlenden Messbarkeit des Einflusses von AR und VR auf die generierten Umsätze der Unternehmen. Die zukünftige Forschung sollte sich mit dem Problem auseinandersetzen und versuchen, Methoden zu entwickeln, die eine entsprechende Messbarkeit der Anwendung von AR und VR ermöglicht. Durch die Entwicklung einer solchen Methode können weitere Rückschlüsse auf den Erfolg der Technologien geschlossen werden. Wenn das Bewusstsein für die Chancen und Möglichkeiten der Technologien von Unternehmen und Kunden steigt, ist mit einer zunehmenden Anwendung zu rechnen. Dafür ist es erforderlich, dass die Forschung ihren entsprechenden Anteil leistet.

Literatur

Atteslander, P. (2010). *Methoden der empirischen Sozialforschung* (13. Aufl.). Berlin, Erich Schmidt Verlag.

Bellini, H., Chen, W., Sugiyama, M., Shin, M., Alam, S. & Takayama, D. (2016). *Goldman Sachs Insights*. https://www.goldmansachs.com/insights/pages/technology-driving-innovation-folder/virtual-and-augmented-reality/report.pdf. (Abfrage: 22.04.2021)

Bogner, A., Littig, B. & Menz, W. (2005). *Das Experteninterview: Theorie, Methode, Anwendung*. Wiesbaden, VS Verlag für Sozialwissenschaften.

Burdea, G. C. & Coiffet, P. (2003). *Virtual Reality Technology* (2. Aufl.). New York, Wiley-IEE Press.

Busch, R., Fuchs, W. & Unger, F. (2008). *Integriertes Marketing. Strategie, Organisation, Instrumente*. Wiesbaden, Gabler.

Churchill, G. A. & Surprenant, C. (1982). An investigation into the determinants of customer satisfaction. *Journal of Marketing Research, 19* (4), 491–504.

Deloitte. (2020). *Deloitte - Technology, Media, and Communication*. https://www2.deloitte.com/. (Abfrage: 20.04.2021)

Deutsch, M., Ebert, D., Gracht, Heiko, L., Petra & Gerards, M. (2016). *KPMG - Home Seite*. https://home.kpmg/content/dam/kpmg/pdf/2016/04/virtual-reality-exec-summary-de.PDF. (Abfrage: 22.04.2021)

Dresing, T. & Pehl, T. (2018). *Praxisbuch Interview, Transkription & Analyse. Anleitungen und Regelsysteme für qualitativ Forschende* (8. Aufl.). Marburg, Dr. Dresing & Pehl GmbH.

Fetscherin, M. & Lattemann, C. (2008). User acceptance of virtual worlds. *Journal of Electronic Commerce Research, 9* (3), 231–242.

Gläser, J. & Laudel, G. (2010). *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen*. Wiesbaden, VS Verlag für Sozialwissenschaften.

Guenzi, P. & Habel, J. (2020). Mastering the Digital Transformation of Sales. *California Management Review, 62* (4), 57–85.

Gutsche, J., Herrmann, A., Huber, F., Kressmann, F. & Algesheimer, R. (2005). Die Wirkung funktionaler, emotionaler und relationaler Nutzendimensionen auf die Markenloyalität. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung, 57* (7), 638–657.

Helfferich, C. & Baur, N. (2019). Leitfaden- und Experteninterviews. In *Handbuch Methoden der empirischen Sozialforschung* (S. 669–686). Wiesbaden, Springer Fachmedien.

Hilken, T., de Ruyter, K., Chylinski, M., Mahr, D. & Keeling, D. I. (2017). Augmenting the eye of the beholder: exploring the strategic potential of augmented reality to enhance online service experiences. *Journal of the Academy Marketing Science, 45* (6), 884–905.

Hofbauer, G. & Hellwig, C. (2016). *Professionelles Vertriebsmanagement. Der prozessorientierte Ansatz aus Anbieter- und Beschaffersicht*. Erlangen, Publicis.

Homburg, C., Jense, O. & Krohmer, H. (2008). Configurations of Marketing and Sales: A Taxonomy. *Journal of Marketing, 72* (2), 133–154.

Kaiser, R. (2014). *Qualitative Experteninterviews: Konzeptionelle Grundlagen und praktische Durchführung*. Wiesbaden, Springer Fachmedien.

Kaplan-Rakowski, R. & Gruber, A. (2019). Low-immersion versus high-immersion virtual reality: Definitions, classification, and examples with a foreign language focus. In *Proceedings of the innovation in language learning international conference*. Florence, Pixel.

Kober, S. (2020). *Digitalisierung im B2B-Vertrieb: Ergebnisse verbessern mit digitalen Tools – Impulse zur Entscheidung und Umsetzung*. Wiesbaden, Springer Fachmedien.

Kruse, J. (2015). *Qualitative Interviewforschung: ein integrativer Ansatz*. Weinheim, Beltz Juventa.

Lässig, R., Leutiger, P., Fey, A., Hentschel, S., Tornier, S. & Hirt, F. (2015). *Die digitale Zukunft des B2B-Vertriebs*. www.rolandberger.com. (Abfrage: 30.04.2021)

LaValle, S. M. (2017). *Virtual reality*. Cambridge University Press.

Lee, E. A.-L. & Wong, K. W. (2014). Learning with desktop virtual reality: Low spatial ability learners aremore positively affected. *Computers & Education, 79*, 49–58.

Maas, M. (2021). *Praxiswissen Vertrieb: Berufseinstieg, Tagesgeschäft und Erfolgsstrategien*. Wiesbaden, Gabler Verlag.

Mayring, P. & Fenzl, T. (2014). Qualitative Inhaltsanalyse. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 543–556). Wiesbaden, Springer Fachmedien.

Milgram, P. & Kishino, F. (1994). A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information Systems, E77-D (12*, 1321–1329.

Rouziès, D., Anderson, E., Kohli, A. K., Michaels, R. E., Weitz, B. A. & Zoltners, A. A. (2005). Sales and Marketing Integration: A Proposed Framework. *Journal of Personal Selling & Sales Management, 25* (2), 113–122.

Sherman, W. R. & Craig, A. B. (2002). *Understanding virtual reality: Interface, application, and design*. San Francisco, CA, Morgan Kaufmann Publishers Inc.

Steffen, J. H., Gaskin, J. E., Meservy, T. O., Jenkins, J. L. & Wolman, I. (2019). Framework of Affordances for Virtual Reality and Augmented Reality. *Journal of Management Information Systems, 36* (3), 683–729.

Steiner, J. (2016). *Praxishandbuch Vertrieb: Sales Management in der Konsumgüterindustrie*. München, UVK Verlag.

Suh, J.-C. & Yi, Y. (2006). When Brand Attitudes Affect the Customer Satisfaction-Loyalty Relation: The Moderating Role of Product Involvement. *Journal of Consumer Psychology, 16* (2), 145–155.

Sweeney, J. C. & Soutar, G. N. (2001). Consumer perceived value: The development of a multiple item scale. *Journal of Retailing, 77* (2), 203–220.

Tse, D. K. & Wilton, P. C. (1988). Models of consumer satisfaction: An Extensive. *Journal of Marketing Research, 25* (2), 204–212.

Universität Bielefeld. (2006). *uni-bielefeld.de*. https://www.uni-bielefeld.de/Universitaet/Einrichtungen/Fakultaeten/Paedagogik/fobika/aqua/forschungslogiken.html. (Abfrage: 20.04.2021)

Vershofen, W. (1959). *Die Marktentnahme als Kernstück der Wirtschaftsforschung*. Berlin, Heymann.

Wedel, M., Bigné, E. & Zhang, J. (2020). Virtual and augmented reality: Advancing research in consumer marketing. *International Journal of Research in Marketing, 37* (3), 443–465.

Westbrook, R. A. (1980). Intrapersonal Affective Influences on Consumer Satisfaction with Products. *Journal of Consumer Research, 7* (1), 49–54.

Winkelmann, P. (2013). *Marketing und Vertrieb: Fundamente für die marktorientierte Unternehmensführung* (8. Aufl.). München Oldenbourg, Oldenbourg Verlag.

Wiswede, G. (2012). *Einführung in die Wirtschaftspsychologie* (5. Aufl.). München, Reinhardt.

Blessing or Curse? The Influence of Neobrokers on the Investment Behavior of Young Investors

Maximilian Janussek

Technische Universität München

Abstract

My thesis addresses the topic related to the impact of neobrokers on the investment behavior of investors. I deal with the questions of which target groups are particularly attracted by neobrokers, for which investment strategies neobrokers are primarily used, and to what extent the design of the neobroker applications plays a significant role in the investment decisions of its users. Based on my online-based questionnaire, it can be determined that neobroker customers are predominantly younger and willing to take more risk compared to customers of branch or direct banks. Moreover, neobrokers are used for short-term investments and not for retirement planning. Here, the design of neobroker applications has a decisive influence on the trading behavior of its users. Not only is the risk of shares assessed differently due to the representations of the stock prices within the neobroker applications but buys as well as sells are carried out more frequently using neobrokers than in comparison to traditional broker providers. Ultimately, I also show that a learning section including a knowledge check of the newly acquired financial expertise within the neobroker applications is perceived as helpful by its users as it is for customers of traditional financial service providers.

Keywords: Neobroker; Trading; Fintech; Attention-grabbing.

1. Introduction

Since the early days of stock markets, mankind has repeatedly had to deal with stock market crashes. Such crises began as early as 1637 in form of the tulip bubble, better known as *Tulipmania* (Garber, 1990). This first speculative bubble was followed by several others, such as the *Twin Crises* of 1720 (Bruner & Miller, 2020) or the *Black Thursday / Black Tuesday* in 1929 (White, 1990). Probably the most famous crash of modern times was, until recently, the *Global Financial Crisis* in 2008, which was triggered by the collapse of the housing bubble (Friedman & Friedman, 2009; Holt, 2009). However, the latest crash at the beginning of 2020 was not due to an economic bubble, but to a global pandemic initiated by the coronavirus, which had an immense impact on the entire world. The peculiarity of this shock compared to all the previous ones is that it resulted from public health concerns rather than economic conditions (Albuquerque, Koskinen, Yang, & Zhang, 2020) and caused financial markets around the world to take rapid and very large falls. But it was not only those markets that had to contend with the pandemic. People, too, have been marked by restrictions and new challenges that entailed changes in

lifestyles and behaviors. On the one hand, vacations had to be canceled, public life had to be scaled back and individuals spent more time at home. On the other hand, however, these restrictions helped savers to have more time and, above all, money at their disposal. This situation was directly reflected in the stock market. Benefiting from the sharp drop in stock market prices, a persistently low interest rate policy, and a lack of alternative options, the German stock market, for example, experienced a real upswing. According to a study by the Deutsches Aktieninstitut (Balonier, Di Dio, & Fey, 2021), around 2.7 million more people in Germany saved in shares, equity funds, or equity-based ETFs in the Corona year 2020 compared to the previous year, bringing the total number of citizens participating in the stock market to almost 12.4 million. But it was not only the stock market that has experienced this veritable boom since the pandemic. So-called neobrokers are gaining in popularity, too, especially among the younger generation. This phenomenon is primarily due to the simplicity of the most mobile applications and the ability to gain low-threshold access to trading financial products for little to no cost (Barber, Huang, Odean, & Schwarz, 2020). The media attention, through oc-

casions such as the short squeeze of GameStop Corporation, promotes the ever-growing demand of neobrokers as well. Internet forums such as the subreddit r/Wallstreetbets are attracting more and more young people, who also want to participate in the stock market and benefit from it. Here, the primary objective is not to invest for the long term, but rather to make quick profits, which entail a high level of risk. Neobrokers offer such high-risk investments as trading warrants or cryptocurrencies, which makes them one of the most significant places to go for exactly these risk-taking and inexperienced investors. The reduced and clear design of the neobroker applications makes trading even more fun, which significantly lowers the barrier of investing. The advent of commission-free trading apps has massively expanded the possibilities of investing and changed its nature as a result (Chaudhry & Kulkarni, 2021).

The playful approach to trading financial products begs the question of what influence neobrokers have on the investment behavior of young investors in particular, and whether this influence should be seen as an advantage or a disadvantage. The purpose of my following thesis is to answer this specific question based on an online questionnaire. With this in mind, the central question motivating my thesis can be broken down into three further sub-questions, all of which are addressed in the following chapters:

1. Which target groups are addressed by neobrokers and what are the decisive reasons for this?
2. For which investment strategies are neobrokers used by their customers and what are the reasons for doing so?
3. Do the neobroker applications and their appearance have an impact on the investment behavior of investors?

These sub-questions form the general framework for the forthcoming elaboration of my thesis. To help me answer them, my paper is divided into six distinct sections. In the first part, I define the terms relevant for the study and present the existing theoretical knowledge related to the topic of neobrokers. This theory then allows me to derive the underlying hypotheses in the next chapter. These hypotheses deal with the influence of neobrokers on the trading behavior of investors and draw first conclusions on whether this influence is to be evaluated positively or negatively. In the following part, I take a closer look at the methodological approach of my thesis. I clearly explain the research environment, the study participants, and the measures related to the survey carried out. Building on this methodological aspect, I then draw conclusions in the succeeding chapter. Here, I present the results of my study and review the hypotheses that were previously formulated to show the extent to which they can be confirmed or not. These findings subsequently enable me to introduce the following discussion. In this chapter, I review both the positive results of the qualitative study as well as the limitations and methodological criticism. Due to these limitations, I additionally address future improvements and suggestions in the concluding part of the penultimate chapter. In the final section of my thesis, I once again present the

key messages and provide an outlook for the future where further research will be necessary.

2. Theoretical Foundation

In the following chapter, I summarize the results of my literature review. I start the section with a brief overview of the development of neobrokers and an introduction to the main terminologies about this subject. Subsequently, I present the theoretical background regarding the previously outlined objectives and thus attempt to answer the research questions based on theoretically developed hypotheses, which I derive in the following chapter. Answering these questions allows me to justify my study described below and achieve the goals I have set.

The existing theory, which I examine in the following sections, is primarily concerned with the issues surrounding Fin-Tech and, in particular, neobrokers. With a pragmatic view of the research objectives, I then discuss the psychological influences of neobrokers on investors based on the existing literature. These influences are represented by topics such as gamification or attention-grabbing. Having found the theoretical path to the research objectives of my thesis, I proceed with my analysis of the most important concepts in the field of neobrokers and their influences on investors. Finally, I summarize the analyses I conducted to provide a systematic review of the literature available and identify research gaps. From this, I can derive the guiding hypotheses for my thesis, presented in chapter 3.

2.1. Development of Neobrokers

Before the development of the internet, investors had to conduct their transactions with their brokers by phone, which was very time-consuming and costly. This changed rapidly with the advent of new technologies and the introduction of the World Wide Web. As personal computers and laptops began to proliferate in the mid to late 1990s, the modern investor, as well as online trading with notably lower commission rates, arose (Barber & Odean, 2001). In response to the lowered costs of brokers like E*Trade and Ameritrade, Charles Schwab Corporation, the trade volume leader at the time with 5.6 million active customer accounts, offered a flat-rate online trading commission. The idea was to create an incentive to trade online with reduced services instead of by phone or in person (Melnick, Nayyar, Pinedo, & Seshadri, 2000). The urge for online trading continued to grow and commission cuts persisted over the years, making competition even more intense. The first neobroker founded in 2013, Robinhood, shook up the entire market with its strategy and changed it once again. As the first broker to offer commission-free trading via a mobile and gamified app, Robinhood transformed the entire investment landscape, becoming a pioneer in the field (Barber et al., 2020). The goal of this neobroker was to provide access to financial markets for everyone through its zero-cost strategy and user-friendly

design. With its anti-Wall Street movement, the online discount broker tried to democratize online trading for the public, as they were left out from traditional brokers due to low and unattractive profit margins. The concept was well-received by the public, giving Robinhood 13 million users by 2020 and a market valuation of $11.2 billion (Tan, 2021). This success prompted other imitators to follow suit. In Germany, providers such as Trade Republic or Scalable Capital have also picked up on this growing trend and are attracting more customers to their business model, each with their own little special features such as copy trading or the purchase of cryptocurrencies.

In my thesis, I mainly focus on two neobrokers. First, I select Robinhood because it is the largest neobroker on the market and thus has the largest share of the existing literature. Second, I choose the German neobroker Trade Republic as most of the participants in my study are based in Germany and a connection can be established between them and the broker.

2.2. Definition of Terms

To gain a better understanding of the subject matter of my thesis, a detailed overview of the most important terminology is required. These technical terms are closely related to the topic and help clarify my underlying questions. I list and define these central concepts in the following subsections.

2.2.1. FinTech

There are several approaches by various authors to define the term FinTech (Zavolokina, Dolata, & Schwabe, 2016). In general, FinTech is an abbreviation and is composed of the terms *financial* and *technology*. According to Gomber, Koch, and Siering (2017), FinTech generally describes the combination of modern, primarily internet-related technologies and well-established businesses in the financial sector. A slightly different approach defines FinTech as an interdisciplinary topic that combines financial, technology, and innovation management. In concrete terms, FinTech applies to all innovative ideas which enhance financial service processes through suggesting technology solutions for various business situations, where the ideas may also lead to the creation of new business models as well as new companies (Leong, 2018).

2.2.2. Neobroker

There is no single valid definition for the term neobroker. It is possible, though, to derive characteristics of neobrokers that distinguish them from classic banks or other online brokers. Neobrokers are generally a new breed of online brokers from the FinTech scene aiming to put pressure on traditional brokers, especially with low-cost smartphone trading. To achieve this, neobrokers significantly narrow the trading spectrum and scope of services, unlike traditional banks, to focus on specific areas such as low cost (Fischer, Hübner, & Bulis, 2020).

2.3. Related Literature

Retail trading is on the rise, which has led to an increase in research interest in recent years. Especially when many small investors can join forces to decisively influence the stock market, particularly in short term, there is a great desire to find explanations for such incidents (Barber, Odean, & Zhu, 2009). Previous research has already focused on the reasons for the occurrence of such phenomena. These theories are of great importance for my study. To analyze the influence of neobrokers on the trading behavior of young investors, I first have to explain the psychological concepts responsible for this. Two of them are particularly important here: Attention-Grabbing and Gamification. In the following two sections, I will discuss these two concepts in more detail to provide a general understanding of these terms.

2.3.1. Attention-Grabbing

Attention is a scarce cognitive resource because each individual's attention is limited (Dukas, 2004). To take advantage of this limited attention, various methods can be used to influence a person's decision in a way that benefits oneself. These kinds of findings are also known from the FinTech scene, as one can benefit from the use of various attention-grabbing methods here as well. For example, Barber and Odean (2008) find that stocks that attract the attention of retail investors, either because they are mentioned in the media or because the stocks exhibit obvious fluctuations in price or trading volume, tend to be purchased preferentially. Here, the literature makes it clear that this primarily affects inexperienced investors. These investors have not yet developed clear criteria by which they base their trading decisions. Thus, they are more influenced by biases (Greenwood & Nagel, 2009) and attention-grabbing (Seasholes & Wu, 2007) that lead them to chase returns. The frequency of Google searches can also be used as an additional indicator of investor attention. On this basis, it can be shown that an increased search volume, also due to attention-grabbing, can predict a short-term increase in stock prices (Da, Engelberg, & Gao, 2011).

This begs the question of how exactly FinTech companies, most notably online brokers, are influencing the investor's attention. On the one hand, brokers such as Robinhood, for example, attempt to target the investor via push notifications in a way that involuntarily diverts attention, regardless of the investor's goals, intentions, or awareness (Theeuwes, 2010). On the other hand, the same broker offers, among other things, a *Top Mover* list, which includes 20 stocks with the highest and lowest daily returns (ten each). These shares have experienced the highest price changes in percentage terms from the previous day compared to all others. With the help of such lists, the broker manages to draw the attention of its investors to the listed companies (Stein, 2020).

2.3.2. Gamification

Gamification is a widely used method to nudge people's perceptions, attitudes, or even behavior in an intentional di-

rection, to enhance their productivity and engagement. Deterding, Dixon, Khaled, and Nacke (2011) describe gamification as the integration of video game elements to improve customer experience and engagement in non-game systems. A shorter, general definition describes gamification as "the use of game design elements in non-game contexts" (Deterding et al., 2011, p. 9). The goal of gamification is not to give rewards for one-time actions, but to improve the user experience to a point where a proper mix of frustration and pride keeps the customer voluntarily engaged for a long time. A good implementation of gamification requires a deep understanding of proper graphical and emotional representation (Sironi, 2016).

Gamification can therefore be an attractive method for FinTech companies to influence individuals and their investment behavior. The use of gamification is no longer a novelty for neobrokers either. Trading apps like Robinhood or Tarde Republic pay special attention to the simple and straightforward design of their applications, as well as a slick user interface, to make trading easy and playful. The low to no costs and the nearly immediate opening of an account further strengthen the playful environment. These neobrokers make no secret of this fact, as Trade Republic makes clear in an advertising slogan. Trading is described as so simple that even while commuting on the bus, you can quickly and easily trade complex financial products (van der Heide & Želinský, 2021). Another example of gamification in neobroker applications is the use of waitlists. Robinhood offers its users the opportunity to move up these lists by actively using the app or inviting friends to get faster access to new products and is currently promoting its latest feature, crypto wallets, with such a campaign (Robinhood, 2021b). Swiping within the neobroker applications to confirm purchases is a further implementation of gamification. This gesture is a ubiquitous movement especially among the younger generation who navigate their mobile life through swipe gestures (Tan, 2021). In addition, Robinhood takes advantage of partly visual features as found in mobile games. Once you deposit your first $10 or refer a friend, you will receive a free share by playing a scratch game. Additionally, a confetti animation appears when you have executed a trade or convinced a friend to use the app as well.

Again, users are rewarded for recommending the application to friends, not with free shares, but with a certain amount of money that both parties receive when the friend makes his first trade (Trade Republic, 2021).

However, this playful design of neobrokers, also caused by the GameStop incident, meets with much criticism and draws consequences. The problem lies mainly in the fact that people enjoy playing games and continue to do so despite negative experiences to improve themselves constantly (Sironi, 2016). Stubborn failure, though, can have devastating effects when it comes to the real world outside of video games. Free investment apps, which make it easier than ever to invest on the go, are no help in getting away from losses that are not just financial in nature.

2.4. Results of Previous Research and Research Gaps

In the previous sections, I introduced and explained the general theories that are relevant in the context of neobrokers. In the following subchapters, I revisit these concepts and present their implications and influences on the investor based on existing research. From this presentation, research gaps emerge, which I will identify in the final part of the section, highlighting the relevance of my thesis.

2.4.1. Attention-Grabbing

Retail traders do not trade based on statistics or fundamentals and valuations, as shown in several studies (Barber & Odean, 2008; de Long, Shleifer, Summers, & Waldmann, 1989), but respond to noises or different signals about financial markets and prices of financial products that catch their attention. Those noise traders engage in trading for non-informative reasons and therefore have no informational advantages (Foucault, Sraer, & Thesmar, 2011). It is precisely these investors that neobrokers are targeting with their attention-grabbing strategies. Neobrokers encourage their customers to actively trade financial products through the above-mentioned Top Mover lists, the significantly reduced and simplified design, and thanks to push notifications with great success. It is no surprise that users of Robinhood trade nine times more stocks than users of comparable brokers like E-Trade and even 40 times more than customers of Charles Schwab. However, since it is not just the attention of an individual that is influenced, but the attention of the entire user base, the likelihood of herding events within these neobrokers increases clearly (Barber et al., 2020). This raises the question of what exactly motivates this bulk of investors to decide for or against a particular investment.

