
ESSAYS ON THE ADVANCEMENT OF SOCIALLY RESPONSIBLE INVESTMENT RESEARCH: A CONCEPTUAL AND METHODOLOGICAL PERSPECTIVE

Universität Hamburg

Fakultät für Wirtschafts- und Sozialwissenschaften

(kumulative) Dissertation

Zur Erlangung der Würde des Doktors der Wirtschafts- und
Sozialwissenschaften

(gemäß der Promotionsordnung vom 24. August 2010)

vorgelegt von

Josua Erik Ottfried Oll

aus Dormagen

Hamburg, den 07. Februar 2019

Prüfungskommission

Vorsitzender: Prof. Dr. Sebastian Spaeth, Universität Hamburg

Erstgutachter: Prof. Dr. Alexander Bassen, Universität Hamburg

Zweitgutachter: Prof. Dr. Dirk Ulrich Gilbert, Universität Hamburg

Drittgutachter: Prof. Dr. Matthias Fifka, Friedrich-Alexander-Universität Erlangen-Nürnberg

Zweitbetreuer: Prof. Dr. Olaf Weber, University of Waterloo

Datum der Disputation: 14. Mai 2019

Acknowledgments

Writing a PhD thesis has been a demanding task but also an exciting and extremely rewarding experience. On this occasion, I would like to express my sincere gratitude to all the people who supported me along the way.

First and foremost, I want to thank my supervisor Professor Alexander Bassen, who provided me with the opportunity to engage in a dissertation project at his chair. Based on a big leap of faith, he granted me a tremendous degree of freedom and allowed me to find and follow my own research interests. At the same time, he provided me with endless support, guidance, and advice, and I benefitted immensely from his experience. For some reason, I have been blessed with happiness and luck throughout my entire life so far. Still, I consider the last four and a half years to be the best years I have had. Surely, several factors contributed to this, but working at Professor Bassen's chair clearly played a leading role.

In addition, I would like to thank Professor Olaf Weber, who I first met in 2015 at the Social and Sustainable Finance Conference at the University of Oxford. As my second supervisor, he provided me with continuous advice and suggestions on how to develop my research further, while always being a very pleasant conversational partner. My gratitude also extends to Professor Sebastian Spaeth and Professor Dirk Ulrich Gilbert, who kindly accepted to take on the roles of dissertation committee chair and second assessor without hesitation.

While working at Professor Bassen's chair, many people contributed to the pleasant and stimulating work environment. In this context, I have to give a big thanks to Katrin Gödker, Florian Lüdeke-Freund, Franziska Sump, and Christine Zöllner. I am particularly indebted to Katrin Gödker and Franziska Sump. Our endless but never boring discussions greatly enhanced the quality of my research, and I hope my input also turned out to be helpful once in a while. I wish you all the best, academically and personally, for your upcoming time at the University of Maastricht and the University of Southern Denmark. I sincerely hope that our academic exchange, discussions, and friendship will continue far beyond our shared time at the University of Hamburg.

Working at Rentzelstraße 7 is a blessing, as many researchers with similar and/or overlapping research interests "reside" under a single roof. In this context, many thanks go to Professor Timo Busch, Professor Kerstin Lopatta, Professor Frank Schiemann, all respective team members, and everyone further associated with the Research Group on Sustainable Finance at the University of Hamburg. I greatly benefitted from the academic exchange across floors but also

enjoyed our social activities. Due to excellent advice from Professor Timo Busch, we shared many unforgettable moments.

As research is rarely a solitary activity, I additionally thank all of my much-appreciated co-authors for their tremendous contributions (and patience). I particularly thank Professor Rüdiger Hahn for putting me on a steep learning curve in my early PhD days. I am also indebted to Professor Martin Meißner, who kindly invited me for a research stay at the University of Southern Denmark. I am looking forward to proceeding with our joint research projects.

Over the years, various parts of this dissertation have been presented at academic conferences, workshops, and doctoral consortia. I thank all participants for the fruitful discussions, as well as helpful suggestions and advice, even on early stage projects. Although it is impossible to mention each and every one of them individually, I would like to thank: Professor Jill Atkins, Professor Nicole Darnall, Professor Matthias Fifka, Professor Edward Freeman, Professor Daniel Geiger, Professor Thomas Kaspereit, Professor Bert Scholtens, Dr. Michael Schröder, and Professor Bryan Stinchfield. In this context, I am very grateful to the Graduate School of the School of Economics and Social Sciences at the University of Hamburg for generously supporting many of my conference visits.

I would also like to express my gratitude to my parents, Hildegund and Bernhard Oll, who sometimes wondered how long it could take to write a dissertation but provided me with tremendous support in all domains of life.

Finally, I thank my beloved partner in life, Nina Schönewolf. You accepted my unusual lifestyle, late working hours, and all of my further irks and quirks without (much) complaints. I believe that most women would have left years ago. You, however, not only stayed but let yourself in for starting a family with me. My delight and gratitude are beyond words.

Table of Contents

	<u>Page</u>
1 Synopsis	1
1.1 Background	1
1.2 Problem Statement	2
1.3 Research Objective and Question	4
1.4 Outline of the Thesis	4
1.5 Contribution	11
2 First Article	22
3 Second Article	59
4 Third Article	96
5 Fourth Article	138
6 Appendices	180
6.1 Short Summary of Articles	181
6.2 Statutory Declaration	186
6.3 Self-Declaration	187

1 Synopsis

1.1 Background

Sustainability challenges, such as climate change, resource depletion, and social inequality, belong to the most fundamental and pressing challenges of our time and impose a severe burden on environmental protection, as well as on economic and social development. In light of these issues, a decisive and ambivalent role can be assigned to financial markets (Busch et al., 2016; Scholtens, 2006; Wiek & Weber, 2014). On the one hand, financial markets significantly contribute to sustainability problems. Although financial intermediaries, such as banks, do not directly “produce hazardous chemicals or discharge toxic pollutants into the air, land or water” (Thompson & Cowton, 2004, p. 199), they channel capital to different actors (e.g., institutions, sectors, regions) whose activities enforce environmental degradation, unsustainable business practices, income disparities, and military conflicts (Addison et al., 2001; Sarokin & Schulkin, 1991; Scholtens, 2006, 2009; Wiek & Weber, 2014). At the same time, however, financial markets possess substantial problem-solving capacities to address sustainability concerns. For instance, recent estimates suggest that an additional annual investment of \$650 to \$900 billion in low-carbon activities is needed over the next few decades to decarbonize the economy and meet the 2 °C threshold (Campiglio, 2016; IEA, 2012; McKinsey & Co, 2010; WEF, 2013). The global fight against climate change thus requires a large-scale investment beyond the public sector’s capacity. Against this backdrop, financial markets have much to contribute to fill the low-carbon (green) investment gap and to direct capital to players that promote sustainability (Campiglio, 2016; Covington, 2017; UNEP FI, 2009).

Many developments document the growing awareness and recognition that financial markets play a decisive role in facilitating sustainability. For instance, voluntary initiatives to foster sustainability in the financial sector have emerged, including the Principles for Responsible Investment (PRI), the Montreal Carbon Pledge, the Portfolio Decarbonization Coalition, or Climate Action 100+. Furthermore, new institutions, such as the Carbon Disclosure Project (CDP) or the Task Force on Climate-related Financial Disclosures (TCFD), have evolved to increase the corporate disclosure of climate-related data. Regulatory bodies have also started to become active. For instance, France established the Energy Transition for Green Growth Act (LTECV) in 2015, mandating that asset management companies and institutional investors report on environmental, social, and governance (ESG) criteria and climate policy on a “comply or explain” principle. In 2016, the European Commission appointed a High-Level Expert Group on Sustainable Finance to develop recommendations for a consistent EU strategy on sustainable

finance. Building on the group's recommendations (HLEG, 2018), the European Commission (2018) published its action plan "Financing Sustainable Growth" in March 2018.

Sustainability considerations are also increasingly incorporated into investment strategies through a broad spectrum of practices (Eurosif, 2016; GSIA, 2016; US SIF, 2018). These practices tend to be "described using overlapping and complementary terms" (Busch et al., 2016, p. 305), including socially responsible investment (SRI), sustainable investment, ethical investment, and others (Busch et al., 2016; Cadman, 2011; Eccles & Viviers, 2011; Schueth, 2003). In general terms, SRI refers to a shift away from purely financial objectives toward the integration of ESG factors into the investment process (Capelle-Blancard & Monjon, 2012). Representing a niche market for a long time, SRI has begun to enter the mainstream of investment practice and has expanded globally in recent years. According to the most recent Global Sustainable Investment Review (GSIA, 2016), the worldwide SRI market comprises \$22.89 trillion of professionally managed assets. In relative terms, SRI captures "26 percent of all professionally managed assets globally" (GSIA, 2016, p. 3). Although SRI experienced tremendous growth overall, SRI markets vary considerably across geographic regions with respect to absolute/relative market sizes, growth rates, the presence of private and institutional investors, asset classes, and investment strategies (Eurosif, 2016; GSIA, 2016; US SIF, 2018). SRI thus constitutes a substantial, but highly heterogeneous market segment.

1.2 Problem Statement

Considering the growing importance of SRI to financial markets, it is no surprise that market growth has been accompanied by increased academic interest and attention. With respect to SRI inquiry, prior research has focused predominantly on the financial performance of SRI and (potential) performance differences between SRI and conventional investment approaches (e.g., Bauer et al., 2005; Bello, 2005; Renneboog, 2008). Capelle-Blancard and Monjon (2012), amongst others (e.g., Hoepner & McMillan, 2009; Von Wallis & Klein, 2015), provide evidence for this assertion based on SRI literature reviews. Yet, Capelle-Blancard and Monjon (2012, p. 240) additionally find that SRI performance studies "are among the most influential papers" in the field, whereas few, and a declining number of papers, "are concerned with ethics, altruism, or moral values." This finding resonates with the observation of other SRI scholars that, despite mounting research, relatively little is known about the underlying drivers and motives for SRI (e.g., Diouf et al., 2016; Glac, 2009; Hong & Kostovetsky, 2012; Scholtens & Sievänen, 2013). The key question of why some investors adopt SRI practices while others refrain from doing so is therefore still largely unanswered.

Considering the current state of the literature on the underlying drivers and motives for SRI, respective research has not only been referred to as fragmented, but also as conceptually and methodologically limited (e.g., Capelle-Blancard & Monjon, 2012; Diouf et al., 2016; Glac, 2009). First, to explain why (some) investors engage in SRI practices, prior research has focused predominantly on socio-demographic characteristics (e.g., age, income, education level) and a narrow set of concepts,¹ particularly pertaining to investors' values, attitudes, norms, and beliefs (e.g., Brodback et al., 2018; Cheah et al., 2011; Hong & Kacperczyk, 2009; Jansson & Biel, 2011, 2014; Junkus & Berry, 2010; McLachlan & Gardner, 2004; Nilsson, 2008, 2009; Rosen, 1991; Tippet & Leung, 2001; Williams, 2007). Second, this stream of research tends to be limited to descriptive and comparative accounts of conventional and socially responsible investors, due to the primary use of surveys.

From a methodological perspective, the dominance of survey research can be viewed as problematic, particularly when considering potential sources of common method biases (Podsakoff et al., 2003). For instance, in the aforementioned studies, survey questions are usually employed to collect data on both independent (e.g., investors' values, beliefs, attitudes) and dependent (e.g., likelihood to invest in SRI, intention to increase SRI in the future) variables. The fact that measures of both variables are provided by the same person and collected through the same method in a common measurement and item context gives rise to potential biases due to respondents' consistency motives and social desirability (Donaldson & Grant-Vallone, 2002; Johns, 1994; Podsakoff et al., 2003; Schmitt, 1994). Furthermore, questions to obtain data on dependent variables tend to be phrased broadly and to draw on participants' implicit understanding of the term SRI (e.g., "I will increase my investments in SRI assets during the next 2–5 years") (Jansson & Biel, 2011, p. 138). Yet, SRI has been referred to as a contested concept (Woods & Urwin, 2010), as definitions and understandings of the term vary (Cheah et al., 2011; Eccles & Viviers, 2011; Von Wallis & Klein, 2015). As such, the concept of SRI remains ambiguous to participants in survey research.

Besides these potential sources of common method biases, further methodological limits of survey research likely constrain the gaining of a deeper understanding of the drivers and motives for SRI. First, although surveys provide insights into the relation between independent and dependent variable(s), they are ill equipped for studying the (implicit) interplay of the mul-

¹ Following Podsakoff et al. (2016, p. 161), this dissertation uses the term "concept(s)" to refer to "cognitive symbols (or abstract terms) that specify the features, attributes, or characteristics of the phenomenon in the real or phenomenological world that they are meant to represent and that distinguish them from other related phenomena."

multiple factors that underlie complex judgments and decision-making processes, such as SRI. Second, conventional and SRI does not take place in a vacuum, but rests on investors' acquisition and processing of investment-relevant (financial and/or non-financial) information (Maines & McDaniel, 2000; Reimsbach et al., 2018). Surveys are poorly suited to open the "black box" between informational inputs (e.g., financial/non-financial information) and decision outcomes (e.g., conventional versus SRI).

1.3 Research Objective and Question

As outlined in the previous section, the question of why some investors adopt SRI practices while others refrain from doing so remains a largely unresolved puzzle. Furthermore, prior research on this question is not only limited, but also predominately grounded in a narrow range of concepts and survey research with associated methodological problems and limits. In light of these shortcomings, the present dissertation seeks to advance SRI research toward a better understanding of the underlying drivers and motives for SRI. More specifically, this dissertation addresses the following research question: how can conceptual and methodological advances contribute to a deeper understanding of the underlying drivers and motives for SRI?

1.4 Outline of the Thesis

This dissertation is grounded in four articles to reach the aforementioned research objective and to answer the stated research question. Article 1 illuminates the potential of factorial surveys for business and society research and illustrates the method's application and benefits, with a showcase example of SRI research. Building on this, Article 2 applies conjoint methodology, a method closely related to factorial surveys,² and investigates how the concepts of intuition and reasoning are associated with the effectiveness of climate labeling in the investment context. The methodologies employed in Articles 1 and 2 are well suited to explore the associations between information inputs and decision outcomes. Moving beyond this pure association, Article 3 suggests a methodological approach to shed light on the underlying processes that occur between the provision of information and decision outcomes.³ The article illustrates the potential of eye tracking to gain insights into the cognitive processes underlying human judgment and decision-making. With regard to advancing SRI research from a conceptual perspective, Articles 2 and 4 contribute to the research objective. Whereas Article 2 contributes a novel

² The key methodological difference is that participants in a conjoint methodology usually compare and rank different alternatives (profile cards), whereas respondents in factorial survey studies evaluate realistic scenarios (vignettes) sequentially on a rating scale (Auspurg & Jackle, 2017; Shooter & Galloway, 2010).

³ This point is further elaborated in the following summary of Article 3.

conceptual perspective (intuition versus reasoning) to SRI research, Article 4 provides a critical review of a more established concept in the SRI literature, namely, the concept of values. The following provides a more detailed summary of each of the four articles.

Article 1: Tackling Complexity in Business and Society Research: The Methodological and Thematic Potential of Factorial Surveys

Understanding coherences and generating knowledge in the domain of business and society (B&S) research often proves challenging, as the field is characterized by dependencies on and the interactions of societal-, organizational-, and individual-level factors (Aguinis & Glavas, 2012; Athanasopoulou & Selsky, 2015). B&S issues also frequently involve judgments and decisions that make it difficult to explicate the underlying factors. Furthermore, B&S scholars inevitably deal with fuzzy and difficult-to-define concepts that are subject to a multiplicity of interpretations (Schwartz & Carroll, 2008). Given the above-described nature of B&S research, scholars are challenged to consider this complexity with regard to the choice and design of methods.

Against this backdrop, Article 1 draws on the discipline of sociology (Rossi, 1979; Rossi & Anderson, 1982; Wallander, 2009) to present the so-called factorial survey as a promising method for research in the field of B&S, as well as SRI in particular. As a hybrid approach, factorial surveys integrate the characteristics of survey research and experimental designs into a single method (Hox et al., 1991; Wallander, 2009), thereby allowing researchers to collect data from a large number of respondents in a controlled setting. Participants in factorial survey studies receive a series of vignettes, i.e., carefully designed descriptions of hypothetical people, social situations, or scenarios. These vignettes contain a large number of factors (dimensions) that vary randomly in their values. Participants are then asked to make judgments about the presented fictive descriptions, and the data of the respondents' judgments generate the dependent (outcome) variable(s) (Jasso, 2006). After completing the rating task, respondent-specific information is collected analogous to a traditional survey.

To carve out the potential of factorial surveys in B&S research, Article 1 illustrates the method's application and benefits based on a showcase example in the realm of SRI. The utilized showcase exemplifies how factorial surveys can help pave the way toward integrating the individual's perspective with the organizational and societal views of SRI. The article additionally illuminates how factorial surveys can add value to research on fuzzy and difficult-to-articulate concepts, such as SRI. In traditional survey research, questions for generating the dependent variable tend to be phrased broadly and usually draw on participants' implicit understanding

of certain concepts. The vignettes of factorial surveys, in contrast, allow for a more granular operationalization of concepts (e.g., SRI), the manipulation of specific conceptual aspects, and the testing of which sub-aspects are empirically relevant from the participants' perspective.

As the methodological literature on factorial surveys is still limited, Article 1 further illustrates the method's implementation process and provides specific recommendations along six main steps. These steps include constructing the research question, constructing the vignette universe, choosing individual-level features, making the sampling decision, collecting respondent-specific information, and analyzing the data. To provide a hands-on illustration, each process step directly relates back to the showcase vignette on SRI. With respect to the analysis of factorial survey data, Article 1 demonstrates the use of multilevel regression analysis based on 200 simulated (virtual) "participants," reflecting 2,400 investment judgments. Before closing with concluding remarks, the article provides an overview of the method's potential advantages and disadvantages and guides the reader to key methodological literature sources.

Article 2: Climate Information in Retail Investors' Decision-Making: Evidence from a Choice Experiment

As meeting the 2 °C threshold requires significant investment, financial markets play a decisive role in financing the low-carbon transition (Campiglio, 2016; Covington, 2017). Although SRI has gained importance in the market and in academia alike, market growth is predominantly driven by institutional investors (Eurosif, 2016; GSIA, 2016). Mobilizing retail investors for climate-friendly investment products therefore represents a promising leverage point to channel more financial capital toward the global fight against climate change.

Individual choices, such as retail investors' investment decisions, often depend on how options and alternatives are presented (Thaler & Sunstein, 2008). Building on Thaler and Sunstein's (2003, 2008) seminal work on the concept of nudge,⁴ Pilaj (2017) recently argued that modifying investors' choice environment has great potential as a means by which to dismantle existing impediments to the adoption of SRI practices. In the recent past, the nudge concept has been increasingly discussed as an attractive policy tool (van Bavel et al., 2013; World Bank, 2015), and considerable evidence indicates that the implementation of nudges can advance more sustainable behavior (e.g., Allcott & Mullainathan, 2010; Costa & Kahn, 2013; Momsen

⁴ The term "nudge" refers to any aspect of the choice environment that "alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives" (Thaler & Sunstein, 2008, p. 6).

& Stoerk, 2014). Up to the present, however, the promise of nudges has not been studied empirically in the realm of SRI. Yet, such research is needed, as the available empirical evidence of the effectiveness of nudges is highly context-dependent and of limited generalizability (Bao & Ho, 2015; Lehner et al., 2015).

Against this backdrop, Article 2 aims to shed light on how the presentation format of climate information influences investment decisions in the retail segment. To reach this objective, Article 2 builds on an online choice experiment among European retail investors in which investors' choice environment is systematically altered. By applying the conjoint methodology, a method closely related to factorial surveys, Article 2 examines the equity fund choices of 953 retail investors from six European countries. These investors are provided with standard financial and climate performance information on investment alternatives. Specifically, the presentation format of climate performance information is varied to test the effectiveness of three distinct label designs to influence investment decisions. The article further utilizes the concepts of intuition and reasoning to analyze the role of investors' cognitive characteristics in the effectiveness of climate labeling. Building on Frederick's (2005) well-established cognitive reflection test, investors are classified along a continuum ranging from intuition (System 1) to reasoning (System 2).

The findings of Article 2 reveal that different climate label designs lead to variation in participants' investment decisions, indicating that these designs differ in their effectiveness to nudge investors toward more climate-friendly investments. Among the three label designs, the star rating label is found to yield the highest emphasis on climate performance information in investors' decision-making. The empirical evidence further indicates that climate labeling is more (less) effective in promoting climate-friendly investing among intuitive (reflective) decision makers. An underlying mechanism of this finding seems to be the individual trading off between a fund's climate performance and its financial performance. Intuitive decision makers tend to place significantly more weight on a fund's climate performance than its financial performance—irrespective of a participant's environmental preference.

Article 3: The Promise of Eye-Tracking Methodology in Organizational Research: A Taxonomy, Review, and Future Avenues

The methodologies employed in Articles 1 and 2, i.e., the factorial survey and conjoint methodology, both belong to the regression-based methods of decision-making research (Aiman-Smith et al., 2002), also referred to as structural modelling (Ford et al., 1989; Glaholt &

Reingold, 2011; Harte & Koele, 1995). These approaches are well equipped to study “the relation between information stimuli (input) and decision responses (outcomes)” (Ford et al., 1989, p. 75), but they do not allow for the exploration of the underlying processes that occur between information input and the final decision outcome (Glaholt & Reingold, 2011). In contrast, process-tracing methods, such as verbal protocols, information boards (search displays), and eye-tracking methodology, “focus more directly on the intervening steps that occur between the introduction of informational inputs and the decision outcomes” (Ford et al., 1989, p. 75).

Technological advances in recent years have greatly lowered the barriers to using eye tracking as a research tool in laboratory and field settings (Wedel, 2015). Not only are the equipment costs on a steady decline, but the data quality and ease of use (e.g., programming, setup, and data analysis) have also improved considerably over the years (Ashby et al., 2016). In an experimental setup, eye-tracking systems allow researchers to record the movements of a participant’s eyes during behavioral processes, thus providing “insights into the cognitive processes underlying a wide variety of human behaviors” (Ashby et al., 2016, p. 96). Researchers most often use eye movements as a proxy for attention that is directed to stimuli. However, eye tracking can also be applied to investigate other psychological constructs, such as arousal, cognitive load, or perceptual fluency.

Although widely employed in other disciplines, such as psychology and marketing (Wedel, 2015; Rayner, 2009), eye tracking is still rarely used in organizational research. The paucity of eye-tracking studies in organizational research is surprising, as other disciplines have broadly applied this methodology in areas of high relevance to organizational research, such as information search and decision-making, learning, training, and expertise. Against this backdrop, Article 3 aims to introduce eye tracking, and thus a new mode of behavioral data, to the field of organizational science.

To reach this objective, Article 3 first provides background information on attention research and eye tracking and then proposes an integrative taxonomy for eye-tracking research based on a synthesis of prior literature. The constituent elements of this taxonomy include drivers of attention, key eye-tracking measures, important psychological constructs, and major areas of application. The proposed taxonomy supports interested scholars by introducing standard eye-tracking vocabulary. The practical classification system for eye-tracking studies further intends to help researchers comprehend the methodological potential of eye tracking and its scope of application. Finally, by posing guiding questions that potential users of eye tracking need to

address, the taxonomy also supports researchers in the process of conducting eye-tracking experiments.

Building on the proposed taxonomy, Article 3 proceeds with a systematic review of eye-tracking studies in leading management journals. The findings of this review indicate that eye-tracking studies are still rare ($n = 15$), although they have been increasingly published in recent years. The review further indicates that technological advances are far from being fully tapped, as researchers predominantly use desktop-based systems in the lab. Furthermore, studies show surprisingly little variation in the psychological constructs under investigation, as all of the identified studies focus on how attention is directed to stimuli. Up to the present, the capability of eye tracking to measure further psychological constructs has thus remained unexploited. Furthermore, eye tracking is predominantly applied in the area of information search and decision-making to address marketing-related research questions.

Keeping the methodological potential and the current use of eye tracking as reflected in leading management journals in mind, the article proceeds with future avenues for eye-tracking research in the substantive domains of organizational science. Illustrative research questions that could be addressed with eye tracking are offered in the domains of entrepreneurship, human resources, and strategic management. Suggested research questions regarding the latter domain particularly relate to investment decision-making. The article further illustrates in more detail exemplary research areas that could benefit from utilizing eye tracking. In this context, Article 3 outlines how eye tracking could be adopted to investigate how investors direct their attention to corporate social responsibility (CSR) and standard financial information in investment decision-making. More specifically, Article 3 exemplifies the benefits of eye tracking by suggesting an extension of the study by Elliott, Jackson, Peecher, and White (2014) titled “The unintended effect of corporate social responsibility performance on investors’ estimates of fundamental value.” The article closes with an outline of key methodological references on which interested researchers are encouraged to draw for methodological guidance and practical advice.

Article 4: The Concept of Values in Socially Responsible Investment Research: A Critical Evaluation and Recommendations for Future Work

By directing attention to the concepts of intuition and reasoning, Article 2 enriches the field of SRI research through the introduction of a novel conceptual perspective. Article 4, in contrast, critically reviews the concept of values in SRI research, thereby putting the emphasis on a con-

cept that is far more widespread in the field of SRI. By reviewing the current state of the literature and offering recommendations for future work, Article 4 provides the foundation for a more rigorous values inquiry and a stronger values-based perspective in SRI research.

Multiple strands of literature have contributed to the understanding of values in the past. Given the dissemination of values research across all social science disciplines, it is no surprise that values have been conceptualized in different ways (Van Quaquebeke et al., 2014). Nevertheless, considerable progress toward conceptual clarity and agreement has been accomplished over the past decades (Agle & Caldwell, 1999; Connor & Becker, 1994; Schwartz, 1994; Schwartz & Bilsky, 1987). As most definitions of values tend to have several key features in common (Rindova & Martins, 2018; Schwartz, 1994; Van Quaquebeke et al., 2014), the concept has been assessed as “fairly well defined” (Agle & Caldwell, 1999, p. 359), and an emergence of “a consensual definition of values” (Connor & Becker, 1994, p. 68) has been asserted. Reviewing and integrating the writings of seminal values theorists (e.g., Allport, 1961; Kluckhohn, 1951; Rokeach, 1973; Williams, 1968), Schwartz (1994, p. 21) provides the following summary definition of values: “desirable transsituational goals, varying in importance, that serve as guiding principles in the life of a person or other social entity.”⁵

Values are widely accepted as one of the most fundamental factors governing human decision-making and behavior (Homer & Kahle, 1988; Meglino & Ravlin, 1998; Rokeach, 1973; Weber, 2015), and they “have long been considered important to explaining action in and around organizations” (Gehman et al., 2013, p. 84). Furthermore, prior research confirms the importance of values in various sustainability-related domains (Fritzsche & Oz, 2007; Gatersleben et al., 2014; Marcus et al., 2015). In this vein, the study of investors’ values could provide crucial insights into the question of why some investors adopt SRI practices while others refrain from doing so.

Despite the prominence of investors’ values in the SRI literature, Article 4 argues that the field’s values perspective is not yet well developed, a situation that impedes scientific progress and the accumulation of a coherent body of knowledge. To substantiate this assertion, the article systematically reviews the conceptualization of values in 47 SRI studies and provides a critique aimed at three major problem areas. First, the analysis indicates that values currently

⁵ Article 4 thus applies the “values-as-criterion” view (Finegan, 2000, p. 150; Williams, 1968, p. 283) as opposed to “the ‘value’ of an object” (Finegan, 2000, p. 150), reflecting the “outcome of an evaluative judgment” (Sánchez-Fernández & Iniesta-Bonillo, 2007, p. 429). Article 4 provides a more detailed elaboration of the two distinctively different meanings of the term “value(s).”

resemble, to a considerable degree, a buzzword rather than a substantive concept. Second, evidence for a substantial lack of conceptual clarity, as well as conceptual confusion, is provided. In the great majority of reviewed articles, the concept of values is neither defined nor investigated on the grounds of established values theories and/or measurement approaches. Furthermore, blurry boundaries between the concepts of values and attitudes are identified. Third, the article detects methodological issues regarding the measurement of values particularly pertaining to the incomprehensive application of values measures.

Overall, Article 4 lends support to the assertion by Hong and Kostovetsky (2009) that the influence of values on investment decision-making is under-researched. As has been pointed out before in the values literature (e.g., Agle & Caldwell, 1999; Connor & Becker, 1994; Dose, 1997; Kirkman et al., 2006; Van der Wal et al., 2006), a lack of conceptual clarity coupled with the constant application of different instruments makes it almost impossible to develop a coherent body of knowledge. As such, Article 4 calls for more coherence in future values research. The article therefore proceeds with recommendations to address the identified conceptual and methodological issues. In this regard, recommendations relate to paying close attention to terminological and conceptual clarity, comprehensively applying established values theories and measurement approaches and considering sources of and remedies for common method biases. To advance the understanding of values as a motive for SRI, the article further specifies future research avenues. Future contributions are likely to come from SRI research adopting a multi-level approach to the study of values, conducting cross-country and longitudinal values studies, demonstrating causality in the value–behavior (SRI) relationship, and exploring the influence of values in the “black box” between information provision and SRI decision-making.

1.5 Contribution

In line with the stated research objective and question, this dissertation contributes to inquiry on the underlying drivers and motives for SRI from a conceptual and methodological viewpoint. A major conceptual contribution of this dissertation is grounded in the introduction of “new” concepts to the field of SRI research. Drawing on the concepts of intuition (System 1) and reasoning (System 2), the effectiveness of climate labeling in the investment context is found to depend critically on investors’ cognitive characteristics: climate labeling is more (less) effective in promoting climate-friendly investing among intuitive (reflective) decision makers. The segmentation of investors has long been of interest in the SRI literature, especially in terms of identifying characteristics that distinguish sustainable from conventional investors. Moving

beyond the criteria commonly used for segmenting investors in the SRI literature, such as attitudes or demographic characteristics, this dissertation contributes by establishing the foundation for novel segmentation analyses according to investors' reliance on intuition and reasoning.

With respect to the concept of values, this dissertation contributes by offering avenues for bringing the "old" concept of values to new light. In this context, specific recommendations are provided to address the identified conceptual and methodological issues of prior SRI research and to foster future research that is more coherent. Respective recommendations relate to paying close attention to terminological and conceptual clarity, comprehensively applying established values theories and measurement approaches, and considering sources of and remedies for common method biases. This dissertation further contributes by specifying future research avenues to advance the understanding of values as a motive for SRI. Promising future avenues pertain to research adopting a multilevel approach to the study of values, conducting cross-country and longitudinal values studies, demonstrating causality in the value-behavior (SRI) relationship, and exploring the influence of values in the "black box" between information provision and SRI decision-making.

From a methodological viewpoint, the present dissertation contributes by expanding the standard toolbox of SRI scholars. First, the dissertation adds a technique from the discipline of sociology, namely, the factorial survey method. In contrast to traditional surveys, factorial surveys collect data on independent and dependent variables through different measures. Factorial surveys further enable a more granular operationalization of concepts (e.g., SRI), thereby reducing item ambiguity among participants. As such, factorial surveys represent a remedy for common method biases associated with traditional surveys. In addition, the method offers the opportunity to provide participants (e.g., investors) with multidimensional scenario descriptions (e.g., SRI products) to manipulate those dimensions systematically (e.g., small/large company) and to determine the importance of dimensions in the decision-making process (e.g., likelihood to invest). Stated differently, factorial surveys are well suited to study the (implicit) interplay of the multiple factors that underlie complex judgments and decision-making processes, such as SRI. By introducing and carving out the potential of factorial surveys, this dissertation contributes by equipping SRI scholars with a methodology to overcome the central deficits of previous, predominantly survey-based research on the drivers and motives for SRI.

Second, the present dissertation provides a methodological contribution that builds on the introduction of eye tracking to the field of organizational research. In this context, an integrative taxonomy is proposed that unravels the methodological potential of eye tracking. The

derived taxonomy also supports interested scholars by posing guiding questions that must be considered when conducting eye-tracking studies. Besides this, the dissertation contributes by specifying illustrative research questions, as well as exemplary research areas that could benefit from utilizing eye tracking, including the domain of SRI. As eye tracking allows the “black box” between informational inputs (e.g., financial/non-financial information) and decision outcomes (e.g., conventional versus SRI) to be opened, the methodology holds great value to understand better the underlying processes in SRI decision-making.

References

- Addison, T., Billon, P. L., & Murshed, S. M. (2001). Finance in conflict and reconstruction. *Journal of International Development*, 13(7), 951-964.
- Agle, B. R., & Caldwell, C. B. (1999). Understanding research on values in business: A level of analysis framework. *Business and Society*, 38(3), 326-387.
- Aguinis, H., & Glavas, A. (2012). What we know and don't know about corporate social responsibility: A review and research agenda. *Journal of Management*, 38(4), 932-968.
- Aiman-Smith, L., Scullen, S. E., & Barr, S. H. (2002). Conducting studies of decision making in organizational contexts: A tutorial for policy-capturing and other regression-based techniques. *Organizational Research Methods*, 5(4), 388-414.
- Allcott, H., & Mullainathan, S. (2010). Behavior and energy policy. *Science*, 327(5970), 1204-1205.
- Allport, G. W. (1961). *Pattern and growth in personality*. New York: Holt, Rinehart & Winston.
- Ashby, N. J. S., Johnson, J. G., Krajbich, I., & Wedel, M. (2016). Applications and Innovations of Eye-movement Research in Judgment and Decision Making. *Journal of Behavioral Decision Making*, 29(2-3), 96-102.
- Athanasopoulou, A., & Selsky, J. W. (2015). The social context of corporate social responsibility: Enriching research with multiple perspectives and multiple levels. *Business & Society*, 54(3), 322-364.
- Auspurg, K., & Jäckle, A. (2017). First equals most important? Order effects in vignette-based measurement. *Sociological Methods & Research*, 46(3), 490-539.
- Bao, J., & Ho, B. (2015). Heterogeneous effects of informational nudges on pro-social behavior. *The BE Journal of Economic Analysis & Policy*, 15(4), 1619-1655.
- Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance*, 29(7), 1751-1767.
- Bello, Z. Y. (2005). Socially responsible investing and portfolio diversification. *Journal of Financial Research*, 28(1), 41-57.
- Brodback, D., Guenster, N., & Mezger, D. (2018). Altruism and egoism in investment decisions. *Review of Financial Economics*, 37(1), 118-148.
- Busch, T., Bauer, R., & Orlitzky, M. (2016). Sustainable development and financial markets: Old paths and new avenues. *Business & Society*, 55(3), 303-329.
- Cadman, T. (2011). Evaluating the governance of responsible investment institutions: An environmental and social perspective. *Journal of Sustainable Finance & Investment*, 1(1), 20-29.

- Campiglio, E. (2016). Beyond carbon pricing: The role of banking and monetary policy in financing the transition to a low-carbon economy. *Ecological Economics*, 121, 220-230.
- Capelle-Blancard, G. and Monjon, S. (2012). Trends in the literature on socially responsible investment: Looking for the keys under the lamppost. *Business Ethics: A European Review*, 21(3), 239-250.
- Cheah, E. T., Jamali, D., Johnson, J. E., & Sung, M. C. (2011). Drivers of corporate social responsibility attitudes: The demography of socially responsible investors. *British Journal of Management*, 22(2), 305-323
- Connor, P. E., & Becker, B. W. (1994). Personal values and management: What do we know and why don't we know more? *Journal of Management Inquiry*, 3(1), 67-73.
- Costa, D. L., & Kahn, M. E. (2013). Energy conservation "nudges" and environmentalist ideology: Evidence from a randomized residential electricity field experiment. *Journal of the European Economic Association*, 11(3), 680-702.
- Covington, H. (2017). Investment consequences of the Paris climate agreement. *Journal of Sustainable Finance & Investment*, 7(1), 54-63.
- Diouf, D. & Hebb, T. (2016). Exploring factors that influence social retail investors' decisions: Evidence from Desjardins fund. *Journal of Business Ethics*, 134(1), 45-67.
- Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology*, 17(2), 245-260.
- Dose, J. (1997). Work values: An integrative framework and illustrative application to organizational socialization. *Journal of Occupational and Organizational Psychology*, 70(3), 219-241.
- Eccles, N. S. and Viviers, S. (2011). The origins and meanings of names describing investment practices that integrate a consideration of ESG issues in the academic literature. *Journal of Business Ethics*, 104(3), 389-402.
- Elliott, W. B., Jackson, K. E., Peecher, M. E., & White, B. J. (2014). The unintended effect of corporate social responsibility performance on investors' estimates of fundamental value. *Accounting Review*, 89(1), 275-302.
- European Commission. (2018). Action Plan: Financing Sustainable Growth. Retrieved from <https://eur-lex.europa.eu/legal-content/ET/TXT/PDF/?uri=CELEX:52018DC0097&from=EN>
- Eurosif. (2016). European SRI study 2016. Brussels: EUROSIF A.I.S.B.L. Retrieved from <https://bit.ly/2guinxX>

- Finegan, J. E. (2000). The impact of person and organizational values on organizational commitment. *Journal of Occupational and Organizational Psychology*, 73(2), 149-169.
- Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. M., & Doherty, M. L. (1989). Process tracing methods: Contributions, problems, and neglected research questions. *Organizational Behavior and Human Decision Processes*, 43(1), 75-117.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4), 25-42.
- Fritzsche, D. J., & Oz, E. (2007). Personal values influence on the ethical dimension of decision making. *Journal of Business Ethics*, 75(4), 335-343.
- Gatersleben, B., Murtagh, N., & Abrahamse, W. (2014). Values, identity and pro-environmental behaviour. *Contemporary Social Science: Journal of the Academy of Social Sciences*, 9(4), 374-392.
- Gehman, J., Trevino, L. K., & Garud, R. (2013). Values work: A process study of the emergence and performance of organizational values practices. *Academy of Management Journal*, 56(1), 84-112.
- Glac, C. (2009). Understanding socially responsible investing: The effect of decision frames and trade-off options. *Journal of Business Ethics*, 87(1), 41-55.
- Glaholt, M. G., & Reingold, E. M. (2011). Eye movement monitoring as a process tracing methodology in decision making research. *Journal of Neuroscience, Psychology, and Economics*, 4(2), 125-146.
- GSIA. (2016). Global sustainable investment review 2016. Retrieved from <http://www.gsi-alliance.org/members-resources/trends-report-2016/>
- Harte, J. M., & Koele, P. (1995). A comparison of different methods for the elicitation of attribute weights: Structural modeling, process tracing, and self-reports. *Organizational Behavior and Human Decision Processes*, 64(1), 49-64.
- HLEG. (2018). Financing a sustainable European economy - Final report. Retrieved from <http://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupDetail-Doc&id=36814&no=1>
- Hoepner, A. G. F., & McMillan, D. G. (2009). Research on 'Responsible Investment': An influential literature analysis comprising a rating, characterisation, categorisation and investigation. *SSRN Electronic Journal*. Retrieved from <https://ssrn.com/abstract=1454793>
- Homer, P. M., & Kahle, L. R. (1988). A structural equation test of the value-attitude-behavior hierarchy. *Journal of Personality and Social Psychology*, 54(4), 638-646.

- Hong, H., & Kacperczyk, M. (2009). The price of sin, the effects of social norms on markets. *Journal of Financial Economics*, 93(1), 15-36
- Hong, H., & Kostovetsky, L. (2012). Red and blue investing: Values and finance. *Journal of Financial Economics*, 103(1), 1-19.
- Hox, J. J., Kreft, I. G. G., & Hermkens, P. L. J. (1991). The analysis of factorial surveys. *Sociological Methods & Research*, 19(4), 493-510.
- IEA. (2012). Energy technology perspectives 2012. International Energy Agency, Paris.
- Jansson, M., & Biel, A. (2011). Motives to engage in sustainable investment: A comparison between institutional and private investors. *Sustainable Development*, 19(2), 135-142.
- Jansson, M., & Biel, A. (2014). Investment institutions' beliefs about and attitudes toward socially responsible investment (SRI): A comparison between SRI and non-SRI management. *Sustainable Development*, 22(1), 33-41.
- Jasso, G. (2006). Factorial survey methods for studying beliefs and judgments. *Sociological Methods & Research*, 34(3), 334-423.
- Johns, G. (1994). How often were you absent? A review of the use of self-reported absence data. *Journal of Applied Psychology*, 79(4), 574-591.
- Junkus, J. C., & Berry, T. C. (2010). The demographic profile of socially responsible investors. *Managerial Finance*, 36(6), 474-481.
- Kirkman, B. L., Lowe, K. B., & Gibson, C. B. (2006). A quarter century of Culture's Consequences: A review of empirical research incorporating Hofstede's cultural values framework. *Journal of International Business Studies*, 37(3), 285-320.
- Kluckhohn, C. (1951). Values and value-orientations in the theory of action: An exploration in definition and classification. In T. Parsons & E. Shils (Eds.), *Toward a general theory of action* (pp. 388-433). Cambridge, MA: Harvard University Press.
- Lehner, M., Mont, O., & Heiskanen, E. (2016). Nudging: A promising tool for sustainable consumption behaviour? *Journal of Cleaner Production*, 134, 166-177.
- Maines, L. A., & McDaniel, L. S. (2000). Effects of comprehensive-income characteristics on nonprofessional investors' judgments: The role of financial-statement presentation format. *The Accounting Review*, 75(2), 179-207.
- Marcus, J., MacDonald, H. A., & Sulsky, L. M. (2015). Do personal values influence the propensity for sustainability actions? A policy-capturing study. *Journal of Business Ethics*, 127(2), 459-478.
- McKinsey&Co. (2010). Impact of the financial crisis on carbon economics: Version 2.1 of the global greenhouse gas abatement cost curve. McKinsey & Company.

- McLachlan, J., & Gardner, J. (2004). A comparison of socially responsible and conventional investors. *Journal of Business Ethics*, 52(1), 11-25.
- Meglino, B. M., & Ravlin, E. C. (1998). Individual values in organizations: Concepts, controversies, and research. *Journal of Management*, 24(3), 351-389.
- Momsen, K., & Stoerk, T. (2014). From intention to action: Can nudges help consumers to choose renewable energy? *Energy Policy*, 74, 376-382.
- Nilsson, J. (2008). Investment with a conscience: Examining the impact of pro-social attitudes and perceived financial performance on socially investment behavior. *Journal of Business Ethics*, 83(2), 307-325.
- Nilsson, J. (2009). Segmenting socially responsible mutual fund investors: The influence of financial return and social responsibility. *International Journal of Bank Marketing*, 27(1), 5-31.
- Pilaj, H. (2017). The choice architecture of sustainable and responsible investment: Nudging investors toward ethical decision-making. *Journal of Business Ethics*, 140(4), 743-753.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879.
- Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2016). Recommendations for creating better concept definitions in the organizational, behavioral, and social sciences. *Organizational Research Methods*, 19(2), 159-203.
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, 62(8), 1457-1506.
- Reimsbach, D., Hahn, R., & Gürtürk, A. (2018). Integrated reporting and assurance of sustainability information: An experimental study on professional investors' information processing. *European Accounting Review*, 27(3), 559-581.
- Renneboog, L., Ter Horst, J., & Zhang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance*, 32(9), 1723-1742.
- Rindova, V. P., & Martins, L. L. (2018). From values to value: Value rationality and the creation of great strategies. *Strategy Science*, 3(1), 323-334.
- Rokeach, M. (1973). *The Nature of Human Values*. New York: Free Press.
- Rosen, B. N., Sandler, D. M., Shani, D. (1991). Social issues and socially responsible investment behavior: A preliminary empirical investigation. *The Journal of Consumer Affairs*, 25(2), 221-234.

- Rossi, P. H. (1979). Vignette analysis: Uncovering the normative structure of complex judgments. In R. K. Merton, J. S. Coleman & P. H. Rossi (Eds.), *Qualitative and quantitative social research: Papers in honor of Paul F. Lazarsfeld* (pp. 176-186). New York, NY: Free Press.
- Rossi, P. H., & Anderson, A. B. (1982). The factorial survey approach: An introduction. In P. H. Rossi & S. L. Nock (Eds.), *Measuring social judgments: The factorial survey approach* (pp. 15-67). Beverly Hills, CA: Sage.
- Sánchez-Fernández, R., & Iniesta-Bonillo, M. Á. (2007). The concept of perceived value: A systematic review of the research. *Marketing Theory*, 7(4), 427-451.
- Sarokin, D., & Schulkin, J. (1991). Environmental concerns and the business of banking. *Journal of Commercial Bank Lending*, 74(5), 6-19.
- Schmitt, N. (1994). Method bias: The importance of theory and measurement. *Journal of Organizational Behavior*, 15(5), 393-398.
- Scholtens, B. (2006). Finance as a driver of corporate social responsibility. *Journal of Business Ethics*, 68(1), 19-33.
- Scholtens, B. (2009). Corporate social responsibility in the international banking industry. *Journal of Business Ethics*, 86(2), 159-175.
- Scholtens, B. and Sievänen, R. (2013). Drivers of socially responsible investing: A case study of four Nordic countries. *Journal of Business Ethics*, 115(3), 605-616.
- Schueth, S. (2003). Socially responsible investing in the United States. *Journal of Business Ethics*, 43(3), 189-194.
- Schwartz, S. H. (1994). Are there universal aspects in the structure and content of human values? *Journal of Social Issues*, 50(4), 19-45.
- Schwartz, S. H., & Bilsky, W. (1987). Toward a universal psychological structure of human values. *Journal of Personality and Social Psychology*, 53(3), 550-562.
- Schwartz, M. S., & Carroll, A. B. (2008). Integrating and unifying competing and complementary frameworks: The search for a common core in the business and society field. *Business & Society*, 47(2), 148-186.
- Shooter, W., & Galloway, S. (2010). The use of factorial surveys in leisure research. *Journal of Leisure Research*, 42(4), 641-652.
- Thaler, R. H., & Sunstein, C. R. (2003). Libertarian paternalism. *American Economic Review*, 93(2), 175-179.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.

- Thompson, P., & Cowton, C. J. (2004). Bringing the environment into bank lending: Implications for environmental reporting. *The British Accounting Review*, 36(2), 197-218.
- Tippet, J., & Leung, P. (2001). Defining ethical investment and its demography in Australia. *Australian Accounting Review*, 11(3), 44-55.
- UNEP FI. (2009). Financing global deal on climate change. United Nations Environment Programme, Geneva.
- US SIF. (2018). Report on US sustainable, responsible and impact investing trends 2018. US SIF, Washington, D.C.
- Van Bavel, R., Herrmann, B., Esposito, G., & Proestakis, A. (2013). Applying behavioural sciences to EU policy-making. Retrieved from <http://ftp.jrc.es/EURdoc/JRC83284.pdf>
- Van der Wal, Z., Huberts, L., Van den Heuvel, H., & Kolthoff, E. (2006). Central values of government and business: Differences, similarities and conflicts. *Public Administration Quarterly*, 30(3), 313-364.
- Van Quaquebeke, N., Graf, M. M., Kerschreiter, R., Schuh, S. C., & van Dick, R. (2014). Ideal values and counter-ideal values as two distinct forces: Exploring a gap in organizational value research. *International Journal of Management Reviews*, 16(2), 211-225.
- Von Wallis, M., & Klein, C. (2015). Ethical requirement and financial interest: A literature review on socially responsible investing. *Business Research*, 8(1), 61-98.
- Wallander, L. (2009). 25 years of factorial surveys in sociology: A review. *Social Science Research*, 38(3), 505-520.
- Weber, J. (2015). Identifying and assessing managerial value orientations: A cross-generational replication study of key organizational decision-makers' values. *Journal of Business Ethics*, 132(3), 493-504.
- Wedel, M. (2015). Attention research in marketing: A review of eye tracking studies. In J. Fawcett, E. F. Risko, & A. Kingstone (Eds.), *The handbook of attention* (pp. 569-588). Cambridge, MA: MIT Press.
- WEF. (2013). The green investment report. World Economic Forum, Geneva.
- Wiek, A., & Weber, O. (2014). Sustainability challenges and the ambivalent role of the financial sector. *Journal of Sustainable Finance & Investment*, 4(1), 9-20.
- Williams, R. M., Jr. (1968). Values. In E. Sills (Ed.), *International encyclopedia of the social sciences* (pp. 283-287). New York: Macmillan.
- Williams, G. (2007). Some determinants of the socially responsible investment decision: A cross-country study. *The Journal of Behavioral Finance*, 8(1), 43-57.

- Woods, C., & Urwin, R. (2010). Putting sustainable investing into practice: A governance framework for pension funds. *Journal of Business Ethics*, 92(1), 1-19.
- World Bank. (2015). World development report 2015: Mind, society, and behavior. Retrieved from <http://www.worldbank.org/en/publication/wdr2015>

2 First Article

Oll, J., Hahn, R., Reimsbach, D. & Kotzian, P. (2018). Tackling Complexity in Business and Society Research: The Methodological and Thematic Potential of Factorial Surveys. *Business & Society*, 57(1), 26-59. [VHB-JOURQUAL: B]

Tackling Complexity in Business & Society Research: The Methodological and Thematic Potential of Factorial Surveys

Abstract

Factorial surveys (FSs) integrate elements of survey research and classical experiments. Using a large number of respondents in a controlled setting, FSs approximate complex and realistic judgment situations through so-called vignettes—that is, carefully designed descriptions of hypothetical people, social situations, or scenarios. Despite being rooted, and predominantly applied, in sociology, FSs are particularly promising for business and society (B&S) scholars. Given the multiplicity, inherent complexity, and sometimes fuzziness of B&S research objects, conventional research methods inevitably reach their limits. This article, therefore, systematically presents methodological and thematic opportunities for FS studies in B&S research. It is argued that FSs are well suited to dealing with the complex interplay of societal-, organizational-, and individual-level factors in B&S research and to studying the principles underlying human perceptions, attitudes, values, social norms, and (anticipated) behavior. The application of the FS method is illustrated based on a showcase example in the realm of socially responsible investments (SRIs). As the literature on the conceptualization of FSs is limited, methodological challenges are addressed to guide B&S researchers past the common methodological pitfalls.

Keywords: Factorial surveys, vignette study, experiment, research methods, multilevel research, socially responsible investments

Introduction

Understanding coherences and generating knowledge in the domain of business and society (B&S) research often proves to be challenging because, unlike many other areas, this field is characterized by reciprocal dependencies and interactions of societal-, organizational-, and individual-level factors (Aguinis & Glavas, 2012; Athanasopoulou & Selsky, 2015). B&S issues also frequently involve judgments for which the underlying factors, as well as their interplay, are hard to make explicit (e.g., what is responsible or moral behavior?). Furthermore, B&S scholars inevitably deal with fuzzy and difficult-to-define concepts, which are subject to a multiplicity of interpretations (Schwartz & Carroll, 2008). Given the above-described nature of B&S research, scholars are challenged to take this complexity into account in regard to the choice and design of methods. In this context, and drawing on the discipline of sociology, this article argues that an interesting and promising role can be assigned to the so-called factorial survey (FS) method, which is also referred to as vignette analysis/study (Dülmer, 2007; Rossi & Anderson, 1982).

As a hybrid approach, FSs integrate characteristics of survey research and experimental designs into a single method (Hox, Kreft, & Hermkens, 1991; Wallander, 2009), thereby allowing researchers to collect data from a large number of respondents in a controlled setting. Therefore, FSs are interesting for B&S researchers, who frequently need to collect primary data. FS participants receive a series of vignettes—that is, carefully designed descriptions of hypothetical people, social situations, or scenarios—and are asked to make judgments about the presented fictive descriptions. The vignettes that are employed in FSs contain a large number of factors (dimensions), which are randomly varied in their valuesⁱ (Killick & Taylor, 2012; Lauder, 2002). In this context, the manipulated dimensions can be understood as “precise references to what are thought to be the most important factors in the decision-making or judgment-making process of the respondents” (Alexander & Becker, 1978, p. 94).

This article argues that FSs are well positioned to deal with the complex nature of B&S research because individual judgments and evaluation processes can be directly related to the societal and organizational levels through a single method. Aguinis and Glavas (2012) argue that “for future research to be most informative, it will require the inclusion of variables from more than one level of analysis” (p. 957). FSs allow for such an inclusion and help researchers to simultaneously address two major challenges of B&S research: studying the underlying principles behind human perceptions, attitudes, values, social norms, and

(anticipated) behavior (Aguinis & Glavas, 2012) and integrating this focus with further perspectives to arrive at a multilevel approach (Athanasopoulou & Selsky, 2015) that mirrors the complexity of real-life situations. While, multilevel issues can and have been studied through other means such as traditional survey-based research, the FS method stands out by providing researchers with the opportunity to unveil the implicit interplay of (multilevel) factors that underlie and drive complex judgments. This article contributes to the field of B&S research by illustrating the methodological and thematic potential of FSs and how the implicit can be made explicit.

The article is structured as follows. The next section introduces the FS technique and makes the case for its broader application in B&S research. Furthermore, a showcase example of research on socially responsible investment (SRI) is introduced to guide the argumentation and illustration in the subsequent section. The following section outlines in detail the implementation and data analysis process of FSs and provides concrete research recommendations. This includes an overview of the potential advantages and disadvantages of the method, existing methodological references, and best practices. The article ends with a brief conclusion.