An explanation for this phenomenon within Robinhood can be provided by the psychological technique known as Fear-of-Missing-Out (FoMO). Almost every investor now uses digital communication channels to make their investment decisions. Being disconnected from the internet can cause investors to fear missing out on important information, so-called I-FoMO (Shiva, Narula, & Shahi, 2020). This fear, however, can also lead to overinvestment. Especially when the market changes rapidly and investors are afraid of missing an opportunity, the fear of unrealized gains becomes greater than the fear of losses. I-FoMO is fueled in particular by the presentation of Top Mover lists at neobrokers like Robinhood, which can trigger the constant feeling of missing a special chance for the investor who wants to quickly jump on the departing train before it is too late (Maciejewski & Lesznik, 2020).

Another reason for the use of attention-grabbing methods by neobrokers is the concept of myopic loss aversion (MLA). For the individual, MLA implies having to accept a loss is worse than realizing a gain, which in absolute terms is equal to the same amount. This indicates a greater perception of losses by an individual than gains, especially when suffering from a narrow time frame, leading to a negative reaction to information about frequent price changes of an asset since the overall picture is not taken into account
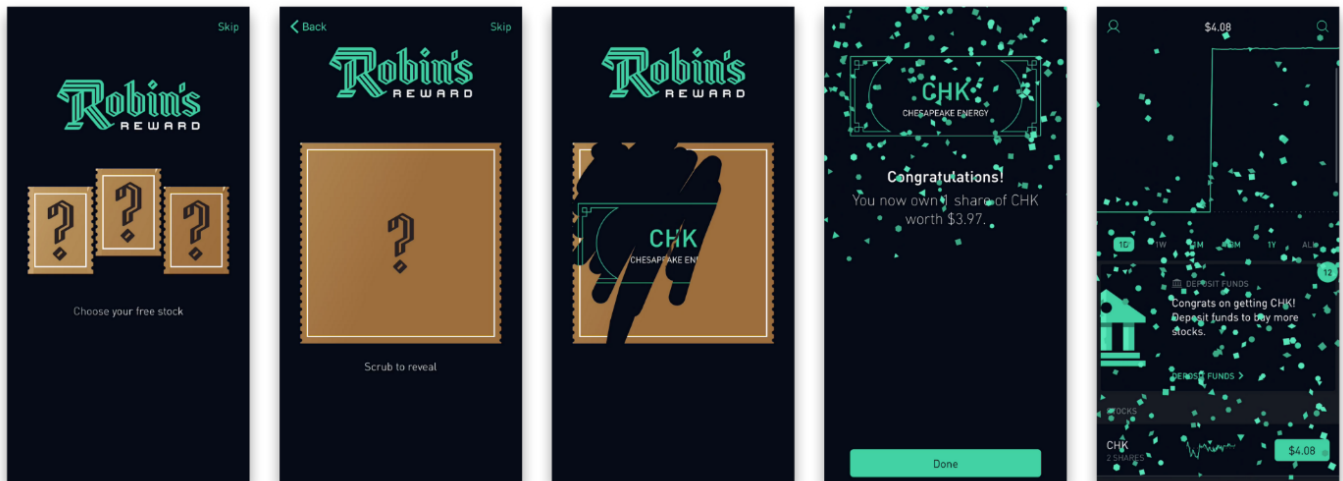
Figure 1: Lottery Scratch Card Interface and Confetti Rain (Hobson, 2018).

This figure presents the scratch card game found in the neobroker Robinhood, in which users are rewarded with a random share in a certain price segment. After receiving the share, the user is also visually rewarded with a shower of confetti.

(Iqbal, Islam, List, & Nguyen, 2021). Investors in this case are too focused on short-term events and lose sight of long-term successes, which can lead to rash selling of their positions if a short-term price drop is imminent (Thaler, Tversky, Kahneman, & Schwartz, 1997). Neobrokers such as Trade Republic are designed to primarily draw the attention of users to the daily price fluctuations of their positions so that they react in case of negative movements in stock prices due to MLA to avoid short-term losses. As a result, the investor interacts more frequently with the application (Barber et al., 2020) from which neobrokers benefit.

2.4.2. Gamification

As described in the previous chapter, online brokers and especially neobrokers use gamification to encourage their users to interact with the application. This gamified experience tempts investors to make rash and impulsive decisions within the neobroker applications. Trading in stocks is more and more perceived as an occupation for spending time by retail investors. Once trading is seen as entertainment and a substitute for boring activities, the number of transactions will steadily increase, as was evident during the Covid-19 crisis (Barber et al., 2020). This is supported by the entire philosophy of neobrokers like Robinhood as well as the increasing individual indebtedness, achieved with easier access to credit (Bernards, 2019), which can also be used for the increasing risky behavior of investors and gambling on the stock market.

The playful nature of investment applications tempts users to make impulsive decisions without concretely informing themselves about their investment choices in advance. However, these decisions do not remain an isolated case and thus the playful and unique affair may turn into a potential addiction (van der Heide & Želinský, 2021) to interact with the broker on an increasingly frequent basis. A reduced and

sleek design, as found in neobroker applications like Robinhood, has a big impact here. Kranz, Teschner, and Weinhardt (2015) show in their studies that such a reduced design not only has a playful influence but also encourages the user to trade. At the same time, they find that increasing the amount of information presented to the user reduces his or her trading behavior (Kranz et al., 2015; Teschner, Kranz, & Weinhardt, 2015). For this reason, the presentation of the application of neobrokers is deliberately chosen to be as minimal and simplified as possible, in order not to overwhelm the investor with a flood of information. In addition, this user interface, including the choice of its eye-catching colors as found in the Robinhood application, offers the possibility to attract especially younger and inexperienced investors, who are more easily influenced by gamification methods.

The swipe gestures already mentioned also encourage users to interact. The familiar and everyday movements of the mobile world greatly reduce the barriers to trading and thus encourage purchases and sales (Tan, 2021). But it is not just the swipe gestures that motivate users to engage with the application, but also the way stock prices are displayed over time in form of price charts. The charts show the share prices at a certain time utilizing a gray line that appears when the user moves his fingers over the display. These small movements are directly associated with stock price fluctuations (Ash, Anderson, Gordon, & Langley, 2018), meaning that active body movements can also trigger I-FoMO. The interactions reduce the inhibition to place a buy order when rising prices are anticipated due to your movements. The same applies to the intention to sell shares when short-term falling prices are indicated (Tan, 2021).

2.4.3. Research Gap

In the following section, I will highlight the gaps in the literature on neobrokers and the topics that have not yet been

sufficiently addressed. Neobrokers in general is still a very under-researched field, resulting in a lack of widely applicable literature on these new types of brokers.

As I indicated in section *2.1 Development of Neobrokers*, most of the literature on retail investor trading behavior is concerned with Robinhood rather than with the general topic of neobrokers. While Robinhood and its users can be used as a representative example of this new generation of brokers, it constitutes only a small fraction of the overall neobroker environment. Therefore, it is necessary to extend the existing literature on Robinhood and leverage its general applicability to the entire field of neobrokers. I do not focus on a specific neobroker or certain investors in my study but use the general term to get a broader picture. Likewise, as already pointed out in the literature review, I use the two brokers Robinhood and Trade Republic as representatives for neobrokers, in general, to get a broader overview of these providers of trading applications and not to focus too much on one particular broker.

Another gap in the literature arises from focusing purely on the U.S. market. For example, since Robinhood is not available outside of the United States, there is no information on the European neobroker market. While it can be assumed that the impact of this neobroker can be applied to other cultures and countries, no studies are addressing this issue yet. This gap is dealt with in my thesis, as the participants of the study are mainly from German-speaking countries. Hereby, the already researched influence of Robinhood on the American market can be extended by the influence of general neobrokers on additional markets.

Moreover, the existing literature primarily highlights attention-grabbing and game-based methods to explain short-term and impulsive buying and selling decisions in neobroker applications. Young and inexperienced traders, who are easily influenced and make their decisions based on such influencing methods, are always referred to. However, little attention is paid to whether and to what extent these investors also use neobrokers for long-term and future-oriented investments. The result for this can only be surmised, which is why I pursue the question as part of my online questionnaire.

As a final consideration, and as a consequence of the general lack of literature available, the poorly treated field of neobroker application design must be addressed. Although there are a few studies on the effects of attention-grabbing elements, for example, these are only available to a limited extent and refer exclusively to Robinhood and its users. In this case, it is also important to take a closer look at the topic of the user interface and validate the influence of its specific design.

3. Hypothesis Development

In the previous subsections, existing empirical results and theories regarding the influence of neobrokers, in particular Robinhood, on the behavior and decisions of retail investors have been discussed. They form the basis for the hypotheses

of my thesis, which build on the questions already addressed in chapter *1. Introduction*. In the following subsection, I formulate and explain the various hypotheses in more detail to be able to answer those questions based on my survey.

The hypotheses are divided into three different areas. The first area will deal with the question of which target groups are addressed by neobrokers and what the decisive reasons for this are. The second part will focus on the question for which investment strategies neobrokers are generally used. It is important to clarify what kind of investment horizon the users of these applications are aiming for and how far they pursue their strategies. The last section delves deeper into the neobroker applications and deals with the question of to what extent neobrokers have an influence on the investment behavior of their users. Here, the focus is on the appearance of the applications and their general user interface.

3.1. Target Group

Trading has become extremely easy for customers thanks to neobrokers and their applications. It often takes only a few clicks and swipes to execute a trade. Furthermore, the low to non-existent costs touted by neobroker providers can seem very enticing to most of society. Robinhood says of itself they are "on a mission to democratize finance for all" (Robinhood, 2021a) and the financial system should be structured so that it works for everyone. Neobrokers advertise their intention to provide access to the stock market for the public and allow profiting from it. Nevertheless, as the previous literature shows, not all people feel the same way about the strategies and methods of neobrokers. This raises the question of who exactly the brokers want to address and entice with their approaches.

Neobrokers need to gain the trust of potential customers and thus familiarize them with their applications because both familiarity and trust influence customers' interactions with the provider (Gefen, 2000). In general, trust can affect people's judgments and beliefs, guiding their behavior in a particular direction (Schunk, 2012). Trust is one of the key factors to win customers for the own e-banking website or application (Yousafzai, Pallister, & Foxall, 2003) and has a positive social influence on online transaction intentions (Wu & Chen, 2005).

For this reason, my first hypothesis also relates to the topic of trust. The higher the trust potential customers have in neobroker applications, the better their attitude towards them and the more likely they are to use them. Formally, my first hypothesis is as follows:

> H1: Trust positively influences the attitude towards neobrokers.

Since trust plays an important role on the internet and in the financial sector, it is even more important for neobrokers to appear as trustworthy providers of financial services. In doing so, most of the neobrokers take a very similar approach to gain people's trust. The goal of these newly established brokers is to develop an application that is easy to

use, does not confuse, and looks familiar. Most neobrokers offer their trading services in form of a mobile application. This application is usually designed in a simple and sleek way, so it does not contain unnecessary information or distractions. This is particularly appealing to the younger generation. These so-called millennials have grown up in the age of the internet and digital revolutions. They are comfortable in a connected and globalized world and keep up with new technologies, which is why they have also acquired the nickname *digital natives*. However, they are not only familiar with the growing field of technologies but also want to gain more first-hand knowledge about financial management (Thompson & Blomquist, 2017). To this end, neobrokers offer just the right combination of new technology and financing, which may explain the rapidly growing number of especially young members. In addition, young investors no longer have confidence in traditional financial institutions. According to Thompson and Blomquist (2017), more than half of the millennials believe their financial advisor is only interested in their own profit and not the client's.

But it is not only young investors who are targeted by neobrokers. New investors who want to gain their first experience with financial investments are also addressed by them. Neobrokers charge little to no transaction fees and offer a clean and simple interface, leaving newcomers unencumbered by fees or complex features, making it very appealing for them to gain their first experience on the stock market here. This lack of trading barriers is also in line with Robinhood's previously mentioned mission of empowering everyone to participate in the stock market.

Relating these findings in turn to the neobroker landscape as a whole, this means that neobrokers preferentially target young and inexperienced people, or, as it is formally expressed:

> H2: Young and/or inexperienced investors are especially addressed by neobrokers.

3.2. Investment Strategy

As highlighted in the previous section, the new generation of investors is focusing on different viewpoints and goals than their predecessors, which is also evident away from the financial markets. Among the most common answers to the question about the life goals of this young and inexperienced generation are things like having a dream job or traveling around the world instead of getting married and starting a family, which were the goals of the previous generation. Thus, it is not surprising that millennials' money is meant to serve short-term goals rather than long-term intentions (Thompson & Blomquist, 2017). This phenomenon can also be found in their investment behavior. As described in section *3.1 Target Group*, neobrokers are the preferred choice for millennials when it comes to investing money. Here, Robinhood serves as a good point of comparison. Most of its users are referred to as retail investors. These investors hardly act according to a defined strategy but are guided by external influences and their own preferences (Barber & Odean, 2008).

Again, the preferences are short-term targets, which can be identified by the high number of trades within the Neobroker application. Robinhood customers trade unusually often compared to users of traditional online brokers: nine times as much as users of E*Trade and 40 times as much as customers of Charles Schwab (Barber et al., 2020). This very high number of trades indicates that these are speculative investments with a short holding period and no intention for long-term investments. Investing is considered part of the overall entertainment, which it is meant to serve in the near future. In addition, other external influences reinforce this amusing aspect. Retail investors, who make up the majority of neobroker users, follow financial information and instructions provided by social media and other channels, which results in day-trading behavior based on betting on riskier options or flipping stocks. This again is supported by the business model of neobrokers, which favor the multitude of short-term investments by convenience and low trading costs (Tan, 2021).

From the information collected above, it can be concluded that users of neobrokers have deliberately chosen them for short-term investments. The low barriers for trading make quick buying and selling decisions very attractive. Based on these assumptions, the following two hypotheses can be derived for my thesis:

> H3: Neobrokers are used for short-term investments.

> H4: Neobrokers are not used for one's own retirement planning.

3.3. Influence on Investment Behavior

In the theoretical foundation of my thesis, I have presented two psychological concepts that can have a significant influence on the investor. Both gamification and attention-grabbing can nudge one's decisions in a certain direction. These two concepts can give investors the feeling of making decisions based on sufficient information, which leads to an increased number of trades. However, this overestimation of the quality of information generally leads to a suboptimal investment strategy (Odean, 1998). To convey this image of qualitative information, neobrokers use various methods, as described in *2.4.1 Attention-Grabbing* and *2.4.2 Gamification*.

Not only has the reduced and sleek design played an important role here, as it is often quoted in the previous literature, but also the well-thought-out structure and presentation of the individual information that make a significant contribution. When it comes to FinTech companies, it is not just about the lowest fees, but also about a flawless platform infrastructure making it as easy as possible for the user to interact with the application and thereby generate more transactions (Ash et al., 2018). The simplified trading on neobroker applications such as Robinhood encourages investors to increase their number of investments (Barber & Odean, 2008), while the deliberate presentation of various pieces of information can also trigger specific individual interactions.

For example, Trade Republic takes advantage of MLA in its representation of stock prices, as it is not possible for the user to choose an annual stock chart as the default representation instead of the daily chart, which is subject to much greater fluctuations. There is no possibility to set this default setting to a longer period like the yearly development. When opening the neobroker application, the investor first sees the daily price trend and can thus deviate from his long-term target based on this daily presentation. Since there may be major fluctuations during the day, even if the price-performance is stable or positive over the years, the daily chart may suggest to the investor that the investment is likely to have a negative outcome. The possibility of a short-term price drop always creates the risk of selling a long-term position due to loss aversion. Part of the reason for this is the general tendency for people to give more weight to negative events than to events with a positive impact. For decision-making, potential costs, especially if they are coupled with losses, have a much greater influence than the possibility of realizing profits (Kahneman & Tversky, 1979). In addition, there are game-based elements that are frequently used in neobroker applications. Whether it is scratching cards to get a free share, interacting with stock charts, or even refer-a-friend campaigns to get additional rewards.

In summary, neobrokers can determine which actions are simplified and promoted for their users and which are made significantly more difficult due to their deliberately reduced design (Chaudhry & Kulkarni, 2021). Based on this theoretical foundation, the following hypothesis can be derived for my thesis:

> H5: The design of neobrokers has an impact on investors' investment behavior.

The retail investors' behavior is altered in several ways. On the one hand, the low barriers, as well as the playful and simple design, entice users of neobroker applications to make riskier investments without having to think too much about it. This allows young and inexperienced investors to consciously apply risky strategies, even if they are rarely successful (Chaudhry & Kulkarni, 2021). At the same time, due to the very low or non-existent costs, this democratization of stock trading leads to significantly more trades being executed on those platforms than on comparable online brokers (Barber et al., 2020). These two phenomena lead to the following hypotheses:

> H6: Neobroker customers take greater risks than customers of traditional banks.

> H7: Neobroker customers trade financial products more frequently than customers of traditional banks.

Investors themselves consciously decide to become customers of neobroker applications. While for many users it is the only way to participate in the stock market without incurring additional costs, traditional online brokers have already followed suit and offer low-cost trading. Thus, the

look and layout of the applications appears to be one of the most important factors for using these new brokers. Above all, the sleek design and the significantly reduced presentation of information is one of the main selling points of neobrokers. This chosen interface is in strong contrast to traditional brokers. Robinhood, for example, provides its clients with only five charting indicators, while TD Ameritrade users have access to 489 (Barber et al., 2020). Customers of Neobroker applications do not want to be overwhelmed with a flood of data and deliberately opt for the reduced presentation of information. It can be concluded that users of neobrokers appreciate not only the appealing design but also the reduced information presentation and do not want to gain more knowledge about their investments within the application. Formally speaking, this means:

> H8: The reduced and simple design of the application is more important for neobroker customers than for customers of traditional banks.

> H9: Neobroker customers do not want more information within the application compared to customers of traditional banks.

4. Methodology

A methodological approach is required to answer the fundamental questions of my thesis and test the hypotheses derived from the previously mentioned literature. The main objective is to reach a large, diverse number of respondents in a pre-defined time frame to be able to derive generally valid conclusions on the subject of neobrokers. For this purpose, a quantitative approach in the form of an online-based questionnaire is the most effective, as it allows me to quickly and inexpensively acquire a variety of participants for my study, whose responses were immediately digitally saved (Diekmann, 2020). By using this approach and employing statistical analysis, it is possible to quantify the data and obtain the general picture (MacDonald & Headlam, 2008) I am aiming for. This allows me to deductively test my already derived hypotheses.

In the following sections, I describe the individual steps of the methodology and explain how the data was obtained. Subsequently, the results relevant for the verification of my hypotheses can be presented based on this data.

4.1. Questionnaire Structure

As previously stated, I utilize a self-generated online questionnaire as part of my thesis to obtain the most diverse and generic data collection possible. This questionnaire is created by using Google Forms and distributed to the participants via an online link. Employing a sequential, comparative literature analysis, the instrumentalization of the survey items and response categories is developed, repeatedly modified, and adapted. My questionnaire is divided into three different thematic blocks. First, eight general demographic questions are asked about the participants, such as their age,

gender, or even their financial knowledge, which is tested by using three questions specifically developed by Lusardi and Mitchell (2011) to measure the financial literacy score. The second part of the survey then focuses on the participants' general interactions with financial institutions. Here, questions are posed on existing experience in trading with financial products, as well as risk tolerance and preferences about the use of online offers from financial service providers. The last part of my questionnaire is related to the main topic of my thesis. In this section, I ask questions about the perception of neobrokers and what influences certain design elements within these applications have on the participants. It is important to mention here that the online questionnaire is published in two different versions, which differ from each other in this section due to a single figure. The first version of the questionnaire displays the daily development of a share price of a given company on a certain day, whereas the second version reflects the annual development of this price. The daily chart shows a declining share price, while the annual one reveals a significant increase in the share performance at the same point in time. Both versions of the online questionnaire are randomly distributed to the respondents to ensure that a representative analysis is still possible. The surveys are additionally written in both English and German to overcome possible language barriers. In total, the two questionnaire versions consist of 37 tick-box questions, a mix of single-choice, multiple-choice, and scaled questions, which can be answered in about ten minutes. The check-all-that-apply (CATA) questions are particularly suitable for a complete description of the most important features of a sample to establish comparability (Jaeger et al., 2015), which is why they are used primarily in the third section of the surveys to contrast the respondents' perceptions of and desires for neobrokers with their actual characteristics. For most items, it is also possible to check a box labeled *Prefer not to say*, if the question cannot be answered, as well as a field titled *Other*, if the respondent wishes to provide additional information.

The following table provides an overview of the questions from my survey, which can be used for verifying each hypothesis. However, not all questions are used for testing, but a selection of the most appropriate for the respective hypothesis is made in order not to go beyond the scope of my thesis.

4.2. Sample Used

To fill one of the research gaps already mentioned, my study needs to obtain participants outside of America. In addition, these respondents should not all belong to a specific target group, but rather represent a field of participants that is as heterogeneous as possible. It is therefore very important to distribute the online questionnaire to the greatest extent feasible. I share the survey with acquaintances and peers, as well as in various forums and online groups which are not directly related to financial subjects to accomplish my goal. In this way, I can collect a total of 251 responses for my questionnaire within two months, starting at the end of July 2021. Among these responses, only two participants complete the

English-language questionnaire, while the remaining participants use the German version. The link through which the respondents fill out my online survey happens to lead to either the daily or the annual chart version of the questionnaire. As a result, 119 responses to the daily and 132 responses to the annual version are obtained. However, not all of these 251 responses can be utilized due to some exclusion criteria. Firstly, all surveys that are not filled out completely are removed. In addition, responses that state the neutral category *Prefer not to say* exclusively or almost always are eliminated as well. The deliberate selection of the same answer over and over, also known as *straightlining*, is considered an indicator of poor response quality (Schonlau & Toepoel, 2015). Here it becomes clear that the survey is not answered conscientiously and carefully, but only clicked through quickly by the participant. Finally, a consistency check is performed to delete any answers that are intentionally given incorrectly or without reference to the study. Thus, after all, adjustment steps have been performed, 224 responses remain. Due to the small number of responses that had to be eliminated, the questionnaire is not adapted during the study. In the following figure, all steps are shown again, including the number of adjusted responses for each step.

Of the remaining 224 responses, 107 participants replied to the daily chart version, and the remaining 117 to the annual version. As a result, 12 questionnaires of the daily version and 15 surveys of the annual version had to be removed from the data set.

4.3. Data Preparation

In the following section, I will go into more detail about the preparation of the data, which is the basis for my qualitative analysis in the first place. To analyze the results of the online questionnaire, it is necessary to properly sort and format the responses received from Google Forms. After completing the survey editing period, I download the bundled raw data in an .xlsx format for subsequent post-processing in Microsoft Excel. For the evaluation of this amount of data, it is essential to use a suitable computer program as a working aid. Here, I decide on the widely used program system IBM SPSS Statistics, or SPSS for short, because of its ability to cover all the necessary functions for my analyses, its user-friendly and graphical interface as well as its particularly wide dissemination in both scientific and practical settings.