The Case for FSs in B&S Research

The origin of the FS method can be traced back to sociologist Peter Henry Rossi in the 1970s, and up to the present day, FSs have been applied in sociological research to study a wide thematic spectrum (Wallander, 2009). In addition to sociology, the method has slowly been implemented in a variety of other contexts (see, e.g., Lauder, 2002), including business studies (e.g., business ethics, leadership, and career development; see Ashill & Yavas, 2006). Although FSs are not (yet) well established in the field of business studies (Martin, 2012), the authors see good reason to argue that the FS approach will provide new opportunities, especially for B&S research, due to the domain's inherent complexity. This section first characterizes and delineates the FS method from related approaches before highlighting its potential in B&S research. Finally, a showcase is sketched to guide the further illustration of FSs in B&S research throughout this article.

The Characterization and Delineation of the FS Method

The FS method applies the basic principles of experimental designs in survey research (Dülmer, 2007; Hox et al., 1991; Lauder, 2002; Taylor, 2006; Wallander, 2009). On the one hand, FSs contain some of the central elements of experiments because causal relationships

can be identified by systematically manipulating dimensions (variables) that potentially have an effect on human judgment. On the other hand, FSs display survey characteristics because respondent-specific information is collected through a traditional survey, researchers control respondent sampling, and a large sample of respondents can be questioned.

The central component of any FS is the vignette, which is a carefully designed description of hypothetical people, social situations, or scenarios that is presented to and judged by respondents (Dülmer, 2007; Wallander, 2009). The upper part of Table 1 provides an exemplary vignette, followed by the corresponding rating task for participants. As outlined in more detail later, the showcase FS was created to address the following topic: investment decisions are driven by the personal characteristics of the investor, the societal background, and the features of the investment object. Thus, our showcase research question is as follows: which of these factors are most relevant to the investment decisions in the context of socially responsible investments? The showcase vignette exemplifies how different dimensions at the external society and organizational level (illustrated in italics in Table 1) can be manipulated using two (or more) values and how they fit into the proposed multilevel approach.

As illustrated in the example, vignettes provide respondents with complex information through the inclusion of a wide range of dimensions (independent variables) (Killick & Taylor, 2012; Ludwick et al., 2004; Taylor, 2006). These dimensions are randomly varied in their values amongst the vignettes.ⁱⁱ The inclusion of numerous dimensions makes it possible to capture the inherent complexity of human evaluation processes and real-life judgment-making situations (Killick & Taylor, 2012; Ludwick et al., 2004; Taylor, 2006). This may lead to a large number of vignette versions, which are also referred to as a “vignette population” or “vignette universe” (Alexander & Becker, 1978). The size of the total vignette population is calculated by multiplying all factor values by each other (Cartesian product) (Wallander, 2009).ⁱⁱⁱ The vignette universe of the exemplary design described in Table 1 thus results in 256 (2^8) different vignettes. With a large vignette universe, participants usually rate a reduced sample from the entire vignette population (Dülmer, 2007; Taylor, 2006; Wallander, 2009). The data on the respondents’ judgments generate the dependent (outcome) variable(s) (Jasso, 2006; Lauder, 2002). In Table 1, this is illustrated through the rating task—asking participants to express on an eleven-point Likert scale how likely they would be to invest in an opportunity described by the vignette.

Table 1: Exemplary FS design for the showcase of SRI^{iv}

Example of a generic FS content	
Manipulations at external level	<p>You currently investigate the following potential investment opportunity: The government of your home country directs [<i>low/high</i>] attention to issues of socially responsible investment. The company you are looking at has its main operations in [<i>a developing country with weak/an industrialized country with strong</i>] social and environmental regulations.</p>
Manipulations at organizational level	<p>It is a [<i>small (revenues \$100 m.) / large (revenues \$10 bn.)</i>] company operating in an industry which is known for being [<i>delicate/unproblematic</i>] with regard to environmental and social issues. Its financial performance in the past few years was [<i>superior/inferior</i>] compared to its peer group. According to widely accepted studies, the company had a [<i>superior/inferior</i>] sustainability-related performance compared to the industry average in the past few years. In the last five years, the company published [<i>a brief sustainability overview based on its own parameters/a comprehensive sustainability report following the guidelines of the Global Reporting Initiative</i>]. This report was [<i>not externally assured/externally assured by a third-party</i>].</p>
Rating task	<p>In your opinion, how likely is it that you undertake this investment? [Scale: (1) very unlikely to (11) very likely]</p>
Measures at individual level	<p>Investor type [professional investor/nonprofessional investor]; Age [years]; Sex [m/f]; Work experience [years]; Income [e.g., in \$]</p> <p>I would describe myself as religious. [Scale: (1) strongly disagree to (7) strongly agree]</p> <p>It is important to me to be rich. I want to have a lot of money and expensive things. [Scale: (1) strongly disagree to (7) strongly agree]</p> <p>I strongly believe that people should care for nature. Looking after the environment is important to me. [Scale: (1) strongly disagree to (7) strongly agree]</p> <p>Conventional investments perform much better than SRI. [Scale: (1) strongly disagree to (7) strongly agree]</p> <p>SRI is associated with a lower risk compared to conventional investments. [Scale: (1) strongly disagree to (7) strongly agree]</p>

After completing the rating task, the FS concludes by collecting respondent-specific information that is analogous to a traditional survey, as illustrated by the exemplary questions in the lower part of Table 1. By including respondent-specific data, FSs also capture the individual level and thus go beyond the study of the effects of vignette dimensions. More precisely, as Martin (2012) summarizes, the FS method enables scholars to investigate “(a) the elements of information used to form judgments, (b) the weight of each of these factors, and (c) how different subgroups of the respondents agree on (a) and (b)” (p. 525).

FSs need to be distinguished from related methods, such as conjoint analysis, for which the primary field of application is the area of marketing (Green, Krieger, & Wind, 2001; Teichert & Shehu, 2010). The FS method has its roots in sociological research (Rossi, 1979), whereas conjoint analysis originated in mathematical psychology (Luce & Tukey, 1964). Due to their distinct origins, both methods differ in regard to their focus. Conjoint analysis is predominantly employed “to assess the utility of product features and willingness to pay for these” (Auspurg & Jackle, 2015, p. 9). In contrast, FSs have a “broad social and psychological focus” beyond marketing strategies and consumer preferences (Shooter & Galloway, 2010, p. 644). Furthermore, operational differences exist regarding the presentation of information and the response task. While the participants in conjoint analyses usually compare and rank different alternatives (profile cards), FS respondents evaluate realistic scenarios (vignettes) sequentially on a rating scale (Auspurg & Jackle, 2015; Furman, Shooter, & Schumann, 2010; Shooter & Galloway, 2010).

Employing FSs for multilevel B&S Issues

FSs open up a promising research avenue for B&S scholars to deal with the multiplicity and inherent complexity of B&S research objects. B&S scholars inevitably confront a high degree of complexity, given the multilevel nature of the research objects in this field. More specifically, the following three central levels are commonly mentioned in the B&S literature: the individual, the organizational, and the institutional/external social level (e.g., Aguinis & Glavas, 2012; Athanasopoulou & Selsky, 2015; Wood, 1991, 2010). As the central element of interest, organizations are inseparably connected to the upper (institutional) and lower (individual) levels of inquiry. Consequently, Wood (1991) observes “that business and society are interwoven rather than distinct entities; therefore, society has certain expectations for appropriate business behavior and outcomes” (p. 695). Moreover, there is a close connection between institutional and organizational issues and the individual level. For example, corporate social responsibility managers have to cope with organizational rules, policies, and culture, as well as external and often conflicting stakeholder expectations, so that the individual level of behavior is nested within the higher-order organizational and institutional levels (e.g., Lülfs & Hahn, 2013, 2014).

Complex multilevel issues are, thus, at the very heart of B&S research, and scholars have acknowledged, for example, that corporate social responsibility “cannot be studied effectively within conceptual silos” (Athanasopoulou & Selsky, 2015, p. 351). Nevertheless, many B&S researchers still tend to conduct single-level analyses that are grounded on single-level

theoretical perspectives (Aguinis & Glavas, 2012; Athanasopoulou & Selsky, 2015). Apart from the often missing connection between different levels in B&S research, Aguinis and Glavas (2012) emphasize that organizations' responsibilities and corresponding activities are "primarily studied at the macro level (i.e., institutional or organizational level) compared to the micro level (i.e., individual level)" (p. 933). The latter level of analysis is "virtually absent" (p. 943), and as a result, the contemporary understanding of the foundations of the role and activities of business in society is still poorly developed. To compound matters further, B&S research at the organizational and institutional levels is usually based on aggregate data, which are unsuitable for providing insights into the individual level. To gain a deeper understanding of the foundations, novel methodological approaches are, therefore, imperative (Aguinis & Glavas, 2012).

Against the background of these two shortcomings (i.e., the missing connection between the levels and the underdeveloped research at the individual level), FSs can be considered a promising method for tackling complex and multilevel B&S issues and for closing the prevalent knowledge gaps regarding the aforementioned individual underpinnings. First, individuals judge vignettes in FSs; that is, FSs allow for the collection and analysis of primary data at the under-researched individual level of B&S inquiry. Second, FS vignettes contain a large number of dimensions that provide researchers with opportunities to simultaneously integrate and manipulate the relevant factors at the organizational and institutional levels. In other words, FSs contribute toward satisfying the need for multilevel approaches in B&S research as they provide scholars with opportunities to identify the factors that flow into a judgment and to reveal the relevance of each factor. The FS method is thus accompanied by the promise of unraveling complex evaluative judgments, which are grounded on a multitude of different factors that the respondents might not even be able to make explicit.^v

However, the complexity of B&S issues not only arises from the interplay between multilevel factors but also emerges from the fuzzy, ambiguous, and contested nature of B&S concepts (Frederick, 1994; Jones, 1983; Wood, 2010). Schwartz and Carroll (2008) make this fuzziness evident by illustrating that the well-established concepts of B&S inquiry, such as business ethics, stakeholder management, sustainability, and corporate citizenship, are not mutually exclusive but rather overlap. Against this background, the question becomes how FSs might add value to fuzzy and difficult-to-articulate concepts. By manipulating specific aspects of broader concepts, such as sustainability, FSs allow researchers to test which sub-aspects are empirically relevant from the perspective of individuals.

Many other B&S topics relate to such complicated judgments regarding concepts that are difficult to define and articulate. In this context, several of Martin and colleagues' (Glac, Elm, Martin, & Painter-Morland, 2014; Martin, 2012, 2015; Martin & Shilton, 2015) recent FS studies on the concept of privacy represent excellent examples illustrating the potential of this method. For instance, Martin (2012) studies the determinants of privacy judgments and finds that privacy is considerably more context dependent and nuanced than was theorized beforehand. This finding can be directly related to the choice of research methods. Instead of employing a static privacy definition, Martin (2012) conducts an FS, which allows for "identifying the privacy factors and their relative importance—the privacy norms—that respondents take into consideration in making a judgment about privacy within particular communities" (Martin, 2012, p. 520). The FS method, therefore, enabled her to shed light on the determinants of privacy judgments without having to define privacy *a priori*.

However, with a few exceptions, the value of the FS method for analyzing fuzzy concepts remains largely untapped in B&S research. To further illustrate the potential of FSs, a study that could have benefitted from using the FS method is thus presented and discussed. Sievänen, Rita, and Scholtens (2013) address the question of what drives responsible investment in pension funds. They survey key financial decision makers at pension funds regarding their funds' characteristics (e.g., legal origin, ownership, and fund size). As a dependent variable, they use a Likert scale to measure whether a fund engages in responsible investment. Depending on the investors' answers, a fund is classified as conventional, neutral, or responsible, and this classification is based on the investors' self-reporting. The investors' views on what responsible investment actually means and encompasses are not considered. However, responsible investment is a contested concept, as there is no consensus—neither about a unifying definition nor about its constitutive characteristics and elements (Cheah, Jamali, Johnson, & Sung, 2011; Eurosif, 2012). In other words, while Sievänen, Rita, and Scholtens (2013) seek to study the drivers of responsible investment, the concept remains ill-defined and is, thus, ambiguous to the study's participants.

An FS could, thus, be an alternative method of addressing the research question posed by Sievänen, Rita, and Scholtens (2013). Several vignettes capturing the controversial debate on the constitutive elements of responsible investment (e.g., investment strategy, processes used, societal outcomes, etc.) could be presented to financial decision makers with a respective rating task, such as the following: To what degree do you perceive the present pension fund as responsible? In conjunction with the survey on fund characteristics, valuable insights could be garnered from uncovering the interdependencies between the drivers of responsible

investment and pension funds' concepts of responsible investment. This general line of reasoning regarding the study by Sievänen, Rita, and Scholtens (2013) can be transferred and applied to other B&S studies that (a) draw on and expose respondents to a fuzzy concept and (b) do not provide respondents with the opportunity to articulate their conceptual understanding.^{vi}

Socially Responsible Investment: A Showcase for FSs

To illustrate the potential of FSs in the B&S domain, the showcase FS in the realm of SRI has already been introduced in Table 1. SRI is a field that is receiving growing attention in B&S research (see, e.g., Renneboog, Ter Horst, & Zhang, 2008; Wallis & Klein, 2014), and it illustrates the need to connect different levels of inquiry and fill knowledge gaps in the individual underpinnings of behavior. Given the steady societal change toward embracing sustainability as a universal development goal, individuals presumably consider environmental and social perspectives along with financial goals and, thus, combine financial and nonfinancial (performance) criteria in their investment decisions (e.g., Berthelot, Coulmont, & Serret, 2012; Dhaliwal, Radhakrishnan, Tsang, & Yang, 2012). Consequently, companies publicly disclose information on their nonfinancial performance in their corporate sustainability reports. Research on SRI and on sustainability reporting is, thus, inseparably intertwined and, accordingly, our showcase vignette incorporates aspects related to the reporting of sustainability information. Interestingly, human perceptions, attitudes, values, social norms, and (anticipated) behavior with regard to the form, content, or quality of sustainability reporting are still under-examined in the extant research (Hahn & Kühnen, 2013). This also applies to the utilization of disclosed information by investors in the realm of SRI, and little is known about the underlying determinants and motives at the individual level of decision making (Glac, 2009; McLachlan & Gardner, 2004). Therefore, FSs can help pave the way toward integrating the individual's perspective with the organizational and societal views of SRI. The vignette presented in Table 1 illustrates how the different dimensions of an FS can be manipulated using two (or more) values and how they fit into the proposed multilevel approach, as will be further illustrated in the following sections.

Methodological Issues for FSs in B&S Research

This section first illustrates the implementation process of FSs along six main steps, and it provides research recommendations in keeping with the above-introduced showcase FS.

Then, the potential advantages and disadvantages of the method are highlighted, and further literature recommendations are given.

The Implementation Process of FSs and Research Recommendations

The process of planning and implementing FSs and analyzing the generated data can be divided into six central steps, as illustrated by the following with reference to the showcase as a concrete example guiding this overview.

Constructing the research question. Research questions generally focus “on vignette factors influencing respondents’ judgments and on respondent characteristics explaining judgment differences between groups of respondents” (Atzmüller & Steiner, 2010, p. 133). They are, therefore, generally presented in the following form: “What is the impact of <factors x, y, z> on decisions by <type of respondents> regarding <focus of decision>?” (Taylor, 2006, p. 1194). To ensure the suitability of the research questions for the FS approach, B&S researchers need to be aware that FSs are based on the assumption of socially and individually structured judgments (Rossi & Anderson, 1982). The social component of judgments implies that there is a certain degree of agreement among respondents “as to the (combination of) factors that it is important to take into consideration when making a specific judgement” (Wallander, 2009, p. 514). The individual component of judgments, in contrast, refers to the assumption that “each individual tends towards consistency in his or her own judgements” (Rossi & Anderson, 1982, p. 17). The showcase-related research question regarding which factors (at the individual, organizational, and societal levels) are most relevant to the investment decisions in the context of SRI explicitly addresses both judgment components.

Constructing the vignette universe. Researchers need to construct the whole vignette universe by determining the “relevant factors and appropriate factor values for the research hypothesis under investigation” (Atzmüller & Steiner, 2010, p. 130). This is crucial because the omission of relevant dimensions reduces the degree of explained variance and, thus, the explanatory power of an FS study (Taylor, 2006). The vignette design also includes the development of the rating task and the respective rating scale. The authors recommend that B&S scholars follow a holistic approach by using relevant theory, extra-theoretical reasoning, prior research, conventional wisdom, practice knowledge, and their own professional judgment to avoid the omission of relevant dimensions (Atzmüller & Steiner, 2010; Jasso, 2006; Taylor, 2006).

When looking at the showcase example, prior research on SRI demonstrates the relevance of the different factors and factor values proposed in the presented vignette (see again Table 1). At the external society level, prior research most notably addresses a company's country of origin and legal requirements (Hahn & Kühnen, 2013; Scholtens & Sievänen, 2013), with archival studies often including dummy variables for companies that are headquartered in countries with a high and in a low social and environmental "conscience" (e.g., Golob & Bartlett, 2007). Furthermore, regulative and legislative developments can be considered a particularly crucial determinant (e.g., Eurosif, 2014; Sievänen et al., 2013), and the effects of mandatory versus voluntary reporting spheres have also been frequently discussed (e.g., Alciatore, Dee, & Easton, 2004; Frost, 2007). The showcase vignette includes the country effects, focusing on the country-specific attention to SRI issues and on country-level social and environmental regulations.

The organizational level is the most intensely researched in the current literature on SRI and sustainability reporting. Among other variables, a company's size, its financial and sustainability performance, and its sector affiliation all stand out. Larger firms are more likely to adopt sustainability reporting and to report extensively on their respective performance (e.g., Fortanier, Kolk, & Pinkse, 2011; Gallo & Christensen, 2011). More profitable firms seem to exhibit different reporting behavior than less profitable firms due to the increased ability and flexibility of a company to bear the costs of sustainability reporting or to cope with the consequences of disclosing potentially damaging information (e.g., Prado-Lorenzo, Rodríguez-Domínguez, Gallego-Álvarez, & García-Sánchez, 2009; Stanny & Ely, 2008). Furthermore, "good" and "poor" sustainability performers seem to differ in their reporting behavior, with good performers being more likely to adopt and extend sustainability reporting (e.g., Belal & Cooper, 2011; Nikolaeva & Bicho, 2011). A company's sector affiliation is typically included in archival studies as a dummy variable differentiating between companies from industries with high and low social and environmental impacts (e.g., Parsa & Kouhy, 2008; Sotorrió & Sánchez, 2010). Finally, the presence or absence of the costly signal of the voluntary external assurance of the reported information is another organizational variable that is linked to the reporting sphere (e.g., Kolk & Perego, 2009; Simnett, Vanstraelen, & Chua, 2009). The vignette mirrors all these aspects and includes manipulations of company size, industry, financial and sustainability performance, and the disclosure and assurance of a sustainability report. Moreover, researchers have to think about what *not* to include in the vignette universe. Depending on the specific research context, for example, the simultaneous

manipulation at all three levels (societal, organizational, and individual) might not always be feasible.

When entering unknown territory, pre-testing might be required for each aspect of the vignette before an FS study can be conducted (for an overview see, e.g., Auspurg & Hinz, 2015). In a first step, general aspects such as comprehensibility, realism, and clarity of the situation described in the vignette should be evaluated through qualitative inquiries. Then, the pre-test quantitatively assesses specific aspects of the situation described in the vignette. The showcase vignette, for example, describes company “operating in an industry which is known for being [delicate/unproblematic] with regard to environmental and social issues”. Here, the pre-test should reveal whether respondents care about this feature and adapt their judgment. If participants substantially differ in what they perceive to be delicate or unproblematic, the description should be more specific. Finally, the vignette has to be pre-tested as a whole. Regarding the dependent variable, in our example the investment decision, the pre-test should analyze whether there is enough variation, that is, not all respondents prefer one investment to the other.

The construction of the vignette population also requires researchers to make informed decisions with respect to the number of included dimensions and factor values. The presented showcase was restricted to eight dimensions with two values each. Adding further dimensions and providing a broader spectrum of factor values could help to better mirror the organizational complexity surrounding professional investors (Juravle & Lewis, 2008, 2009). However, the desire to provide realistic scenario descriptions via the inclusion of many dimensions can easily clash with respondents’ limited processing capacity; furthermore, lengthy and complex scenarios lead to the risk of boredom and fatigue effects (Auspurg, Hinz, & Liebig, 2009; Wallander, 2009). Unfortunately, no definite rule for the appropriate number of dimensions exists, and the few empirical studies on this matter provide rather inconclusive findings (e.g. Auspurg, Hinz, & Liebig, 2009; Auspurg et al., 2009; Sauer, Auspurg, Hinz, & Liebig, 2011). Reference is, therefore, commonly made to cognitive psychological findings suggesting that humans’ short-term memory capacity is limited to 7 (± 2) elements (Auspurg et al., 2009; Auspurg, Hinz, Liebig, & Sauer, 2010).

Within the vignette design phase, researchers also have to decide on the number and variation of factor values. The number of factors and their values matter in FSs because they determine the size of the vignette population. While Beck and Opp (2001) state that the number of factor values should be kept as low as possible to simplify subsequent data

analysis, they also stress that no definite rule exists. In the presented showcase, the number of factor values was held constant among all vignette factors because variations in the number of factor values might trigger investors to react more strongly to vignette dimensions with a wider value differential (Auspurg et al., 2010; Highhouse, Luong, & Sarkar-Barney, 1999; Wilde, Cooke, & Janiszewski, 2008). B&S researchers are, therefore, advised to keep factor value variation to a minimum.

The vignette design stage further includes the development of the rating task and the respective rating scale. The respondents can be asked to make one or several judgments per vignette. While the illustrative showcase included only one rating task, additional tasks could be integrated. An example of another question of interest could be “To what extent do you perceive the investment opportunity as socially responsible?” Adding this question would allow for the identification of those vignette features that are perceived by investors as constitutive elements and characteristics of SRI. Besides differentiating between features with and without relevance to SRI, further light could be shed on the level of conceptual (dis)agreement among different types of investors. Ludwick et al. (2004) recommend using no more than three rating tasks per vignette to help ensure that respondents do not lose focus and attention.

Choosing individual-level features. In the showcase, the individual level is measured rather than manipulated (see again Table 1) and, thus, does not extend the vignette universe. The reason for this is that, in the showcase, the variables related to the individual level are certain personal characteristics of decision makers (i.e., the participants of the study) and, thus, cannot be manipulated.^{vii} These (non-manipulated) variables are organized around the interdependencies between three main issues: financial aspects, nonfinancial aspects, and investors’ sociodemographics (Nilsson, 2008). In terms of sociodemographics, studies frequently reveal differences between SRI and conventional investors (e.g., Rosen, Sandler, & Shani, 1991; Tippet & Leung, 2001). The nonfinancial dimension of SRI is usually addressed by studying investors’ values, attitudes, religiosity, or ethics, while the financial dimension mainly captures issues such as investors’ beliefs about financial performance and risk (e.g., Iyer & Kashyap, 2009; Jansson & Biel, 2011; 2014; Lenssen et al., 2014; Pasewark & Riley, 2010). In terms of individual values, the showcase FS draws on the Schwartz (1992, 1994) theory of basic human values but only covers two exemplary questions (items) along the self-transcendence versus self-enhancement values dimension. Beyond illustrative purposes, B&S scholars are advised to make use of standardized value measurement scales (see, e.g.,

Cieciuch & Davidov, 2012) and follow a comprehensive approach, as SRI can be driven by more than one value priority (Lönqvist, Verkasalo, Wichardt, & Walkowitz, 2013).

Making the sampling decision. With regard to the sampling decisions, both the type and number of respondents have to be considered. A widely discussed issue regarding the type of respondents is the use of student samples (e.g., Wallander, 2009; Wason, Polonsky, & Hyman, 2002). Due to validity concerns, B&S scholars are urged to make use of student samples with care. For example, finance students might be used to approximate professional investors' decision making. When looking at sample size, a lack of guidance can be identified in the literature, and this might, at least to some extent, explain why sample size decisions are rarely justified in FS studies (Lauder, 2002). Furthermore, the scholars who have addressed this issue tend to restrict themselves to the formulation of rather imprecise statements. Beck and Opp (2001), for example, state that the minimum sample size depends, among other factors, on the number of dimensions and values and that sample size decisions are difficult to make. Other researchers claim that FSs require fewer respondents than traditional forms of survey research, as multiple vignettes are rated per respondent without providing specific decision guidelines or further suggestions (e.g., Auspurg et al., 2009). As a rule of thumb, some researchers propose using 20 times more cases than independent variables, while others recommend conducting statistical power calculations to estimate the required sample size a priori (Lauder, 2002).

In the context of sampling, researchers also have to consider the number of vignettes to be evaluated by each respondent. Similar to the number of included dimensions, the presentation of too many vignettes might lead to fatigue/boredom effects, information overload, and inconsistent ratings (Auspurg et al., 2009; Auspurg et al., 2010; Sauer et al., 2011). Here again, no precise rule can be put forth (Beck & Opp, 2001). Sauer, Auspurg, Hinz, and Liebig (2011) were amongst the first who shed some light on the topic and revealed that a low educational level among the respondents has a negative impact on the consistency of the evaluations when using a large number of vignettes. Furthermore, they recommend using no more than 20 vignettes. Their study is based on general population samples, and, thus, B&S researchers should cautiously scrutinize their own respondents. With regard to the showcase, around 12 or fewer vignettes might be recommended when studying retail investors because of the relatively long text (see again Table 1), whereas in the professional investment context, a somewhat larger number of vignettes might be used per respondent, given that a higher processing capacity can be expected for these investors.

Generally, the respondents in FS studies are usually given only a sample of the entire vignette population. Thus, researchers also have to choose a vignette-sampling strategy. In general, vignette samples can be drawn by using either random or quota designs. With randomly drawn samples (applied with or without replacements), each respondent receives a selection of vignettes that represents “a unique random sample of the same size” (Dülmer, 2007, p. 383). The basic rationale for using random designs is to ensure a wide coverage of the vignette population. A variation is the clustered random design in which a randomly drawn sample of unique vignettes is evaluated by multiple respondents “to obtain multiple ratings per vignette, allowing not only respondent-specific but also vignette-specific analyses” (Dülmer, 2007, p. 385). The quota designs instead rely on one specific vignette set and are recommended when FS studies employ relatively small vignette sets (Dülmer, 2007).

Another issue is that vignettes may affect the evaluation of subsequent vignettes. For instance, if a vignette that describes a major environmental incident is followed by a vignette that describes a less severe incident, the latter might be evaluated as less problematic compared to the major incident. However, if the less severe incident were presented as the first incident, the evaluation may be different. To counter the effects arising from a particular sequence, the presentation of the vignettes to a respondent is, at times, kept constant or randomized.^{viii}

Collecting respondent-specific information. Respondent-specific information that cannot be incorporated into vignettes (e.g., demographic, social, or general information) is usually collected through a traditional survey to allow for the effect analysis of respondent characteristics on judgment differences between (groups of) respondents (Atzmüller & Steiner, 2010; Killick & Taylor, 2012; Taylor, 2006). The showcase illustrates how age, gender, and value priorities, among others, can be put forth exemplarily as important investor-specific characteristics in the context of SRI. While B&S researchers might identify and wish to include further aspects, questions regarding the expected time required for completing the survey, respondents’ availability time, and potential fatigue effects demand careful consideration (Beck & Opp, 2001).

Analyzing the data. FSs are grounded on the assumption of socially and individually structured judgments (Rossi & Anderson, 1982). Thus, interested B&S scholars are urged to analyze the influence of vignette dimensions and respondent variables on judgments (Hox et al., 1991). In FS studies, multivariate OLS regression analysis represents the most frequently applied method for analyzing the collected judgments (Lauder, 2002; Wallander, 2009), and

many researchers argue that this technique is the appropriate statistical tool (e.g., Ludwick & Zeller, 2001; Taylor, 2006). However, when several ratings made by each of the respondents are pooled together for statistical analyses, the observed judgments can no longer be assumed to be independent, which is the assumption of OLS regression^{ix} (Aguinis & Culpepper, 2015; Hox et al., 1991); thus, the aggregation of judgments is exposed to the potential risk of intra-rater correlation (Wallander, 2009). Therefore, the use of hierarchical/multilevel models is recommended because FSs produce complex hierarchically structured (nested) data by design (e.g., Auspurg & Hinz, 2015; Hox et al., 1991; Lauder, 2002).

To illustrate the estimation and interpretation of the results of an FS, a data set of 200 virtual “participants” was constructed, and each had randomly chosen personal features, such as values and demographics. With this sample, we simulated the FS as described in the showcase presented in Table 1. Each participant rated 12 randomly chosen vignettes, resulting in 2,400 decisions constituting the data set.^x Personal features were measured using the scales described in Table 1, and the features of the vignette were included as dummy variables in the data set. Descriptive statistics of and correlations among the variables used are given in Table 2. As the decisions were nested within the participants, the statistical analysis has to use a multilevel regression model and setting the individuals as the grouping variable. This assumes that properties of the individual, regardless of being measured or latent, affect all decisions the individual makes.

The dependent variable, that is the rating task, was measured using an eleven-point-scale, asking explicitly for a probability to invest in the situation presented and did so in terms of percentages. This allows treating the variable as metric and to use linear models, such as regression, for the statistical analysis. The issue of when a Likert scale with a certain range allows use of an OLS regression or whether an ordinal logit model is the more appropriate method must always be decided depending on the scale used by the researcher. Oftentimes, Likert scales with a range of five to seven points are used, which are then analyzed using OLS regression. However, in the case that the dependent variable is non-metric (e.g., binary, ordinal, truncated, or count data), the analysis must use the multilevel versions of the regression models appropriate to the nature of the dependent variable. Examples are the mixed routine in STATA or the xt-versions of other routines of models for binary, count, or truncated data.

Table 2: Descriptive Statistics and Correlations^{xi}

Variable	Min.	Max.	M.	SD.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 InvestmentDecision	0	10	4.83	1.719																
2 AttentionSRI	0	1	0.49	0.5	-0.03															
3 CompanyLocation	0	1	0.49	0.5	0.27**	0.01														
4 CompanyLarge	0	1	0.5	0.5	0.31**	-0.02	0.02													
5 IndustryProblematic	0	1	0.5	0.5	-0.33**	0.02	0.02	-0.02												
6 CompanyPerformance	0	1	0.49	0.5	0.30**	-0.04*	0.02	0.00	0.00											
7 CompanySustainability	0	1	0.51	0.5	0.34**	0.02	-0.02	0.01	0.03	0.02										
8 GRIReport	0	1	0.51	0.5	0.06**	-0.02	0.00	-0.01	0.00	-0.02	0.01									
9 ReportExternallyValidated	0	1	0.5	0.5	0.08**	0.00	0.00	0.01	-0.01	0.01	0.03	-0.03								
10 ProfessionalInvestor	0	1	0.69	0.463	0.30**	0.00	-0.01	0.03	0.01	0.00	-0.01	0.05*	-0.03							
11 SexMale	0	1	0.74	0.439	-0.15**	-0.02	0.00	-0.03	0.03	-0.02	-0.02	-0.06**	-0.02	0.00						
12 WorkExperience	2	34	17.51	9.635	-0.05*	-0.02	0.00	0.02	-0.01	-0.03	-0.02	-0.03	0.02	-0.08**	0.06**					
13 Religiosity	1	7	4.26	2.052	0.01	0.03	0.01	0.00	0.00	0.00	0.00	-0.01	0.01	-0.01	-0.18**	-0.08**				
14 ImportantRich	1	7	3.81	1.933	0.00	-0.02	0.02	0.01	-0.01	0.01	0.00	-0.04	-0.02	0.02	-0.01	0.05*	0.01			
15 Environmentalism	1	7	2.92	1.779	0.15**	0.01	0.01	-0.01	0.01	0.00	-0.02	0.04*	0.00	0.01	-0.39**	-0.04	0.09**	-0.06**		
16 PerformanceBelief	1	7	4.12	2.023	0.02	-0.02	0.00	0.01	-0.03	0.00	-0.01	0.03	-0.01	0.21**	0.01	0.10**	-0.01	0.02	-0.07**	
17 RiskBelief	1	7	3.67	1.95	0.02	-0.01	0.00	0.02	0.04*	-0.01	-0.04	0.00	-0.01	0.01	0.05*	-0.04*	-0.11**	-0.03	0.03	-0.45**

The initial, exploratory step of the analysis would be to test for the effects of the experimental manipulations—that is, the vignette features—which are typically the researcher’s central concern. In the showcase, these features were randomly assigned to the participants so that they were uncorrelated and could be included simultaneously (see Model 1 in Table 3 below). The interpretation of Model 1 would be that there are some highly significant effects for the features of the vignettes. While the attention that the home country government pays to issues of SRI is irrelevant to the participant’s decision in our simulation, the features of the company in question—for example, financial and sustainability performance—are highly relevant.

Next, the lower five rows in Table 3 provide information about the model ($N=2,400$ rating decisions from 200 “groups”, which are the participants in the presented showcase) and the model’s fit (R^2 ’s). Unlike conventional regressions, multilevel regressions differentiate between two sources of variation: between the participants and between the vignettes within the participants. Due to personal properties, a participant may show a tendency to rate the investment in a particular way in all 12 vignettes. Personal features, such as professionalism or environmentalism, may explain this tendency, and “ R^2_{between} ” is the information related to how well the variation *between* participants can be explained by the personal features used. The features of the vignettes will explain why a particular company described by the vignette is less attractive as an investment. “ R^2_{within} ” gives the information about how well differences among the ratings given by *one* participant can be explained by the features of the vignettes. Finally, “ R^2_{overall} ” gives the overall fit—that is, how well the features of the vignettes and the participants taken together explain the ratings given in the study (see Snijders & Bosker, 2012, for a detailed introduction).

In the simulation, the explanatory contribution of the vignette features is 0.78 (R^2_{within}), which is substantial. However, “ R^2_{between} ” is expected to be zero because no information whatsoever about the participants was used in this regression model. This would be the case if the effects of the vignette features and the effects of the personal features were fully independent of each other. Contradicting this presumption, an “ R^2_{between} ” of 0.08 indicates that the relationships are not that simple and that there might be conditional effects, where the effects of participant-level variables depend on vignette-level variables.

Table 3: Results of a simulated multi-level regression analysis^{xii}

	Model 1	Model 2	Model 3	Model 4
<i>Vignette Features</i>				
AttentionSRI	-0.038 (0.030)	-0.038 (0.030)	-0.017 (0.020)	-0.028 (0.039)
CompanyLocation	0.999*** (0.036)	0.998*** (0.036)	0.482*** (0.033)	0.409*** (0.069)
CompanyLarge	0.976*** (0.037)	0.974*** (0.037)	0.433*** (0.035)	0.359*** (0.069)
IndustryProblematic	-1.264*** (0.032)	-1.264*** (0.032)	-1.263*** (0.021)	-1.179*** (0.039)
CompanyPerformance	0.952*** (0.040)	0.952*** (0.040)	0.387*** (0.042)	0.415*** (0.069)
CompanySustainability	1.206*** (0.058)	1.207*** (0.058)	-0.011 (0.038)	-0.003 (0.074)
GRIReport	0.195*** (0.028)	0.192*** (0.028)	0.183*** (0.020)	0.152*** (0.039)
ReportExternallyValidated	0.211*** (0.028)	0.212*** (0.028)	0.194*** (0.020)	0.215*** (0.039)
<i>Personal Features</i>				
ProfessionalInvestor		1.139*** (0.142)	-0.021 (0.151)	-0.094 (0.082)
SexMale		-0.277 (0.141)	-0.297* (0.137)	-0.303*** (0.048)
WorkExperience		-0.001 (0.006)	-0.001 (0.006)	-0.000 (0.002)
Religiosity		-0.009 (0.032)	-0.009 (0.031)	-0.009 (0.010)
ImportantRich		-0.002 (0.031)	-0.002 (0.030)	-0.002 (0.010)
Environmentalism		0.121** (0.037)	-0.085* (0.038)	-0.087*** (0.016)
PerformanceBelief		-0.032 (0.037)	-0.027 (0.036)	-0.027* (0.011)
RiskBelief		0.017 (0.040)	0.017 (0.038)	0.016 (0.011)
<i>Interaction Effects</i>				
PerformanceXProfessional			0.837*** (0.048)	0.864*** (0.083)
LocationXProfessional			0.747*** (0.040)	0.772*** (0.083)
LargeXProfessional			0.770*** (0.043)	0.870*** (0.084)
SustainabilityXEnvironmtalism			0.414*** (0.011)	0.417*** (0.022)
_cons	3.215***	2.416***	3.813***	3.838***

	(0.083)	(0.385)	(0.368)	(0.145)
<i>Model Statistics</i>				
R ² _within	0.777	0.777	0.898	
R ² _between	0.082	0.380	0.408	
R ² _overall	0.496	0.617	0.700	0.702
N	2.400	2.400	2.400	2.400
N_groups	200	200	200	

In Model 2, the personal features of the participants were added. The result indicates that there is a strong and highly significant effect for the variable “ProfessionalInvestor” (1 = professional investor, 0 = nonprofessional investor). Further, “Environmentalism” affects the ratings, but not to a degree that is comparable with professionalism. Including personal features improves the explanatory power by explaining about 38% of the variation between participants, leading to an overall R² of 0.62.

The next step consists of taking into account the interaction effects between the vignette level and the participant level. Constructing interaction effects can either be oriented at empirical results (constructing and testing more or less all possible interactions and discussing the significant ones) or toward the theoretical background of the study (which interaction effects make sense given the theoretical arguments made?). The showcase takes the second option, and, thus, two basic interactions are hypothesized: first, it is presumed that professional investors have strong preferences for large companies from industrialized countries with strong financial track records (Ferreira & Matos, 2008; Pinnuck, 2004). Second, the authors presume that the participants for whom the environment matters strongly pay more attention to the company’s sustainability performance, while this information is of little relevance for the participants who do not care about the environment (Getzner & Grabner-Kräuter, 2004; Nilsson, 2009).

To test the validity of these arguments about interaction effects, four multiplicative interaction variables are created: (1) between financial company performance and professionalism, (2) between company location and professionalism, (3) between company size and professionalism, and (4) between environmentalism and the information about the company’s sustainability-related performance compared to the industry average in the past few years.^{xiii} For (1) PerformanceXProfessional, (2) LocationXProfessional, and (3) LargeXProfessional, each of the constitutive variables is dummy coded, as are the resulting interaction variables.^{xiv} The substance of these interaction variables concerns the question of whether professionals place more emphasis on financial performance, location, and size. If,

for example, the coefficient for the performance and professionalism interaction variable is significant, it implies that financial performance is indeed more important for professional investors. If it is insignificant, there are no differences between professional and nonprofessional investors regarding the importance they assign to the company's financial performance in their rating. The same argument applies to the interaction between the participant's environmentalism and the company's sustainability-related performance (here, "0" indicates inferior and "1" indicates superior performance). The interaction effect (4) SustainabilityXEnvironmentalism constitutes an additional weight of sustainability, which becomes higher as the participant's environmentalism becomes stronger.

Including the interaction variables in the regression yields the results given in Model 3. The explanatory power (R^2_{overall}) increased once more to 0.7, indicating that the inclusion of these effects did make sense in terms of an explanatory contribution.^{xv} The coefficients of those vignette features for which no interaction effects were constructed are largely unchanged in magnitude and significance. However, the remaining coefficients of the vignette features changed substantially. They are much smaller and, in the case of CompanySustainability, are no longer even significant. This is typical for models that include interaction effects: the features of the vignette do not matter for all participants alike, nor do the features of the participants matter to the same degree in all vignettes. For those vignette features for which interaction variables involving professionalism (ProfessionalInvestor) were constructed, the coefficients of the original vignette features in the upper section of Table 3 change compared to Model 2. They now represent the effect of this particular vignette feature for a nonprofessional investor. It can be seen that these vignette features still matter significantly for nonprofessionals, but much less than for professionals. Looking at personal features, those features for which no interaction variables were constructed still have approximately the same coefficients. Those personal features constituting the other component of an interaction effect, however, provide a different picture. The coefficient of ProfessionalInvestor is now much smaller and is insignificant, indicating that after taking the interactions into account, professional investors are no different from nonprofessionals.

When looking at the interaction effects themselves, all four effects are of substantial magnitude and significance. For a nonprofessional investor, superior financial performance of the company shifts the rating upward by on average 0.387 points, compared to inferior financial performance. For professional investors, the effect is higher by an additional 0.837 points (PerformanceXProfessional). The interpretation of the other interaction effects is equivalent. The insight from the simulated showcase data is that professional investors do not

differ per se in regard to how they rate; rather, they differ regarding how much importance they assign to certain information (in the showcase: financial performance, location, and size of the company). In Model 2, the insight was restricted to the fact that professional investors differ from nonprofessionals. With Model 3, one knows more precisely that the difference lies in the relevance assigned to specific information when giving a rating.

For environmentalism, the story is equivalent. Neither environmentalism nor CompanySustainability matters much when the interaction is taken into account. Indeed, regardless of the information given in the vignette, the participants with high levels of environmentalism tend to rate the likelihood of investing slightly lower. However, for the participants with environmentalist attitudes, a company with a strong sustainability performance is an attractive investment. For this effect to occur, both aspects must be present: a participant must have a strong environmentalist attitude, and the company must have superior sustainability-related performance. Again, it is the different perception of identical information that matters.

A further strategy of conducting the analysis of interactions would be to differentiate the models by the company. One could, for example, run one regression for those vignettes for which the company has superior sustainability-related performance and compare the results with the model using only vignettes for which the company has inferior sustainability-related performance. While this strategy is more elaborate in terms of how many models have to be specified and tested, the causalities and effects, which are conditional on the company's features, are easier to identify. Finally, Model 4 shows the results as obtained using a conventional OLS regression. As can be seen, the coefficients are fairly robust, and the story to be learned is the same.

Including interaction effects offers interesting insights that go beyond what can be gained from using only individual and vignette features. However, it also reveals practical problems, two of which we will discuss in more detail here. First, using multiplicative interaction terms results in variables which are, "by construction", highly correlated. In the showcase, this concerns, for example, the interaction variable PerformanceXProfessional, which is derived from two constitutive variables, ProfessionalInvestor and CompanyPerformance. When including all three variables in the regression simultaneously, the problem of multicollinearity may arise. The showcase features an "optimal" data set in which multicollinearity is no problem. In practice, however, multicollinearity is an issue of insufficient information in the data, which is best avoided in the first place. In the setting of a FS, multicollinearity can be

reduced by making sure that (1) the experimental manipulations are uncorrelated, (2) personal features are uncorrelated with the experimental manipulations, and (3) there is enough variation in personal features. The first point is achieved by a research design where the manipulations are chosen fully at random for each case. The second point is achieved by making sure that the randomization is successful, that is, researchers need to assure that the assignment of participants with certain features to certain experimental constellations is perfectly random. Fully automated assignments procedures, which are available in most web-based survey tools, help to minimize these two issues. The third aspect, the most difficult in practice, is assured by choosing a sample which is heterogeneous enough. For the showcase, this translates, for example, to not using a sample consisting predominantly of professional investors when the professional status shall be used as an explanatory variable and for creating interaction effects.

Second, as with all research designs, the degree of complexity increases with the number of manipulated features included. This problem becomes particularly virulent, however, when the effects are conditional, such as in the simulated showcase in which the differences between the participants occur only in specific situations. The construction and inclusion of interaction effects is an issue whereby the researcher can either proceed exploratively, by constructing and testing all interaction effects, or confirmatively, whereby there are theoretical reasons to expect certain interactions. In any case, the number of potential interactions increases dramatically as the number of manipulated features increases. Researchers should be aware that the simplicity of constructing complex FSs may pose challenges when analyzing and, in particular, interpreting FS results.

The Potential Advantages and Disadvantages of the FS method

As shown in Table 4, a multitude of potential advantages and disadvantages has been put forth in the FS literature. However, the broad spectrum of methodological pros and cons is often not adequately reflected in FS studies—that is to say, potential benefits tend to be stressed, whereas potential drawbacks are frequently mentioned only in passing. This phenomenon might be explained by fragmented methodological research, a lack of empirical knowledge on methodological implications, and partly inconclusive findings (Auspurg & Hinz, 2015; Barter & Renold, 2000; Sauer et al., 2011).^{xvi}

Table 4: Advantages and disadvantages of FSs

Potential advantages	
Argument	<i>Discussed, for example, by the following researchers</i>
<p>Larger samples compared to traditional experimental designs</p> <ul style="list-style-type: none"> Stronger claim of representativeness 	Auspurg, Hinz, and Liebig (2009); Hox, Kreft, and Hermkens (1991)
<p>Capability to communicate contextual information and concrete descriptions in support of vignettes' realism</p> <ul style="list-style-type: none"> Higher external validity Reduced cognitive effort required by respondents Moderating effect on the probability of unreliable and biased responses 	Ashill and Yavas (2006); Atzmüller and Steiner (2010); Auspurg, Hinz, and Liebig (2009); Dülmer (2007); Hox, Kreft, and Hermkens (1991); Wason, Polonsky, and Hyman (2002)
<p>Systematic variation of factor values</p> <ul style="list-style-type: none"> Enables decomposition of the structure of judgements and accurate assessment of the relative importance of determinants, especially when people are not aware of what influences their judgements 	Alexander and Becker (1978); Atzmüller and Steiner (2010); Auspurg, Hinz, and Liebig (2009); Caro et al. (2012); Dülmer (2007); Wason, Polonsky, and Hyman (2002)
<p>Usually there is no interviewer and no self-report of measures of behavior; manipulation of dimensions is usually subtle and therefore unlikely to be fully recognized by respondents</p> <ul style="list-style-type: none"> Judgements less susceptible to social approval and desirability biases compared to traditional surveys 	Alexander and Becker (1978); Auspurg, Hinz, and Liebig (2009); Wallander (2009); Wason, Polonsky, and Hyman (2002)
Potential disadvantages	
Argument	<i>Discussed, for example, by the following researchers</i>
<p>Vague recommendations on specific design issues of FSs based on plausible reasoning rather than empirical evidence</p> <ul style="list-style-type: none"> Lack of methodological guidance on how to conceptualize FS studies 	Auspurg, Hinz, and Liebig (2009); Barter and Renold (2000); Wallander (2009)
<p>Vignettes may not capture the entire complexity of social life; imagining a social situation and real-life engagement may be experienced differently</p> <ul style="list-style-type: none"> Restricted external validity if vignettes do not duplicate the complexity of social life Respondents may use heuristics that differ from real-life judgment-making processes in repetitive rating of similar vignettes 	Auspurg, Hinz, and Liebig (2009); Barter and Renold (2000); Collett and Childs (2011); Hughes (1998)
<p>FSs (only) enable measuring behavioral intentions, which do not necessarily correspond with actual behavior</p> <ul style="list-style-type: none"> External validity potentially restricted 	Barter and Renold (2000); Caro et al. (2012); Eifler (2007); Eifler (2010); Hughes (1998)

Probably the most severe criticism against FSs relates to the method's external validity—that is, the degree to which the results of FS studies can be generalized. The first line of critique points out that FSs measure only hypothetical judgments and decisions (behavioral intentions), which do not necessarily correspond with actual behavior (Caro et al., 2012). The second objection draws on the question of whether an imagined social situation and real-life engagement are experienced in the same way; in other words, “vignettes may fail to capture important nuances of social experience” (Collett & Childs, 2011, p. 514). Interestingly, and as shown in Table 4, external validity is also cited as a potential methodological advantage. Dülmer (2007), for example, argues that FSs allow researchers to study “judgment behavior under concrete conditions that are much closer to real-life judgment-making situations than relatively abstract questions that are more typical for opinion surveys” (p. 382). Despite opposing views, few empirical studies investigate this issue. Overall, the results are mixed, with findings speaking for (e.g., Carlson, 1996; Rahman, 1996) and against (e.g., Eifler, 2007, 2010; Pager & Quillian, 2005) FSs' external validity, which thus remains subject to controversial debate and further research (Auspurg & Hinz, 2015; Barter & Renold, 2000; Hughes, 1998).

In light of the scattered methodological research and inconclusive findings, the controllability of FS studies needs to be highlighted. Researchers can take active measures to at least partially mediate the potential methodological drawbacks. For example, the risk of heuristic-based decision making can be reduced by limiting the number of rating tasks, while the realism of vignettes can be enhanced by carefully selecting factors and factor values. Against this backdrop, interested scholars are advised to draw on the methodological recommendations and guidelines outlined above and to consult the literature briefly outlined in the following section.

Further Literature and Guidance on FSs

While FSs are often not included in social science methodology textbooks, other sources still allow interested scholars to learn about and delve deeply into the method. These sources, however, differ considerably with regard to the scope of content covered and the level of (statistical) skill required. For example, basic and accessible introductions are provided by Lauder (2002), Ludwick et al. (2004), Taylor (2006), and Wallander (2009). In contrast, the comprehensive work by Jasso (2006) uses highly technical language that requires advanced statistical knowledge. Complementing this literature, a second strand places emphasis on focused methodological questions. For instance, valuable literature exists on specific issues,

such as vignette sampling approaches (Atzmüller & Steiner, 2010; Dülmer, 2007) or the analysis of FSs (Hox et al., 1991). Apart from article-based publications, Auspurg and Hinz (2015) provide an accessible and application-oriented guide, which, to the best of our knowledge, represents the first book-length engagement with FSs since the milestone publication edited by Rossi and Anderson (1982).

Conclusion

Due to conceptual fuzziness, a multiplicity of interpretations and the strong interdependencies of societal-, organizational-, and individual-level factors, B&S issues are inherently complex. Multilevel research approaches seek to address this complexity but pose severe methodological challenges for researchers. In this context, FSs constitute a promising method for the collection and analysis of primary data and for coping with the complex and multilevel nature of B&S phenomena. By studying decision-making processes at the individual level, FSs enable scholars to advance the limited understanding of the foundations of B&S inquiry. Furthermore, as multiple dimensions are integrated into FS vignettes, the individual level can be directly related to critical factors at the organizational and societal levels. The FS method thus allows researchers to tease apart what factors flow into complex evaluative judgments and the extent to which these factors matter, even when respondents are unable to make these factors explicit.

This article makes the case for implementing FSs, which are rooted in sociology, in B&S research by illustrating a showcase example in the realm of SRI and by presenting methodological recommendations and best practices. While the authors see great potential for FSs in multilevel B&S research, this method is, of course, no panacea for all the challenges in B&S research. FSs specifically attempt to approximate complex and realistic judgment situations via the inclusion of a large number of dimensions that are open to scientific scrutiny. This is especially relevant in case, for example, there are limited substantive theories that would allow researchers to reduce empirical complexity (Atzmüller & Steiner, 2010). In general, FSs are most valuable when little is known about the underlying factors of the decision-making process, which is particularly likely in novel or underdeveloped research fields. Other methods, however, might be more suitable techniques in judgment situations that have proven to be affected by a very limited set of factors.

In 1979, Rossi expressed the wish that FSs would become “a part of standard repertory for social researchers” (p. 186); however, he anticipated that it would take some time for FSs

to become a common technique. This has proven correct for the source discipline of FSs—sociology (Wallander, 2009)—and even more so for adjacent disciplines, such as business studies. In light of the outlined thematic and methodological potential of FSs in B&S research, this article will hopefully inspire interested scholars and contribute to the future dispersion of the method.

References

- Aguinis, H., & Culpepper, S. A. (2015). An Expanded Decision-Making Procedure for Examining Cross-Level Interaction Effects With Multilevel Modeling. *Organizational Research Methods, 18*(2), 155–176.
- Aguinis, H., & Glavas, A. (2012). What We Know and Don't Know About Corporate Social Responsibility: A Review and Research Agenda. *Journal of Management, 38*(4), 932–968.
- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Alciatore, M., Dee, C. C., & Easton, P. (2004). Changes in environmental regulation and reporting: The case of the petroleum industry from 1989 to 1998. *Journal of Accounting and Public Policy, 23*(4), 295–304.
- Alexander, C. S., & Becker, H. J. (1978). The use of vignettes in survey research. *Public Opinion Quarterly, 42*(1), 93–104.
- Ashill, N. J., & Yavas, U. (2006). Vignette Development: An Exposition and Illustration. *Innovative Marketing, 2*(1), 28–36.
- Athanasopoulou, A., & Selsky, J. W. (2015). The Social Context of Corporate Social Responsibility: Enriching Research With Multiple Perspectives and Multiple Levels. *Business & Society, 54*(3), 322–364.
- Atzmüller, C., & Steiner, P. M. (2010). Experimental Vignette Studies in Survey Research. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 6*(3), 128–138.
- Auspurg, K., & Hinz, T. (2015). *Factorial survey experiments. Quantitative applications in the social sciences: Vol. 175*. Thousand Oaks, Calif.: Sage.
- Auspurg, K., Hinz, T., & Liebig, S. (2009). Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey. *Methoden-Daten-Analysen, 3*(59-96).
- Auspurg, K., Hinz, T., Liebig, S., & Sauer, C. (2010). Wie unplausibel darf es sein? Der Einfluss von Designmerkmalen auf das Antwortverhalten in 59 Faktoriellen Surveys. In H.-G. Soeffner (Ed.), *Unsichere Zeiten: Herausforderungen gesellschaftlicher Transformationen* (pp. 1–13). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Auspurg, K., & Jackle, A. (2015). First Equals Most Important?: Order Effects in Vignette-Based Measurement. *Sociological Methods & Research*, Online First.