The processing of the raw data is done in several steps, which I perform with the help of the software application Microsoft Excel. First, the CATA questions have to be restructured in a way that they can be evaluated for the analysis. Thus, all CATA questions are reformulated as multiple binary variables, also called dummy variables. For example, *Where have you already purchased financial products?* becomes the three binary questions of whether the respondents have already purchased financial products from a local branch bank, a direct bank, or a neobroker. This results in a variable count of 92 at the beginning, which all are imported into SPSS and used for the analysis. In the next step, the binary questions are recoded such that *yes* becomes 1 and *no* becomes 0. The

Table 1: Testing Hypotheses by using Questionnaire.

This table presents all possible questions of my online survey that can be used for the verification of the respective hypotheses. From these questions, all those that are most suitable for testing these hypotheses are selected.

Hypothesis	Questions that can be used to answer the hypothesis
H1	Questions 17-19 & Question 29
H2	Question 3 & Questions 7-14 & Questions 17-19
H3	Question 16 & Questions 17-19 & Question 24 & Questions 26-27
H4	Question 16 & Questions 17-19 & Questions 26-27
H5	Questions 30-36
H6	Questions 17-19 & Questions 20-21 & Question 25
H7	Questions 15-16 & Questions 17-19 & Question 24
H8	Questions 17-19 & Question 22 & Question 28
H9	Questions 17-19 & Questions 37-38
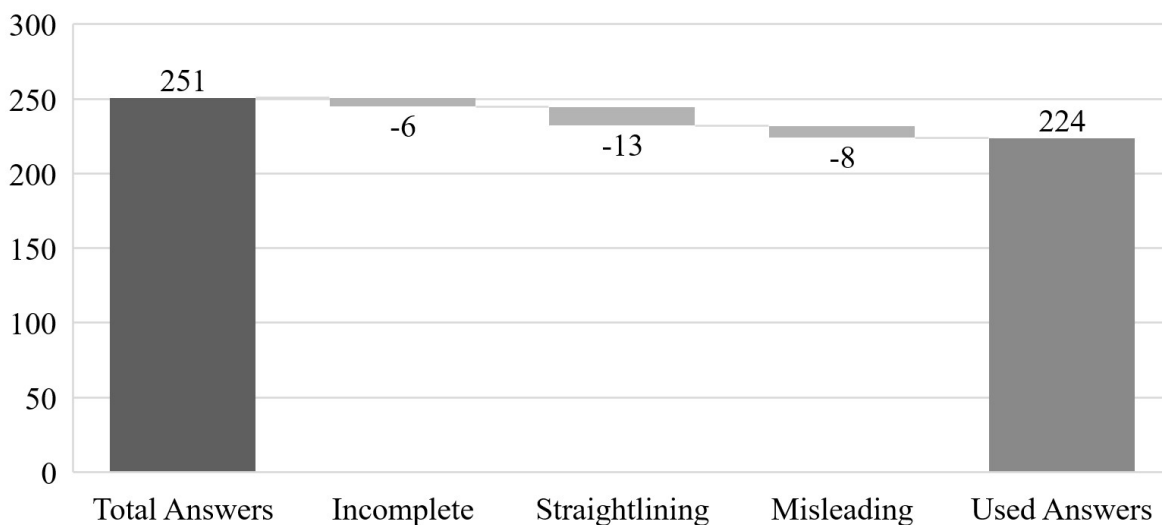
Figure 2: Presentation of Answers used from Online Survey.

This figure presents the number of respondents who participated in my survey. After the three revision steps (removing the incomplete, straightlining, and misleading responses), the remaining number of responses used for the evaluation of my thesis is also shown at the end.

same conversion from string variables to numerical variables is also performed for the remaining questions. With the interval scales used in the online questionnaire, recoding of the data obtained is not required because they already reflect scales with sections of the same size. The nominal scales used for my survey, on the other hand, also require a recording of the responses, with the statement *Prefer not to say* being coded as -99 to subsequently exclude it from the analysis. The free text fields would have also been coded in a meaningful way during the preparation, but after eliminating all questions that were not to be used, no answers to the free text fields remained. In a final step, all derived binary questions of a CATA question are combined into a variable set to evaluate some of the hypotheses.

4.4. Evaluation Methods

The evaluation begins with a descriptive presentation of the sociographic characteristics, which are illustrated using frequency distributions. The most important factors here are gender, age, education, employment status, income, and financial literacy.

The remaining questions are employed to test the hypotheses formulated in my thesis, using a combination of univariate and multivariate analyses (Everitt, 2010). Univariate analyses are performed to test the hypotheses three and four as they are accurately captured from the questionnaire. In this way, it is possible to create frequency distribution tables representing the breakdown of the responses. To show the relationship between two variables, multivariate analyses are performed. With this analytical method, it is possible to create cross-tabulations and perform regression analyses, thus mapping the correlations between the variables of interest.

Two variants of regression analyses are particularly suitable for testing my hypotheses. Binary logistic regression is used in all cases where the dependent variable has a binary nature. The groups are usually indexed with 0 and 1 in these scenarios. The following applies to the probabilities of these binary cases: $P(Y = 0) = 1 - P(Y = 1)$. Compared to the linear regression, the logistic regression function has an S-shaped course and can be interpreted as a distribution function similar to the normal distribution. For all other cases, where the dependent variable does not correspond to a binary variable, regression is performed using the Ordinary Least Squares (OLS) method. This method searches for those values of the parameters for which the sum of squared residuals becomes minimal. This gives the regression line an optimal fit to the observed values. (Backhaus, Erichson, Plinke, & Weiber, 2016)

It should be noted once again that the two versions of the online questionnaire are evaluated collectively. Only when testing the fifth hypothesis both versions are considered separately to highlight any differences in responses to questions 30 and 31 between the two versions.

5. Results

In the following sections, I present the results obtained from the analysis of my online questionnaire. The chapter is divided into four parts. First, the descriptive statistics are evaluated and presented. Subsequently, my established hypotheses are tested based on univariate and multivariate analyses. Building on these findings, the final section of this chapter contains a discussion of the results.

5.1. Descriptive Statistics

At the beginning of the descriptive statistics, I first highlight the sociological factors. As already described in section *4.2 Sample Used*, the total data set for my analysis consists of 224 responses after adjustments. Out of these evaluated participants, 55.8 percent are male, and 44.2 percent are female. The majority of the respondents are between 25 and 34 years old (51.3 percent), closely followed by 18- to 24-year-olds, who account for 38.4 percent. The remaining participants are divided between under 18 years old (0.4 percent), 35 to 44 years old, 45 to 54 years old (2.7 percent each), and over 55 years old (4.5 percent). In terms of educational attainment, the majority of the respondents have a bachelor's degree (52.7 percent), followed by high school diplomas and master's degrees. Only one respondent did not want to comment on this question (0.4 Percent). 57.1 percent of the participants are still studying, while 35.7 percent are already employed full-time. Part-time employment (4.0 percent), self-employment (2.2 percent), and unemployment (0.9 percent) make up the minority. On average, the household income of the participants is between € 10,000 and € 50,000 (33.0 percent), closely followed by below € 10,000 (28.6 percent) and € 50,000 to € 100,000 (22.8 percent). Out of the 224 participants, 18 (8.0 percent) did not want to

comment on this question. Regarding the financial literacy of the survey participants, the absolute majority (77.2 percent) is educated, as they answered all 3 questions about their financial literacy correctly. Only 15.2 percent of the respondents had less than 3 questions correct, whereas 7.6 percent did not even try to answer the questions in the first place.

Looking at the two survey versions separately, no major differences are observed concerning the sociological factors. In the daily chart version, the percentage of male respondents is slightly lower compared to the second version (52.3 percent vs. 59.0 percent). This difference between the two versions increases slightly when the age of the respondents is considered. In each case, roughly 45 percent of participants in the first version are either between 18 and 24 years old or between 25 and 34 years old, while 57.3 percent of those in the second version are in the 25 and 34 age group, followed by 18 to 24 years with 31.6 percent. As far as the education factor is concerned, hardly any difference can be detected between the two versions. Both reflect the overall result of the entire data set. In the daily chart version, almost two out of three participants are still students, whereas, in the annual chart version, only one out of two is. The household income of the respondents of the first version is evenly distributed as well. 31.8 percent of the respondents report having an estimated annual household income of less than € 10,000, 30.8 percent choose the answer between € 10,000 and € 50,000, succeeded by between € 50,000 and € 100,000 with 21.5 percent. The data set of the second version provides similar results. Here, 35.0 percent of the respondents report having an annual household income between € 10,000 and € 50,000, followed by below € 10,000 (25.6 percent) and between € 50,000 and € 100,000 (23.9 percent). The two groups also perform comparably in terms of financial education. 76.6 percent and 77.8 percent of respondents in the first and second versions can correctly answer all three questions related to the financial literacy score.

As shown above, no serious differences can be found between the data set of the first and the second version. Due to the quite similar field of participants of both versions, the data sets can on the one hand be considered and put together as a whole, and on the other hand also be separated and compared with each other, as it is necessary for testing the fifth hypothesis. In the following sections, the underlying hypotheses of my thesis will be examined and tested based on the data set.

5.2. Univariate Analysis

Univariate analysis is suitable for testing the third and fourth hypotheses. This allows the generation of frequency distribution tables for verification of the two hypotheses after the responses to all CATA questions have been defined into sets of variables. To validate these two assumptions, the participants of my questionnaire are asked for which scenarios and investment strategies they would consider a branch bank, a direct bank, or a neobroker. A total of 204 responses (91.1 percent) can be collected to the question of whether the

Table 2: Separate Descriptive Statistics for both Versions.

The following table presents the descriptive representation of the respondents divided by the questionnaire version that was answered. It shows separately the number and percentage distribution of respondents according to their gender, age, education, employment status, income, and financial literacy score. Both versions combined result in the total descriptive evaluation of the online questionnaire. All 224 responses are shown below.

		Daily chart version		Annual chart version	
		Total	Percent	Total	Percent
Gender	Female	51	47.7	48	41.0
	Male	56	52.3	69	59.0
Age	Under 18 years old	1	0.9	-	-
	18-24 years old	49	45.8	37	31.6
	25-34 years old	48	44.9	67	57.3
	35-44 years old	2	1.9	4	3.4
	45-54 years old	2	1.9	4	3.4
	over 55 years old	5	4.7	5	4.3
Education	Secondary School Diploma	2	1.9	3	2.6
	High School Diploma	21	19.6	28	23.9
	Bachelor's Degree	59	55.1	59	50.4
	Master's Degree	22	20.6	21	17.9
	Trade School Diploma	2	1.9	6	5.1
	Prefer not to say	1	0.9	-	-
Employment	Student	70	65.4	58	49.6
	Employed full-time	34	31.8	46	39.3
	Employed part-time	1	0.9	8	6.8
	Self-employed	2	1.9	3	2.6
	Unemployed	-	-	2	1.7
Income	Below € 10,000	34	31.8	30	25.6
	€ 10,000 - € 50,000	33	30.8	41	35.0
	€ 50,000 - € 100,000	23	21.5	28	23.9
	€ 100,000 - € 150,000	7	6.5	6	5.1
	Over € 150,000	2	1.9	2	1.7
	Prefer not to say	8	7.5	10	8.5
Score	0	9	8.4	8	6.8
	1	4	3.7	7	6.0
	2	12	11.2	11	9.4
	3	82	76.6	91	77.8

respondent would prefer a branch or direct bank over neobrokers. A clear trend can be identified here: compared to neobrokers, branch and direct banks are favored especially for long-term investments. 168 participants vote for long-term investments for their own wealth accumulation (82.4 percent of respondents), 157 participants for long-term investments for their own retirement provision (77.0 percent), and 117 participants for setting up savings plans (57.4 percent). Only a small percentage of respondents prefer branch

and direct banks over neobrokers for short-term or speculative investments (9.3 percent and 5.9 percent, respectively). Looking at the answers to the question in which scenarios the respondents would prefer neobrokers, very similar results are obtained. Among the total 202 responses (90.2 percent) to this question, the selection options are dominated by short-term (87.6 percent of responses) and speculative investments (78.2 percent). Only 17.8 percent and 12.4 percent of respondents, respectively, opt for long-term invest-

ments for asset accumulation or for their own retirement provision, which represents the lowest proportion of participants. Unfortunately, the two hypotheses cannot be tested for statistical significance due to lack of information, which is why only the descriptive analysis is available for the evaluation. However, the descriptive results indicate that both hypotheses can be confirmed. On the one hand, neobrokers are preferred for short-term investments and, on the other hand, they are not used for retirement planning, not even for long-term asset accumulation, by the participants.

5.3. Multivariate Analysis

Regression analysis is chosen as a suitable method for testing the remaining hypotheses to show dependencies between several variables. Ordinary Least Squares regression cannot be applied to examine the first and second hypothesis since the dependent variable under investigation, in this case, the possible intention to use a neobroker for purchasing financial products, has only two manifestations: *yes* or *no*.

To further investigate the impact of trust in neobrokers on this particular variable (see H1), I perform a binary logistic regression using SPSS. For this analysis, 183 responses can be utilized because 41 respondents do not want to provide any information for either of the two related questions. Using the *Omnibus Test of Model Coefficient*, which compares the current model with the baseline model, it can be shown that the current model with the inclusion of the independent variable *trust in neobrokers* is a better fit than the baseline model without any independent variables, as the *p-value* of 0.000 is below the typical rejection threshold of 0.05 (O'Keefe, 2007). The new model explains 48.3 percent (*Nagelkerke* R^2) of the variance in the variable *possible intention to use a neobroker* and correctly classifies 83.6 percent of the cases. If the respondent indicates trusting a neobroker, it is 23.342 (*odds ratio*) times more likely that the participant can imagine using a neobroker compared to those who do not trust a neobroker at all. Applying binary logistic regression, the first hypothesis, which states that trust positively influences the attitude towards neobrokers, can thus be confirmed.

Using the second binary logistic regression, I test the next hypothesis, which deals with the target groups especially addressed by neobrokers (see H2). To evaluate whether inexperienced investors are approached by neobrokers apart from younger ones, the independent variables *financial literacy*, as well as *investment experience*, are evaluated in addition to *age*. For this purpose, I can draw on 223 responses for my analysis. Including the independent variables, the new model is found to have a better fit than the initial one, as the *p-value* of 0.000 is once again below the typical rejection threshold of 0.05. The new model can explain 21.5 percent (*Nagelkerke* R^2) of the variance of the dependent variable. Since this time, I include more than one categorical predictor in the new model, I can additionally output the *Hosmer-Lemeshow-Test*. This test indicates how much the observed and predicted cases differ from each other. It is important to note that the significance should not be less than 0.05, as

this would imply a poor fit (Hosmer, Lemeshow, & Sturdivant, 2013). The *p-value* here is 0.704, indicating a good fit of the model. Overall, 68.6 percent of the cases can be classified. As the respondents' age increases, the likelihood that they can imagine using a neobroker changes by 0.580 (*odds ratio*). In other words, the probability of not using a neobroker is 1.724 times, or 72.4 percent, higher with increasing age, which confirms the first part of my second hypothesis. However, this cannot be determined for both the *financial literacy score* and the *investment experience*. For example, a respondent who had prior investment experience is 3.738 (*odds ratio*) times more likely to imagine using a neobroker than an inexperienced participant of my study. In addition, the higher the score of those respondents, the more likely they are to use a neobroker (*odds ratio* = 1.419). This last result, in turn, is not statistically significant (*p-value* = 0.054). The second hypothesis can therefore only be partially confirmed: younger investors are more likely to be addressed by neobrokers, which is not the case for respondents with little financial experience.

For testing the fifth hypothesis, the entire data set is first split into its two different versions and considered separately. In doing so, I want to find out whether the design of neobrokers influences the investment behavior by using two different illustrations within the questionnaire. With the help of these presentations of the same stock, once as a daily chart version and once as an annual chart version, the perceived risk of an investment for the subject is queried. The average associated risk for the first version is 3.58 (1 corresponds to *low risk*, 5 to *very risky*), with 4 being the most selected risk level with 41.1 percent of responses. The same stock is rated as less risky in the second version with a mean risk score of 2.94. This time, risk level 2 is the most frequently selected with 35.0 percent. Also, when asked to what extent the respondents would consider trading this stock, 44.9 percent of participants in the first version indicate, that the stock is too risky and only 30.8 percent show interest in trading the stock. In contrast, 51.3 percent of respondents from the second version are interested in trading the stock and only 8.5 percent consider it too risky. For testing the hypothesis, all participants are additionally asked how they would react to an advertisement of popular stocks, as found in many neobrokers. For this evaluation, the data set is once again considered as a whole, with 5 participants not wanting to give an answer. Based on the presentation alone, 35.7 percent would decide to take a closer look at the listed companies, 22.8 percent would follow the price performance of the respective stock to find a good entry position, and just under 30 percent of the respondents would completely ignore the presentation. Conversely, it means that more than two-thirds of the respondents would look more deeply into the listed stocks and 46 participants would even consider trading them. When asked how the participants would react to an Invite-your-Friends campaign, 38.0 percent of the 204 respondents would at least show it to their friends, if not try to convince them to make a trade. However, it is worth mentioning that 34.4 percent would completely ignore the campaign.

Table 3: Frequency Distribution Tables & Bar Charts for Testing H3 & H4.

This table presents the descriptive distribution of the given answers according to the questions: *In which of the following scenarios would you prefer a branch or direct bank over a neobroker?* and *Which of these scenarios would you consider a neobroker for?*. In total, 204 responses can be used for the CATA question regarding the preferences for branch and direct banks and 202 responses for the preferences for neobrokers.

	Branch or Direct Bank		Neobroker	
	Total	Percent of Cases	Total	Percent of Cases
Short-term Investment	19	9.3	177	87.6
Speculative Investment	12	5.9	158	78.2
Long-term Investment for Asset Accumulation	168	82.4	36	17.8
Stock Portfolio without Diversification	63	30.9	67	33.2
Long-term Investment for Retirement	157	77.0	25	12.4
Monitoring of Share Prices	27	13.2	101	50.0
Savings Plans	117	57.4	41	20.3
ETFs	71	34.8	86	42.6

Table 4: Binary Logistic Regression for Testing H1.

This table shows the results of the binary logistic regression used to test the first hypothesis. Here, I examine the influence of trust in neobrokers (independent variable) on the possible intention to use a neobroker to purchase financial products (dependent variable). 183 responses are available for this analysis. *, **, and *** represent significance below the 10%, 5%, and 1% level, respectively.

		Results
P-Value	Omnibus Tests	0.000
R^2	Nagelkerke	0.483
Overall	Classification Outcome	0.836
Odds Ratio	Trustworthy	23.342***

Dependent Variable: Imagine Neobroker
Independent Variable: Trustworthy

To perform a statistically significant evaluation, I run an OLS regression in addition to the descriptive presentation. The regression is based on the *risk score* of the stock as the dependent variable and the *survey version* as the independent variable achieving an *adjusted R^2* value of 8.7 percent and a *p-value* of 0.000. The estimated average increase in risk score when looking at the daily chart version of the stock price is 0.605 points (*unstandardized beta*), or 12.1 percent. Based on the linear regression, a statistically relevant influence of the design of and representations provided by neobrokers on the investment behavior can be identified, which confirms the fifth hypothesis of my thesis.

Table 5: Binary Logistic Regression for Testing H2.

This table shows the results of the binary logistic regression used to test the second hypothesis. Here, I examine the influence of financial literacy, investment experience, and age (independent variables) on the possible intention to use a neobroker to purchase financial products (dependent variable). 223 responses are available for this analysis. $*$, $**$, and $***$ represent significance below the 10%, 5%, and 1% level, respectively.

		Results
P-Value	Omnibus Tests	0.000
R^2	Nagelkerke	0.215
P-Value	Hosmer-Lemeshow-Test	0.704
Overall	Classification Outcome	0.686
Odds Ratio	Age	0.580***
	Financial Literacy Score	1.419*
	Investment Experience	3.738***

Dependent Variable: Imagine Neobroker
Independent Variables: Age, Financial Literacy Score, Investment Experience

Table 6: Comparison of two Versions of Questionnaire for Testing H5.

This table shows the descriptive results of the risk assessment on a scale of 1 to 5, with 1 representing *low risk* and 5 representing *very risky*, and the influence on the respondents' decision, depending on which version of the questionnaire they answered. In particular, the second question asks the respondents to what extent they would consider trading the above-mentioned stock in the questionnaire.

		Daily Version		Yearly Version	
		Total	Percent of Cases	Total	Percent of Cases
Risk Score	1	1	0.9	4	3.4
	2	14	13.1	41	35.0
	3	31	29.0	29	24.8
	4	44	41.1	40	34.2
	5	17	15.9	3	2.6
Influence Decision	Thinking about buying	6	5.6	15	12.8
	Thinking about selling	8	7.5	17	14.5
	Interested in stock	33	30.8	60	51.3
	Too risky	48	44.9	10	8.5
	Ignoring	12	11.2	15	12.8

The set of variables created for the CATA question regarding the current use of financial service providers is utilized for the first time to test the sixth and seventh hypotheses. Using this set of variables, it is possible to create cross-tabulations, with which a correlation between the CATA question and the respective variables under consideration can be examined.

164 responses are available to me for the analysis of the sixth hypothesis. Looking at the risk that the respondent is willing to take for the chance of higher profits, it is clear that neobroker customers would, on average, take a significantly higher risk than customers of traditional banks. While respondents who are at least a customer of a branch bank give an average risk score of 2.8 out of 5 (2.7 if they are solely customers of branch banks), customers of neobrokers give an

Table 7: Responses to *Popular Stocks* and *Invite your Friends* Representations.

This table shows the descriptive results to the questions on how people would react to the presentation of *Popular Stocks* on the one hand and to the *Invite your Friends* campaign on the other hand. In total, all 224 responses can be used for the breakdown.

		Total	Percent of Cases
Popular Stocks	Investigating in Stock	80	35.7
	Quick Investment	2	0.9
	Happy because of Presentation	22	9.8
	Investigating Chart	51	22.8
	Ignoring	64	28.6
	Prefer not to say	5	2.2
Friends Invite	Showing Friends	60	26.8
	Convince Friends	25	11.2
	Interesting	62	27.7
	Ignoring	77	34.4

Table 8: OLS Regression for Testing H5.

This table shows the results of the OLS regression used to test the fifth hypothesis. Here, I examine the influence of the survey version (independent variable) on the respondents' risk assessment for the same stock (dependent variable) presented differently in both versions. 224 responses are available for this analysis. *, **, and *** represent significance below the 10%, 5%, and 1% level, respectively.

	Results
Adjusted R^2	0.087
P-Value	0.000
Unstandardized Beta	0.605***

Dependent Variable: Risk Score
Independent Variable: Survey Version

average score of 3.4. By comparison, the average risk score for direct bank customers is 3.1, or 3.0 if the respondents have an account exclusively with that provider. In addition, 47 percent of the respondents, who are at least customers of a neobroker, report a score of 4 or higher, compared to only 30 percent of branch bank customers and 32.7 percent of direct bank customers.

Following the cross-tabulation, I run an OLS regression to test the sixth hypothesis. Only those participants who selected at least one response option on the CATA question will be considered to obtain a meaningful result (164 respondents). For the OLS regression, the dummy variable *UNeobroker*, representing *current user of a neobroker*, is chosen as the reference category to illustrate the relation with the other dummy variables. The *adjusted R^2* value of 8.7 percent, which can be extracted from the model summary, is considered respectable, as humans are much more difficult to predict than physical processes (Roslan & Shafri, 2018). Furthermore, based on the *p-value* of 0.001, which again is

below the typical rejection threshold of 0.05, it can be concluded a significant impact of the choice of a financial service provider on the risk tolerance of the respondents. Compared with the reference category *UNeobroker*, the subjects' willingness to take risk decreases by -0.748 (*unstandardized beta*) if they use a branch bank, corresponding to 14.96 percent, or by 9.7 percent (*unstandardized beta* = -0.485) if they use a direct bank exclusively as their financial service provider. If the respondent is a customer at multiple financial service providers but not at a neobroker, risk tolerance decreases by 14.96 percent (*unstandardized beta* = -0.748), as well. In addition, the tendency to take risks decreases by 0.045 (*unstandardized beta*), representing 0.9 percent, if the participant is a user of several financial service providers, including a neobroker. However, this result is not significant (*p-value* = 0.844). Conversely, it can be said that risk-taking is highest among participants who are exclusively users of neobrokers. Based on these results, my sixth hypothesis, stating that neobroker customers are willing to take more risk compared to

users of traditional banks, can be confirmed.

The number of responses for testing the seventh hypothesis is additionally reduced by four participants who, compared to the previous ones, did not want to answer the question about how often they trade in financial products. The remaining responses indicate that on average customers of branch banks buy or sell financial products once in six months, customers of direct banks once in three months, and users of neobrokers even once a month. Looking again at only those participants who are exclusively customers of one type of broker, the results are as follows: on average, branch bank customers trade less than once in six months, direct bank customers range between once in three and once in six months but tend towards the former, and neobroker users trade once in three months.

Running an OLS regression with the same selected observation group and dummy variables as in the OLS regression before, a correlation of trading frequency with using a neobroker is also observed. An *adjusted R^2* value of 27.1 percent and a *p-value* of 0.000 form the basis for the following analysis. Both when using a branch bank (*unstandardized beta* = 1.691) or a direct bank (*unstandardized beta* = 0.352), the period for trading frequency increases compared to users of neobrokers, although the result related to direct banks is not statistically significant (*p-value* = 0.354). The time interval between trades also increases if the respondent is a customer of both a branch bank and a direct bank, but not a user of a neobroker (*unstandardized beta* = 0.920). However, this result is not statistically significant as well due to the *p-value* of 0.077. If the respondent is not exclusively a customer of a branch or direct bank, but also a user of a neobroker, the trading frequency decreases (*unstandardized beta* = -0.964). Based on the previous results, I can only speculate that the frequent trading and interaction is carried out with neobrokers and the long-term investing with few transactions is done using the traditional brokers. Nevertheless, it can be concluded that customers of neobrokers have a shorter investment horizon compared to customers of traditional banks, thus confirming the seventh hypothesis of my thesis.

The eighth hypothesis indicates that neobroker customers want the reduced and simple design compared to customers of traditional banks. To verify this assumption, I first look at the evaluation of the 22nd and 28th CATA questions of my survey. I consider two groups here: first, all respondents who report being a customer of neobrokers (68 respondents), and second, all who are customers of other banks but not of neobrokers (96 respondents). *Fast buying and selling options* (53 percent of responses from all neobroker customers), *intuitive handling* (51 percent), as well as a *user-friendly application environment,* and *low costs* (38 percent each) are particularly important to neobroker customers when using financial institutions' online services. However, the *appealing design* of an application only ranks in a divided fifth place with 29 percent of the mentions. Looking at respondents who are not customers of neobrokers but of other broker providers, similar results can be obtained. *Intuitive handling* (56 per-

cent), *data security* (49 percent), and a *user-friendly application environment* (46 percent) are the most important aspects for non-neobroker customers. *Appealing design*, on the other hand, ranks seventh at 20 percent, one position behind the responses of neobroker customers. When asked what aspects participants associate with neobrokers, *appealing design* again finishes in a split fifth (neobroker customers) and seventh place (non-neobroker customers).

The evaluation of the descriptive analysis does not enable me to already draw a statistically relevant conclusion about my previously stated hypothesis. Therefore, I conduct a binary logistic regression using *user neobroker* as the dependent variable to test the influence of whether the appealing design of a financial service provider's application is one of the most important features for a neobroker customer. Consequently, I consider all responses for the analysis. Based on the *p-value* of 0.077, it can be concluded that the new model with the inclusion of the independent variable is not an improvement over the baseline model. The new model has a poor fit between model and data (*Nagelkerke R^2* = 0.020), indicating low explanatory power, and can correctly classify 69.6 percent of the cases. If the design of a financial service provider's application is important to the respondent, the probability that the participant is a neobroker user increases by a factor of 1.825 (*odds ratio*), or 82.5 percent. However, this result is not statistically significant (*p-value* = 0.074). As a consequence, based on the obtained results, I cannot confirm the eighth hypothesis in a statistically significant way.

For the ninth hypothesis to be tested, it is important to consider only those participants of my study who are active customers of at least one type of financial service provider. Therefore, the 164 responses already mentioned remain for consideration again. Since the ninth hypothesis is about the assumption that neobroker customers do not want to be provided with further information within the application compared to customers of traditional brokers, I evaluate the last two questions of my survey at this point. In response to the first question about how helpful the respondents find a learning section based on a representation in my survey on a scale from 1 *not helpful* to 5 *very helpful*, the mean scores of 4.26 (customers of branch banks only), 3.98 (customers of direct banks only), and 4.24 (customers of neobrokers only) are obtained. If the respondent is a customer of both a branch bank and a direct bank, but not of a neobroker, the mean score for the helpfulness of a learning section is exactly 4. The average score of 3.98 is achieved if the respondent is not only a customer of a neobroker but also of another provider of financial services. The following mean values result from the question of whether it would be useful to test the previously acquired knowledge in the form of a quiz: 3.72 (customers of branch banks only), 3.66 (customers of direct banks only), 4.08 (customers of neobrokers only), 3.46 (customers of branch and direct banks, but not of neobrokers) and 3.86 (customers of neobrokers and other providers of financial products). Based on the descriptive analysis, the hypothesis presented above cannot be confirmed. Additional information and knowledge checks seem to be just as desirable for neobroker customers

Table 9: Willingness to take Risk.

This table shows the descriptive results to the question of how much risk the respective respondent is willing to take if there is a chance of higher profits. The results are first divided into respondents who are at least customers of a financial service provider and then those who are exclusively customers of a financial service provider. The risk is rated on a scale of 1 to 5, with 5 representing high risk.

		Willingness to take Risk					
		1	2	3	4	5	Mean
At least	Branch Bank	4	26	19	20	1	2.8
Customer of	Direct Bank	2	22	40	26	5	3.1
	Neobroker	0	13	23	23	9	3.4
Exclusively	Branch Bank	3	15	12	9	0	2.7
Customer of	Direct Bank	1	10	24	8	1	3.0
	Neobroker	0	4	10	7	4	3.4

Table 10: OLS Regression for Testing H6.

This table shows the results of the OLS regression used to test the sixth hypothesis. Here, I examine the impact of using different financial service providers (independent variables) on respondents' willingness to take risk for the chance of higher profits (dependent variable). 164 responses are available for this analysis. *, **, and *** represent significance below the 10%, 5%, and 1% level, respectively.

		Results
Adjusted R^2		0.087
P-Value		0.001
Unstandardized Beta	UBranchBank	-0.748***
	UDirectBank	-0.485**
	UNoNeobroker	-0.748**
	UNeobrokerPlus	-0.045

Dependent Variable: Willingness to take Risk
Independent Variables: UBranchBank, UDirectBank, UNoNeobroker, UNeobrokerPlus
Reference Category: UNeobroker

Table 11: Investment Frequency of Respondents (at least one Financial Service Provider).

This table shows the descriptive results for the question of how often the respective respondent trades in financial products. The results are related to all those respondents who are at least customers of one financial service provider.

	How often do you buy or sell financial products?							
	Almost Daily	2-3 times a week	Once a week	Once a month	Once in 3 months	Once in 6 months	Less than once in 6 months	Never
Branch Bank	1	2	3	19	1	6	17	17
Direct Bank	1	3	9	39	8	10	18	7
Neo-broker	1	3	9	33	7	4	9	2

as for customers of traditional banks.

In addition to the analysis presented above, I also run two OLS regressions to obtain statistically relevant results. The dummy variable *UNeobroker* is again chosen as the ref-

Table 12: Investment Frequency of Respondents (exactly one Financial Service Provider).

This table shows the descriptive results for the question of how often the respective respondent trades in financial products. The results are related to all those respondents who are customers of exactly one financial service provider.

	How often do you buy or sell financial products?							
	Almost Daily	2-3 times a week	Once a week	Once a month	Once in 3 months	Once in 6 months	Less than once in 6 months	Never
Branch Bank	0	0	1	5	0	3	12	14
Direct Bank	0	0	1	17	7	4	11	4
Neo-broker	0	0	1	12	5	0	5	2

Table 13: OLS Regression for Testing H7.

This table shows the results of the OLS regression used to test the seventh hypothesis. Here, I examine the impact of using different financial service providers (independent variables) on the frequency of trading financial products of the respondents (dependent variable). 160 responses are available for this analysis. *, **, and *** represent significance below the 10%, 5%, and 1% level, respectively.

		Results
Adjusted R^2		0.271
P-Value		0.000
Unstandardized Beta	UBranchBank	1.691***
	UDirectBank	0.352
	UNoNeobroker	0.920*
	UNeobrokerPlus	-0.964**

Dependent Variable: HowOftenBuyOrSell
Independent Variables: UBranchBank, UDirectBank, UNoNeobroker, UNeobrokerPlus
Reference Category: UNeobroker

erence category for the independent variables and both the *learning section* and *quiz* scores are used as dependent variables according to the OLS regression. *Adjusted R^2* values of 0.001 and -0.006, respectively, and *p-values* of 0.378 and 0.544 indicate no significant relationship between customers of different financial service providers and the usefulness of a learning section or a quiz. If the participant is exclusively a customer of a branch bank (learning section: *unstandardized beta* = -0.104, *p-value* = 0.681; quiz: *unstandardized beta* = -0.362, *p-value* = 0.250), of a direct bank (learning section: *unstandardized beta* = -0.383, *p-value* = 0.121; quiz: *unstandardized beta* = -0.421, *p-value* = 0.172), or of both (learning section: *unstandardized beta* = -0.360, *p-value* = 0.285; quiz: *unstandardized beta* = -0.618, *p-value* = 0.141), the respondents' perceived usefulness of a learning section or a quiz decreases relative to neobroker customers. Likewise, the perceived usefulness of both options decreases if the respondent is a customer not only of a neobroker but of additional banks as well (learning section: *unstandardized beta* = -0.383, *p-*

value = 0.122; quiz: *unstandardized beta* = -0.220, *p-value* = 0.477). However, since all obtained results are above the typical rejection threshold of 0.05, they are all not statistically significant and are therefore neglected. In other words, no significant difference in the evaluation of the learning section and the quiz questions between the users of neobrokers and the customers of traditional banks can be observed. From the lack of significant results, I cannot conclude that additional information and also knowledge checks are less desirable for neobroker users than for customers of other broker types. In consequence, the ninth hypothesis cannot be confirmed.

6. Discussion

In the following sections, I review the positive and negative results of my study. Afterward, I discuss the limitations and the methodological criticism before I present the possibilities for further improvements in the last subsection of this chapter.

Table 14: Descriptive Presentation of Evaluation of 22nd and 28th Question.

This table shows the descriptive results to the CATA questions, which aspects are important to the respondents regarding an application of a financial service provider, and which aspects they associate with neobrokers. The results are divided into two groups. One group reflects the responses of neobroker customers, which is the case for 68 people, and the other group includes the responses of participants who are not neobroker users, which is the case for 96 people.

| | What is important in an application? | | What are aspects of Neobroker? | |
	Number of Neobroker User (N = 68)	Number of Non-Neobroker User (N = 96)	Number of Neobroker User (N = 68)	Number of Non-Neobroker (N = 96)
Intuitive	35	54	50	48
Design	20	19	50	37
Pictures	14	21	29	32
Fast	36	35	62	76
User-friendly	26	44	44	36
Short Loading	9	12	40	46
Digital	9	4	53	66
Safety	20	47	10	10
Costs	26	30	59	57
Consultation	3	10	1	4
Easy Account opening	3	4	54	45
Personal	1	5	1	3

Table 15: Binary Logistic Regression for Testing H8.

This table shows the results of the binary logistic regression used to test the eighth hypothesis. Here, I examine the influence of the importance of the design to the respondent (independent variable) on the use of a neobroker (dependent variable). 224 responses are available for this analysis. *, **, and *** represent significance below the 10%, 5%, and 1% level, respectively.

		Result
P-Value	Omnibus Tests	0.077
R^2	Nagelkerke	0.020
Overall	Classification Outcome	0.696
Odds Ratio	Important Design	1.825*

Dependent Variable: User Neobroker
Independent Variables: Important Design

Table 16: Usefulness of Learning Section and Quiz within Neobrokers.

This table shows the descriptive results on the questions to what extent learning sections and a query of what has been learned in the form of a quiz are considered useful by the respondents, divided into the different user groups of different financial service providers. Usefulness is rated on a scale from 1 *not helpful* to 5 *very helpful*.

| | Number of Responses | | | | | | | | | |
| | Usefulness of Learning Section | | | | | Usefulness of Quiz | | | | |
	1	2	3	4	5	1	2	3	4	5
UBranchBank	0	3	2	16	18	2	4	7	16	10
UDirectBank	2	2	5	21	14	5	4	5	17	13
UNeobroker	0	0	3	10	12	2	0	2	11	10
UNoNeobroker	1	1	0	6	5	2	2	1	4	4
UNeobrokerPlus	1	5	4	17	16	2	5	6	14	16

Table 17: OLS Regression for Testing H9.

This table shows the results of the OLS regression used to test the ninth hypothesis. Here, I examine the impact of using different financial service providers (independent variables) on the perception of the usefulness of both a learning section and a knowledge query in form of a quiz (dependent variables). 164 responses are available for this analysis. *, **, and *** represent significance below the 10%, 5%, and 1% level, respectively.

		Learning Section	Quiz
		Results	Results
Adjusted R^2		0.001	-0.006
P-Value		0.378	0.544
Unstandardized Beta	UBranchBank	-0.104	-0.362
	UDirectBank	-0.383	-0.421
	UNoNeobroker	-0.360	-0.618
	UNeobrokerPlus	-0.383	-0.220

Dependent Variable: HowOftenBuyOrSell
Independent Variables: UBranchBank, UDirectBank, UNoNeobroker, UNeobrokerPlus
Reference Category: UNeobroker

6.1. Summary of Results

Based on my online questionnaire and the following analysis, most of my hypotheses already derived from the existing literature regarding neobrokers can be confirmed. To begin with, it can be observed that trust in the new generation of online brokers has a positive impact on the acceptance of neobrokers. Respondents who have trust in neobrokers are 23.342 times more likely to imagine using a neobroker than participants who do not. The assumption that neobrokers have particularly targeted the younger generation is also

confirmed, as the probability of not using a neobroker increases by 72.4 percent with rising age. No statistically significant results, though, can be obtained about people with a low level of financial education as a target group for neobrokers. Therefore, no conclusions can be drawn in this regard. Neobrokers are primarily used for short-term and speculative investments, as the data obtained from my survey suggest based on descriptive statistics. Long-term asset accumulation for one's own retirement provision is not a priority. Unfortunately, these results can only be presented descriptively and

cannot be tested for significance due to a lack of data. It is precisely this type of investment behavior that neobrokers are targeting. By means of their design and the presentation of information, such as stock charts as a daily version by default, neobrokers aim to influence the trading behavior of their users. Based on my results, I can confirm this assumption. A different presentation of the same stock (daily vs. annual chart) leads to an increase in the perceived risk of 12.1 percent, which in turn triggers an adjustment of the respondents' investment behavior. In addition, the outcome of the questionnaire indicates that customers of neobrokers are willing to take a significantly higher risk (between 9.7 percent and 14.96 percent) compared to customers of other financial service providers, as they assume the chance of higher profits. It can also be shown that these neobroker customers have a shorter investment horizon and thus trade financial products more frequently compared to non-neobroker customers. On average, they engage in trading at least once every three months (if not monthly), while customers of branch banks become active once every six months.

However, two of my hypotheses derived from the existing theory cannot be confirmed due to a lack of statistical significance. On the one hand, no significant correlation can be found between neobroker customers and their desire for a reduced and simple design of an application. On the other hand, no significant relationship between neobroker customers and the perceived usefulness of learning sections as well as knowledge checks can be elaborated. These knowledge transfers seem to be just as intentional for users of neobrokers as for customers from other broker providers.

6.2. Limitations and Methodological Criticism

In this section, I address the limitations of my thesis and critically examine the methodology I conducted. The first factor to mention is the limited number of participants due to a tight time frame for the development of my study. A larger field of participants would allow me to make more general statements than it can be made with my 224 responses available. In addition, a longer period of time could also enable a stronger international orientation of the sample beyond the German-speaking region. In context of my thesis, only two participants out of the entire sample used the English-language version of the online questionnaire, representing less than 1 percent of the total responses collected. Furthermore, the age groups *18 to 24 years* and *25 to 34 years* dominate among the participants (just under 90 percent). This can also be attributed to the short processing period, as no additional time was available to expand the subject group to include, for example, older participants. Finally, it is also critical to see that every tenth answer is not suitable for the evaluation, as these were either incomplete, filled in with meaningless answers, or clicked through quickly. Although these answers were removed, it cannot be ruled out that non-seriously completed questionnaires still remain in the evaluation. This is supported above all by the existing anonymity of a quantitative study like mine. No questionnaire can be traced back to the person who filled it out, which lowers the

inhibition threshold for a quick click-through or giving wrong answers.

Another limitation of my thesis is the structure of the online questionnaire itself. Due to an inaccurate construction of the 26th and 27th questions, I cannot test the third and fourth hypotheses regarding statistical significance, but only to evaluate them based on descriptive data. This could have been avoided if the related questions had been asked differently, so that reference could be made to the respective respondents. Unfortunately, however, since the questionnaire was created before the exact hypotheses were generated, this could not be avoided.

These limitations are primarily a matter of quantitative research per se. In principle, rather general findings are obtained, and complex problems and their causes are not comprehensively recorded. Topics such as a respondent's willingness to take risk has several influencing factors and requires consideration of the individuality of each person. However, these factors cannot be adequately captured using a standardized questionnaire with predetermined choices.

6.3. Possibilities for Improvements

Some of the limitations I mentioned in the previous section can be improved or even eliminated for further studies, starting with the online-based questionnaire. To exclude any responses that were simply clicked through quickly, it is beneficial to record the respondents' processing time. In this way, all questionnaires answered too quickly can be excluded with the help of a minimum processing time that would be necessary for a serious answering of the survey. The use of filter questions can also help exclude such non-serious responses to the questionnaire. Questions such as *Please select option b)* can be used to exclude all respondents who answered the filter question incorrectly due to a quick click-through or careless completion. A longer processing period can also be useful in both gathering a larger sample size for the study and paying more attention to the diversification of respondents to get the most heterogeneous field of participants possible. Likewise, it is advisable to know exactly all the hypotheses before creating the questionnaire. This enables designing the survey precisely according to the requirements necessary for testing the hypotheses.

To minimize the general limitations of a quantitative study, it is advisable to use a mixture of quantitative and qualitative research instruments for further studies. Hereby, the advantages of both study elements can be combined, and thus the respective weaknesses are reduced or eliminated. However, the use of such a mixed research method is also associated with disadvantages, since this research design entails, among other things, a considerable amount of additional work (additional costs, the time factor, etc.), which can hardly be managed by a single researcher and therefore requires a research group (Johnson & Onwuegbuzie, 2004).

7. Conclusion

As already presented in chapter *2.1 Development of Neobrokers*, the publicly proclaimed goal of neobrokers, especially that of the two considered within my study, is to democratize access to stock markets and make trading available for everyone. One of the ways this is achieved is by providing mobile trading applications that feature a sleek and appealing design. In addition, other factors such as low to no trading costs and easy opening of one's own account within the brokers support the declared goal of attracting as many new, especially young customers as possible. This allows users of neobrokers to profit from the stock market even with small capital and they are not deterred by high hurdles, as it is the case with larger and long-established brokers.

I claim that gamified elements and reduced information within the neobroker applications intentionally entice their own, mostly young customers to participate more actively and frequently in the stock market, even tempting them to make imprudent and risky investments. This is where representations such as the Top Mover lists, the standardized displays of daily charts, or even the game-like interactions of the users within the application come into play. This playful and influencing approach of neobrokers not only tempts further interactions with the broker itself but can also have disastrous consequences, as the example of a 20-year-old who became overindebted through highly speculative investments and later committed suicide (van der Heide & Želinský, 2021) shows. According to my study, the dangerous nature of neobrokers can be confirmed by the fact that they are preferred for speculative purposes and not for long-term investments which would help to avoid high market fluctuations, such as those currently triggered by the Corona pandemic.

From these results, no clear statement can be made as to whether neobrokers are a blessing or a curse for individual users. Rather, they are a mixture of both in this respect. On the one hand, neobrokers provide the opportunity for anyone, rich or poor, young or old, experienced or inexperienced, to participate in the stock market without major hurdles and thus the chance to also benefit from the general upswing of market economies over the years. On the other hand, however, neobrokers do not act out of pure charity, as the late medieval heroic figure from literature, Robin Hood, once did. Rather, they aim to entice their users to interact more frequently with the stock markets within their applications, as they receive a percentage of every transaction. It does not matter whether the investment is successful or not, as long as buying and selling is done a lot and often, from which, as shown in my study, a danger for the user is generated.

Due to major events such as the aforementioned short squeeze of GameStop Corporation, the negative aspects of neobrokers are increasingly becoming a main focus of the media and society. This raises the question of how neobrokers can continue to grow while reducing the negative influences on the curse side. Based on my study, I propose to not only utilize the already established gamified approach of the neobroker applications for influencing the trading behavior of its users, but also to make a beneficial implementation at the same time. I argue that new ideas such as integrating a learning section and playfully testing the newly acquired knowledge are not rejected by neobroker customers, as the literature suggests, but are found to be as useful as by customers of well-established brokers with long-term investment goals. By adding an educational aspect to the offering, the mostly young investors can delve deeper into the important topic of investing themselves and thus make their own decisions based on their recently gained in-depth knowledge.

The topic of neobrokers, in general, is still an underresearched field. Although some studies have already dealt with the broker Robinhood, it is difficult to draw general conclusions from them, as this broker is only available in the American market and thus covers just a small part of the entire neobroker market. Looking to the future and more general statements, further studies are therefore required in this regard, also to strengthen the recommendations I have made concerning a learning section. The topic of *educational mediation within neobroker applications* needs to be explored in more detail before a meaningful and universally valid suggestion can be made. For this purpose, it is important to have a larger share of neobroker customers in the field of participants to be able to reflect their actual interests and opinions. Furthermore, in my study, I did not question the individual motivations for respective decisions. Since my survey instrument is a predefined, online-based questionnaire, only the given answers are analyzed and not elaborated in depth. As a result, important information, for example why neobrokers are not used for one's own retirement planning, is lost. I leave such a more detailed evaluation of individual motivations to future research.