- Barter, C., & Renold, E. (2000). 'I wanna tell you a story': Exploring the application of vignettes in qualitative research with children and young people. *International Journal of Social Research Methodology*, 3(4), 307–323.
- Beck, M., & Opp, K.-D. (2001). Der faktorielle Survey und die Messung von Normen. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 53(2), 283–306.
- Belal, A. R., & Cooper, S. (2011). The absence of corporate social responsibility reporting in Bangladesh. *Critical Perspectives on Accounting*, 22(7), 654–667.
- Berthelot, S., Coulmont, M., & Serret, V. (2012). Do Investors Value Sustainability Reports? A Canadian Study. *Corporate Social Responsibility and Environmental Management*, 19(6), 355–363.
- Brambor, T., Roberts Clark, W., & Golder, M. (2006). Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis*, 14(1), 63–82.
- Carlson, B. E. (1996). Dating Violence: Student Beliefs About Consequences. *Journal of Interpersonal Violence*, 11(1), 3–18.
- Caro, F. G., Yee, C., Levien, S., Gottlieb, A. S., Winter, J., McFadden, D. L., & Ho, T. H. (2012). Choosing Among Residential Options: Results of a Vignette Experiment. *Research on Aging*, 34(1), 3–33.
- Cheah, E.-T., Jamali, D., Johnson, J. E., & Sung, M.-C. (2011). Drivers of Corporate Social Responsibility Attitudes: The Demography of Socially Responsible Investors. *British Journal of Management*, 22(2), 305–323.
- Cieciuch, J., & Davidov, E. (2012). A comparison of the invariance properties of the PVQ-40 and the PVQ-21 to measure human values across German and Polish Samples. *Survey Research Methods*, 6(1), 37–48.
- Collett, J. L., & Childs, E. (2011). Minding the gap: Meaning, affect, and the potential shortcomings of vignettes. *Social Science Research*, 40(2), 513–522.
- Dhaliwal, D. S., Radhakrishnan, S., Tsang, A., & Yang, Y. G. (2012). Nonfinancial Disclosure and Analyst Forecast Accuracy: International Evidence on Corporate Social Responsibility Disclosure. *The Accounting Review*, 87(3), 723–759.
- Doran, C. J. (2009). The Role of Personal Values in Fair Trade Consumption. *Journal of Business Ethics*, 84(4), 549–563.

- Dülmer, H. (2007). Experimental Plans in Factorial Surveys: Random or Quota Design? *Sociological Methods & Research*, 35(3), 382–409.
- Eifler, S. (2007). Evaluating the Validity of Self-Reported Deviant Behavior Using Vignette Analyses. *Quality & Quantity*, 41(2), 303–318.
- Eifler, S. (2010). Validity of a Factorial Survey Approach to the Analysis of Criminal Behavior. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 6(3), 139–146.
- Eurosif. (2012). *European SRI Study 2012*. Brussels, Belgium. Retrieved from www.eurosif.org/publication/download/european-sri-study-2012
- Eurosif. (2014). *European SRI Study 2014*. Brussels, Belgium. Retrieved from <http://www.eurosif.org/wp-content/uploads/2014/09/Eurosif-SRI-Study-20142.pdf>
- Ferreira, M. A., & Matos, P. (2008). The colors of investors' money: The role of institutional investors around the world. *Journal of Financial Economics*, 88(3), 499–533.
- Fortanier, F., Kolk, A., & Pinkse, J. (2011). Harmonization in CSR Reporting. *Management International Review*, 51(5), 665–696.
- Franzese, R., & Kam, C. (2007). *Modeling and Interpreting Interactive Hypotheses in Regression Analysis*. Ann Arbor: University of Michigan Press.
- Frederick, W. C. (1994). From CSR1 to CSR2: The Maturing of Business-and-Society Thought. *Business & Society*, 33(2), 150–164.
- Frost, G. R. (2007). The Introduction of Mandatory Environmental Reporting Guidelines: Australian Evidence. *Abacus*, 43(2), 190–216.
- Furman, N., Shooter, W., & Schumann, S. (2010). The Roles of Heuristics, Avalanche Forecast, and Risk Propensity in the Decision Making of Backcountry Skiers. *Leisure Sciences*, 32(5), 453–469.
- Gallo, P. J., & Christensen, L. J. (2011). Firm Size Matters: An Empirical Investigation of Organizational Size and Ownership on Sustainability-Related Behaviors. *Business & Society*, 50(2), 315–349.
- Getzner, M., & Grabner-Kräuter, S. (2004). Consumer preferences and marketing strategies for “green shares”. *International Journal of Bank Marketing*, 22(4), 260–278.
- Glac, K. (2009). Understanding Socially Responsible Investing: The Effect of Decision Frames and Trade-off Options. *Journal of Business Ethics*, 87(S1), 41–55.

- Glac, K., Elm, D. R., Martin, K. E., & Painter-Morland, M. (2014). Areas of Privacy in Facebook. *Business and Professional Ethics Journal*, 33(2), 147–176.
- Golob, U., & Bartlett, J. L. (2007). Communicating about corporate social responsibility: A comparative study of CSR reporting in Australia and Slovenia. *Public Relations Review*, 33(1), 1–9.
- Green, P. E., Krieger, A. M., & Wind, Y. (2001). Thirty Years of Conjoint Analysis: Reflections and Prospects. *Interfaces*, 31(3), S56-S73.
- Hahn, R., & Kühnen, M. (2013). Determinants of sustainability reporting: a review of results, trends, theory, and opportunities in an expanding field of research. *Journal of Cleaner Production*, 59, 5–21.
- Hamilton, B. H., & Nickerson, J. A. (2003). Correcting for Endogeneity in Strategic Management Research. *Strategic Organization*, 1(1), 51–78.
- Highhouse, S., Luong, A., & Sarkar-Barney, S. (1999). Research Design, Measurement, Andeffects of Attribute Range on Job Choice: More Than Meets the Eye. *Organizational Research Methods*, 2(1), 37–48.
- Hox, J. J., Kreft, I. G. G., & Hermkens, P. L. J. (1991). The Analysis of Factorial Surveys. *Sociological Methods & Research*, 19(4), 493–510.
- Hughes, R. (1998). Considering the Vignette Technique and its Application to a Study of Drug Injecting and HIV Risk and Safer Behaviour. *Sociology of Health and Illness*, 20(3), 381–400.
- Iyer, E. S., & Kashyap, R. K. (2009). Noneconomic goals of investors. *Journal of Consumer Behaviour*, 8(5), 225–237.
- Jansson, M., & Biel, A. (2011). Motives to engage in sustainable investment: a comparison between institutional and private investors. *Sustainable Development*, 19(2), 135–142.
- Jansson, M., & Biel, A. (2014). Investment Institutions' Beliefs About and Attitudes Toward Socially Responsible Investment (SRI): A Comparison Between SRI and Non-SRI Management. *Sustainable Development*, 22(1), 33–41.
- Jasso, G. (2006). Factorial Survey Methods for Studying Beliefs and Judgments. *Sociological Methods & Research*, 34(3), 334–423.
- Jones, T. M. (1983). An Integrating Framework for Research in Business and Society: A Step Toward the Elusive Paradigm? *Academy of Management Review*, 8(4), 559–564.

- Juravle, C., & Lewis, A. (2008). Identifying impediments to SRI in Europe: a review of the practitioner and academic literature. *Business Ethics: A European Review*, 17(3), 285–310.
- Juravle, C., & Lewis, A. (2009). The Role of Championship in the Mainstreaming of Sustainable Investment (SI): What Can We Learn From SI Pioneers in the United Kingdom? *Organization & Environment*, 22(1), 75–98.
- Killick, C., & Taylor, B. J. (2012). Judgements of Social Care Professionals on Elder Abuse Referrals: A Factorial Survey. *British Journal of Social Work*, 42(5), 814–832.
- Kolk, A., & Perego, P. (2009). Determinants of the adoption of sustainability assurance statements: an international investigation. *Business Strategy and the Environment*, n/a.
- Lauder, W. (2002). Factorial survey methods: A valuable but under-utilised research method in nursing research? *Journal of Research in Nursing*, 7(1), 35–43.
- Lenssen, G., Nijhof, A., Roger, L., Kievit, H., Sreekumar Nair, A., & Ladha, R. (2014). Determinants of non-economic investment goals among Indian investors. *Corporate Governance: The international journal of business in society*, 14(5), 714–727.
- Lönnqvist, J.-E., Verkasalo, M., Wichardt, P. C., & Walkowitz, G. (2013). Personal values and prosocial behaviour in strategic interactions: Distinguishing value-expressive from value-ambivalent behaviours. *European Journal of Social Psychology*, n/a-n/a.
- Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, 1(1), 1–27.
- Ludwick, R., Wright, M. E., Zeller, R. A., Dowding, D. W., Lauder, W., & Winchell, J. (2004). An Improved Methodology for Advancing Nursing Research. *Advances in Nursing Science*, 27(3), 224–238.
- Ludwick, R., & Zeller, R. A. (2001). The Factorial Survey: An Experimental Method to Replicate Real World Problems. *Nursing Research*, 50(2), 129–133.
- Lülf, R., & Hahn, R. (2013). Corporate Greening beyond Formal Programs, Initiatives, and Systems: A Conceptual Model for Voluntary Pro-environmental Behavior of Employees. *European Management Review*, 10(2), 83–98.
- Lülf, R., & Hahn, R. (2014). Sustainable Behavior in the Business Sphere: A Comprehensive Overview of the Explanatory Power of Psychological Models. *Organization & Environment*, 27(1), 43–64.

- Martin, K. E. (2012). Diminished or Just Different? A Factorial Vignette Study of Privacy as a Social Contract. *Journal of Business Ethics*, 111(4), 519–539.
- Martin, K. E. (2015). Privacy Notices as Tabula Rasa: An Empirical Investigation into How Complying with a Privacy Notice Is Related to Meeting Privacy Expectations Online. *Journal of Public Policy & Marketing*,
- Martin, K. E., & Shilton, K. (2015). Why experience matters to privacy: How context-based experience moderates consumer privacy expectations for mobile applications. *Journal of the Association for Information Science and Technology*,
- McLachlan, J., & Gardner, J. (2004). A Comparison of Socially Responsible and Conventional Investors. *Journal of Business Ethics*, 52(1), 11–25.
- Nikolaeva, R., & Bicho, M. (2011). The role of institutional and reputational factors in the voluntary adoption of corporate social responsibility reporting standards. *Journal of the Academy of Marketing Science*, 39(1), 136–157.
- Nilsson, J. (2008). Investment with a Conscience: Examining the Impact of Pro-Social Attitudes and Perceived Financial Performance on Socially Responsible Investment Behavior. *Journal of Business Ethics*, 83(2), 307–325.
- Nilsson, J. (2009). Segmenting socially responsible mutual fund investors. *International Journal of Bank Marketing*, 27(1), 5–31.
- Pager, D., & Quillian, L. (2005). Walking the Talk?: What Employers Say Versus What They Do. *American Sociological Review*, 70(3), 355–380.
- Parsa, S., & Kouhy, R. (2008). Social Reporting by Companies Listed on the Alternative Investment Market. *Journal of Business Ethics*, 79(3), 345–360.
- Pasewark, W. R., & Riley, M. E. (2010). It's a Matter of Principle: The Role of Personal Values in Investment Decisions. *Journal of Business Ethics*, 93(2), 237–253.
- Pinnuck, M. (2004). Stock preferences and derivative activities of Australian fund managers. *Accounting and Finance*, 44(1), 97–120.
- Prado-Lorenzo, J., Rodríguez-Domínguez, L., Gallego-Álvarez, I., & García-Sánchez, I. (2009). Factors influencing the disclosure of greenhouse gas emissions in companies world-wide. *Management Decision*, 47(7), 1133–1157.
- Rahman, N. (1996). Caregivers' Sensitivity to Conflict: The Use of the Vignette Methodology. *Journal of Elder Abuse and Neglect*, 8(1), 35–47.

- Renneboog, L., Ter Horst, J., & Zhang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance*, 32(9), 1723–1742.
- Rosen, B. N., Sandler, D. M., & Shani, D. (1991). Social Issues and Socially Responsible Investment Behavior: A Preliminary Empirical Investigation. *Journal of Consumer Affairs*, 25(2), 221–234.
- Rossi, P. H. (1979). Vignette Analysis: Uncovering the Normative Structure of Complex Judgements. In R. K. Merton, J. S. Coleman, & P. H. Rossi (Eds.), *Qualitative and quantitative social research. Papers in honor of Paul F. Lazarsfeld* (pp. 176–186). New York: Free Press.
- Rossi, P. H., & Anderson, A. B. (1982). The factorial survey approach: An introduction. In P. H. Rossi & S. L. Nock (Eds.), *Measuring social judgments. The factorial survey approach* (pp. 15–67). Beverly Hills, CA: Sage; Sage Publications.
- Sauer, C., Auspurg, K., Hinz, T., & Liebig, S. (2011). The Application of Factorial Surveys in General Population Samples: The Effects of Respondent Age and Education on Response Times and Response Consistency. *Survey Research Methods*, 5(3), 89–102.
- Scholtens, B., & Sievänen, R. (2013). Drivers of Socially Responsible Investing: A Case Study of Four Nordic Countries. *Journal of Business Ethics*, 115(3), 605–616.
- Schwartz, M. S., & Carroll, A. B. (2008). Integrating and Unifying Competing and Complementary Frameworks: The Search for a Common Core in the Business and Society Field. *Business & Society*, 47(2), 148–186.
- Schwartz, S. H. (1992). Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries. In *Advances in Experimental Social Psychology. Advances in Experimental Social Psychology Volume 25* (pp. 1–65). Elsevier.
- Schwartz, S. H. (1994). Are There Universal Aspects in the Structure and Contents of Human Values? *Journal of Social Issues*, 50(4), 19–45.
- Sethi, S. P. (1979). A Conceptual Framework for Environmental Analysis of Social Issues and Evaluation of Business Response Patterns. *Academy of Management Review*, 4(1), 63–74.
- Shooter, W., & Galloway, S. (2010). The Use of Factorial Surveys in Leisure Research. *Journal of Leisure Research*, 42(4), 641–652.

- Sievänen, R., Rita, H., & Scholtens, B. (2013). The Drivers of Responsible Investment: The Case of European Pension Funds. *Journal of Business Ethics*, 117(1), 137–151.
- Simnett, R., Vanstraelen, A., & Chua, W. F. (2009). Assurance on Sustainability Reports: An International Comparison. *The Accounting Review*, 84(3), 937–967.
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2. ed.). Los Angeles, Calif.: Sage.
- Sotorrió, L. L., & Sánchez, J. L. F. (2010). Corporate social reporting for different audiences: the case of multinational corporations in Spain. *Corporate Social Responsibility and Environmental Management*, 17(5), 272–283.
- Stanny, E., & Ely, K. (2008). Corporate environmental disclosures about the effects of climate change. *Corporate Social Responsibility and Environmental Management*, 15(6), 338–348.
- Taylor, B. J. (2006). Factorial Surveys: Using Vignettes to Study Professional Judgement. *British Journal of Social Work*, 36(7), 1187–1207.
- Teichert, T., & Shehu, E. (2010). Investigating Research Streams of Conjoint Analysis: A Bibliometric Study. *BuR - Business Research*, 3(1), 49–68.
- Tippet, J., & Leung, P. (2001). Defining Ethical Investment and its Demography in Australia. *Australian Accounting Review*, 11(25), 44–55.
- Wallander, L. (2009). 25 years of factorial surveys in sociology: A review. *Social Science Research*, 38(3), 505–520.
- Wallis, M. von, & Klein, C. (2014). Ethical requirement and financial interest: a literature review on socially responsible investing. *Business Research*,
- Wason, K. D., Polonsky, M. J., & Hyman, M. R. (2002). Designing Vignette Studies in Marketing. *Australasian Marketing Journal (AMJ)*, 10(3), 41–58.
- Wilde, E. de, Cooke, A. D., & Janiszewski, C. (2008). Attentional Contrast During Sequential Judgments: A Source of the Number-of-Levels Effect. *Journal of Marketing Research*, 45(4), 437–449.
- Wood, D. J. (1991). Corporate Social Performance Revisited. *Academy of Management Review*, 16(4), 691–718.
- Wood, D. J. (2010). Measuring Corporate Social Performance: A Review. *International Journal of Management Reviews*, 12(1), 50–84.

Zenisek, T. J. (1979). Corporate Social Responsibility: A Conceptualization Based On Organizational Literature. *Academy of Management Review*, 4(3), 359–368.

ⁱ Instead of the term “values,” FS scholars often synonymously use the term “levels.” However, in this article, “levels” refer to the societal-, organizational-, and individual-level factors in B&S research. Therefore, this article uses the term “values” to describe the factor variations in FSs.

ⁱⁱ This random variation of factor values creates a unique set of vignettes and separates the vignettes employed in FS studies from other vignette-based designs (Lauder, 2002). To account for this fact, the vignettes in FS studies are sometimes referred to as “Rossi vignettes” (Wallander, 2009, p. 508).

ⁱⁱⁱ For example, a vignette description comprising 10 dimensions with four values per factor results in a vignette population of $4^{10} = 1,048,576$ vignettes.

^{iv} The left column (different levels of FS) is included for illustrative purposes but would not be presented to subjects in a real FS.

^v The authors thank one of the anonymous reviewers for bringing up this important aspect.

^{vi} A similar case can be made, for example, for fair trade research (e.g., Doran, 2009).

^{vii} The individual level could, however, be easily manipulated, particularly in studies with an exclusive focus on *retail* investors. For example, a financial advisor could be integrated into the scenarios and be manipulated with respect to age, gender, trustworthiness, or attitude toward SRI.

^{viii} The issue of sequence also applies to the order in which dimensions are presented within the vignettes (Auspurg & Jackle, 2015).

^{ix} The assumptions of whether the statistical model used are met in the available data should always be considered critically. In our showcase, the problem of nested and therefore non-independent observations is obvious, but this is not always the case. A prominent example of an article critically discussing a traditional approach to analyze a certain type of data and to investigate the problem of endogeneity in the data is the work by Hamilton and Nickerson (2003).

^x The ratings were simulated by an assumed decision rule, which includes not only systematic (personal and situational) but also random components.

^{xi} N = 2400 (2400 decisions, 12 decisions nested within 200 participants. M: Mean, SD: Standard Deviation. * Correlation is significant at the 0.01 level (2-tailed); ** Correlation is significant at the 0.05 level (2-tailed).

^{xii} Unstandardized regression coefficients. Models 1 to 3 are multilevel regressions, model 4 simple OLS regression. Robust standard errors are given in parenthesis, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

^{xiii} Here, only significant variables were used. However, this is not a clear-cut criterion for deciding which interaction effects to test. It may well be the case that a variable exerts an effect that is fully conditional on another variable. This shows up only when the interaction effect is included but not if the variable is included as it is. Imagine that a variable has a positive effect for men and a negative effect for women—on average, the effect is zero. Only the interaction effect with gender indicates the true situation.

^{xiv} “CompanyPerformance” (financial) is coded “1” for superior performance and “0” for inferior performance, “ProfessionalInvestor” is coded “0” for a nonprofessional and “1” for a professional investor, thus the resulting in interaction variable “PerformanceXProfessional” is “1,” if the company presented in the vignette had superior financial performance and the participant who did the rating was a professional investor. All other constellations are coded as “0.” The other interactions also follow this pattern.

^{xv} The typical measure to compare the explanatory power across models with different numbers of explanatory variables is the adjusted R^2 , which takes into account the number of variables used and the R^2 achieved. In terms of adjusted R^2 , including many more variables to achieve only little improvement in terms of explanatory power is flagged as being not worth the increase in model complexity.

^{xvi} For a detailed overview and discussion of methodological issues and research, see Auspurg and Hinz (2015).

3 Second Article

Bassen, A., Gödker, K., Lüdeke-Freund, F. & Oll, J. (2018). Climate Information in Retail Investors' Decision-Making: Evidence from a Choice Experiment. *Organization & Environment*, 32(1), 62-82. [VHB-JOURQUAL: B]

Climate Information in Retail Investors' Decision-Making: Evidence from a Choice Experiment

Abstract

Financial markets play a decisive role in the transition to a low-carbon economy. This study investigates the role of climate information presentation for climate-friendly investing among retail investors. We conduct a choice experiment in which we vary the presentation format of climate information by means of three label designs to test their influence on investment practices. We provide empirical evidence for the effectiveness of climate labeling as a potential nudge for climate-friendly investing. Further, we find heterogeneity in the influence of climate information across different label designs and cognitive characteristics of investors. Intuitive (reflective) decision-makers tend to place significantly more (less) weight on funds' climate performance compared to financial performance — irrespective of a participant's environmental preference.

Keywords: Low-carbon economy, sustainable investing, retail investors, nudge, cognitive reflection, climate information, choice experiment, conjoint analysis

1 Introduction

There is wide international consensus that limiting global temperature increase to 2°C above pre-industrial levels is crucial to mitigate the most adverse effects of anthropogenic climate change (UNFCCC, 2015). As meeting the 2°C threshold requires significant investment, a decisive role devolves upon financial markets to finance the low-carbon transition (Campiglio, 2016; Covington, 2016; UNEP FI, 2009). In the financial sector, climate change is addressed through a broad range of sustainable investing (SI) practices, which combine financial objectives with nonfinancial concerns. While SI has gained importance in the market and academia alike, market growth is predominantly driven by institutional investors (Paetzold & Busch, 2014). In the European market, for example, only 3% of total SI assets under management are held by retail investors (Eurosif, 2014). Against the backdrop of this untapped potential, mobilizing retail investors for climate-friendly investment products represents a promising leverage point to channel more financial capital toward the global combat against climate change.

Individual choices often depend upon how they are presented (Thaler & Sunstein, 2008). Building on Thaler and Sunstein's (2003, 2008) seminal work on the concept of *nudge*,¹ Pilaj (2015) recently argued that modifying investors' choice environment has great potential as a means by which to dismantle existing impediments to the adoption of SI practices. In the recent past, the concept of nudge has been increasingly discussed as an attractive policy tool (European Commission, 2013; World Bank, 2015), and considerable evidence indicates that the implementation of nudges can advance more sustainable behavior (e.g., Allcott & Mullainathan, 2010; Costa & Kahn, 2013; Momsen & Stoerk, 2014). To the best of our knowledge, the promise of nudges has not been studied empirically in the realm of climate-friendly investing. However, such research is needed, as the available empirical evidence about the effectiveness of nudges is highly context-dependent and of limited generalizability (Bao & Ho, 2015; Lehner, Mont, & Heiskanen, 2015).

This article therefore aims to shed light on how the presentation format of climate information influences investment decisions in the retail segment. We conduct an online choice

¹ In this context, the term *nudge* refers to any aspect of the choice environment that "alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives" (Thaler & Sunstein, 2008, p. 6).

experiment among European retail investors in which we systematically alter investors' choice environment. Applying conjoint methodology, we examine the equity fund choices of 953 retail investors from six European countries who are provided with standard financial *and* climate performance information on investment alternatives. Specifically, we vary the presentation format of climate performance information to test the effectiveness of three distinct label designs to influence investment decisions. We further analyze the role of investors' cognitive characteristics in the effectiveness of climate labeling.

We find that different climate label designs lead to variation in participants' investment decisions, indicating that these designs differ in their effectiveness to nudge investors toward more climate-friendly investments. Among the three label designs, we find that the star rating label yields the highest emphasis on climate performance information in investors' decision-making. We further find that climate labeling is more (less) effective in promoting climate-friendly investing among intuitive (reflective) decision-makers. An underlying mechanism of this finding seems to be the individual trading off between the fund's climate performance and financial performance. Intuitive decision-makers tend to place significantly more weight on funds' climate performance compared to financial performance in our choice setting — irrespective of a participant's environmental preference.

This paper contributes to the literature by subjecting the potential of climate-friendly nudges in the retail investment sector to a rigorous empirical test. First, we provide empirical evidence for the effectiveness of climate labeling as a potential nudge for climate-friendly investing. We thus identify a promising leverage point to promote climate-friendly investments in the retail segment and to redirect capital toward the transition to a low-carbon economy. Second, by revealing the importance of investors' cognitive reflection abilities in the adaptation of SI practices, we establish the basis for novel segmentation analyses and future research. Further, our results provide important practical insights by identifying an effective climate label design that is feasible for both mandatory and voluntary labeling.

2 Background and Hypothesis

One of the central concepts in many economic theories is the *homo economicus*, in which humans are modeled as rational and self-interested utility maximizers with perfect information processing capacities (Frank, 1987; Persky, 1995). Based on the *homo economicus* model, it is

purely the prospect of superior risk-return relationships that inhibits investors from preferring SI to conventional investing (Beal, Goyen, & Phillips, 2005). Yet, alternative explanations emerge if the homo economicus assumptions are relaxed. Behavioral economists, for instance, have long argued that embracing humans' bounded rationality is critically important to understanding and explaining deviations between rational choice models and actual human behavior (Simon, 1955; Tversky & Kahneman, 1974).

A common cognitive model used to explain human behavior is the dual process theory (e.g., Kahneman, 2003). According to dual process models, human decision-making is based on two distinct modes of cognitive functioning, which are related to intuition (system 1) and reasoning (system 2) (Stanovich, 1999; Stanovich & West, 2000). System 1 is involved in fast, automatic, associative, emotional, and effortless actions, whereas system 2 operates slowly, neutrally, in a controlled and rule-governed way, and with effort (e.g., Epstein, 1994; Kahneman & Frederick, 2002). According to Kahneman and Frederick (2002, 2005), the two systems describe a continuum rather than discrete processes. Opportunities to promote behavior change via nudges are particularly likely to emerge when system 1 affects human decision-making (Lehner et al., 2015).

Humans' partial reliance on system 1 implies that the mere provision of information may suffice when it is intuitively appealing and salient. Against this backdrop, carefully designed labels are widely considered promising nudges (Lehner et al., 2015; Ölander & Thøgersen, 2014). Labels usually provide simplified and attention-grabbing information at the point of purchase and thereby alter people's choice environment, which may in turn affect their behavior (Ölander & Thøgersen, 2014).