Neobrokers are becoming increasingly popular and provide general access to the stock markets. This can boost the global economy and offer everyone the chance to benefit from it. However, all that glitters is not gold. As recent events have shown, a merger of numerous retail investors through the same broker can not only drive-up stock prices but also pose great danger. It is the responsibility of neobrokers to educate and warn their customers about these precise risks. After all, financial education and conscious investing are of tremendous importance, not only to help neobroker customers succeed in the stock market, but also to save users' lives by preventing them from getting into serious trouble.

References

Albuquerque, R., Koskinen, Y., Yang, S., & Zhang, C. (2020). Resiliency of Environmental and Social Stocks: An Analysis of the Exogenous COVID-19 Market Crash. *The Review of Corporate Finance Studies*, 9(3), 593–621.

Ash, J., Anderson, B., Gordon, R., & Langley, P. (2018). Digital interface design and power: Friction, threshold, transition. *Environment and Planning D: Society and Space*, 36(6), 1136–1153.

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2016). *Multivariate Analysemethoden*. Berlin, Heidelberg, Springer.

Balonier, S., Di Dio, D., & Fey, G. (2021). Deutschland und die Aktie: Eine neue Liebesgeschichte? *Deutsches Aktieninstitut*.

Barber, B. M., Huang, X., Odean, T., & Schwarz, C. (2020). Attention Induced Trading and Returns: Evidence from Robinhood Users. *SSRN Electronic Journal*.

Barber, B. M., & Odean, T. (2001). The Internet and the Investor. *Journal of Economic Perspectives*, 15(1), 41–54.

Barber, B. M., & Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *Review of Financial Studies*, 21(2), 785–818.

Barber, B. M., Odean, T., & Zhu, N. (2009). Do Retail Trades Move Markets? *Review of Financial Studies*, 22(1), 151–186.

Bernards, N. (2019). The poverty of fintech? psychometrics, credit infrastructures, and the limits of financialization. *Review of International Political Economy*, 26(5), 815–838.

Bruner, R. F., & Miller, S. C. (2020). The First Modern Financial Crises: The South Sea and Mississippi Bubbles in Historical Perspective. ., 32(4), 17–33.

Chaudhry, S., & Kulkarni, C. (2021). Design Patterns of Investing Apps and Their Effects on Investing Behaviors. In W. Ju, L. Oehlberg, S. Follmer, S. Fox, & S. Kuznetsov (Eds.), *Designing interactive systems conference 2021* (pp. 777–788). New York, NY, USA, ACM.

Da, Z., Engelberg, J., & Gao, P. (2011). In Search of Attention. *The Journal of Finance*, 66, 1461–1499.

de Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1989). The Size and Incidence of the Losses from Noise Trading. *The Journal of Finance*, 44(3), 681–696.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness. In A. Lugmayr, H. Franssila, C. Safran, & I. Hammouda (Eds.), *Proceedings of the 15th international academic mindtrek conference on envisioning future media environments - mindtrek '11* (p. 9). New York, New York, USA, ACM Press.

Diekmann, A. (2020). *Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen*. Reinbek bei Hamburg, Rowohlt Taschenbuch Verlag. (13. Auflage, vollständig überarbeitete und erweiterte Neuausgabe August 2007, Orginalausgabe)

Dukas, R. (2004). Causes and consequences of limited attention. *Brain, Behavior and Evolution*, 63(4), 197–210.

Everitt, B. (2010). *The Cambridge dictionary of statistics* (4th edition ed.). Cambridge and New York, Cambridge University Press.

Fischer, R., Hübner, M., & Bulis, P. (2020). *Online-Wertpapier-Brokerage: Der Kampf um die nächste Milliarde im deutschen Markt*. https://www.oliverwyman.de/content/dam/oliver-wyman/v2-de/publications/POV_Oliver%20Wyman_Online%20Brokerage_web.pdf. (Accessed: 2021-12-18)

Foucault, T., Sraer, D., & Thesmar, D. J. (2011). Individual Investors and Volatility. *The Journal of Finance*, 66(4), 1369–1406.

Friedman, H. H., & Friedman, L. W. (2009). The Global Financial Crisis of 2008: What Went Wrong? *SSRN Electronic Journal*.

Garber, P. M. (1990). Famous First Bubbles. *Journal of Economic Perspectives*, 4(2), 35–54.

Gefen, D. (2000). E-commerce: the role of familiarity and trust. *Omega*, 28(6), 725–737.

Gomber, P., Koch, J.-A., & Siering, M. (2017). Digital Finance and FinTech: current research and future research directions. *Journal of Business Economics*, 87(5), 537–580.

Greenwood, R., & Nagel, S. (2009). Inexperienced investors and bubbles. *Journal of Financial Economics*, 93(2), 239–258.

Hobson, T. (2018). *Gamification in the Most Delightful Way*. https://medium.com/pinch-pull-press/gamification-in -the-most-delightful-way-504caf72c1bc. Pinch Pull Press. (Accessed: 2021-11-29)

Holt, J. (2009). A Summary of the Primary Causes of the Housing Bubble and the Resulting Credit Crisis: A Non-Technical Paper. *The Journal of Business Inquiry*, 8(1), 120–129.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression*. Wiley.

Iqbal, K., Islam, A., List, J., & Nguyen, V. (2021). *Myopic Loss Aversion and Investment Decisions: from the Laboratory to the Field*. Cambridge, MA, National Bureau of Economic Research.

Jaeger, S. R., Beresford, M. K., Paisley, A. G., Antúnez, L., Vidal, L., Cadena, R. S., ... Ares, G. (2015). Check-all-that-apply (CATA) questions for sensory product characterization by consumers: Investigations into the number of terms used in CATA questions. *Food Quality and Preference*, 42, 154–164.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, 33(7), 14–26.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263.

Kranz, T. T., Teschner, F., & Weinhardt, C. (2015). Beware of Performance Indicators. *Business & Information Systems Engineering*, 57(6), 349–361.

Leong, K. (2018). FinTech (Financial Technology): What is It and How to Use Technologies to Create Business Value in Fintech Way? *International Journal of Innovation, Management and Technology*, 74–78.

Lusardi, A., & Mitchell, O. S. (2011). Financial Literacy around the world: An Overview. *Journal of pension economics & finance*, 10(4), 497–508.

MacDonald, S., & Headlam, N. (2008). *Research Methods Handbook: Introductory guide to research methods for social research*. London, CLES.

Maciejewski, G., & Lesznik, D. (2020). An outline of Polish investors' and stock traders' profiles. *Marketing i Rynek. Journal of Marketing and Market Studies*, 2020(2), 25–34.

Melnick, E. L., Nayyar, P. R., Pinedo, M. L., & Seshadri, S. (Eds.). (2000). *Creating Value in Financial Services: Strategies, Operations and Technologies*. Boston, MA, Springer US.

Odean, T. (1998). Volume, Volatility, Price, and Profit When All Traders Are Above Average. *The Journal of Finance*, 53(6), 1887–1934.

O'Keefe, D. J. (2007). Brief Report: Post Hoc Power, Observed Power, A Priori Power, Retrospective Power, Prospective Power, Achieved Power: Sorting Out Appropriate Uses of Statistical Power Analyses. *Communication Methods and Measures*, 1(4), 291–299.

Robinhood. (2021a). *About Us*. https://robinhood.com/us/en/about-us/.

Robinhood. (2021b). *Invite Friends, Move Up the Wallets Waitlist*. https://robinhood.com/us/en/support/articles/invite-friends-move-up-the-crypto-wallets-waitlist/.

Roslan, S. N. A., & Shafri, H. Z. M. (2018). Developing building comfort index from building climate factors in a tropical urban environment. *IOP Conference Series: Earth and Environmental Science*, 169, 012014.

Schonlau, M., & Toepoel, V. (2015). Straightlining in Web survey panels over time. Advance online publication.

Schunk, D. H. (2012). Social cognitive theory. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *Apa educational psychology handbook, vol 1: Theories, constructs, and critical issues* (pp. 101–123). Washington: American Psychological Association.

Seasholes, M. S., & Wu, G. (2007). Predictable behavior, profits, and attention. *Journal of Empirical Finance*, 14(5), 590–610.

Shiva, A., Narula, S., & Shahi, S. (2020). What drives retail investors' investment decisions? Evidence from no mobile phone phobia (nomophobia) and Investor Fear of Missing Out (I - FoMO). *Journal of Content, Community and Communication*, 11(10), 2–20.

Sironi, P. (2016). *FinTech Innovation: From robo-advisors to goal based investing and gamification*. Chichester, UK, John Wiley & Sons, Ltd.

Stein, R. (2020). The Top 5 Predictable Effects of New Entries in Robinhood's '100 Most Popular' List. *SSRN Electronic Journal*.

Tan, G. (2021). Democratizing finance with Robinhood: Financial infrastructure, interface design and platform capitalism. *Environment and*

Planning A: Economy and Space, 1–17.

Teschner, F., Kranz, T. T., & Weinhardt, C. (2015). The Impact of customizable Market Interfaces on Trading Performance. *Electronic Markets*, *25*(4), 325–334.

Thaler, R. H., Tversky, A., Kahneman, D., & Schwartz, A. (1997). The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test. *The Quarterly Journal of Economics*, *112*(2), 647–661.

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*(2), 77–99.

Thompson, K., & Blomquist, E. (2017). *Millennails & Money: The Millennial Investor Becomes a Force.* https://www.accenture.com/_acnmedia/PDF-68/Accenture-Millennials-and-Money-Millennial-Next-Era-Wealth-Management.pdf. (Accessed: 2021-12-12)

Trade Republic. (2021). *Wie kann ich am Refer-a-Friend-Programm teilnehmen?* https://support.traderepublic.com/de-de/174-Wie-kann-ich-am-Refer_a_Friend_Programm-teilnehmen. (Accessed: 2021-12-20)

van der Heide, A., & Želinský, D. (2021). 'Level up your money game': an analysis of gamification discourse in financial services. *Journal of Cultural Economy*, 1–21.

White, E. N. (1990). The Stock Market Boom and Crash of 1929 Revisited. *Journal of Economic Perspectives*, *4*(2), 67–83.

Wu, I.-L., & Chen, J.-L. (2005). An extension of Trust and TAM model with TPB in the initial adoption of on-line tax: An empirical study. *International Journal of Human-Computer Studies*, *62*(6), 784–808.

Yousafzai, S. Y., Pallister, J. G., & Foxall, G. R. (2003). A proposed model of e-trust for electronic banking. *Technovation*, *23*(11), 847–860.

Zavolokina, L., Dolata, M., & Schwabe, G. (2016). *FinTech – What's in a Name?*

Junior Management Science

Rethinking Digital Governance – How Collaborative Innovation Strategies Advance the Development of Digital Innovations in Public Organisations

Laura John

Westfälische Wilhelms-Universität Münster

Abstract

Digital innovations bear the potential to increase the efficiency and transparency of governments and create more accessible and user-centric public services. However, public organisations are facing several challenges in the development of digital innovations and the unique democratic requirements imply that digital services cannot simply be procured from private contractors. Hence, a new strategy called collaborative innovation appears to be a possible solution, but few institutional designs have been found to sustain collaborative innovation in the present governance system. Therefore, this thesis investigates the unexplored phenomenon of innovation fellowship programmes by conducting a diagnostic case study about Tech4Germany with a focus on the research question: How does the fellowship programme Tech4Germany contribute to the development of digital innovations in German public sector organisations? Interviews with participants of Tech4Germany revealed that fellowship programmes are a suitable institutional design for collaborative innovation as they create an attractive setting for tech experts and provide public employees a unique room to experience agile and user-centric approaches. In particular, mutual learning is stimulated, and the implementation resistance reduced.

Keywords: Collaborative innovation; Fellowship programmes; Public innovation; Digital transformation.

1. Introduction

The interest in public innovation has intensified among public administration scholars. The public sector is facing new challenges such as ageing society, skilled labour shortage and various 'wicked' problems, hence a growing number of scholars agree that innovation is central to the public organisation's capacity to deal with those challenges (DeVries, Bekkers, & Tummers, 2016; Hartley, Sørensen, & Torfing, 2013; Lindsay et al., 2017). Particularly, the emergence of advanced technologies like robotic process automation, artificial intelligence (AI), or machine learning offer the potential to increase the efficiency, transparency, and effectiveness of governments (Gil-Garcia, Dawes, & Pardo, 2018). Furthermore, digital innovations can potentially be a response to rising citizen expectations towards faster, more accessible, and user-centric public service delivery (Mergel, 2019). However, public organisations are encountering several challenges in the development of digital innovations. It has been well-rehearsed in the pertinent literature that traditional organisational characteristics of public organisations such as hierarchy, silo structures, and red tape are severe barriers to public innovation (Bommert, 2010; Damanpour, 1991; DeVries,

Tummers, & Bekkers, 2018; Rainey, 2014). Moreover, the political environment and multitude of stakeholders tend to increase the complexity of public services and public organisations have a comparatively small innovation budget as well as long-term financial planning horizons (Borins, 2001; Cinar, Trott, & Simms, 2019). Of particular relevance for digital innovations is the lack of personnel with the required technical skills and a deficient usage of modern working methods like design thinking and agile project management (Coglianese, 2020; Mergel, 2016).

At the same time, it has been shown that public employees are not intrinsically less innovative and there are also distinctive drivers of innovation in the public sector, such as the political and normative pressure for improvement (Hartley et al., 2013; Sørensen & Torfing, 2011). Moreover, it is acknowledged that the unique rules and procedures in the public sector usually serve a democratic purpose, for instance, equal opportunities (Neumann, Matt, Hitz-Gamper, Schmidthuber, & Stürmer, 2019). Therefore, digital innovations are only justifiable when they create public value and it is emphasised that practices and solutions from the private sector cannot simply be transferred to the public sec-

tor (Heartley, 2005). Additionally, scholars point to the risk of an overreliance on external information technology (IT) providers as private contractors "may not be sufficiently sensitive to the particular demands on government for explainability, due process, or avoidance of bias" (Coglianese, 2020, p. 49). Given this tension between potential benefits and risks, it is a key question in the research of public innovation how to develop digital innovations *within* public sector organisations.

In recent years, a new strategy to spur public innovation has been developed which is termed *collaborative innovation* and can be defined as "multi-actor collaboration that [. . .] may foster innovation by bringing together public and private actors with relevant innovation assets, facilitating knowledge sharing and transformative learning, and building joint ownership to new innovative visions and practices" (Sørensen & Torfing, 2012, p. 1). This new research field combines theoretical knowledge about collaborative governance (Ansell & Gash, 2008) with findings and theories of innovation (Borins, 2001; Chesbrough, 2003; Greenhalgh, Robert, Macfarlane, Bate, & Kyriakidou, 2004; Moore & Hartley, 2008). As such, collaboration is not new to the public administration literature but is closely related to the concepts of co-production (Verschuere, Brandsen, & Pestoff, 2012) and co-creation (Voorberg, Bekkers, & Tummers, 2015). However, co-production has primarily been used to encourage user participation in public service delivery (Fledderus, Brandsen, & Honingh, 2014) and has only recently been extended to the question of public sector innovation. Here, the work of Sørensen and Torfing (2011) and Torfing (2016) are frequently referred to as the key theoretical foundation (DeVries et al., 2018).

The emergence of *collaborative innovation* can be embedded in the wider paradigm shift towards the New Public Governance (NPG) model (Osborne, 2006). Hence, it can be distinguished from *hierarchical innovation strategies* which have been predominant in the Traditional Public Administration (TPA) paradigm and *competitive innovation strategies* that emerged in the New Public Management (NPM) era (Hartley et al., 2013; Heartley, 2005; Torfing, 2019). As hierarchical and competitive strategies inhere some limitations, collaborative innovation has been suggested as one possible approach for public organisations to mitigate the emerging challenges (Hartley et al., 2013; Torfing, 2019). A growing number of studies found empirical evidence for a positive impact of multi-actor collaboration (McGann, Wells, & Blomkamp, 2021; Neumann et al., 2019). At the same time, recent literature is concerned with the challenges and limitations of collaborative innovation (Agger & Sørensen, 2018; Torfing, Sørensen, & Røiseland, 2019; Wegrich, 2019).

However, the question of how collaborative innovation can be supported and sustained in the present governance system has not yet been resolved since there is no history or tradition of systematically opening up public bureaucracies to citizens and third sector organisations (Bommert, 2010; Hartley et al., 2013; Torfing, 2019). Mostly, collaborative innovation appears either in form of one-time projects (see for

instance Neumann et al., 2019 and Lindsay et al., 2017) or in institutionalised and permanent innovation labs (Gascó, 2017; McGann, Blomkamp, & Lewis, 2018; Whicher & Crick, 2019). The former induces the problem that learning effects about the collaborative process itself cannot be sustained and the scope is relatively limited (Torfing, 2016). The latter faces the challenge that innovations are again developed in a separate setting outside the public organisation and often constricted to the development of ideas, thus facing challenges in reaching the implementation stage (McGann et al., 2021). A very new phenomenon that has not been covered by the peer-reviewed literature so-far are *innovation fellowship programmes*. Hereby, citizens with relevant skills and expertise spend a short-term stay in the federal government to collaboratively create innovative solutions for specific problems together with the government employees. Mergel (2016, p. 520) mentioned these innovation fellowship programmes as "innovative HR policies" in the context of agile innovation management in public organisations, yet there is a crucial research gap since this new practice has neither been systematically analysed nor connected with the existing scientific knowledge about collaborative innovation. As a growing number of scholars emphasize the untapped potential of collaborative innovation (Hartley et al., 2013) and fellowship programmes might be a suitable approach to overcome the outlined barriers, it is of high scientific importance to analyse the mechanisms, potentials, and limitations of this new practice. This thesis aims to fill that gap by conducting a diagnostic case study about Tech4Germany, a fellowship programme that brings citizens with relevant technical skills into the German national ministries for three months to develop digital innovations for prevailing problems. To systematically investigate the new phenomenon of innovation fellowship programmes, the following research question will be answered in the course of this research:

> *How does the fellowship programme Tech4Germany contribute to the development of digital innovations in German public sector organisations?*

From the main research question, three sub-questions are derived to guide the analysis of this case study in alignment with the existing knowledge about collaborative innovation. Each sub-question investigates one phase of the innovation cycle. Thereby, it is aimed to develop an in-depth understanding of the fellowship programme by analysing whether the theoretical mechanisms of collaborative innovation are observable. It is asked:

(1) In what ways does the fellowship programme Tech4Germany enable the collaboration of empowered actors to contribute to the generation of innovative ideas?

(2) To what extent does the fellowship programme Tech4Germany stimulate mutual and transformative learning to contribute to the development of digital innovations?

(3) To what extent does the fellowship programme Tech4Germany stimulate the creation of joint ownership of

the innovation among the involved actors to facilitate the implementation of digital innovations?

By answering the proposed research questions, the study adds to the existing scholarship in several ways. Firstly, since fellowship programmes are emerging as a third space between one-time projects and innovation labs, this theory-guided case study provides highly relevant scientific insights to the question of how collaborative innovation can be initiated and sustained in public organisations. Secondly, by connecting the theorised mechanisms of collaborative innovation with empirical observations of a so-far unexplored form of collaboration, the study adds to the existing knowledge about the causal relationship between collaboration and innovation. Lastly, the study provides empirical findings about digital innovations which have only recently begun to receive attention from public administration researchers (DeVries et al., 2016).

Therefore, the study is also of high social relevance. The digital transformation is one of the most pervasive transformations of the public sector (Dunleavy, Margetts, Bastow, & Tinkler, 2006). Governments increasingly use new forms of data analysis and emerging technologies that no longer solely automate existing processes but instead induce entirely new forms of governing societies and running public organisations (Gil-Garcia et al., 2018; Peeters & Schuilenburg, 2020). This development comes with a severe tension between the potential benefits of digital innovations and various identified risks (Wirtz, Weyerer, & Geyer, 2019). One concern is that public employees rely on technology without having insights and control about how the algorithms produce their results (Peeters & Schuilenburg, 2020). Thus, it is called for new ways to deal with digital innovations and to strengthen the understanding of new technologies among the employees (Wirtz et al., 2019). Since collaboration in form of fellowship programmes may be one approach to increase the in-house capacity of public organisations to develop, implement and oversee new technologies in line with democratic principles (Coglianese, 2020), this thesis entails socially important findings and provides relevant practical implications for policy and governance regarding new digital innovation strategies.

To answer the proposed research questions, the thesis is structured into five major sections. The next chapter provides the theoretical foundation for the case study. Here, the concept and mechanisms of collaborative innovation are granularly elaborated to develop the analytical framework that will guide this study. Subsequently, the methodology will be outlined, including a description of the investigated case. In the fourth chapter, the case Tech4Germany will be analysed by applying the analytical framework to the empirical evidence. Hereafter, a critical discussion of the results and the limitations of the research is presented to answer the three proposed sub-questions. Finally, the thesis concludes with an answer to the overall research question and practical implications as well as suggestions for future research.

2. Theoretical foundation

This chapter provides the theoretical foundation for the research. It starts with conceptualising digital public sector innovation, followed by an elaboration of the concept of collaborative innovation. Subsequently, the evolution of innovation strategies in the public sector is outlined. The fourth section explains the theoretical mechanisms of collaborative innovation, complemented by a description of the key limitations and challenges of collaborative innovation. The chapter concludes with connecting collaborative innovation and digital innovations and a summary of the developed analytical framework.

2.1. Digital public sector innovation: a definition

Digital innovation is a social construct that has different meanings in different contexts. Since there is not one commonly used definition in the public administration literature, this section clarifies how digital innovation is conceptualised in the study. On the most general level, innovation can be understood as "an idea, practice, or object that is perceived as new by an individual or other unit of adoption" (Rogers, 2003, p. 12). Based on a systematic literature review of studies about public sector innovation, DeVries et al. (2016, p. 152) define three facets as necessary and sufficient conditions for public sector innovation: the perceived novelty, the first adoption of an idea by a given organisation, and a discontinuity with the past. Together, these dimensions differentiate innovation from reform, change, or new ideas (Sørensen & Torfing, 2011, p. 849).

As this definition still includes a wide range of different innovation forms, DeVries et al. (2016) distinguish four types of public sector innovation: administrative or technological process innovation, product or service innovation, governance innovation, and conceptual innovation. This study investigates digital innovation, thus the focus lies on technological process innovation that is defined as the "creation or use of new technologies, introduced in an organization to render services to users and citizens" (DeVries et al., 2016, p. 153). Simultaneously, literature from the private sector refers to digital innovation as "innovating products, processes, or business models using digital technology platforms as a mean or end within and across organizations" (Ciriello, Richter, & Schwabe, 2018, p. 565). Taken together the outlined dimensions, this study defines digital public sector innovation as *the first adoption of a technology that is perceived as new by the given public organisation and produces a significant change in the specific context.*

Importantly, the presented definition does not include anything about whether the innovation produces good or bad consequences (Heartley, 2005). Even though the term has a positive connotation and the goal in most cases is to create an improvement, the risk of failure is usually very high and the perception of whether an innovation is an improvement or debasement is a subjective valuation (McIvor, 2020; Torfing, 2019). Therefore, critical innovation theories point to a "pro-innovation bias" (Godin & Vinck, 2017, p. 8) as several

scholars and practitioners tend to assume that innovation is always the solution and always leads to improvement. Since the study does not aim to evaluate whether the collaboratively developed digital innovations are an improvement, this debate will not be further elaborated. However, the critical view will be considered in the discussion of the results.