The promise of labeling may also hold true in the specific context of SI, as insufficient or overly complex information and inadequate information processing and transfer have been identified as major barriers to more engagement in SI (e.g., Benson & Humphrey, 2008; Hummels & Timmer, 2004; Rhodes, 2010; Schrader, 2006). In the investment context, however, labeling research is still fragmented and usually limited to standard financial information (e.g., Drescher, Roosen, & Marette, 2014; Hüsser, 2015). Rare exceptions exist in the form of studies on the impact of ethical or sustainability labels on investment decisions, but these do not address the issue of label design effects (e.g., Døskeland & Pedersen, 2016; Gutsche, Dai, & Zwergel, 2015). Nevertheless, ample evidence indicates that retail investors'

decision-making is not only influenced by the content of available information, such as risk-return relationships, but also by the way in which information is presented (e.g., Foster, Ng, & Wee, 2015; Kozup, Howlett, & Pagano, 2008; Maines & McDaniel, 2000). Furthermore, with respect to the effectiveness of labeling, design aspects such as label complexity, salience, and color have long been identified as key influencing factors in prior research (e.g., Banerjee & Solomon, 2003; Hieke & Taylor, 2012; Sacks, Rayner, & Swinburn, 2009; Teisl, 2003). Thus, the first part of the paper explores whether the level of relative importance of climate performance information in the investment decision is associated with the design of the climate label used.

Moving beyond the general impact of climate label design, we explore whether the relative importance of the presented climate performance information is heterogeneous among retail investors. As pointed out by Lehner et al. (2015), most existing policy tools rest on the assumption that a prevalent lack of information must be overcome to allow for changes in behavior. The great majority of policy tools thus invoke system 2 (reflective) and not system 1 (intuitive). In contrast, labeling changes the choice environment of individuals by providing additional information; however, it does so in a simplified and condensed manner to minimize the cognitive effort of information acquisition and processing. In opposition to purely textual information, for instance, graphically salient and intuitively appealing labels thus tend to resonate with decision-making dominated by system 1 (Lehner et al., 2015; Ölander & Thøgersen, 2014; Thaler & Sunstein, 2008). It is expected that the effectiveness of labeling can be influenced by participants' cognitive dual processes; that is, labeling may be more effective in an intuitive cognitive mode. Considering heterogeneity in cognitive reflection ability across retail investors, our hypothesis is therefore: *The effectiveness of a climate label is higher for intuitive (system 1) decision-makers than for reflective (system 2) decision-makers.*

3 Methodology

We conduct an online choice experiment, applying conjoint methodology to test our hypothesis. Introduced into marketing by Green and Rao (1971), conjoint analyses are increasingly used in the social sciences to examine how people make multidimensional choices and elicit their underlying preferences for specific attributes of objects, such as products and services (Lohrke, Holloway, & Wooley, 2010; Green, Krieger, & Wind, 2000). Participants' stated preferences

are derived from an observed series of their choices between, or ratings of, hypothetically described objects. Attributes that are considered important determinants of the specific decision under examination thereby vary with regard to their levels. This allows estimating part-worth utilities and weighing the relative importance of attributes for the specific choice or rating reflecting participants' individual preferences.

Using the conjoint methodology has certain advantages. First, the conjoint approach offers a direct assessment of participants' underlying preferences and judgments, increasing internal validity compared to post hoc methods, which collect data on self-reported preferences and decisions through surveys (Arnold & Feldman, 1981; Schwarz, 1999). This is especially important with regard to a potentially high social desirability bias when participants receive climate performance information. Second, conjoint analyses generate results that are more robust when participants lack sufficient insight into their own decision processes to report them accurately (Bertrand & Mullainathan, 2001), which can be the case for investment decisions in capital markets. By testing several attribute-level combinations, we are able to observe participants' willingness to invest in a climate-friendly manner. We collect data on how they deal with trade-offs in reaching their decision, such as trading off the financial performance of an equity fund against its climate performance.

3.1 Experimental Design and Variables

We use a factorial experimental design to examine participants' hypothetical investment choices. Correspondingly, we vary five characteristics of equity funds within subjects: the fund provider, financial performance, risk class, fund costs, and climate performance. The attributes that constitute the conjoint task and the levels of each are shown in Table 1. Both the fund attributes and the corresponding levels were developed in cooperation with fund market experts.² The ordering of presented equity fund attributes is fixed. The first four conventional fund attributes follow the common order used in information documents, such as Key Investor Information Documents (KIID), and climate performance is added as a further attribute. Climate performance represents the contribution of the fund's investment strategy to global

² We had access to several fund experts through the Climate-KIC-funded research project within which this study was conducted. See <http://www.climate-kic.org/projects/the-first-climate-impact-rating-for-investment-funds/>.

climate change mitigation (see Appendix A for detailed information). The choice scenarios presented to participants apply a randomized choice-based conjoint setting. Each choice task includes three randomly drawn equity fund options and a “none” option.

Table 1. Conjoint attributes and levels

Attributes and their levels used to describe hypothetical equity funds.

Attributes	Levels		
Provider	High name recognition Medium name recognition Low name recognition		
Financial performance	9.5% (better than benchmark) 7.5% (similar to benchmark) 5.5% (worse than benchmark)		
Risk class	Risk level 7 (high) Risk level 4 (medium) Risk level 1 (low)		
Fund costs	2.25% 1.5% 0.75%		
Climate performance	Impact scale treatment: Green area Yellow area Red area	Star rating treatment: 5 stars 3 stars 1 star	Award treatment: Award No award -

To manipulate our key variable of interest, we vary the way in which the climate performance information is presented between subjects. We test three different climate label designs: a climate award, a star rating, and an impact scale. These three designs were identified in the wider literature on label designs as well as by considering popular designs used in practice, such as the Morningstar rating, eco-labels for consumer goods, and awards of different kinds (e.g., “eco-awards”) (e.g., Døskeland & Pedersen, 2016; Drescher et al., 2014; Heinzle & Wüstenhagen, 2012). Figure 1 displays the three different label designs. The impact scale design displays funds’ climate performance information negatively (red), neutrally (yellow), or positively (green) and covers all investment funds.³ The star rating design, in contrast, only provides neutral or positive information, covering all investment funds with average or better-

³ The benchmark for assigning a particular climate performance level is a fictive ranking ranging from 1 to 100 climate performance points. We label funds with 1–33 rating points negatively, funds with 34–66 rating points neutrally, and funds with 67–100 rating points positively. The survey instructions are provided in Appendix A.

than-average climate performance.⁴ The award design presents positive climate performance information (awarded) and covers only a portion of all investment funds.⁵ Participants are randomly assigned to one of the three label design treatment conditions.

We also collect data on participants' cognitive reflection abilities (Frederick, 2005) and common confounding variables in the investment context, such as risk preferences (Dohmen et al., 2011), financial literacy (van Rooij, Lusardi, & Alessie, 2011), prior pro-environmental investment behavior, and sociodemographic characteristics. Appendix B displays the composition of subsamples by country and treatment condition. Participants are roughly evenly distributed over the different countries and treatments.

Figure 1. Label designs

Climate label designs tested.

Impact scale:



Star rating:



Award:



3.2 Sample

We base our study on a sample of 953 retail investors from six European countries with large conventional funds and sustainable investment markets: Germany, France, the United Kingdom, Italy, Sweden, and Switzerland. A professional panel provider, Lightspeed GMI Germany, administered recruitment and access to the participants. The sample consists of participants who, at the time of the experiment, owned company shares directly or indirectly

⁴ This includes funds with 50 rating points or above.

⁵ This includes only funds that achieve 70–100 rating points.

through a mutual fund, pension plan, or other retirement fund in the respective countries. The sample composition is based on regional and demographic characteristics of the respective population.⁶ Table 2 presents the descriptive statistics of the full sample. Approximately 39% of the participants are female. The average age is 47 years, with ages ranging from 18 to 85. The median cognitive reflection score is 1. Approximately 71% of the participants are in a relationship, and 24% of them indicate that environmental issues have already influenced their investment decisions in the past.

Table 2. Summary statistics

This table contains the summary statistics of the participants. *Age* is participants' age, measured in years; *Female* is a dummy variable that is equal to one if a participant is female; *Share ownership* is a dummy variable that is equal to one if a participant owns equity shares, including shares held directly or indirectly through a mutual fund, pension plan, or other retirement fund; *Living standard* is a participant's self-assessed standard of living, measured on a scale from 1 (highest) to 6 (lowest); *In a relationship* is a dummy variable that is equal to one if a participant is married or in a relationship; *Cognitive reflection score* represents a participant's number of correct answers in the cognitive reflection test, and the score ranges from 0 (no correct answer in the cognitive reflection test) to 3 (correct answers to all three questions in the cognitive reflection test). *Environmental preference* is a dummy variable that is equal to one if environmental issues have already influenced a participant's investment decisions in the past.

Variable	N	Mean	Median	SD	Min	Max
Age	952	47.43	47	13.67	18	85
Female	953	0.39	0	0.49	0	1
Share ownership	953	1.00	1	0.00	1	1
Living standard	938	3.04	3	0.99	1	6
In a relationship	953	0.71	1	0.45	0	1
Cognitive reflection score	953	0.95	1	1.02	0	3
Environmental preference	953	0.24	0	0.43	0	1

3.3 Experimental Protocol

The survey experiment was conducted online in April 2016. After an introduction to the study, participants completed their conjoint tasks. Participants were presented with a hypothetical situation in which they were invited to consider equity fund profiles as investment opportunities. They were provided with 12 different choice sets, one at a time, each describing three equity funds based on random attribute-level combinations. Participants were then asked

⁶ The scope of the panel ranges from 43,000 members in Switzerland to 457,000 members in the United Kingdom. The provider uses diverse recruitment strategies, such as mailing, newsletters, e-mailing, affiliate networks, and social media. Participants are remunerated based on survey reward points programs and lotteries.

which of the three funds, if any, they would most likely choose if they were to consider investing in a fund. The choice task was followed by a questionnaire with the cognitive reflection test and control questions. After finishing the questionnaire, participants were debriefed and remunerated for participating in the study. The experimental session lasted about 15 minutes for each participant. The survey experiment was programmed and conducted with the Sawtooth Software module SSI Web 8. The online sessions were organized and administrated by Lightspeed GMI, Germany. Survey instructions in English are provided in Appendix A.

4 Results

In this section, we first outline how the presentation format of climate information influences investment decisions. We analyze the role of climate performance information in investors' fund choices and whether the tested label designs affect investment decisions differently (4.1). Subsequently, we present results regarding heterogeneity in investors' behavior, focusing on cognitive reflection abilities as an important determinant (4.2).

4.1 Relative Importance of Climate Performance Information

Table 3 displays the computed relative importance values of each attribute using the hierarchical Bayes (HB) estimation method. Importance values are based on estimated individual-level part-worth utilities. The relative importance values for each attribute are calculated by taking the difference between the highest and lowest part-worth utility estimated for the levels of each attribute and scaling the sum total differences of all attributes to 100 percent. The relative importance values for each attribute are calculated for participants individually and then averaged across the sample. Thus, each relative importance value indicates the respective attribute's contribution to the average participant's choice. The estimation results are provided separately for each label design. The results reveal that common financial market information is of the highest importance for retail investment decisions across the various label design treatments. This includes information representing the weighing of risk and return according to standard finance theories: the funds' risk class, costs, and financial performance. Taken together, the relative importance of risk class, financial performance, and costs total roughly 73% for the award subsample, 67% for the impact scale subsample, and 63% for the star rating subsample. The fund provider, varied by name recognition, is less important

for retail investors' choices; its relative importance amounts to 16%, 15%, and 14% in the respective subsamples.

Table 3. Relative importance of attributes across different climate label designs

This table contains estimated relative importance values based on zero-centered part-worth utilities at the attribute level and corresponding standard deviations. Individual part-worth utilities are drawn from HB estimations. The zero-centered part-worth utilities are provided in Appendix C. Results are shown separately for the different information designs tested. The total sample size is 953 participants.

Attributes	Award subsample (n=311)		Impact scale subsample (n=322)		Star rating subsample (n=320)	
	Relative importance	SD	Relative importance	SD	Relative importance	SD
Provider	15.99	12.33	15.06	11.30	14.04	11.41
Financial performance	21.83	12.60	20.17	12.04	19.10	11.58
Risk class	24.19	12.60	24.64	12.79	22.46	12.18
Fund costs	26.97	14.19	22.09	12.03	21.68	12.50
Climate performance	11.02	9.57	18.03	11.38	22.72	13.46

The relative importance of the climate performance attribute differs strongly among the three label design conditions and reaches a remarkably high level in the star rating subsample. The estimation results indicate a relative importance for climate performance of 11% for the average participant presented with a climate award and 18% for the average participant provided with a climate impact scale. When a star rating of the funds' climate performance is shown to participants, the relative importance increases to 23%. A comparison of the estimation values across the investment attributes indicates that in the star rating label condition, climate performance seems to be of considerably higher importance for individuals' investment choices than financial performance (19%) and around the same importance as risk class (22%) and fund costs (22%).

Different climate label designs lead to variation in the relative importance of climate performance for participants' choices and can lead to shifts in the importance rank of the considered fund attributes. Building on the findings by Hainmueller, Hangartner, and Yamamoto (2015), we interpret our results derived from hypothetical choices to be predictors of actual behavior in the retail investment context. Correspondingly, we argue that in practice, different climate label designs would lead to variation in real-world investment decisions and so would differ in their potential to influence climate-friendly investing. In particular, among the three popular label designs tested, a star rating label would yield the highest emphasis on climate performance information in investors' decision-making.

4.2 Cognitive Reflection and Heterogeneity in Behavior

In this section, we turn to heterogeneity in the observed effects of the climate label on individual investment decisions. In particular, we investigate whether heterogeneity in cognitive reflection abilities leads to systematic differences in the relative importance of climate performance information across individuals.⁷

Figure 2 shows the attributes' relative importance values categorized by individuals' measured cognitive reflection score when participants are provided with a star rating climate label. The individual cognitive reflection score indicates the participant's total number of correct answers in the cognitive reflection test (Frederick, 2005). The cognitive reflection test is designed to assess participants' cognitive dual processes. It evaluates individuals' ability to suppress an intuitive and wrong system 1 answer in favor of a reflective and correct system 2 answer (see Frederick, 2005, for further details). The standard test includes three questions to assess participants' cognitive reflection capabilities. Thus, the possible score ranges from 0 (no question answered correctly; that is, a markedly intuitive system 1 decision-maker) to 3 (all questions answered correctly; that is, a markedly reflective system 2 decision-maker).

Figure 2 reveals substantial heterogeneity in the observed relative importance values across participants with different cognitive reflection abilities. First, the relative importance of financial performance information seems to increase with cognitive reflection abilities; that is, higher cognitive reflection scores are associated with a higher relative importance of financial performance information in individuals' investment choices. Second, the relative importance of climate performance information seems to be negatively associated with cognitive reflection abilities and decreases with higher cognitive reflection scores.

⁷ In this section, we base our analyses on the subsample of participants provided with the star rating climate label, since comparisons across the different information designs are not feasible and the star rating label shows the highest potential to emphasize climate performance information in our decision task (see Section 4.1).

Figure 2. Relative importance of attributes by investors' cognitive reflection scores

This figure displays estimated relative importance values based on zero-centered part-worth utilities at the attribute level. Individual part-worth utilities are drawn from HB estimations. Results are shown for participants provided with a star rating design, separated by cognitive reflection score. *Cognitive reflection score* represents a participant's number of correct answers in the cognitive reflection test, and the score ranges from 0 (no correct answers in the cognitive reflection test) to 3 (correct answers to all three questions in the cognitive reflection test). The sample size is 320 participants.

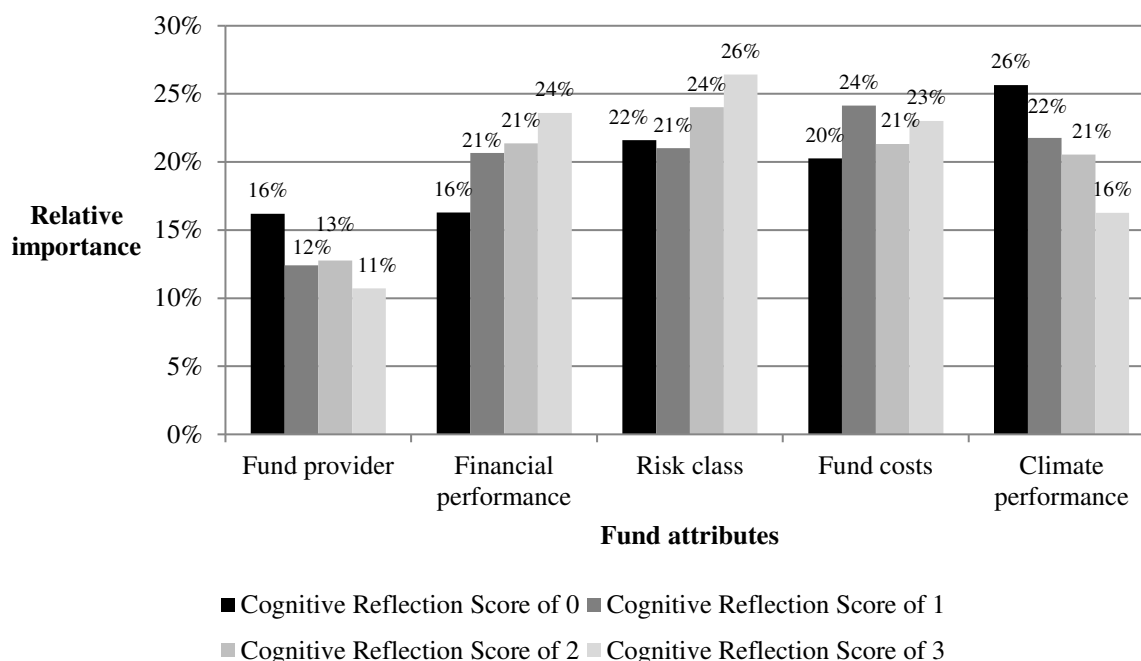


Table 4 reports in greater detail on the attributes' relative importance values categorized by individuals' cognitive reflection scores. On average, participants with a score of 0 ascribe a relative importance of 16% to financial performance in their decision-making, whereas financial performance holds a relative importance value of 24% for participants with a score of 3. The importance of climate performance seems to be higher for participants with lower cognitive reflection scores. This is in line with our hypothesis. The climate label tends to be most effective in increasing the importance of climate performance for the average individual with a cognitive reflection score of 0 (26%), followed by participants with scores of 1 (22%) and 2 (21%). Individuals who have a strong ability to suppress intuitive decision-making seem to be less affected (lowest relative importance score of 16%).

Table 4. Relative importance of attributes by investors' cognitive reflection scores

This table contains estimated relative importance values based on zero-centered part-worth utilities at the attribute level and corresponding standard deviations. Individual part-worth utilities are drawn from HB estimations. The zero-centered part-worth utilities are provided in Appendix D. Results are shown for participants provided with a star rating design, separated by cognitive reflection score. *Cognitive reflection score* represents a participant's number of correct answers in the cognitive reflection test, and the score ranges from 0 (no correct answers in the cognitive reflection test) to 3 (correct answers to all three questions in the cognitive reflection test). The sample size is 320 participants.

Attributes and levels	Cognitive reflection score of 0 (n=149)		Cognitive reflection score of 1 (n=75)		Cognitive reflection score of 2 (n=59)		Cognitive reflection score of 3 (n=37)	
	Relative importance	SD	Relative importance	SD	Relative importance	SD	Relative importance	SD
Provider	16.20	13.27	12.42	9.26	12.76	9.13	10.72	9.01
Financial performance	16.29	9.86	20.67	11.48	21.36	13.05	23.59	13.29
Risk class	21.59	12.24	21.01	12.55	24.02	11.05	26.41	12.27
Fund costs	20.27	12.16	24.13	13.61	21.32	11.45	23.02	12.74
Climate performance	25.65	14.49	21.77	12.73	20.54	12.21	16.27	9.03

Table 5. Differences between the relative importance of attributes by investors' cognitive reflection scores

This table contains differences between estimated relative importance values based on zero-centered part-worth utilities at the attribute level and corresponding mean values and standard deviations. Individual part-worth utilities are drawn from HB estimations. The zero-centered part-worth utilities are provided in Appendix D. Results are shown for participants provided with a star rating design, separated by cognitive reflection score. Kruskal-Wallis test statistics are provided to test for differences between groups based on participants' cognitive reflection score. *Cognitive reflection score* represents a participant's number of correct answers in the cognitive reflection test, and the score ranges from 0 (no correct answers in the cognitive reflection test) to 3 (correct answers to all three questions in the cognitive reflection test). The sample size is 320 participants.

Difference	Cognitive reflection score	N	Mean	SD	Kruskal-Wallis test
Climate performance minus financial performance	0	149	9.36	20.12	$p = 0.0001$
	1	75	1.10	20.11	
	2	59	-0.82	21.30	
	3	37	-7.31	18.79	
Climate performance minus risk class	0	149	4.06	22.08	$p = 0.0004$
	1	75	0.76	20.87	
	2	59	-3.47	17.28	
	3	37	-10.13	15.95	

Climate performance minus fund costs	0	149	5.38	21.25	$p = 0.0045$
	1	75	-2.35	21.82	
	2	59	-0.78	20.43	
	3	37	-6.75	17.48	
Climate performance minus fund provider	0	149	9.45	22.07	$p = 0.3866$
	1	75	9.36	14.72	
	2	59	7.78	15.19	
	3	37	5.56	12.11	

The SI literature relies on the dual nature of investor motives to argue that investors weigh their desire for financial returns and non-wealth returns (Beal et al., 2005; Renneboog, Ter Horst, & Zhang, 2011; Statman, 2004). The varying relative importance values of financial and climate performance information for different levels of cognitive reflection ability in our data might stem from this mechanism. Thus, we test whether the difference between the relative importance of climate performance and of monetary fund information (financial performance, risk class, costs, and provider) differs significantly across investors' cognitive reflection scores. Table 5 reports the mean values of the difference between the relative importance of the respective attributes, standard deviations, and Kruskal-Wallis test statistics. The results reveal that investor segments distinguished according to their cognitive reflection scores show significantly different relationships between the relative importance of climate performance and of financial performance, risk class, and fund costs, respectively, at the 1% level. A positive difference in relative importance reflects a dominance of climate performance information, a difference of around zero indicates equal weighting, and a negative difference demonstrates a dominance of the monetary attribute information in investors' decision-making. The findings document heterogeneity in these differences across investors' cognitive reflection abilities. The estimations reveal positive differences for low cognitive reflection abilities (score of 0) and negative differences for high cognitive reflection abilities (scores of 2 and 3) between the relative importance of climate performance and monetary fund information.

The strongest heterogeneity across investors' cognitive reflection scores seems to exist in the individual trading off between the fund's climate performance and financial performance, which represents the largest difference in relative importance (from 9.36 to -7.31). Table 6 displays the results of an OLS regression in which the dependent variable is the difference between an investor's individual relative importance of climate performance and financial performance. The main explanatory variable is an investor's intuitive choice score, which is a transposed variable of the participant's cognitive reflection score for a more stringent interpretation of the estimation results. We control for participants' risk preferences in general matters using a scale from 0 (lowest) to 10 (highest) (Dohmen et al., 2011), financial literacy via quiz-like questions (van Rooij et al., 2011), investment horizon, and self-stated environmental preferences. The coefficient on the intuitive choice variable is positive and significant at the 1% level, indicating that decision-makers who are more intuitive show a significantly larger positive difference between the relative importance values (i.e., place

significantly more weight on funds' climate performance compared to financial performance) (column 1). We further observe that participants who are more risk-tolerant and financially literate, as well as participants with longer investment horizons, place less weight on funds' climate performance compared to financial performance. In addition, participants with environmental preferences place significantly more relative weight on climate performance. We also control for participants' age, gender, living standard⁸, education, relationship status, occupation, and country (column 2). The inclusion of these sociodemographic variables does not change the conclusions drawn from our estimation results.

Interestingly, the findings indicate that individuals who do not yet invest in an environmentally friendly manner or have not even considered it nevertheless show higher relative importance of climate performance compared to financial performance if they are intuitive decision-makers. This further underlines the important role of cognitive reflection ability in our choice context irrespective of a participant's environmental preference. In support of this notion, participants' cognitive reflection abilities and environmental preferences seem to be rather uncorrelated (see Appendix E).

Table 6. Difference between the relative importance of climate and financial attributes

This table contains the coefficients and t-statistics (in parentheses) of OLS regressions in which the dependent variable is a participant's difference between the relative importance of climate performance and financial performance, calculated by subtracting the relative importance value of financial performance from the relative importance value of climate performance. Estimated relative importance values are based on zero-centered part-worth utilities at the attribute level. Individual part-worth utilities are drawn from HB estimations. The zero-centered part-worth utilities are provided in Appendix D. *Intuitive choice score* denotes a participant's tendency to decide intuitively, and the score ranges from 0 (three correct answers in the cognitive reflection test) to 3 (no correct answers in the cognitive reflection test); *Risk tolerance* is subjects' self-assessed general risk tolerance, measured on a scale from 0 (lowest) to 10 (highest); *Financial literacy* represents a participant's number of correct answers in a financial literacy test, measured on a scale from 0 to 11, where 0 indicates no correct answers and 11 indicates correct answers to all questions; *Investment horizon* indicates a participant's investment horizon, measured on a scale from 1 (one day) to 8 (more than 10 years); *Environmental preference* is a dummy variable that is equal to one if environmental issues have already influenced a participant's investment decisions in the past; *Age* is participants' age, measured in years; *Female* is a dummy variable that is equal to one if a participant is female; *Living standard* is a participant's self-assessed standard of living, measured on a scale from 1 (highest) to 6 (lowest); *University degree* is a dummy variable that is equal to one if a participant has completed a university degree; *In a relationship* is a dummy variable that is equal to one if a participant is married or in a relationship. Model 2 also controls for dummies of categorical variables indicating participants' *Occupation* and *Country*. Results are shown for participants provided with a star rating design. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Difference (Climate performance minus financial performance)	(1)	(2)
	3.559 *** (3.340)	3.088 *** (2.860)

⁸ This variable is intended to control for potential wealth effects in a participant's choice behavior.

Intuitive choice score					
Risk tolerance	-1.283 ***	(-2.890)	-1.055 **	(-2.200)	
Financial literacy	-1.471 ***	(-3.470)	-1.869 ***	(-4.090)	
Investment horizon	-1.427 **	(-2.170)	-1.149 *	(-1.680)	
Environmental preference	11.169 ***	(4.460)	12.076 ***	(4.800)	
Age			0.320 ***	(2.990)	
Female			2.113	(0.880)	
Living standard			-1.142	(-0.950)	
University degree			0.012	(0.010)	
In a relationship			4.020 *	(1.700)	
Constant	18.745 ***	(3.050)	10.791	(1.030)	
Occupation				Yes	
Country				Yes	
N	320		313		
Adj. R ²	0.20		0.26		

5 Discussion and Conclusion

This study aimed to shed light on how the presentation format of climate information influences investment decisions in the retail segment. Our findings clearly reveal that labeling can be effective, not only providing great potential from a theoretical perspective but living up to its potential in rigorous empirical testing. Thus, we provide empirical evidence for the effectiveness of climate labeling as a potential nudge for climate-friendly investing.

While we identify climate labeling as a powerful nudge, ethical concerns have been raised against the concept of nudging (e.g., Hausman & Welch, 2010; Schnellenbach, 2012; Sunstein, 2015). One major concern put forth is the question of whether nudges implemented by private and public institutions with the intention of affecting people's behavior also respect people's freedom of choice, as postulated by the concept of libertarian paternalism (Sunstein & Thaler, 2003).⁹ With respect to our climate label nudges, we argue that they seek to appeal to, and activate, system 1 to affect people's behavior, but they do not restrict investors' decision-making. In contrast to default settings, which also form a widely used nudging measure, labeling interventions still require people's active choice (Sunstein & Reisch, 2014). Another potential criticism is that nudges are manipulative if they lack transparency (their motivation is

⁹ Criticism is usually directed at the concept of nudge and the broader concept of libertarian paternalism, upon which the concept of nudge builds (Lehner et al., 2015). Libertarian paternalism is defined as "an approach that preserves freedom of choice but that encourages both private and public institutions to steer people in directions that will promote their own welfare" (Sunstein & Thaler, 2003, p. 1201).

hidden or concealed) (Sunstein, 2015). However, labels represent a transparent disclosure of factual information (Lehner et al., 2015; Sunstein, 2015), and climate labels in particular do not conceal the motivation of invoking public concern.

As a second contribution, we reveal the importance of investors' cognitive reflection abilities in the adaptation of SI practices and thereby establish the basis for novel segmentation analyses and future research. We find that investors' cognitive characteristics are associated with the effectiveness of climate labeling as a potential nudge. More specifically, our findings indicate that climate labeling is particularly effective among intuitive decision-makers and less effective among reflective decision-makers. The segmentation of investors has long been of interest in the SI literature, especially in terms of identifying characteristics that distinguish sustainable from conventional investors. Moving beyond criteria commonly used for segmenting investors in the SI literature, such as past behavior, attitudes, or demographic characteristics, we make the case for clustering investors according to their cognitive reflection abilities.

By establishing the foundation for novel segmentation analyses in the SI literature, we provide the starting point for exciting future research on noise trading and bubble formation in financial markets. As argued by Orlitzky (2013), investors might experience difficulties with interpreting public corporate sustainability information accurately, due to two underlying forces. First, the link between reported sustainability information and firms' economic fundamentals might still be unclear for market participants. Second, company managers have an incentive to report distorted sustainability information and may exaggerate the scope and scale of firms' actual corporate sustainability. Noise trading and excess market volatility is therefore argued to increase among all listed companies and excess market valuations among companies that are widely considered as being sustainable. However, recent experimental asset market research shows that investors' cognitive sophistication influences the occurrence of bubbles and crashes. In a study by Bosch-Rosa, Meissner, and Bosch-Domènech (2017), for instance, bubbles and crashes were only observed in experimental markets populated by subjects with low levels of cognitive sophistication. Excess market valuations of socially responsible firms may not only stem from the nature of companies' sustainability disclosure, as argued by Orlitzky (2013), but also from the cognitive characteristics of sustainable investors.

Our results, like results from experimental studies on human decision-making in general, are particularly vulnerable to the research method's external validity. Yet, a recent study by Hainmueller et al. (2015) provides an external validation test of vignette and conjoint analyses by making use of naturalization referendums in Switzerland as a natural experiment. They find that these survey experiments performed well in capturing the structural effects under examination. In particular, the study shows that the paired conjoint analysis with a non-student sample, like our study, precisely uncovers the relative importance patterns of the behavioral benchmark in the real world.¹⁰

Given the external validity of conjoint methodology, this study also provides valuable practical insights. Our experiment shows that a star rating seems to work best as a nudge for investment decisions that are more climate-friendly. By identifying differences in effectiveness among the three label designs tested, we provide important insights to policy-makers as well as other parties involved in the development and launch of climate labels. For instance, beginning in 2018, a new regulation on KIIDs for packaged retail and insurance-based investment products (PRIIPs) will take effect in the European Union (European Union, 2014). The new regulation is aimed at improving the quality of information provided and increasing the comparability of investment products for European retail investors by means of a standardized, simple, and accessible KIID. While the revised KIID does not capture social or environmental goals or outcomes up to the present, the potential inclusion of such aspects is explicitly mentioned in the legislative act. Furthermore, it is emphasized that “the Commission thoroughly considers developments relating to social and environmental investment products” (European Union, 2014, p. 5). Considering the possibility of upcoming regulative change, our analysis indicates that a star rating could be a promising way to communicate sustainability-related performance aspects in the KIID, at least for equity funds. In addition to the potential implications for mandatory labeling, this study may also provide practical support to existing and evolving voluntary labeling initiatives by identifying a feasible and effective label design option. For these initiatives, the impact scale is practically infeasible as funds are unlikely to implement a negative (red) rating by choice. Yet, as the star rating clearly outperformed in our

¹⁰ Hainmueller et al. (2015) argue that participant engagement is a key mechanism for external validity in such studies. In the paired condition, participants were more engaged in the survey and thus less prone to questionnaire satisficing.

study, communicating negative climate performance information does not appear to form a prerequisite for promoting climate-friendly investments.

Finally, our study also points toward important future research avenues. Specifically, we believe that two aspects of the potential effects of nudging require further research. First, our current understanding of the long-term effects of nudging is still fragmented (Croson & Treich, 2014). Regarding this study's findings, the question arises of whether labeling can nudge retail investors toward climate-friendly investments in the long run or whether this effect fades over time. Second, due to humans' tendency for moral self-regulation and self-licensing (Merritt, Effron, & Monin, 2010; Sachdeva, Iliev, & Medin, 2009), nudges may backfire by triggering unintended spillover effects. For example, as outlined by Merritt et al. (2010), past good deeds (e.g., pro-environmental behavior) can liberate (license) people to engage in bad deeds (e.g., environmentally harmful behavior) if individuals sense that they have done "enough." This raises the question of whether climate nudges in the investment context induce counterproductive climate effects in other behavioral domains of retail investors' lives, such as purchasing a new car with high fuel consumption. We therefore encourage future scholarly inquiry to explore the long-term effects of nudges through longitudinal studies and build on research settings capturing more than one behavioral domain. Furthermore, future research could focus on investors' underlying processes that occur between information input and the final investment decision. Methodologies for conducting decision-making research, such as conjoint analyses, are well suited to explore which aspects are most influential in determining decisions (Aiman-Smith, Scullen, & Barr, 2002; Oll, Hahn, Reimsbach, & Kotzian, 2016). However, conjoint analyses cannot provide comprehensive insights into the "black box" of underlying processes (Glaholt & Reingold, 2011). Scholars seeking to investigate this issue could, for instance, combine conjoint analyses with eye tracking (e.g., Meißner, Musalem, & Huber, 2016) or employ other process-tracing techniques, such as MouseLab or information display boards.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding Statement

Financial funding was provided by Climate-KIC, supported by the EIT, a body of the European Union.

References

- Aiman-Smith, L., Scullen, S. E., & Barr, S. H. (2002). Conducting studies of decision making in organizational contexts: A tutorial for policy-capturing and other regression-based techniques. *Organizational Research Methods*, 5, 388-414.
- Allcott, H., & Mullainathan, S. (2010). Behavioral science and energy policy. *Science*, 327, 1204-1205.
- Arnold, H. J., & Feldman, D. C. (1981). Social desirability response bias in self-report choice situations. *Academy of Management Journal*, 24, 377-385.
- Banerjee, A., & Solomon, B. D. (2003). Eco-labeling for energy efficiency and sustainability: A meta-evaluation of US programs. *Energy Policy*, 31, 109-123.
- Bao, J., & Ho, B. (2015). Heterogeneous effects of informational nudges on pro-social behavior. *The B.E. Journal of Economic Analysis & Policy*, 15, 1619-1655.
- Beal, D. J., Goyen, M., & Philips, P. (2005). Why do we invest ethically? *The Journal of Investing*, 14, 66-78.
- Benson, K. L., & Humphrey, J. E. (2008). Socially responsible investment funds. Investor reaction to current and past returns. *Journal of Banking & Finance*, 32, 1850-1859.
- Bertrand, M., & Mullainathan, S. (2001). Do people mean what they say? Implications for subjective survey data. *American Economic Review*, 91, 67-72.
- Bosch-Rosa, C., Meissner, T., & Bosch-Domènech, A. (2017). Cognitive bubbles. *Experimental Economics*. Advance online publication. doi: 10.1007/s10683-017-9529-0
- Campiglio, E. (2016). Beyond carbon pricing: The role of banking and monetary policy in financing the transition to a low-carbon economy. *Ecological Economics*, 121, 220-230.
- Costa, D. L., & Kahn, M. E. (2013). Energy conservation “nudge” and environmentalist ideology: evidence from a randomized residential field experiment. *Journal of the European Economic Association*, 11, 680-702.
- Covington, H. (2016). Investment consequences of the Paris climate agreement. *Journal of Sustainable Finance & Investment*. Advance online publication. doi: 10.1080/20430795.2016.1196556

- Croson, R., & Treich, N. (2014). Behavioral environmental economics: Promises and challenges. *Environmental and Resource Economics*, 58, 335-351.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9, 522-550.
- Dolan, P., Hallsworth, M., Halpern, D., King, D., Metcalfe, R., & Vlaev, I. (2012). Influencing behaviour: The mindspace way. *Journal of Economic Psychology*, 33, 264-277.
- Døskeland, T., & Pedersen, L. J. T. (2016). Investing with brain or heart? A field experiment on responsible investment. *Management Science*, 62, 1632-1644.
- Drescher, L. S., Roosen, J., & Marette, S. (2014). The effects of traffic light labels and involvement on consumer choices for food and financial products. *International Journal of Consumer Studies*, 38, 217-227.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, 49, 709-724.
- European Commission. (2013). *Applying behavioural sciences to EU policy-making*. Retrieved from <http://ftp.jrc.es/EURdoc/JRC83284.pdf>
- European Union. (2014). *Regulation (EU) No 1286/2014 of the European Parliament and of the Council of 26 November 2014 on key information documents for packaged retail and insurance-based investment products (PRIIPs)*. Retrieved from <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2014:352:FULL&from=EN>
- EUROSIF. (2014). *European SRI Study 2014*. Retrieved from <http://www.eurosif.org/wp-content/uploads/2014/09/Eurosif-SRI-Study-20142.pdf>
- Foster, F. D., Ng, J., & Wee, M. (2015). Presentation format and financial literacy: Accessibility and assessability of retirement savings statements. *Journal of Consumer Affairs*, 49, 519-549.
- Frank, R. (1987). If Homo Economicus could choose his own utility function, would he want one with a conscience? *The American Economic Review*, 77, 593-604.
- Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, 19, 25-42.

- Glaholt, M. G., & Reingold, E. M. (2011). Eye movement monitoring as a process tracing methodology in decision making research. *Journal of Neuroscience, Psychology, and Economics*, 4, 125-146.
- Green, P. E., Krieger, A. M., & Wind, Y. (2000). Thirty Years of Conjoint Analysis: Reflections and Prospects. *Technical Report*, Wharton School.
- Green, P. E., & Rao, V. R. (1971). Conjoint measurement for quantifying judgmental data. *Journal of Marketing Research*, 8(3), 355-363.
- Gutsche, G., Dai, J., & Zwergel, B. (2015). *The importance of sustainability and transparency labels in investment decisions between conventional and sustainable mutual funds*. Retrieved from: http://www.webmeets.com/files/papers/eaere/2015/1096/Full_Paper-EAERE_April.pdf
- Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences*, 112, 2395-2400.
- Hausman, D. M., & Welch, B. (2010). Debate: To nudge or not to nudge. *Journal of Political Philosophy*, 18, 123-136.
- Heinzle, S. L., & Wüstenhagen, R. (2012). Dynamic adjustment of eco-labeling schemes and consumer choice: The revision of the EU energy label as a missed opportunity? *Business Strategy and the Environment*, 21, 60-70.
- Hieke, S., & Taylor, C. R. (2012). A critical review of the literature on nutritional labeling. *Journal of Consumer Affairs*, 46, 120-156.
- Hummels, H., & Timmer, D. (2004). Investors in need of social, ethical, and environmental information. *Journal of Business Ethics*, 52, 73-84.
- Hüsser, A. (2015). The role of investors' objective financial knowledge on the assessment of risk disclosures in mutual fund advertisements. *Journal of Financial Services Marketing*, 20, 5-22.
- Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, 93, 1449-1475.

- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). New York: Cambridge University Press.
- Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 267-293). New York: Cambridge University Press.
- Kozup, J., Howlett, E., & Pagano, M. (2008). The effects of summary information on consumer perceptions of mutual fund characteristics. *Journal of Consumer Affairs*, 42, 37-59.
- Lehner, M., Mont, O., & Heiskanen, E. (2015). Nudging: A promising tool for sustainable consumption behaviour? *Journal of Cleaner Production*, 134, 166-177.
- Lohrke, F. T., Holloway, B. B., & Woolley, T. W. (2010). Conjoint analysis in entrepreneurship research: A review and research agenda. *Organizational Research Methods*, 13, 16-30.
- Maines, L., & McDaniel, L. (2000). Effects of comprehensive-income characteristics on nonprofessional investors' judgments: The role of financial-statement presentation format. *Accounting Review*, 75, 179-207.
- Meißner, M., Musalem, A., & Huber, J. (2016). Eye tracking reveals processes that enable conjoint choices to become increasingly efficient with practice. *Journal of Marketing Research*, 53, 1-17.
- Merritt, A. C., Effron, D. A., & Monin, B. (2010). Moral self-licensing: When being good frees us to be bad. *Social and Personality Psychology Compass*, 4, 344-357.
- Momsen, K., & Stoerk, T. (2014). From intention to action: Can nudges help consumers to choose renewable energy? *Energy Policy*, 74, 376-382.
- Ölander, F., & Thøgersen, J. (2014). Informing versus nudging in environmental policy. *Journal of Consumer Policy*, 37, 341-356.
- Oll, J., Hahn, R., Reimsbach, D., & Kotzian, P. (2016). Tackling complexity in business and society research: The methodological and thematic potential of factorial surveys. *Business & Society*. Advance online publication. doi: 10.1177/0007650316645337
- Orlitzky, M. (2013). Corporate social responsibility, noise, and stock market volatility. *The Academy of Management Perspectives*, 27, 238-254.

- Paetzold, F., & Busch, T. (2014). Unleashing the powerful few: Sustainable investing behaviour of wealthy private investors. *Organization & Environment*, 27, 347-367.
- Persky, J. (1995). Retrospectives: The ethology of homo economicus. *The Journal of Economic Perspectives*, 9, 221-231.
- Pilaj, H. (2015). The choice architecture of sustainable and responsible investment: Nudging investors toward ethical decision-making. *Journal of Business Ethics*. Advance online publication. doi: 10.1007/s10551-015-2877-9
- Renneboog, L., Ter Horst, J., & Zhang, C. (2011). Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds. *Journal of Financial Intermediation*, 20, 562-588.
- Rhodes, M. J. (2010). Information asymmetry and socially responsible investment. *Journal of Business Ethics*, 95, 145-150.
- Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science*, 20, 523-528.
- Sacks, G., Rayner, M., & Swinburn, B. (2009). Impact of front-of-pack 'traffic-light' nutrition labeling on consumer food purchases in the UK. *Health Promotion International*, 24, 344-352.
- Schnellenbach, J. (2012). Nudges and norms. On the political economy of soft paternalism. *European Journal of Political Economy*, 28, 266-277.
- Schrader, U. (2006). Ignorant advice – customer advisory service for ethical investment funds. *Business Strategy and the Environment*, 15, 200-214.
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, 54, 93-105.
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69, 99-118.
- Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and Brain Sciences*, 23, 701-717.

- Statman, M. (2004). What do investors want? *The Journal of Portfolio Management*, 30, 153-161.
- Sunstein, C. R. (2015). Nudges and choice architecture: Ethical considerations. *Yale Journal on Regulation*. Forthcoming.
- Sunstein, C. R., & Reisch, L. A. (2014). Automatically green: Behavioral economics and environmental protection. *Harvard Environmental Law Review*, 38, 127-158.
- Sunstein, C., & Thaler, R. (2003). Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review*, 70, 1159-1202.
- Teisl, M. F. (2003). What we may have is a failure to communicate: Labeling environmentally certified forest products. *Forest Science*, 49, 668-680.
- Thaler, R. H., & Sunstein, C. R. (2003). Libertarian paternalism. *The American Economic Review*, 93, 175-179.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge. Improving decisions about health, wealth, and happiness*. New Haven & London: Yale University Press.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- UNEP FI. (2009). *Financing global deal on climate change*. Retrieved from <http://www.unepfi.org/fileadmin/documents/FinancingGlobalDeal.pdf>
- UNFCCC. (2015). *Adoption de l'Accord de Paris*. Retrieved from <http://unfccc.int/resource/docs/2015/cop21/fre/l09f.pdf>
- Van Rooij, M., Lusardi, A., & Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, 101, 449-472.
- World Bank. (2015). *Mind, society, and behavior. World development report 2015*. Retrieved from <http://www.worldbank.org/en/publication/wdr2015>

APPENDIX A

Instructions in English

In the following investment decision task, you will be provided with a selection of different equity funds. Each fund will be described along five categories.

Fund provider

The fund provider is the investment company that issues and manages the fund.

Financial performance

The financial performance is based on the fund's historical annual return (in percentage of invested capital) over the past three years. Furthermore, the fund's financial performance will be shown in comparison to a benchmark. A benchmark is a standard against which the performance of a fund can be measured. Generally, broad market stock indexes are used for this purpose (e.g., S&P 500, DAX).

Risk class

The fund's risk class reflects the volatility of its returns over the past three years. It is measured by the historical standard deviation of the fund's returns and will be expressed in risk classes from 1 (low risk) to 7 (high risk).

Fund costs

The equity fund's costs include total annual costs (all fees and expenses). Costs will be described relative to invested capital (in %).

Climate performance

Funds can invest in companies that differ in terms of their contribution to global climate change. They can invest in more climate-friendly or less climate-friendly companies. For the following decision task, the funds are rated according to this climate performance.

[Label description varied across treatments; the star rating treatment is described below.]

A star rating is used to show the respective performance for the equity funds. Imagine that a fund can score a maximum of 100 points for its positive contribution to mitigating climate change (e.g., by investing in renewable energies). This score is assigned in comparison to an industry benchmark of 50 points. For a climate rating below 50 points, no star is awarded.

The star rating can be interpreted according to the following:

- 1 star: 51–60 rating points
- 2 stars: 61–70 rating points
- 3 stars: 71–80 rating points

- 4 stars: 81–90 rating points
- 5 stars: 91–100 rating points

[The award treatment is described below.]

A climate award is used to show you the respective performance of the equity funds. Imagine a fund can score a maximum of 100 points for its positive contribution to mitigating climate change (e.g., by investing in renewable energies). This score is assigned in comparison to an industry benchmark of 50 points. The climate award honors funds scoring 70 points or more. Funds with a climate rating below 70 points do not receive this award.

[The impact scale treatment is described below.]

A climate scale is used to show the respective performance of the equity funds. Imagine that a fund can score a maximum of 100 points for its positive contribution to mitigating climate change (e.g., by investing in renewable energies). This score is assigned in comparison to an industry benchmark of 50 points. The climate scale will show you how many points the respective fund scored (between 0 and 100).

The scale uses a visualization with the following color coding:

- Red: 1–33 rating points
- Orange: 34–66 rating points
- Green: 67–100 rating points

APPENDIX B

Table B.1. Subsample statistics

A. By country	N	Percent	Cum.
France	157	16.47	16.47
Germany	164	17.21	33.68
Italy	162	17.00	50.68
Switzerland	161	16.90	67.58
Sweden	159	16.68	84.26
United Kingdom	150	15.74	100
Total	953	100	
B. By treatment	N	Percent	Cum.
Award	311	32.63	66.21
Star rating	320	33.58	33.58
Impact scale	322	33.79	100
Total	953	100	

APPENDIX C**Table C.1. Relative importance values and average utilities for the award label design**

This table contains estimated relative importance values and zero-centered part-worth utilities at the attribute level and corresponding standard deviations. Individual part-worth utilities are drawn from HB estimations. Results are shown for participants provided with an award design. The sample size is 311 participants.

Attributes and levels	Relative importance	SD	Lower 95% CI	Upper 95% CI	Average utilities	SD	Lower 95% CI	Upper 95% CI
Provider	15.99	12.33	14.62	17.36				
high					20.86	47.75	15.55	26.16
medium					-6.22	36.87	-10.32	-2.12
low					-14.64	37.54	-18.81	-10.47
Financial performance	21.83	12.60	20.43	23.23				
9.5					36.53	47.33	31.27	41.79
7.5					6.75	29.96	3.42	10.08
5.5					-43.28	46.30	-48.42	-38.13
Risk class	24.19	12.60	22.79	25.59				
high					-50.18	49.49	-55.68	-44.68
medium					21.49	41.26	16.91	26.08
low					28.69	46.72	23.50	33.88
Fund costs	26.97	14.19	25.40	28.55				
2.25					-30.82	72.10	-38.84	-22.81
1.5					6.98	32.13	3.41	10.55
0.75					23.84	67.71	16.31	31.36
Climate performance	11.02	9.57	9.96	12.09				
award					19.99	30.56	16.60	23.39
no award					-19.99	30.56	-23.39	-16.60
None					-33.99	179.80	-53.97	-14.01

Table C.2. Relative importance values and average utilities for the impact scale label design

This table contains estimated relative importance values and zero-centered part-worth utilities at the attribute level and corresponding standard deviations. Individual part-worth utilities are drawn from HB estimations. Results are shown for participants provided with an impact scale design. The sample size is 322 participants.

Attributes and levels	Relative importance	SD	Lower 95% CI	Upper 95% CI	Average utilities	SD	Lower 95% CI	Upper 95% CI
Provider high	15.06	11.30	13.83	16.29	13.13	45.7	8.13	18.12
medium					1.69	38.86	-2.55	5.94
low					-14.82	32.99	-18.42	-11.21
Financial performance	20.17	12.04	18.86	21.49				
9.5					32.33	44.94	27.42	37.24
7.5					4.68	26.90	1.74	7.61
5.5					-37.01	46.82	-42.12	-31.89
Risk class	24.64	12.79	23.24	26.04				
high					-56.26	47.28	-61.42	-51.09
medium					20.68	39.11	16.40	24.95
low					35.58	43.29	30.85	40.31
Fund costs	22.09	12.03	20.78	23.41				
2.25					-21.98	59.24	-28.45	-15.51
1.5					10.05	31.18	6.64	13.46
0.75					11.93	57.55	5.65	18.22
Climate performance	18.03	11.38	16.79	19.27				
Green area					18.51	41.02	14.03	22.99
Yellow area					15.96	30.32	12.65	19.27
Red area					-34.47	42.87	-39.16	-29.79
None					-30.98	163.2	-48.8	-13.15

Table C.3. Relative importance values and average utilities for the star label design

This table contains estimated relative importance values and zero-centered part-worth utilities at the attribute level and corresponding standard deviations. Individual part-worth utilities are drawn from HB estimations. Results are shown for participants provided with a star rating design. The sample size is 320 participants.

Attributes and levels	Relative importance	SD	Lower 95% CI	Upper 95% CI	Average utilities	SD	Lower 95% CI	Upper 95% CI
Provider	14.04	11.41	12.79	15.29				
high					16.97	43.31	12.22	21.71
medium					-1.66	35.22	-5.52	2.20
low					-15.31	31.97	-18.81	-11.80
Financial performance	19.10	11.58	17.83	20.37				
9.5					33.43	40.72	28.97	37.90
7.5					2.31	27.95	-0.76	5.37
5.5					-35.74	43.01	-40.46	-31.03
Risk class	22.46	12.18	21.13	23.79				
high					-46.43	46.12	-51.48	-41.38
medium					19.14	35.62	15.24	23.05
low					27.28	47.43	22.09	32.48
Fund costs	21.68	12.50	20.32	23.05				
2.25					-30.61	56.75	-36.83	-24.40
1.5					10.50	25.42	7.72	13.29
0.75					20.11	55.02	14.08	26.14
Climate performance	22.72	13.46	21.24	24.19				
5 stars					46.47	40.27	42.06	50.88
3 stars					7.79	28.09	4.71	10.87
1 star					-54.26	42.40	-58.91	-49.61
None					-27.44	146.4	-43.48	-11.39

APPENDIX D**Table D.1. Relative importance values and average utilities by investors' cognitive reflection scores**

This table contains estimated relative importance values and zero-centered part-worth utilities at the attribute level and corresponding standard deviations. Individual part-worth utilities are drawn from HB estimations. Results are shown for participants provided with a star rating design separated by cognitive reflection score.

Cognitive reflection score represents a participant's number of correct answers in the cognitive reflection test, and the score ranges from 0 (no correct answers in the cognitive reflection test) to 3 (correct answers to all three questions in the cognitive reflection test). The sample size is 320 participants.