2.2. The concept of collaborative innovation as a new strategy for the public sector

The concept of *collaborative innovation* presents the key theoretical foundation for this study. At its core, the strategy is characterised by the feature that "the private and third sector and citizens are integrated into the innovation cycle (idea generation, selection, implementation and diffusion) from the earliest stage onwards" (Bommert, 2010, p. 16). This definition emphasises that collaborative innovation does not refer to the output but to the process that (potentially) leads to innovation. Thereby, innovations are developed in a complex, nonlinear, and iterative process (Eggers & Singh, 2009). Hence, to reveal the mechanisms of collaborative innovation, four analytical phases of the so-called innovation cycle can be depicted (Figure 1). The first stage, the *generation of ideas*, typically starts with the identification and analysis of the problem, followed by the clarification of the goal. Then, a set of possible solutions and creative ideas is developed. Subsequently, the *selection of ideas* evolves around the decision of which ideas should be further pursued by designing, testing, and redesigning prototypes. The *implementation of new ideas* refers to the conversion of ideas into concrete products, procedures, practices, or services (Eggers & Singh, 2009).[1]

Secondly, the collaborative element remains to be clarified. Torfing (2016, p. 64) defines *collaboration* as "a temporal process through which a plurality of actors work together in an organized way to transform problems and opportunities into joint solutions that rest on provisional agreements that are formed despite the persistence of various forms of dissent." In the context of public innovation, those actors can on the one hand include politicians, public managers and government employees, and on the other hand private companies, civil society organisations, and citizens (Lopes & Farias, 2020). In sum, collaborative innovation in the public sector can be conceptualised as *a temporal process in which nonstate actors are integrated into the innovation cycle from the first stage onwards whereby the involved groups - public and external actors - engage in collaborative activities to jointly develop and implement public innovation.*

2.3. The evolution of innovation strategies in the public sector

In order to illustrate how the strategy of collaborative innovation may help to spur innovation in the public sector, it is contrasted with hierarchical and competitive innovation

strategies in the following. Hereby, it is acknowledged that the outlined strategies are not mutually exclusive but co-exist and might develop hybrid forms of innovation approaches (Wegrich, 2019).

Hierarchical strategies tend to favour in-house innovation initiated in a top-down approach by public managers or leaders (Hartley et al., 2013). Beyond the already mentioned organisational barriers, this approach seldomly produces innovative ideas that break with the past because public leaders have few incentives for change, carry the whole responsibility for possible failure and solely rely on their own ideas which limits the scope of developed solutions and tends to create group-thinking and blind spots (Cinar et al., 2019; Eggers & Singh, 2009; Wegrich, 2019). Competitive strategies are strongly characterised by the adoption of modern business practices in the public sector during the NPM era and aim to counteract the deficiencies of TPA strategies with the "creation of quasi markets and the adoption of new forms of strategic leadership and performance management" (Hartley et al., 2013, p. 824). Even though the new management practices can enhance public innovation in some dimensions, new barriers like a tendency towards standardisation, control and a reluctance to share knowledge can be induced by the strong focus on competition and performance (Torfing, 2019). Therefore, collaborative innovation has been suggested as one possible approach for public organisations to mitigate the challenges and limitations of hierarchical and competitive strategies by including empowered actors with diverse knowledge, skills and perceptions in a collaborative process (Hartley et al., 2013; Torfing, 2019).

Importantly, this new strategy "requires a reformulation of the traditional roles of public and private actors" which resembles a wider trend in the shift towards NPG (Hartley et al., 2013, p. 827). Whereas citizens have been mainly perceived as passive clients in the TPA model and as customers in line with the market orientation of the NPM era, they are required to take on an active role as co-creators in the collaborative innovation strategy (Hartley et al., 2013). At the same time, the role of public managers shifts from providing standardised public services as professional bureaucrats (TPA) or running public organisations like a business in the role of managers (NPM) to taking the role of mediators in the NPG model (Heartley, 2005; Sicilia, Guarini, Sancio, Andreani, & Ruffini, 2016). In the context of collaborative innovation, this includes encouraging and empowering different actors and constructively managing interdependencies (Sørensen & Torfing, 2015). Concluding, the roles and key characteristics of the outlined innovation strategies are summarised in Table 1.

2.4. From collaboration to innovation: the underlying mechanisms

Based on the previously outlined conceptualisation, Torfing (2016) theorised the causal relationship between collaboration and public innovation by connecting interdisciplinary theoretical building blocks with empirical findings. Hereby, it must be noted that it is not a fully developed theory of

---

[1]Due to the limited scope of this study, the dissemination phase will not be analysed.
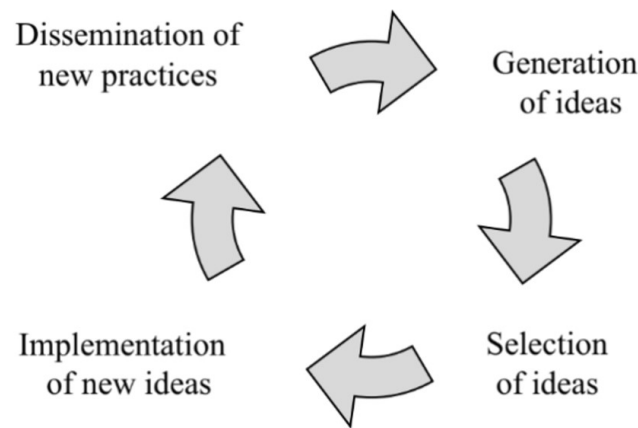
Figure 1: The cycle of innovation

Source: Sørensen and Torfing (2011, p. 851)

Table 1: The evolution of innovation strategies

	Hierarchical innovation	*Competitive innovation*	*Collaborative innovation*
Related Paradigm	Traditional Public Administration	New Public Management	New Public Governance
Organisational values	Hierarchy, control, bureaucracy	Market orientation, focus on performance, contracting in/out	Multi-actor networks, inter-organisational relationships
Innovation	In-house innovation by public managers (top-down)	Innovation through competition between actors	Innovation through collaboration of diverse actors
Role of public managers	Providers	Managers	Mediators
Role of population	Clients	Customers	Co-creators

Source: Adapted from Heartley (2005) and Sicilia et al. (2016)

collaborative innovation. Nonetheless, it provides a solid and suitable foundation for this study as it allows scrutinising fellowship programmes as a new form of collaboration in the light of the existing scholarship. The key proposition is that collaborative innovation between actors with diverse experiences, skills, and knowledge positively influences all phases of the innovation cycle through three underlying mechanisms: empowered actors, mutual and transformative learning, and joint ownership (Torfing, 2016). The following sections outline each of these causal relationships to develop the analytical framework for this theory-guided case study.

2.4.1. Collaboration of empowered actors

The first theoretical mechanism proposes that collaboration has a positive effect on the first phase of the innovation cycle, the idea generation, when *empowered actors* - that are affected and relevant actors - engage in collaborative problem-solving by exchanging different experiences and challenging prevailing opinions (Torfing, 2016). The *"affected actors"* are social or political actors who are directly impacted by the benefits or inconveniences induced by the innovation and can therefore provide a comprehensive understanding of the problem and first-hand experiences with the circumstances (Torfing, 2016, p. 131). *"Relevant actors"*

are actors with knowledge, skills and ideas that are essential for developing and implementing an innovation (Torfing, 2016, p. 132). Depending on the problem to be solved, the form of expertise can differ. In the context of digital innovations, there is a strong demand for technical skills, agile project management, design thinking, interdisciplinary problem-solving and building (software) prototypes (Mergel, 2016). To empower these actors to collaborate, the theory suggests including a third group in the collaboration, the *"boundary spanners",* who are capable of translating and connecting the diverse knowledge of affected and relevant actors in order to leverage potential synergies (Torfing, 2016, p. 133). Importantly, collaboration does not occur inadvertently. It must be initiated, facilitated, and organised in a way that enables the actors to produce the desired outcome (Agger & Sørensen, 2018; Ansell & Gash, 2008). Hence, empowerment requires *clear ground-rules* including the definition of roles, responsibilities, and procedures to prevent conflicts and provide room for creative thinking and constructive discussions (Sørensen & Torfing, 2017). Also, it is essential that all involved actors develop a *shared understanding of the goal* to prevent misunderstandings (Neumann et al., 2019). Concluding, the first theoretical argument proposes that collaborative innovation allows public organisations to include affected and relevant actors into the innovation cycle. That in turn is positively related to the likelihood of innovation as the empowered collaboration between both is expected to create a group that is most capable of developing an innovative solution that meets the requirements of the specific context (Bommert, 2010; McGann et al., 2021).

2.4.2. Mutual and transformative learning

The second causal relationship proposes that collaboration stimulates learning processes through which the actors acquire new skills and expand or revise their knowledge, and these learning processes can spur the development of innovations (Torfing, 2016; Voorberg, Bekkers, Timeus, Tonurist, & Tummers, 2017). The first dimension of the mechanism is grounded in the assumption that the collaborative innovation strategy encourages learning when empowered actors with diverse knowledge and opinions engage in iterative rounds of the outlined innovation cycle and "participate in a joint and cross-disciplinary assessment of the content, feasibility, and potential gains and risks of competing ideas" (Torfing, 2016, p. 65). The second dimension of the mechanism is based on constructivist learning theories which theorise that learning from and with diverse actors has a positive effect on the development of innovations (Sørensen & Torfing, 2011; Voorberg et al., 2017). To elaborate on this relationship, it can be distinguished between mutual and transformative learning.

Mutual learning refers to a continuous dialogue and mutual exchange of experiences and ideas that stimulate collective, creative, and experimental problem-solving (Lindsay et al., 2017; Torfing, 2016). Hereby, it is assumed that one individual usually does not have all the needed knowledge and abilities to solve a complex problem (Ansell & Torfing, 2015). Thus, every actor has some relevant expertise

and through the circulation of that knowledge, collaboration contributes to overcoming information asymmetries between public agencies and public service users (McGann et al., 2021). Further, learning theories emphasize that collaboration broadens the repertoire of solutions through the communicative search for new ways of doing things and can spur innovation through inspiration and imitation (McGann et al., 2021; Torfing, 2016). Hence, this type of learning has an instrumental focus that facilitates the acquisition of practical skills and knowledge to produce the desired innovation (Lindsay et al., 2017; Torfing, 2016).

Transformative learning goes one step further as the actors not only get to know new approaches but also new ways of thinking and reframing the problem (Voorberg et al., 2017). Whereas individually acquired knowledge and skills are often continuous with former assumptions, mindsets, and habits, collective learning of social actors can disrupt bounded forms of thinking through critical and collective reflection (Neumann et al., 2019; Torfing, 2016). Furthermore, exchanges of actors who are different in terms of their cognition and culture are likely to prevent group-thinking and blind spots that are often a problem in closed circles (Wegrich, 2019). Since innovation requires by definition a disruption of old ways of doing things, transformative learning is expected to positively influence public innovation as it enables creative ways of understanding a complex problem and exploring new opportunities that break with the past (Agger & Sørensen, 2018; Torfing, 2016). Concluding, it is theorised that multi-actor collaboration stimulates mutual as well as transformative learning and therefore increases the likelihood of public innovation.

2.4.3. Joint ownership of innovations

The third proposition states that collaborative innovation has a positive effect on innovation because it can create a "joint ownership" of new ideas that reduces the *implementation resistance* (Sørensen & Torfing, 2011, p. 852). Joint ownership is understood as the shift of the *decision-making authority* from the public agency to the collective of the involved actors (Ansell & Gash, 2008). Moreover, it implies a form of commitment since ownership creates a *shared responsibility* for the project (Ansell & Gash, 2008; Neumann et al., 2019). The theoretical mechanism is based on sociological planning theories stating that not sufficiently taking the stakeholder dynamics into account can lead to implementation failures because the involved actors develop a severe resistance when they do not have ownership of the plans and new developments (Cinar et al., 2019; Sørensen & Torfing, 2011). Given the uncertain and destructive character of innovation, this risk is particularly high for innovations as the implementation phase bears potential conflicts, power struggles, and failures (Wegrich, 2019). Hence, enabling an active exchange between the affected public stakeholders and the actors responsible for developing an innovation can provide the government employees with the opportunity to actively control and shape the outcome which increases the chances of a successful implementation (McGann et al., 2021). In

sum, the theory proposes that through the creation of joint ownership of the innovation among the involved actors, collaboration enhances the implementation stage of the innovation cycle and is therefore positively related to the likelihood of public innovation.

2.5. Limitations and challenges of collaborative innovation

The core value of collaborative innovation is the interactive engagement of actors with highly diverse skills, knowledge, and perceptions. At the same time, this characteristic creates severe challenges as collaboration in such a constellation faces the risk of losing the capacity to constructively work together. Hence, the existing knowledge in this aspect is presented in the following to critically investigate innovation fellowship programmes.

According to Wegrich (2019), the collaborative innovation strategy inherits two key *limitations*. First, a divergent understanding during a collaborative process is likely to persist when the commitment towards the goal of the innovation project is superficial or weak. As a consequence, the diverse actors may have different interpretations of the objectives which might lead to *misunderstandings and conflicts*. In strong contrast, the second limitation is that collaboration might lead to too much alignment of the involved actors, ultimately inducing the risk "that one particular world view and approach to doing things becomes dominant" (Wegrich, 2019, p. 17). In this case, the previously outlined mechanisms would not work properly, and the benefits of collaboration disappear. Furthermore, several studies reveal the potential problem that the process can be manipulated by influential interests (Ansell & Gash, 2008), and collaboration tends to have high transaction costs in terms of time and resources (Hartley et al., 2013).

Lastly, the notion of collaborative governance itself is not undisputed. McIvor (2020) points to some dangers and unintended consequences of the normative and political stakes in collaborative governance. He outlines how collaboration policies can serve as an ideological justification of government actions and that the outcome of collaboration in terms of its success is likely to be perceived quite differently by public managers and politicians than by civic groups or citizens. Moreover, "interest group pluralism can too easily slide into corporatist models of governance" and in that case rather serve business value than public value (McIvor, 2020, p. 512). In conclusion, these limitations illustrate that collaborative innovation is not a suitable strategy for all contexts and social actors will not necessarily collaborate constructively.

2.6. Connecting the collaborative innovation strategy with digital innovations

The outlined theoretical foundation of collaborative innovation is applicable to many different types of public sector innovation. Thus, the specific arrangement "should be determined by the problem or challenge at hand" (Torfing, 2019, p. 4). The challenge investigated in this study is the lacking capacity of public organisations to develop and implement *digital* innovations. Therefore, the question remains why this strategy may be suitable for the unique setting of digital public sector innovations. The first reason is that governments need personnel with technological skills to create and develop digital solutions (Coglianese, 2020; Wirtz et al., 2019). Particularly, because the NPM era incentivised outsourcing of digital public service delivery, technical skills like coding and user-centric web design are mostly not available among civil servants (Dunleavy et al., 2006; Wirtz et al., 2019). Therefore, the inclusion of empowered actors, in this case empowered by technical skills, is a prerequisite to creating digital innovations.

Secondly, the development of digital innovations requires very different working methods than what is traditionally established in public organisations. Especially, design thinking and agile project management are crucial in software development and digital innovation projects (Mergel, 2016). Simultaneously, private actors are presumably not conversant with the unique organisational structure, processes, and requirements of a public organisation (Coglianese, 2020). Thus, the process of mutual and transformative learning is expected to have a positive influence on digital innovations as the diverse actors might help each other to acquire the necessary skills for the successful development of digital innovations in the specific setting of public organisations. Lastly, resistance towards transformative change is expected to be particularly strong in the case of new technologies because the affected government employees might not fully understand their functionality and possible consequences (Wirtz et al., 2019). Therefore, creating joint ownership by including them in the innovation process has the potential advantage, compared to outsourcing it to external IT-providers, that the public employees may be offered an opportunity to understand the new technologies and their potential risks and benefits.
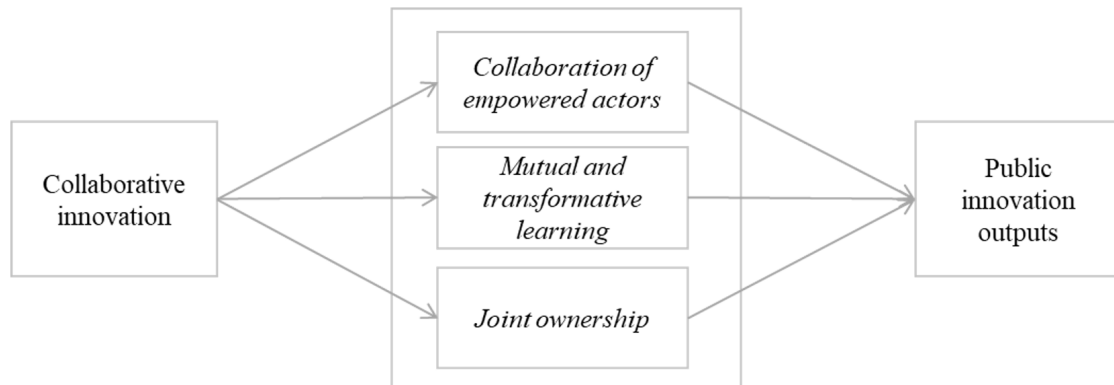
2.7. Analytical Framework

Concluding, the theorised mechanisms of collaborative innovation serve as the analytical framework for this study as summarised in Figure 2. Thus, it allows connecting the existing knowledge about collaborative innovation with the so-far unexplored phenomenon of innovation fellowship programmes. In all, it is expected that the three outlined mechanisms potentially explain how the investigated fellowship programme contributes to the development of digital innovations in public organisations.

3. Methodology

This section clarifies and justifies the chosen methods for answering the beforementioned research question. It starts with elaborating on the research design of this study, followed by a justification of the case selection and a description of the case. Hereafter, the unit of analysis and units of observation are specified. Lastly, the data collection method and data analysis are described.

Figure 2: Analytical framework: Mechanisms of collaborative innovation

Source: Own representation, based on Torfing (2016)

3.1. Research design

In line with the main research question that aims to uncover the underlying mechanisms of how Tech4Germany contributes to the development of digital innovations, this study is designed as an explanatory single-case study (Swanson & Holton, 2005; Yin, 2003). This research design arises directly from the proposed question because the thesis investigates one unique case, Tech4Germany, and "[t]he case study method is most likely to be appropriate for 'how' and 'why' questions" (Yin, 2003, p. 22). Moreover, since the objective of this study is not to analyse one specific variable but to develop a comprehensive understanding of the causal mechanisms, a case study is "uniquely predisposed to taking into account a broad and diverse set of explanatory factors" (Blatter & Haverland, 2012, p. 5). Further, this case study follows a diagnostic approach. Generally, this type of applied research is "concerned with using the knowledge acquired through research to contribute directly to the understanding or resolution of a contemporary issue" (Ritchie, 2003, p. 24). As outlined in the introduction, fellowship programmes are a so-far unexplored form of collaborative innovation. At the same time, the theory and scholarship about collaborative innovation are well established. Therefore, a diagnostic case study is the most suitable research design because it allows the researcher to develop an in-depth understanding of a unique phenomenon by connecting the empirical observations with the theoretical concepts and existing knowledge (Blatter & Haverland, 2012; Swanson & Holton, 2005). Hence, the study will be theoretically-guided by the analytical framework that was deductively derived from the theory of collaborative innovation in chapter 2 (see Figure 2).

Given the uniqueness of the case and the outlined diagnostic case study design, this study does not attempt to generalise its findings to a wider population. However, the generated knowledge may be transferable to other cases, namely other fellowship programmes (Lewis & Ritchie, 2003). In regard to the internal validity, multiple sources (contracts, project reports, and semi-structured interviews) are collected and triangulated to ensure that the findings conform with reality (Lewis & Ritchie, 2003; van Thiel, 2007). Furthermore, triangulation mitigates the subjectivity in the research process and therefore strengthens the overall reliability (van Thiel, 2007).

3.2. Case selection

As outlined above, the case is not assumed to be statistically representative of a wider population but was chosen for "intrinsic" reasons (Swanson & Holton, 2005, p. 330). The selection of this particular case is justified by three arguments. First, the case is very "information-rich" (Patton, 1990, p. 169) and scientifically relevant because it applies the strategy of collaborative innovation in an unexplored form. The analysis of this case is therefore predestined to fill the identified research gap. Secondly, Tech4Germany is the only fellowship programme of its kind in terms of the organisation in project groups that actively work together for a short term of three months and its explicit focus on digital innovations. Even though there are comparable fellowship programmes in other countries, for instance, the Presidential Fellowship Programme in the USA (Obama White House Archives, 2021) and the No. 10 Innovation Fellowship Programme in Great Britain (HM Government, 2021), they take place over a longer period and are not organised in unchanging groups of citizens and public employees. The project characteristic of Tech4Germany therefore provides an especially bounded and clear setting to investigate the collaborative process (Swanson & Holton, 2005). Additionally, the accessibility of Tech4Germany is very high as relevant documents and names of former participants are provided online (van Thiel, 2007). Lastly, Germany offers a striking setting because it ranks comparatively low, on place 21 in the European Union, in terms of the provision of digital public services (European Commission, 2020). Hence, it is particularly interesting to analyse how German public employees experience the collaboration with tech experts from the civil society and the potential clash of cultures.

3.3. Case description

The fellowship programme Tech4Germany started as an initiative of a young German citizen in 2018 with the goal to "expedite the digitalisation of Germany and thereby learn from and with each other" (Tech4Germany, 2021c, p. 1). The core idea of the fellowship programme is to bring talented citizens in the fields of software engineering, product management and User Experience (UX) or User Interface (UI) design into the national ministries or agencies where they work together with public employees to solve a concrete problem or improve public services by using new technologies (Tech4Germany, 2021b). Each year, about 30 citizens collaboratively work on digital innovation projects with the government employees for twelve weeks. The project teams consist of four citizens, the so-called *Fellows*, and two to four public employees, the so-called *Digitallotsen* (Tech4Germany, 2021b). Hereby, the Fellows get a small monetary contribution in form of a scholarship by Tech4Germany to cover their living costs (DigitalService4Germany, 2021d). The projects explicitly focus on user-centric and agile approaches to develop a technological prototype (Tech4Germany, 2021c). Examples are a chatbot for the German Ministry of Family Affairs that navigates families through a large amount of available information and an online tool for pensioners to file their tax returns (Tech4Germany, 2021b).

From the very beginning, the initiative operated under the patronage of the head of the federal chancellery which underpins its high political relevance. After the successful pilot year, the initiator founded the non-profit start-up 4Germany GmbH together with two other young citizens which received pilot government funding in 2019 (DigitalService4Germany, 2021a). Hereafter, Tech4Germany presented the developed innovations of the second cohort to the German chancellor and federal cabinet, and was called the technology taskforce for the German government (Presse- und Informationsamt der Bundesregierung, 2020). In September 2020, the non-profit start-up was bought by the German government and turned into the DigitalService4Germany GmbH. It is now fully funded and owned by the state (DigitalService4Germany, 2021b) and can therefore be categorised as a "[g]overnment enterprise" (Rainey, 2014, p. 75). Next to the fellowship programme Tech4Germany and a second fellowship called Work4Germany, the DigitalService4Germany GmbH currently builds a digital service team that operates as a permanent in-house coding force (DigitalService4Germany, 2021a; see Mergel (2019) for a scientific analysis of digital service teams).

3.4. Unit of analysis and units of observation

Derived from the main research question, the unit of analysis is Tech4Germany. The research objective to develop a comprehensive understanding implies that the case is not divided into smaller sub-units or specific variables (Yin, 2003). Instead, the unit of analysis can be specified to be the whole process that was undergone during the collaborative innovation projects (Yin, 2003). The "[s]ampling within a case should be guided by the research questions and by the theory that underlies the initial conceptualization of the case" (Swanson & Holton, 2005, p. 336). In accordance with the analytical framework (Figure 2), information about the collaboration and the experiences of the involved actors needs to be gathered. Therefore, actors who have actively participated in the collaborative innovation process are the units of observation. This translates into conducting interviews with the participating Fellows and Digitallotsen. Given the highly individual character of collaboration, for instance the experience of learning, it is indisputable to directly talk to former participants of the fellowship programme to fully understand its mechanisms and dynamics. Further, it is essential to talk to both groups since the perception might differ significantly among the actors (Torfing, 2019). As both groups are very limited in their number, the selection of the interviewees was "purposive" based on three criteria (Ritchie, Lewis, & Elam, 2003, p. 78): the respondents participated in different projects in 2019 or 2020, took on different roles (Engineering, Design or Product Fellow or Digitallotse), and are diversified in terms of their gender. These criteria allow to gain insights from diverse perspectives and increase the internal validity of the study (a list of the interviewees is provided in Appendix B).

3.5. Data collection method and data analysis

The study draws on two sources of evidence. First, a mix of materials was collected whereby the website of Tech4Germany (Tech4Germany, 2021b) was the main source. Precisely, the documents include project reports, contracts, information brochures, and guidelines (a list of documents is provided in Appendix A). Further, secondary data such as personal experience reports of Fellows and press statements were investigated to enrich the perspective and mitigate the potential subjectivism of information published by Tech4Germany itself (Finnegan, 2006). Additionally, podcasts with the founders of Tech4Germany were consulted which provide relevant information as they reflect the position and experiences of the third involved group in the collaborative process, namely the core team of Tech4Germany. The analysis of these documents sets the foundation for the conducted interviews and serves as supplementary information for the analysis.

The second and main source of evidence are in-depth interviews with former participants of Tech4Germany. In total, six interviews were conducted, three with former Fellows and three with Digitallotsen. The method of semi-structured interviews was chosen which "involves the implementation of a number of predetermined questions and special topics" while the interviewer has the freedom to ask follow-up questions and go beyond the prepared standardised questions (Berg, 2009, p. 107). This is the most suitable method because it allows to structure the interviews in accordance with the theoretical framework and it simultaneously provides the flexibility to dive deeper into a topic that occurs to be particularly relevant for the investigated case (van Thiel, 2007). Following the outlined research design, the interview guideline was

deductively derived from the analytical framework. Thereby, each theoretical dimension (empowered actors, mutual and transformative learning, and joint ownership) was operationalised into concrete questions (presented in Appendix C), whereby the abstract concepts were translated into everyday language to ensure the interviewees can understand what is asked (Kvale, 2007). While using the same dimensions in all interviews, the specific questions have been slightly adopted for the Fellows and Digitallotsen to take the different functions and perspectives into account. The interviews took on average 35 minutes and were conducted via a video meeting on Zoom with one exception via phone.

The interviews were recorded, transcribed and anonyvmised to prepare the data for the qualitative content analysis which is understood as the "nonnumerical examination and interpretation of observations, for the purpose of discovering underlying meanings and patterns of relationships" (Babbie, 2013, p. 390). As the diagnostic research design implies, the coding scheme was developed based on the deductive category assignment method whereby the analysis is systematic, theory-driven and rule-bound (Mayring, 2015). Hence, the analytical framework was operationalised into specific codes in order to apply the theoretical frame to the interview transcripts. After a first familiarisation with the data, the software "ATLAS.ti" was used to apply the theoretical codes to the empirical data. In a third step, it was scrutinised whether the deductively derived codes capture all data or further inductively developed codes are needed (Mayring, 2015; Swanson & Holton, 2005). To increase the reliability of the qualitative data analysis and interpretations, the coding scheme is provided in Appendix D.

4. Analysis

In this chapter, the findings that were derived from the content analysis are presented in the light of the analytical framework whereby each section provides the results of one code group. In accordance with the theoretical foundation, the chapter starts with presenting the data regarding the collaboration of empowered actors, followed by mutual and transformative learning, and joint ownership of the digital innovations. Lastly, the identified limitations and challenges are presented.

4.1. Findings regarding the collaboration of empowered actors

Upon exploring in what ways Tech4Geramany enables the collaboration of empowered actors, the first key finding is that Tech4Germany is responsible for the *initiation of the collaboration* (Fellow 2, 2021). One of the co-founders describes Tech4Germany in an interview as "an independent intermediary who lowers the obstacles for both sides so that [the collaboration] is attractive" (Lang, 2020, l.73f.). Particularly, they address professions such as designers who usually do not find job advertisements in German public organisations and provide a setting that is attractive for young tech

talents who otherwise would have not considered working for a public organisation (Fellow 1, 2021; Fellow 3, 2021). Simultaneously, Tech4Germany creates a unique room for public employees who are willing to approach their digital innovation project with new methods but do not have the available resources, knowledge, or opportunity to do so (Anton & Hupperth, 2020b; Digitallotse 3, 2021). To ensure the feasibility of the digital innovation project, Tech4Germany selects the participants through an application process. The Fellows are chosen based on their technical skills, methodological expertise, and motivation (DigitalService4Germany, 2021d). The Digitallotsen apply with a concrete project (Digitallotse 3, 2021) which is assessed by an external jury (including representatives of the administration, science, civil society and Tech4Germany alumni) based on three criteria: impact, open-endedness, and feasibility (Tech4Germany, 2021a). However, it appeared in one case that a Digitallotse was not involved in the application process but appointed to join the fellowship because another department led the application process and nobody else from his/her department wanted to participate (Digitallotse 2, 2021).

To analyse whether the projects of Tech4Germany meet the concept of collaborative innovation, it was investigated whether the selected participants qualify as the in chapter 2.4.1 outlined roles of empowered actors. Hereby, it was identified in the document study that the Digitallotsen suit the theoretical concept of affected actors and the Fellows the role of relevant actors (see for instance Rodríguez (2021); project reports, Appendix A). In the interviews, these roles were clearly confirmed. The Digitallotsen described that they were directly *affected* by the problem to be solved in their working life as they work in the responsible department and therefore had detailed knowledge of the problem and its context. This coincides with the perception of all interviewed Fellows who experienced that the Digitallotsen had "many direct points of contact" (Fellow 1, 2021, l.64) with the problem and their expertise was "very detailed" (Fellow 3, 2021, l.60). During the project, Digitallotse 2 perceived his/her *role* to be a "mediator [. . .] between the Fellows and the department or the house [ministry] in general" (ll.38ff.). Hereby, the main responsibilities spanned providing support in regard to special requirements like data protection and accessibility (Fellow 2, 2021), connecting the Fellows with other important stakeholders (Digitallotse 1, 2021), putting their political opinion into context (Digitallotse 3, 2021), and providing direct access to the end-users of the developed innovation (Fellow 1, 2021). Therefore, all Fellows stated that "it would have not been possible to realise [the project] without the Digitallotsen" (Fellow 3, 2021, ll.208f.).

All interviewees reported that the Fellows had relevant and necessary *skills* to develop the digital innovation which confirms their characterisation as relevant actors. Particularly, it was emphasised that the teams were interdisciplinary, and their expertise complemented one another well (Digitallotse 1, 2021; Digitallotse 2, 2021). Besides others, these skills include UX-/UI-design, coding, design thinking, agile software development and user-centric problem-solving (Fel-

low 1, 2021; Fellow 2, 2021; Fellow 3, 2021). What stands out is that the methodological expertise is generally more emphasised than purely technical skills. Hence, the Fellows mainly took on the *role* of project and product managers who structured the projects as an iterative innovation cycle and enabled the collaborative development of the prototype with the outlined working methods (Fellow 2, 2021; Fellow 3, 2021). Particularly, identifying the actual problem from a user-perspective and investigating the problem with an un-prepossessed view was stated to be a key value (Digitallotse 1, 2021; Digitallotse 2, 2021). Thus, all Digitallotsen described that they could not have achieved the same results without the Fellows.

Secondly, it is a key premise that the projects can indeed be characterised as *collaboration* between the Digitallotsen and Fellows. Overall, all interviewees reported an active and intense collaboration throughout all project phases. Hereby, the Fellows did a large part of the analytical work like conducting and synthesising user interviews whereby the Digitallotsen helped the Fellows to understand the problem and were available the whole time to answer questions or provide additional information. One day a week, the Fellows and Digitallotsen collaboratively worked on the digital innovation in form of workshops which included open discussions, ideation sessions, brainstorming, prioritising of ideas and testing prototypes (Digitallotse 1, 2021; Digitallotse 2, 2021; Fellow 1, 2021). Especially because the Digitallotsen participated in the innovation activities without hierarchical structures (Fellow 2, 2021) and as full members of the team (Digitallotse 2, 2021; Digitallotse 3, 2021; Fellow 3, 2021), these workshop activities accord with the theoretical understanding of collaboration. Hereby, Digitallotse 3 emphasised the "co-creative momentum" (l.102) which strongly differentiated the fellowship from the relationship to an external IT-provider. Therefore, the findings suggest that the collaboration of diverse actors with interdisciplinary skills contributed to the *generation of innovative ideas* (Digitallotse 3, 2021; Fellow 1, 2021; Fellow 2, 2021).

Having found that the projects are indeed a form of collaborative innovation, the question remains how Tech4Germany enables the relevant and affected actors to collaborate. Beyond the initiation of the collaboration, the data indicates that Tech4Germany meets the theoretical description of a *boundary spanner*. One of the co-founders described her role in a podcast as "enabler and problem-solver *[Probleme-aus-dem-Weg-Räumer]*" (Lang, 2021, l.86). Precisely, Tech4Germany facilitated the start of the collaboration by providing each group with relevant information and setting the *ground-rules,* for instance, the meeting and time schedule (Digitallotse 1, 2021; Fellow 1, 2021) and the legal foundation (DigitalService4Germany, 2021c). Since all participants apply for a specific role (Engineering, Design or Product Fellow or Digitallotse), Tech4Germany also implicitly prescribed the responsibilities (Fellow 2, 2021). Moreover, the core team organised an onboarding week and workshops whereby the Fellows learned about the processes and vocabulary of public organisations and the Digitallot-

sen about the working methods and vocabulary of the digital economy (Anton & Hupperth, 2020a; Fellow 2, 2021; Tech4Germany, 2021c). This coincides with the outlined theoretical role of a boundary spanner to translate and link the diverse knowledge of the actors (Torfing, 2016). Both groups, Fellows and Digitallotsen, perceived those measures to be helpful for the collaboration as they also allowed to get to know each other in a safe setting (Digitallotse 2, 2021; Fellow 1, 2021).

Regarding a *shared goal,* Tech4Germany did not prescribe a concrete objective for each project but only guiding principles like a strong focus on user-centricity and the aim to finish the project with a prototype (Fellow 2, 2021). Hereby, the Fellows and Digitallotsen had a divergent understanding of the objectives in the beginning but developed a shared goal over time (Digitallotse 2, 2021; Fellow 3, 2021). However, in line with the design thinking and agile approach, it was not a fixed goal but reframed constantly (Fellow 2, 2021). During the course of the project, Tech4Germany was not actively involved in the collaboration (Digitallotse 1, 2021; Fellow 1, 2021), but "they provided the frame and [the participants] could fill this frame completely free and independently" (Fellow 2, 2021, ll.157f.). Tech4Germany regularly asked whether everything was going well (Fellow 1, 2021), provided additional feedback and help when the project teams had problems (Fellow 3, 2021) and moderated between the Fellows and Digitallotsen in cases of conflict which was perceived as very helpful (Digitallotse 2, 2021). The key findings are summarised in Table 2. Additionally, the code frequencies and paraphrased key messages per interviewee which serve as the basis for the analysis are presented in Appendix E.

4.2. Findings regarding mutual and transformative learning

The analytical framework ascribes an important role to mutual and transformative learning for the development of innovations through collaboration. In the case of Tech4Germany, this importance is reflected in the official contracts, stating that the DigitalService4Germany GmbH "runs the programme Tech4Germany [...] for the development of digital competencies *[Digitalkompetenzen]* of the public employees" (DigitalService4Germany, 2021c, p. 1) and "provides the talents *[Fellows]* a platform for open-ended thinking in regard to digitalisation processes through an open-ended knowledge transfer in an 'experimental' space" (DigitalService4Germany, 2021d, p. 1). In the interviews, it was investigated whether these official statements conform with reality.

For the dimension of *mutual learning*, all interviewed Digitallotsen reported they learned new methods, particularly agile project management skills and design thinking techniques. Additionally, some got to know new workshop formats and tools for digital collaboration (Digitallotse 1, 2021). The learning experience is underpinned as the Digitallotsen continued using the acquired methods and tools after the project (Digitallotse 1, 2021; Digitallotse 2, 2021). However, all interviewees did not acquire new technical skills,

Table 2: Collaboration of empowered actors: Key findings

Key finding	Number of responding interviewees who confirmed key finding*	
	Fellows	*Digitallotsen*
Digitallotsen are directly affected by problem/innovation	3/3	3/3
Digitallotsen take on the role of mediators	3/3	3/3
Fellows have relevant skills for the innovation project	3/3	3/3
Fellows take on the role of project/product managers	3/3	3/3
Tech4Germany takes on the role of a boundary spanner	3/3	3/3
Tech4Germany initiates the collaboration	3/3	2/3
Active collaboration takes place in form of weekly workshops	3/3	3/3
A shared goal is developed throughout the project	3/3	3/3
Tech4Germany sets the ground-rules	3/3	3/3
The collaboration contributes to the generation of innovative ideas	3/3	2/2

*In the illustration x/y, y indicates the number of interviewees who gave a relevant response to the attribute and x indicates the number of interviewees who confirmed the key finding. A finding counts as confirmed if any indication that the attribute applies was made.

either because they already had advanced IT-skills (Digitallotse 1, 2021), they were not interested in technical issues (Digitallotse 2, 2021), or the time was not sufficient to develop advanced technical skills (Digitallotse 3, 2021). From the perspective of the Fellows, the willingness to learn differed among the participating Digitallotsen. Whereas most were eager to learn new working methods and it was observable that they acquired new techniques, some were sceptical because they either assumed to already know these working methods (Fellow 2, 2021) or they did not acknowledge the value of the new approaches (Fellow 1, 2021).

Concerning the *transformative learning*, the findings are mixed. On the one hand, the attitude towards digital innovations in general did not change (Digitallotse 1, 2021; Digitallotse 3, 2021). To a large extent, this can be explained by the fact that the participating public employees already had a positive attitude towards digital innovations before the collaboration (Fellow 1, 2021). On the other hand, two Digitallotsen described that their thinking on how to approach a digitalisation project changed significantly. Especially, Digitallotse 2 reported that s/he thought little about whether a product or service even makes sense for the citizens beforehand, and then learned how to approach problems from a user-centric perspective from the Fellows. Hereby, it was emphasized as a key value of Tech4Germany that new working methods are not taught in a training but can be experienced in a real digital innovation project (Digitallotse 3, 2021; Fellow 2, 2021; Fellow 3, 2021). The importance of this real-life experience is underpinned by the statement of Digitallotse 2 that the degree of transformative learning appeared to be "related to [. . .] the degree how strongly one was involved in the process" (ll.109f.). Additionally, the Digitallotsen "disregarded existing rules to imagine how it could look like in an ideal world" (Digitallotse 2, 2021, ll.315f.) and the used methods in the workshops aimed to "break with the grid-locked administrative thinking" (Digitallotse 2, 2021, l.94), which again indicates the experience of transformative learning. At the same time, it was mentioned that it was rather a first encounter and "it takes more to really learn the mindset of design thinking" (Fellow 2, 2021, ll.226f.).

On the side of the Fellows, all interviewees described that they experienced *mutual learning* during the project. Partic-

ularly, because the public administration was perceived as a "Blackbox" (Fellow 2, 2021, l.206) before the project, they acquired new knowledge about the requirements, decision-making and administrative processes in public organisations. Furthermore, some learned new skills from the other participating Fellows and during the workshops organised by Tech4Germany (Fellow 2, 2021; Fellow 3, 2021). However, there were few findings that suggest *transformative learning*. Only Fellow 2 experienced a mindset shift away from career and monetary success towards purposeful work to such an extent that s/he started working for the public sector after the Tech4Germany project. In sum, it was reported to be highly important that the collaborative learning allowed for a "more realistic" (Fellow 2, 2021, l.354) *development of suitable digital innovations* because without the expertise of the Digitallotsen, the Fellows would have not been able to understand the specific context of the public organisation (Fellow 3, 2021). To conclude, the key findings are summarised in Table 3.

4.3. Findings regarding joint ownership of digital innovations

Concerning the third theoretical dimension, the data clearly indicates that joint ownership of the digital innovation was created among the actively involved participants. Firstly, a *shared responsibility* is observable since all interviewees felt strongly responsible for the success of the digital innovation project. Fellow 1 explained that the implementation of the developed prototype "was only possible because [the] project partners were very committed" (l.221). Also, the Fellows and Digitallotsen equally contributed to the success of the project (Fellow 3, 2021). The positive impact of the collaboration is confirmed by the statement of Digitallotse 3 that s/he would have not cared about the project's success that much if it was a normal relationship with an external IT-provider. Secondly, it can be derived that all important decisions were made together in the team and in many cases even consensual (Digitallotse 1, 2021; Fellow 2, 2021). Thus, the *decision-making authority* moved from the public agency to the collective of the involved actors.

In line with the theoretical mechanism, this joint ownership led to a reduced *implementation resistance* among the actively involved Digitallotsen as they all personally advocated the implementation of the developed digital innovation and took action to make the implementation possible after the project ended (Digitallotse 1, 2021; Digitallotse 3, 2021). The importance of the collaboration is underpinned by the finding that the implementation resistance of the public employees who were not involved partly remained. For instance, some rejected the changes that would have come along with the digital innovation as they "got the feeling to do everything wrong and now the young people come and want to tell [them] how public administration works" (Digitallotse 2, 2021, ll.100ff.). In contrast, those who regularly noticed the progress and observed the applied methods became more open-minded and did not oppose digital

innovations and modern approaches anymore (Digitallotse 2, 2021).

Another important aspect is that the collaboration with Tech4Germany in some projects directly helped to get the necessary support for a digital innovation the Digitallotsen have unsuccessfully tried to implement for many years because the Fellows were able to present a concrete prototype to the important actors and illustrate the benefits more precisely (Fellow 1, 2021). Further, it was essential that the projects with Tech4Germany were often supported and promoted by high political decision-makers who were involved through presentations of the progress and results, for instance the state secretary or federal minister (Digitallotse 3, 2021; Fellow 1, 2021). Concluding, the key findings are summarised in Table 4.

4.4. Findings regarding limitations and challenges of fellowship programmes

As the past three sections outlined in what ways the three theoretical mechanisms of collaborative innovation are observable, this section presents the key challenges and limitations of collaborative innovation in form of fellowship programmes to allow for a critical assessment of the concept of Tech4Germany.

Firstly, several *conflicts and misunderstandings* were reported. At the beginning of the collaborative innovation process, it was perceived as a challenge that the public employees and digital experts used different vocabulary and talked about processes neither of them knew of the other (Fellow 1, 2021). Additionally, there was disagreement about the chosen communication tools (Digitallotse 2, 2021). Moreover, a conflict occurred because the Fellows decided the aspired solution does not make sense for the specific context which was met with disappointment by some Digitallotsen (Fellow 2, 2021). In one case, a Digitallotse "even rejected to take part in workshops" (Fellow 1, 2021, l.184) because s/he feared ridicule and did not see the value of those methods. However, all interviewees stated that these were minor conflicts that could either be solved or did not strongly impair the overall project.

Beyond that, some circumstances caused *limitations*. Many of the tools used by the Fellows were not compatible with the IT-equipment of the public employees or not allowed due to security regulations (Digitallotse 2, 2021; Fellow 1, 2021; Fellow 2, 2021). Furthermore, time was described to be a key limitation factor because the Digitallotsen participated in the Tech4Germany project on top of their daily work, resulting in a very limited time frame and a reduced learning opportunity (Digitallotse 3, 2021; Fellow 1, 2021; Fellow 2, 2021). In addition to that, three months is a very short time frame for a digital innovation so the end product usually was a good prototype but not a functioning minimum viable product (MVP) and not all prototypes were implemented afterwards (Digitallotse 1, 2021; Fellow 2, 2021). Moreover, the findings show that even if all three outlined mechanisms are observable, there are still *barriers to the implementation* that cannot be changed by

Table 3: Mutual and transformative learning: Key findings

Key finding	Number of responding interviewees who confirmed key finding*	
	Fellows	*Digitallotsen*
Digitallotsen experience mutual learning	3/3	3/3
Fellows experience mutual learning	3/3	n.a.
Digitallotsen experience transformative learning	2/3	2/3
Fellows do *not* experience transformative learning	2/3	n.a.
The collaboration contributes to the development of a suitable innovation	3/3	3/3

* In the illustration x/y, y indicates the number of interviewees who gave a relevant response to the attribute and x indicates the number of interviewees who confirmed the key finding. A finding counts as confirmed if any indication that the attribute applies was made.

Table 4: Joint ownership: Key findings

Key finding	Number of responding interviewees who confirmed key finding*	
	Fellows	*Digitallotsen*
Fellows and Digitallotsen share the responsibility	3/3	3/3
Decisions are made collectively	2/2	1/1
The degree of implementation resistance is reduced by the collaboration	3/3	3/3

* In the illustration x/y, y indicates the number of interviewees who gave a relevant response to the attribute and x indicates the number of interviewees who confirmed the key finding. A finding counts as confirmed if any indication that the attribute applies was made.

Tech4Germany, for instance, limitations due to data protection and accessibility requirements, technical issues in the larger IT-system of the public organisation, and procedural regulations (Digitallotse 1, 2021; Fellow 2, 2021). Additionally, the Digitallotsen usually do not have the full authority to decide about the implementation (Digitallotse 1, 2021). Lastly, the contact *restrictions due to the Covid-19 pandemic* strongly impaired the digital innovation projects in 2020 as collaboration is more fruitful in presence than in digital formats (Digitallotse 1, 2021; Fellow 1, 2021; Fellow 2, 2021).

In sum, two key points regarding the *scope and impact of Tech4Germany* can be derived from the findings. On the one hand, the concept is limited in its scope as the above-outlined mechanisms mainly apply to the comparatively small num-

ber of active participants. In this aspect, the difficulty to communicate these new approaches to the entire ministry and to really "live" the mindset was described (Digitallotse 2, 2021, l.153). Thus, it will take a relatively long time until Tech4Germany has an impact beyond the directly involved actors. Nonetheless, all interviewees strongly emphasized that the fellowship programme is a valuable concept that should be continued. Hereby, the identified key value is that Tech4Germany opens-up the public organisations and provides a setting where the public employees can experience an interdisciplinary, cross-sectional, and agile way of working (Fellow 2, 2021). Therefore, Tech4Germany is described by the participants as a very important first step initiating a process of change and thus paving the way for further dig-

ital innovations projects that are approached from an agile and user-centric perspective (Digitallotse 2, 2021; Fellow 2, 2021; Fellow 3, 2021). Hence, the impact of Tech4Germany was summarised with the words: "Constant dripping wears away the stone" (Digitallotse 2, 2021, l.358).

5. Discussion

5.1. Results

The presented findings serve to answer the sub-questions of this study. Regarding the ways the fellowship programme enables the collaboration of empowered actors, it became clear that Tech4Germany meets the role of a boundary spanner by bringing together the Digitallotsen in their position as affected actors and Fellows whose technical and methodological skills make them relevant actors. Therefore, Tech4Germany provides a space where both groups can collaboratively work together on a concrete digital innovation by defining the basic structure, rules and responsibilities. Beyond that, the workshops and onboarding week facilitate the collaboration by linking and translating the diverse knowledge and vocabulary. What must be noted is that the active collaboration between the affected and relevant actors only takes place one day a week, whereas the rest of the time the interaction can rather be described as mutual support. Furthermore, Tech4Germany does not operate as a boundary spanner throughout the whole collaboration but is available upon request, for instance in cases of conflict.

The findings for the second sub-question are mixed, indicating that mutual and transformative learning was experienced in some cases but highly depends on the individual motivation, previous knowledge, and extent to which the actor was actively involved. Hereby, it stands out that the Digitallotsen did not learn new technical skills which must be noted as a limitation as this study investigates the development of *digital* innovations. On the other hand, the public employees clearly acquired new methodological competencies and changed the way they approach digital innovation projects towards user-centric and agile thinking. Since it is widely acknowledged that these approaches are crucial for digital innovation projects (Mergel, 2016), it is concluded that the actors overall learned relevant skills for the development of digital innovations in the specific context of public organisations.

Concerning the third sub-question, it was found that the actively involved actors evidently created joint ownership as they shared the responsibility and decision-making authority. Therefore, the collaborative innovation contributes to a low level of implementation resistance among the involved public employees. Similar to transformative learning, this effect seems to depend on the degree of involvement and gets weaker the fewer contact points a public employee had with the Tech4Germany project. At the same time, the sheer presence of Tech4Germany helped to reduce the implementation resistance of decision-makers by illustrating potential benefits and increasing the political attention.

All in all, the results show that the theoretical mechanisms of collaborative innovation apply to a very large extent and the theory of Torfing (2016) provides relevant propositions to explain how a fellowship programme contributes to the development of digital innovations in public organisations. At the same time, it appears that even if all mechanisms apply, the developed prototypes were not necessarily transformed into digital innovations. Hence, there are most likely relevant antecedents affecting the implementation of collaboratively developed digital innovations beyond the proposed mechanism of joint ownership among the involved actors. The findings of this study indicate the organisational design and support of decision-makers as important factors. Therefore, a possible explanation for an unsuccessful implementation is a hierarchical decision-making structure in the public organisation. Although the public employees who are affected by the innovation in their daily work are included in the collaboration process and advocate the implementation of the developed prototype, political leaders or public managers with the decision-making authority to determine the continuation of the digital innovation project can still impede the implementation. Consequently, the organisational design and particularly the decision-making structure of the public organisation should be considered in the analysis of collaborative innovation projects.

5.2. Limitations of the research

With regard to the limitations of the thesis, this single-case study does not allow generalising the findings to other fellowship programme but only provides transferable indications and starting points for further studies. Additionally, the small sample size implies that it was not possible to empirically validate the extent to which the theoretical mechanisms apply in this case. Since not all participants have been interviewed, some perspectives might be missing, and no conclusion can be drawn about the overall impact of Tech4Germany. Particularly, it stands out that all interviewees generally had a positive attitude towards the fellowship. Thus, the selection process might have induced a bias as not all persons replied to the interview request and it may be that advocates of Tech4Germany were more willing to share their experience than sceptics. Moreover, since all participants are publicly named on the website of Tech4Germany, they might have an intrinsic motivation to present the programme in a positive light.

Beyond that, the exceptional circumstances during the Covid-19 pandemic infer that the experiences of the respondents who participated in 2020 might differ in comparison to other years and contain elements that are not representative for the fellowship programme. However, it was found that the key findings coincide between the participants of 2019 and 2020 which mitigates this limitation. Lastly, all interviews have been conducted in German so the translated quotations may have different connotations and thus can lose some of their meaning.

Table 5: Limitations and challenges: Key findings

Key finding	Number of responding interviewees who confirmed key finding*	
	Fellows	*Digitallotsen*
Conflicts or misunderstandings occur	3/3	1/3
The limited time capacities of the Digitallotsen and/or insufficient IT-equipment impair the collaboration	3/3	2/2
External barriers impede the implementation	2/2	2/2
The Covid-19 pandemic impaired the collaboration	2/2	1/1
Tech4Germany has a positive impact on the development of digital innovations, but the scope is limited	3/3	3/3

* In the illustration x/y, y indicates the number of interviewees who gave a relevant response to the attribute and x indicates the number of interviewees who confirmed the key finding. A finding counts as confirmed if any indication that the attribute applies was made.

6. Conclusion

While it is known that collaborative innovation adheres key advantages compared to hierarchical and competitive innovation strategies, no setting in which collaborative innovation can be supported and sustained in the present governance system had been found yet. Therefore, this study aimed to analyse the so-far unexplored concept of innovation fellowship programmes by answering the research question of how the fellowship programme Tech4Germany contributes to the development of digital innovations in German public organisations. The core finding of this study is that the scientific phenomenon of fellowship programmes qualifies as a form of collaborative innovation as the theoretical mechanisms are observable to a large extent. Therefore, fellowship programmes can be described as a third space between one-time collaboration projects and innovation labs. In theoretical terms, a fellowship programme thus contributes to the development of digital innovations by providing a new institutional design for collaborative innovation. Hence, this study fills the identified research gap by contributing important findings to the discussion on suitable ways to integrate collaborative innovation in the current governance system.

Precisely, fellowship programmes enable collaborative innovation by lowering the transaction costs for the public organisations, creating a setting that attracts digital innovation experts and providing the basic structure and rules for constructive collaboration. Importantly, the new institutional design simultaneously provides a unique room in which public employees can experience new ways of working, particularly agile and user-centric approaches that are crucial in digital innovation projects. Experiencing new methods in a real project stimulates mutual and transformative learning, thus implying a positive impact on the development of digital innovations beyond the single project. However, the findings in this aspect are mixed, indicating that the experience of learning depends on individual characteristics and the degree of involvement in the collaboration. In addition to that, fellowship programmes create joint ownership of the developed prototype, leading to support for the implementation of the digital innovation among the involved public employees. As current research found that reaching the implementation stage is a key challenge in separated institutional designs such as innovation labs (McGann et al., 2021), this finding emphasises the relevance of the embeddedness of innovation fellowship programmes within the public organisation.

The outlined mechanisms strongly differentiate the collaboration of public employees with digital experts from contracting-out the development of digital innovations to external IT-providers. The study therefore aligns with the wider theoretical paradigm shift towards New Public Governance. Particularly, the results illustrate how installing an intermediate instance like the management team of Tech4Germany helps to overcome former barriers of collaborative forms of governance. By initiating the first step towards opening-up public organisations for interdisciplinary and cross-sectional expertise and allowing the public organisations to experiment with agile and user-centric approaches in a secure setting, fellowship programmes hence contribute an important building block to the research about new forms of governance in the digital age. Additionally, the unique institutional design points to the theoretical importance of overcoming simplified discussions about contracting-in or

contracting-out digital public services, instead looking beyond established organisational forms.

Having systematically analysed the fellowship programme in the light of the existing scholarship of collaborative innovation, this study sets the initial groundwork for future research on the scientific phenomenon of innovation fellowship programmes and provides promising starting points for further studies. First, quantitative studies are needed to measure the impact of fellowship programmes. Here, an interrupted time series design based on a questionnaire answered by all participants in the beginning and at the end of each project appears to be a valuable research design. Given the bounded setting of Tech4Germany, the case also provides opportunities for different data collection methods, for instance focus groups or observations of behaviour. This would allow to further investigate the before outlined theoretical implications, particularly the analysis of additional antecedents like the organisational design and decision-making hierarchy. Moreover, a study that evaluates the long-term impact and the degree of improvement through the developed digital innovations is recommended. Beyond that, case studies of different fellowship programmes and especially cross-country comparative studies would increase the validity of these findings and would allow to analyse which institutional design of fellowship programmes provides the best setting for collaborative innovation.

In addition to the implications for research, the study provides recommendations for practitioners in politics, public administration, and governance. Concerning the specific setting of Tech4Germany, the findings clearly suggest that the Digitallotsen should be granted more time to actively collaborate with the Fellows and thus need to be freed from some of their daily responsibilities. This would increase the positive effect of mutual and transformative learning and therefore also the long-term impact on the development of digital innovations. Particular attention should also be paid to including public employees in the collaboration who are sceptical about digital innovations and modern working methods. Even though this might increase the transaction costs of the collaborative projects, it is essential to establish a cultural change and ultimately reach the goal of Tech4Germany to "expedite the digitalisation of Germany and thereby learn from and with each other" (Tech4Germany, 2021c, p. 1).

From a wider perspective, fellowship programmes appear to be a suitable alternative or supplement to innovation labs and one-time collaboration projects as they not only create digital innovation prototypes but may also increase the in-house capacity of public organisations to develop and oversee digital services and products. Hence, it is recommended to establish more fellowship programmes, for instance on the federal state level. This seems to be realistic as this study shows that experts from the digital economy are generally willing to contribute their technical knowledge to a trustful relationship with the public entities and the number of applications by far exceeds the number of projects that Tech4Germany can implement (Anton & Hupperth, 2020a). Beyond the unique concept of fellowship programmes, this study revealed the importance of actively leveraging the knowledge and experiences from outside the public sector without fully outsourcing digital services to IT-providers. Therefore, it is recommended to overall work towards a stronger emphasis on collaborative interactions between digital experts and public employees since opening-up public organisations to new forms of collaboration bears the potential to develop digital innovations in line with democratic principles.

References

Agger, A., & Sørensen, E. (2018). Managing collaborative innovation in public bureaucracies. *Planning Theory*, *17*(1), 53–73.

Ansell, C., & Gash, A. (2008). Collaborative Governance in Theory and Practice. *Journal of Public Administration Research and Theory*, *18*(4), 543–571.

Ansell, C., & Torfing, J. (2015). How does collaborative governance scale? *Policy & Politics*, *43*(3), 315–329.

Anton, S., & Hupperth, A. (2020a). *Episode 11: Tech4Germany - wie macht man den Staat fit für die Digitalisierung, Sonja Anton & Anna Hupperth? (Teil 2)*. https://open.spotify.com/episode/69SzsRr5rZnPf88toyXJiX. (Accessed: 2021-05-04)

Anton, S., & Hupperth, A. (2020b). *Episode 9: Tech4Germany – Wie digitalisiert man die öffentliche Verwaltung? (Teil 1)*. https://open.spotify.com/episode/5Xs8gMeDUbMuiOLdtDGQQZ. (Accessed: 2021-05-04)

Babbie, E. R. (2013). *The practice of social research* (13th ed.). Cengage Learning.

Berg, B. (2009). *Qualitative Reserach Methods for the Social Sciences*. Boston, Allyn & Bacon.

Blatter, J., & Haverland, M. (2012). *Designing case studies: Explanatory approaches in small-N research*. New York, Palgrave Macmillan.

Bommert, B. (2010). Collaborative Innovation in the Public Sector. *International Public Management Review*, *11*(1), 15–33.

Borins, S. (2001). Encouraging innovation in the public sector. *Journal of Intellectual Capital*, *2*(3), 310–319.

Chesbrough, H. W. (2003). *Open innovation: The new imperative for creating and profiting from technology*. Boston, Harvard Business School Press.

Cinar, E., Trott, P., & Simms, C. (2019). A systematic review of barriers to public sector innovation process. *Public Management Review*, *21*(2), 264–290.

Ciriello, R. F., Richter, A., & Schwabe, G. (2018). Digital Innovation. *Business & Information Systems Engineering*, *60*(4), 563–569.

Coglianese, C. (2020). Algorithmic regulation: Machine learning as a governance tool. In M. Schuilenburg & R. Peeters (Eds.), *The algorithmic society. Technology, power, and knowledge* (pp. 35–52). Andover, Routledge.

Damanpour, F. (1991). Organizational Innovation: A Meta-Analysis of Effects of Determinants and Moderators. *The Academy of Mangagement Journal*, *34*(3), 555–590.

DeVries, H., Bekkers, V., & Tummers, L. G. (2016). Innovation in the public sector: a systematic review and future research agenda. *Public Administration*, *94*(1), 146–166.

DeVries, H., Tummers, L. G., & Bekkers, V. (2018). The Diffusion and Adoption of Public Sector Innovations: A Meta-Synthesis of the Literature. *Perspectives on Public Management and Governance*, *1*(3), 159–176.

Digitallotse 1. (2021). (2021, May 10). Interview 1.

Digitallotse 2. (2021). (2021, May 12). Interview 4.

Digitallotse 3. (2021). (2021, May 21). Interview 6.

DigitalService4Germany. (2021a). *Company Website*. https://digitalservice4germany.com/. (Accessed: 2021-04-20)

DigitalService4Germany. (2021b). *Corporate Governance Bericht 2020: nach Ziffer 6.1 des Public Corporate Governance Kodex des Bundes*. https://digitalservice4germany.com/assets/corporategovernance-bericht-2020.pdf. (Accessed: 2021-04-20)

DigitalService4Germany. (2021c). *Kooperationsvertrag Tech4Germany*. https://tech.4germany.org/wp-content/uploads/2021/02/Kooperationsvertrag_Tech4Germany-2021.docx.pdf. (Accessed: 2021-04-20)

DigitalService4Germany. (2021d). *Stipendiumsrichtlinie Tech4Germany*. https://tech.4germany.org/wp-content/uploads/2021/02/Stipendiumsrichtlinien_Tech4Germany_2021.docx.pdf. (Accessed: 2021-04-20)

Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2006). New Public Management Is Dead - Long Live Digital-Era Governance. *Journal of Public Administration Research and Theory*, *16*(3), 467–494.

Eggers, W. D., & Singh, S. K. (2009). *The public innovator's playbook: Nurturing bold ideas in government*. Washington, DC, Harvard Kennedy School of Government.

European Commission. (2020). *Digital Economy and Society Index (DESI) 2020: Digital public services*. https://ec-europa-eu.ezproxy2.utwente.nl/digital-single-market/en/digital-public-services. (Accessed: 2021-05-03)

Fellow 1. (2021). (2021, May 10). Interview 2.

Fellow 2. (2021). (2021, May 11). Interview 3.

Fellow 3. (2021). (2021, May 17). Interview 5.

Finnegan, R. (2006). Using Documents. In R. Sapsford & V. Jupp (Eds.), *Data collection and analysis* (pp. 138–150). London, SAGE.

Fledderus, J., Brandsen, T., & Honingh, M. (2014). Restoring Trust Through the Co-Production of Public Services: A theoretical elaboration. *Public Management Review*, *16*(3), 424–443.

Gascó, M. (2017). Living labs: Implementing open innovation in the public sector. *Government Information Quarterly*, *34*(1), 90–98.

Gil-Garcia, J. R., Dawes, S. S., & Pardo, T. A. (2018). Digital government and public management research: finding the crossroads. *Public Management Review*, *20*(5), 633–646.

Godin, B., & Vinck, D. (Eds.). (2017). *Critical studies of innovation: Alternative approaches to the pro-innovation bias*. Cheltenham, Edward Elgar.

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: systematic review and recommendations. *The Milbank Quarterly*, *82*(4), 581–629.

Hartley, J., Sørensen, E., & Torfing, J. (2013). Collaborative Innovation: A Viable Alternative to Market Competition and Organizational Entrepreneurship. *Public Administration Review*, *73*(6), 821–830.

Heartley, J. (2005). Innovation in Governance and Public Services: Past and Present. *Public Money and Management*, *25*(1), 27–34.

HM Government. (2021). *No.10 Innovation Fellowships*. https://no10innovationfellows.campaign.gov.uk/. (Accessed: 2021-05-03)

Kvale, S. (2007). *Doing interviews*. London, SAGE.

Lang, C. (2020, July). *Frauen digitalisieren Deutschland*. https://open.spotify.com/episode/5N2Wm8Tyfn9CZT7Z3jN8Ii. (Accessed: 2021-04-29)

Lang, C. (2021, April). *Episode #7 – Christina Lang*. https://open.spotify.com/episode/7i3XDBuZsJI2Rpqmsodmjn?si=AHbqLIKLQzqvY9WQZGvfbQ. (Accessed: 2021-04-29)

Lewis, J., & Ritchie, J. (2003). Generalising form Qualitative Research. In J. Ritchie & J. Lewis (Eds.), *Qualitative research practice* (pp. 263–286). London, SAGE.

Lindsay, C., Findlay, P., McQuarrie, J., Bennie, M., Corcoran, E. D., & van der Meer, R. (2017). Collaborative Innovation, New Technologies, and Work Redesign. *Public Administration Review*, *78*(2), 251–260.

Lopes, A. V., & Farias, J. S. (2020). How can governance support collaborative innovation in the public sector? A systematic review of the literature. *International Review of Administrative Sciences*, 1–17.

Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12th ed.). Basel, Beltz Verlag.

McGann, M., Blomkamp, E., & Lewis, J. (2018). The rise of public sector innovation labs: experiments in design thinking for policy. *Policy Science*, *51*, 249–267.

McGann, M., Wells, T., & Blomkamp, E. (2021). Innovation labs and co-production in public problem solving. *Public Management Review*, *23*(2), 297–316.

McIvor, D. W. (2020). Toward a Critical Theory of Collaborative Governance. *Administrative Theory & Praxis*, *42*(4), 501–516.

Mergel, I. (2016). Agile innovation management in government: A research agenda. *Government Information Quarterly*, *33*(3), 516–523.

Mergel, I. (2019). Digital service teams in government. *Government Information Quarterly*, *36*(4), 1–16.

Moore, M., & Hartley, J. (2008). Innovations in governance. *Public Management Review*, *10*(1), 3–20.

Neumann, O., Matt, C., Hitz-Gamper, B. S., Schmidthuber, L., & Stürmer, M. (2019). Joining forces for public value creation? Exploring collaborative innovation in smart city initiatives. *Government Information Quarterly*, *36*(4), 1–14.

Obama White House Archives. (2021). *Presidential Innovation Fellows*. https://obamawhitehouse.archives.gov/innovationfellows. (Accessed: 2021-05-03)

Osborne, S. P. (2006). The New Public Governance? *Public Management*

Review, *8*(3), 377–387.

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Newsbury Park, SAGE.

Peeters, R., & Schuilenburg, M. (2020). The algorithmic society: An introduction. In M. Schuilenburg & R. Peeters (Eds.), *The algorithmic society. Technology, power, and knowledge* (pp. 1–16). Andover, Routledge.

Presse- und Informationsamt der Bundesregierung. (2020, October). *Tech4Germany. Technologie-Taskforce für die Bundesministerien.* https://www.bundesregierung.de/breg-de/suche/tech4germany-1797948. (Accessed: 2021-05-03)

Rainey, H. G. (2014). *Understanding and managing public organizations* (5th ed.). San Francisco, Jossey-Bass.

Ritchie, J. (2003). The Applications of Qualitative Methods to Social Research. In J. Ritchie & J. Lewis (Eds.), *Qualitative research practice* (pp. 24–46). London, SAGE.

Ritchie, J., Lewis, J., & Elam, G. (2003). Designing and Selecting Samples. In J. Ritchie & J. Lewis (Eds.), *Qualitative research practice* (pp. 77–108). London, SAGE.

Rodríguez, J. E. (2021). *Mit Tech4Germany Dienstleistungen gestalten, die besser für alle funktionieren (2021, February 12)*. https://jerdesign.medium.com/mit-tech4germany-dienstleistungen-gestalten-die-besser-f%C3%BCr-alle-funktionieren-26a456a211d6. (Accessed: 2021-04-29)

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York, Free Press.

Sicilia, M., Guarini, E., Sancio, A., Andreani, M., & Ruffini, R. (2016). Public services management and co-production in multi-level governance settings. *International Review of Administrative Sciences*, *82*(1), 8–27.

Sørensen, E., & Torfing, J. (2011). Enhancing Collaborative Innovation in the Public Sector. *Administration & Society*, *43*(8), 842–868.

Sørensen, E., & Torfing, J. (2012). Introduction: Collaborative innovation in the public sector. *Innovation Journal*, *17*(1), 1–14.

Sørensen, E., & Torfing, J. (2015). Enhancing Public Innovation through Collaboration, Leadership and New Public Governance. In A. Nicholls, J. Gundelfinger, & M. Gabriel (Eds.), *New Frontiers in Social Innovation Research* (pp. 145–169). Basingstoke, Palgrave Macmillan.

Sørensen, E., & Torfing, J. (2017). Metagoverning Collaborative Innovation in Governance Networks. *The American Review of Public Administration*, *47*(7), 826–839.

Swanson, R. A., & Holton, E. F. (2005). *Research in organizations: Foundations and methods of inquiry*. San Francisco, Berrett-Koehler.

Tech4Germany. (2021a). *Checkliste: Eignen wir uns mit unserem Projekt für das Fellowship?* https://tech.4germany.org/wp-content/uploads/2020/12/Checkliste-Eignet-sich-mein-Projekt-fu%CC%88r-das-Tech4Germany-Fellowship_-2-2.pdf. (Accessed: 2021-04-20)

Tech4Germany. (2021b). *Company Website.* https://tech.4germany.org/. (Accessed: 2021-04-20)

Tech4Germany. (2021c). *Tech4Germany Kurzinfo: Für digitale Vorreiter:innen in der Bundesverwaltung.* https://tech.4germany.org/wp-content/uploads/2021/02/Tech4Germany-Kurzinfo.pdf. (Accessed: 2021-04-20)

Torfing, J. (2016). *Collaborative innovation in the public sector*. Washington, DC, Georgetown University Press.

Torfing, J. (2019). Collaborative innovation in the public sector: the argument. *Public Management Review*, *21*(1), 1–11.

Torfing, J., Sørensen, E., & Røiseland, A. (2019). Transforming the Public Sector Into an Arena for Co-Creation: Barriers, Drivers, Benefits, and Ways Forward. *Administration & Society*, *51*(5), 795–825.

van Thiel, S. (2007). *Research Methods in Public Administration and Public Management: An Introduction*. New York, Routledge.

Verschuere, B., Brandsen, T., & Pestoff, V. (2012). Co-production: The State of the Art in Research and the Future Agenda. *Voluntas*, *23*(4), 1083–1101.

Voorberg, W., Bekkers, M., & Tummers, L. G. (2015). A Systematic Review of Co-Creation and Co-Production: Embarking on the social innovation journey. *Public Management Review*, *17*(9), 1333–1357.

Voorberg, W., Bekkers, V., Timeus, K., Tonurist, P., & Tummers, L. (2017). Changing public service delivery: learning in co-creation. *Policy and*

Society, *36*(2), 178–194.

Wegrich, K. (2019). The blind spots of collaborative innovation. *Public Management Review*, *21*(1), 12–20.

Whicher, A., & Crick, T. (2019). Co-design, evaluation and the Northern Ireland Innovation Lab. *Public Money and Management*, *39*(4), 290–299.

Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial Intelligence and the Public Sector - Applications and Challenges. *International Journal of Public Administration*, *42*(7), 596–615.

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, SAGE.