Attributes and levels	Cognitive reflection score of 0		Cognitive reflection score of 1		Cognitive reflection score of 2		Cognitive reflection score of 3		Cognitive reflection score of 0		Cognitive reflection score of 1		Cognitive reflection score of 2		Cognitive reflection score of 3	
	Rel. import.	SD	Rel. import.	SD	Rel. import.	SD	Rel. import.	SD	Av. util.	SD	Av. util.	SD	Av. util.	SD	Av. util.	SD
Provider	16.20	13.27	12.42	9.26	12.76	9.13	10.72	9.01								
high									18.36	49.37	21.43	34.88	7.57	40.44	17.3	35.71
medium									0.95	41.45	-6.07	26.96	2.18	31.66	-9.33	25.77
low									-19.31	37.13	-15.36	26.59	-9.74	27.89	-7.97	22.54
Financial performance	16.29	9.86	20.67	11.48	21.36	13.05	23.59	13.29								
9.5									22.09	36.06	37.65	40.57	42.71	41.81	55.80	43.88
7.5									3.25	28.61	2.61	28.46	2.73	27.37	-2.75	25.56
5.5									-25.33	40.95	-40.25	44.57	-45.44	44.50	-53.05	35.74
Risk class	21.59	12.24	21.01	12.55	24.02	11.05	26.41	12.27								
high									-40.11	49.12	-40.91	44.33	-57.44	40.33	-65.51	38.33
medium									18.26	36.52	17.84	36.13	22.85	31.92	19.44	37.55
low									21.85	46.56	23.06	49.95	34.60	44.35	46.06	45.98
Fund costs	20.27	12.16	24.13	13.61	21.32	11.45	23.02	12.74								
2.25									-20.76	57.12	-29.02	64.85	-47.58	43.16	-46.48	48.43
1.5									10.82	25.28	9.00	27.23	15.29	23.24	4.63	25.08
0.75									9.94	54.02	20.01	63.05	32.29	44.43	41.85	48.20
Climate performance	25.65	14.49	21.77	12.73	20.54	12.21	16.27	9.03								
5 stars									54.38	43.29	40.46	39.45	43.75	36.13	31.12	27.78
3 stars									10.06	28.18	7.56	31.97	3.18	25.16	6.45	23.39

1 star	-64.44	41.21	-48.02	46.30	-46.94	40.83	-37.57	31.15
None	-42.15	159.88	-36.13	153.16	-2.26	106.26	9.29	122.26

APPENDIX E

Table E.1. Correlation of participants' cognitive reflection score and environmental preference

This table contains pairwise correlations based on the full sample (953 participants). *Cognitive reflection score* represents a participant's number of correct answers in the cognitive reflection test, and the score ranges from 0 (no correct answers in the cognitive reflection test) to 3 (correct answers to all three questions in the cognitive reflection test); *Environmental preference* is a dummy variable that is equal to one if environmental issues have already influenced a participant's investment decisions in the past. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Cognitive reflection score	Environmental preference
Cognitive reflection score	1.000	
Environmental preference	-0.083	1.000

4 Third Article

Meißner, M. & Oll, J. (2019). The Promise of Eye-Tracking Methodology in Organizational Research: A Taxonomy, Review, and Future Avenues. *Organizational Research Methods*, 22(2), 590-617. [VHB-JOURQUAL: A]

The Promise of Eye-Tracking Methodology in Organizational Research: A Taxonomy, Review, and Future Avenues

Abstract

Technological advances in recent years have greatly lowered the barriers for using eye tracking (ET) as a research tool in laboratory and field settings. However, despite its potential and widespread application in other disciplines, the use of ET in organizational research remains sparse. This article therefore aims to introduce ET, and thus a new mode of behavioral data, to the field of organizational research. Based on a synthesis of prior literature, we propose an integrative taxonomy that unravels the methodological potential of ET as well as its scope of application. Building on our proposed taxonomy, we systematically review the use of ET in leading management journals and reflect on the current state of research. We further illustrate future avenues for ET in the domains of strategic management, entrepreneurship, and human resources to contribute to the method's future dissemination and to the advancement of organizational science as well.

Keywords: Eye tracking, attention, taxonomy, literature review, behavioral data

Introduction

Although openness to adopting methodologies from other disciplines is often seen as a characteristic or strength of organizational science (Ray & Smith, 2012; Uhlmann et al., 2012), the standard toolbox of organizational researchers has not changed substantially in recent decades (Aguinis, Pierce, Bosco, & Muslin, 2009; Podsakoff & Dalton, 1987). With respect to the measurement of constructs, self-report methods (e.g., surveys, interviews) continue to be part of the common repertoire on which organizational scholars draw extensively (Chaffin et al., 2015; Waldman, Wang, & Fenters, 2016). The limitations and problems associated with self-reported measures are, however, well-documented (e.g., social desirability bias, halo and leniency effects, nonconscious activation) (Donaldson & Grant-Vallone, 2002; Nisbett & Wilson, 1977; Podsakoff & Organ, 1986). Inside and outside the field of organizational research, explicit calls have therefore been made for more frequent utilization of behavioral data and building on multi-method data sources (Baumeister, Vohs, & Funder, 2007; Bono & McNamara, 2011; Chaffin et al., 2015; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003).

In this context, eye tracking (ET) represents one promising source of behavioral data for organizational researchers. In an experimental set-up, ET systems allow researchers to record the movements of a participant's eyes during behavioral processes, thus providing "insights into the cognitive processes underlying a wide variety of human behaviors" (Ashby, Johnson, Krajbich, & Wedel, 2016, p. 96). ET researchers most often use eye movements as a proxy for attention that is directed to stimuli (as further discussed in the following section). However, ET can also be applied to investigate other psychological constructs, such as arousal, cognitive load, or perceptual fluency.

Although widely employed in other disciplines, such as psychology and marketing (see, for example, the review papers by Wedel, 2015 and Rayner, 2009), ET is still rarely used in organizational research. We examined the same 30 journals examined in previous reviews (e.g., Aguinis & Bradley, 2014; Podsakoff, MacKenzie, Podsakoff, & Bachrach, 2008), as these journals represent influential outlets across the principal domains of management, including top-tier micro-organizational journals. Overall, we found only 15 empirical ET studies. The identified articles were published in four different journals and predominantly pertain to the field of

marketing. The great majority of studies were published in the very recent past, which may be interpreted as a first sign of ET slowly gaining ground in the organizational sciences.¹

Still, the paucity of ET studies in the reviewed journals is surprising as other disciplines have broadly applied ET in areas of high relevance to organizational research, such as information search and decision-making, learning, training, and expertise. Furthermore, rapid advances in information and communication technology have fundamentally increased the amount of information that is available to organizational stakeholders. As individuals' attention determines which pieces of information will be recognized, prioritized, and processed, the question of how organizational actors direct their attention has never been more salient and pressing (van Knippenberg, Dahlander, Haas, & George, 2015). Given the capability of ET to unravel attention processes, the methodology has much to contribute to organizational research in the "information revolution age" (Kuvaas, 2002, p. 977). In addition, technological advances in recent years have greatly lowered the barriers for using ET as a research tool in laboratory and field settings (Wedel, 2015). Not only are the costs for ET equipment on a steady decline, but the data quality and ease of use (e.g., programming, set-up, and data analysis) have also improved considerably over the years (Ashby et al., 2016).

In light of the above, we argue that the time is right to expand the standard methodological tool kit of organizational scholars by bringing ET to their minds and hands. The objective of our article is therefore to introduce ET, and thus a new mode of behavioral data, to the field of organizational science. The contribution of this article unfolds as follows. We begin by providing background information on attention research and ET, and then propose an integrative taxonomy for ET research based on a synthesis of prior literature. Our taxonomy contributes by unraveling the methodological potential of ET as well as its scope of application. Our taxonomy further contributes by posing guiding questions that interested researchers need to consider when conducting ET-based research. Building on our proposed taxonomy, we proceed with a systematic review of ET studies in leading management journals and illustrate future avenues for ET in strategic management, entrepreneurship, and human resources. By stimulating organizational readers' imagination and motivation for (novel) uses of ET, we hope to contribute to the method's future dissemination and to the advancement of organizational science alike. We

¹ Details about the application of ET, both inside and outside the field of organizational research, are provided at a later stage in the article.

close with an outline of key methodological references that interested researchers are encouraged to draw on for methodological guidance and practical advice.

Basics on Attention Research and Eye Tracking

Visual attention is regularly defined as selectivity in perception (Orquin & Mueller Loose, 2013) and is the psychological construct of key interest in eye-movement research (Wedel & Pieters, 2008a). Human visual perception consists of three parts: foveal, parafoveal, and peripheral vision. The fovea is the central part of the retina and has the highest density of sensory neurons and therefore the greatest acuity. The acuity decreases in the area surrounding the fovea, i.e., in the parafoveal area, and even more in the peripheral area. In order to see a stimulus more clearly, humans have to move their eyes to bring the respective stimulus into the foveal region, which has the greatest visual resolution. In many situations, humans will bring objects into the foveal region because they want to focus on them (Duchowski, 2017). If the eyes are relatively stable and “rest on” a certain stimulus that is in the foveal region, that instance is called a *fixation*. A rapid movement of the eyes between two consecutive fixations is called a *saccade*. Research has shown that humans can only acquire information during a fixation and not during a saccade (Rayner, 1998), as the brain blocks visual processing during eye movements in a way that neither the motion of the eye nor the gap in visual perception is noticeable to the individual.

When humans bring their attention to stimuli that are initially outside the fovea, the process is called *overt visual attention*. While overt visual attention is likely the default state in many situations, it is also possible that humans pay attention to objects currently outside the fovea (Orquin & Mueller Loose, 2013). This process is then called *covert visual attention*. Researchers often make the assumption that the number of fixations on an area of interest (AOI) is an indicator for how much attention the individual directs to the respective AOI. When researchers study the psychological construct of attention and count the number of fixations to quantify attention, they base their analysis on the “eye-mind assumption” (Just & Carpenter, 1980), which assumes that information that is seen is also cognitively processed. Although the assumption is “generally validated” (Orquin & Mueller Loose, 2013, p. 193), there are also empirical findings indicating that eye movements do not necessarily reflect cognitive processes in certain decision contexts (see Anderson, Bothell, & Douglass, 2004). In simple discrimination tasks, the locus of attention and eye location can easily be decoupled (Posner, 1980), and some

researchers (e.g., Charness, Reingold, Pomplun, & Stampe, 2001; Reingold, Charness, Pomplun, & Stampe, 2001) have shown that experts can extract information from widely distanced and parafoveal regions. Moreover, in repeated tasks, participants might get better at extracting information outside of the fovea over time. Because the focus of the gaze is not always a mirror of what is being processed, Ashby et al. (2016, p. 98) conclude: “while eye movements can be used to infer a great deal about information uptake and cognitive processes, they do not necessarily have a one-to-one relationship with specific cognitive processes.” The authors suggest combining ET with neural measures, such as electroencephalography and fMRI to more holistically measure attention. As the eye-mind assumption is often taken for granted, we encourage ET researchers to critically reflect on whether making the assumption is valid. Especially in reading tasks (Lai et al., 2013) and for complex stimuli (Rayner, 1998), research has shown that eye movement and attention are closely linked.

In order to measure eye movements, modern ET systems use video images of the eyes to determine the so-called point of regard. Most systems use the corneal reflection method: infrared light from LEDs embedded in cameras is directed into the eyes. The LED light enters the participant’s retinas, and a large proportion of the non-visible light is reflected. Based on the reflected light, the image-processing software of the eye tracker identifies the center of the pupil and the corneal reflection and subsequently determines the point of regard using trigonometric calculations.

Two basic ET systems can be distinguished: desktop-based (also called remote or stationary) and mobile ET systems. Desktop-based ET systems record participants’ eye movements from a fixed position. In most experimental situations, participants sit in front of a computer monitor on which the stimuli are presented. Mobile ET systems, in contrast, allow participants to freely move around as eye movements are recorded using lightweight ET glasses that are worn like regular glasses. Mobile ET is therefore better suited to investigating the interaction of users in natural and virtual environments and thus to examining gaze behavior *in situ* (Wästlund, Otterbring, Gustafsson, & Shams, 2015). To learn more about the technical differences between desktop and mobile ET, we refer the reader to Duchowski (2017).²

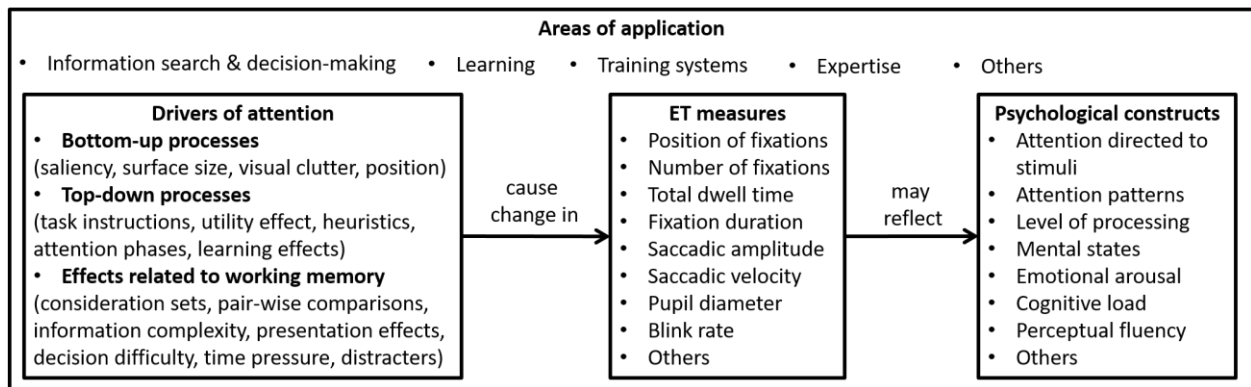
² We further discuss the advantages and disadvantages of using mobile ET in the section Methodological Guidance on Eye Tracking.

Toward an Integrative Taxonomy of Eye-Tracking Research

Although ET research is predominantly published outside organizational journals, the broad application of ET has provided important insights into organizationally relevant areas. Taking stock of the state of ET literature thus forms a logical first step in carving out the methodological potential of ET in organizational research. Synthesizing past ET literature, several major taxonomies can be identified that have been utilized for structuring and analyzing the existing body of ET research. Up to the present, however, the respective reviews frequently tend to approach ET from a specific angle.

Duchowski (2002), for example, reviewed ET literature pertaining to scientific disciplines such as marketing, psychology, and information systems and identified central areas of application. In psychology, for instance, researchers have used ET to investigate reading and scene perception as exemplary areas of application. Holmqvist et al. (2011), in contrast, provided a systematic overview of ET measures, such as the number of fixations or pupil diameter. The number of fixations, for example, is supposed to give an indication of how much attention individuals direct to certain stimuli. Reviewing such measures aids the reader's understanding of how ET data can be analyzed. A third angle is to review psychological constructs that make use of ET measures. Rosch and Vogel-Walcutt (2013), for example, reviewed the literature to determine which ET measures had previously been adapted to quantify the psychological construct of cognitive load. A fourth angle is to review the ET literature with a focus on drivers of attention that cause changes in psychological constructs, as carried out by Orquin and Mueller Loose (2013).

Overall, the review approaches outlined above focus on ET measures, psychological constructs, drivers of attention, or areas of application. While current taxonomies provide valuable insights into the specifics of ET research, the applied foci come at the expense of methodological scope. We therefore argue that an integrative taxonomy is needed to fully grasp (and tap) the methodological promise of ET in organizational research—particularly among scholars with little or no prior ET experience. More specifically, we posit that a robust taxonomy should allow for reflection on all four of the aforementioned foci. Extending prior work, we thus advance an integrative ET taxonomy as depicted in Figure 1 and outlined in more detail below.

Figure 1 *A taxonomy for eye tracking research*

Our proposed taxonomy first supports interested scholars by introducing standard ET vocabulary. Our practical classification system for ET studies is also intended to help researchers comprehend the methodological potential of ET and its scope of application. Figure 1 illustrates how the different dimensions of the taxonomy are interconnected. The drivers of attention cause respective changes in ET measures which may reflect changes in psychological constructs. Alternative ways of investigating these psychological constructs may exist, implying that ET measures represent just one possible methodological approach to investigating a given construct. Our integrative taxonomy further captures several major areas of ET application to highlight the benefits of cross-disciplinary literature search. For example, the question of how learning affects gaze behavior is relevant in several different disciplines, including marketing, psychology, or pedagogy. Thinking about whether these areas of application are relevant for a particular research project should lead to an extended literature search outside the immediately relevant discipline and may point toward new research prospects and questions. Finally, we consider the illustrated framework helpful because it supports researchers in the process of conducting ET experiments by posing guiding questions that potential users need to address. These questions are:

- In which *area of application* is the research situated?
- Which *drivers of attention* are investigated? Does the researcher manipulate these drivers of attention? If yes, how is the manipulation carried out?
- Which *ET measures* are used? Which psychological constructs do the ET measures reflect?
- Which *psychological constructs* are investigated, if any?

The framework illustrated in Figure 1 also serves as a tool kit that allows scholars to reflect on existing ET studies. Potential questions are:

- Do researchers from other disciplines investigate similar research questions using ET in their respective *area of application*?
- Which alternative explanations (*drivers of attention*) exist that could explain observed differences in gaze behavior?
- Would the empirical results change if alternative *ET measures* were used?

In the following, we elaborate in greater detail the constituent elements of our taxonomy, starting with an overview of drivers of attention. We subsequently introduce key ET measures and important psychological constructs that can be investigated using ET. We close with an overview of major areas of application in which ET has been used across academic disciplines up to the present.

Drivers of Attention

Two basic research questions lay the foundation for identifying drivers of attention in the ET literature (developed by Orquin & Mueller Loose, 2013): 1) Which processes influence the direction of attention to stimuli? and 2) Which working memory effects influence the direction of attention? To answer the first research question, attention research distinguishes two basic processes: *goal-driven* and *stimulus-driven attention* (Corbetta & Shulman, 2002; Ocasio, 2011). Goal-driven attention can be described as a top-down process. Depending on their search goal, humans will direct their attention differently. For example, a shopper will pay more attention to price information in a supermarket if her personal goal is to save money. Stimuli of higher task relevance will thus receive increased attention. Different factors can be used to alter or manipulate the perceived task relevance. Orquin and Mueller Loose (2013) mentioned task instructions, utility effects, heuristics, attentional phases, and learning effects as potential factors.

In contrast to goal-driven attention, stimulus-driven attention can be described as a bottom-up process. With stimulus-driven attention, it is not a particular goal that brings attention to the stimulus; it is the visual characteristics of the stimulus that draw attention to it. Orquin and Mueller Loose (2013) mentioned saliency, surface size, visual clutter, and position of the

stimulus as examples of bottom-up factors. Visual saliency, for instance, includes various features of visual conspicuity, such as color, contrast, and movement. For example, a red-colored product on a supermarket shelf stands out visually if all neighboring packages are black and white. Color as a visual characteristic will thus bring attention to a product package simply because its visual saliency guides and directs attention. The interested reader is referred to Itti, Koch, and Niebur (1998) who present a conceptual framework and computational model for saliency-driven visual attention.

Attention research also investigates how *working memory effects* influence the direction of attention; information complexity, presentation effects, decision difficulty, time pressure, distracters, consideration sets, and pairwise comparisons are central factors that have been studied. Given the very limited capacity of working memory (also known as short-term memory, see Ballard, Hayhoe, & Pelz, 1995) the above-mentioned factors can easily exceed demands on working memory resources which as a consequence influences participants' gaze behavior. An intuitive example for a working memory effect is a study that manipulates the difficulty of a decision situation, for instance, by increasing the number of decision alternatives. The study by Just and Carpenter (1976) demonstrated that increasing task difficulty led to an increase in working memory load which then increased the number of fixations in a task. Orquin and Mueller Loose (2013) emphasized that most empirical studies testing these effects involve both top-down and bottom-up processes but that assigning the effects to either of these processes would be inappropriate. Time pressure, for example, is likely to influence participants' gaze behavior through both top-down and bottom-up processes: if participants change their goal orientation because of time pressure, then top-down processes will affect how attention is directed to stimuli. If participants under time pressure, however, increasingly focus on visually salient information, the respective change in attention directed to stimuli would be driven by bottom-up processes.

Eye-Tracking Measures

Analyzing ET data most often comes down to examining the sequence of fixations and saccades. ET systems, however, provide a wealth of different measures. We recommend the book by Holmqvist et al. (2011), which provides an excellent and comprehensive list of more than 100 possible measures researchers can use. In the following, we focus on central measures that are frequently used in ET studies and employ the same terminology as Holmqvist et al. (2011).

The first measure to consider is what is in the center of the field of vision, i.e., the *position of a fixation* (also known as fixation location). The *position of a fixation* is the most important measure because other ET measures can only be reasonably interpreted when it is known which stimuli the participant has looked at. The x- and y-coordinates of the fixations indicate the center of the gaze at a particular point in time. In a desktop study, for example, the x- and y-coordinates will indicate the exact pixel of the computer screen that was at the center of the participant's gaze. Knowing the position of the fixations, we can calculate multiple measures that quantify how attention is directed to the AOIs. The *number of fixations*, for example, quantifies how often an AOI was fixated and the *total dwell time* quantifies how much time was spent in the AOI over the whole trial.

Second, it can be of particular interest to analyze the *fixation duration*. The duration of a fixation quantifies how long a participant's eyes were still in a position, for example, when looking at a particular AOI. Depending on the viewed stimuli and their characteristics (Rayner, 1998), fixation duration generally ranges from 100 to 500 milliseconds (ms). In simple reading tasks, the average fixation duration is about 250 ms (Lai et al., 2013). As explained in more detail below, longer fixations might indicate an increased level of processing.

Third, researchers can analyze the *saccadic amplitude* (also known as saccade distance or saccade length) between two consecutive fixations. The saccadic amplitude describes the distance that the eyes have traveled from one fixation to a consecutive fixation. The *saccadic velocity* (also known as speed of the eye movements) can also be determined by calculating the ratio of saccadic amplitude and elapsed time.

Fourth, besides fixations and saccades, ET devices also record *pupil diameter* (also known as pupil dilation or pupil size). Two muscles are responsible for changes in pupil diameter. The dilator pupillae, in the outer part of the iris, dilates the pupil, and the sphincter pupillae,

located in the central parts, constricts it. Both pupillary muscles receive inputs from brain systems involved in cognitive and autonomic functions (Samuels & Szabadi, 2008). That is why changes in cognitive and autonomic activity influence pupil diameter. Human pupils may therefore dilate for a number of different reasons, including lighting, the luminance of the items on a computer screen, cognitive difficulty, arousal, pain, working memory load (Beatty, 1982; Wang, Spezio, & Camerer, 2010), or surprise (Preuschoff, 't Hart, & Einhäuser, 2011). When analyzing pupil diameter data, it is important to correct for gaze position (Hayes & Petrov, 2016). Moreover, when observing changes in pupil diameter, researchers must use 1 second inter-screen-intervals in order to allow for the lag in pupil diameter. We recommend the papers by Aston-Jones and Cohen (2005) as well as Moore and Zirnsak (2017) to those readers who are interested in learning more about the neural mechanisms leading to changes in pupil diameter (and selective visual attention in general).

Fifth, *blink rate* (also known as number of eye blinks) is another frequently used ET measure. A blink of the eye is the rapid closing and reopening of the eyelid. A partially closed eye is called an *incomplete blink*. Eye blinks generally last from 150 to 300 ms. Blink rate counts the number of blinks per minute. While there is a physiological need to blink for humans (to guarantee, for example, the necessary lubrication of the eyeballs), research by Stern, Walrath, and Goldstein (1984) has shown that eye blinks are also related to cognitive state variables, such as processing flow, as further explained below.

Psychological Constructs

In this section, we review psychological constructs that are frequently operationalized by ET measures. Given the scope of this paper, we exclusively focus on ET measures and do not elaborate on alternative psychophysiological measures. For example, emotional arousal can be measured using pupillary changes, as explained in the following, but it could also be assessed by monitoring heart rate or skin conductance (Bradley, Miccoli, Escrig, & Lang, 2008).

Attention directed to stimuli. Attention is a psychological construct, as it is not directly observable. Yet, researchers can use ET measures as indicators of attention. Many ET experiments study how attention is directed to different AOIs. The question of how much attention participants direct to one piece of information compared to other pieces is often

interpreted as an indicator of relevance. For example, researchers count the number of fixations to an AOI to quantify how much attention is directed to the respective AOI. The line of thinking is that attention directed to stimuli is the outcome of a goal-directed (search, choice, or learning) process that led to the respective gaze behavior. Different ET measures can be used as indicators. Meißner, Musalem, and Huber (2016), for example, count the number of fixations to attributes and alternatives in repeated multi-attribute choices.

Attention patterns. Researchers often investigate the sequence in which respective stimuli are examined. For example, sequence analysis can be used to investigate in which order participants explore visual stimuli most frequently (Çöltekin, Fabrikant, & Lacayo, 2010). In decision-making research, the search sequence is often investigated when participants see a decision matrix, as the sequence can provide clues about which strategy was used when processing the respective information (Ball, 1997). The search index developed by Payne (1976), for example, is often applied in decision-making research to examine whether participants process information primarily within objects (alternative-wise search) or directly compare several objects using multiple characteristics (attribute-wise search).

Level of processing. The level of processing as a psychological construct is based on the idea that the fixation duration indicates how deeply the information that the participant looks at is processed. The assumption is that short fixations (up to 250 ms) indicate scanning and automatic processes, whereas longer fixations (longer than 500 ms) are indicative of deeper processing, such as deliberate consideration of information (Velichkovsky, Rothert, Kopf, Dornhöfer, & Joos, 2002). When examining the structure of a new task, for example, fixations are shorter and primarily used for orientation purposes, whereas in a later phase, higher levels of cognitive processing are indicated by longer fixation durations (Velichkovsky, 1999). Similarly, Rayner (1998) found in a reading context that fixations are shorter when individuals read silently compared to more taxing tasks that involve typing.

Mental states (local and global search). Based on previous research, van der Lans et al. (2008) postulated that individuals switch between two latent attention states in target searches: a local and a global state. While in a global state, the individual's goal is to locate the target stimulus, whereas in an identification (local) state, the goal is to evaluate whether the located stimulus has been identified correctly. The search process will end in the identification state when sufficient evidence has been accumulated or otherwise continue in the localization state. The

authors' key contribution is the development of a statistical model to determine the switches between the two attentional states. The discrimination between the two states is based on ET measures, such as the saccadic amplitude. In the local state, participants extract information from adjacent fixation locations, a process characterized by shorter saccades. In the global state, participants seek to identify those locations that are most promising to process. This orientation process is characterized by longer saccades. Analyzing the saccadic amplitude is thus helpful when researchers seek to determine the respective mental state of participants. This stream of research has been further extended, for example, by Wedel, Pieters, and Liechty (2008), who explored switching between local and global states during scene perception.

Emotional arousal. Bradley et al. (2008) used a large set of well-validated pictures to test the effect of emotional arousal on pupillary responses. Their empirical results confirmed that pupil diameter increased when participants processed emotionally engaging stimuli. Additionally, they did not find that the hedonic valence, i.e., whether participants saw pleasant or unpleasant stimuli, changed the effect of emotional arousal. The results also revealed that pupillary changes covaried with skin conductance reactions, supporting “the hypothesis that pupil diameter during picture viewing predominantly reflects sympathetic nervous system activity” (p. 6).³ While other psychophysiological measures are frequently used to measure emotional arousal, the paper by Bradley et al. (2008) shows that ET measures are suitable for that purpose, too.

Cognitive load. Cognitive load refers to the total amount of mental effort being used in working memory (Sweller, 1994). Rosch and Vogel-Walcutt (2013) reviewed papers utilizing ET measures as indicators of cognitive load and provided support for changes in cognitive load being reflected by changes in pupil response. Increases in pupil diameter thus indicate an increase in cognitive load. Marshall (2002) developed an index of cognitive activity that is also based on pupil diameter. The index measures the abrupt discontinuities in pupil diameter over a designated period of time. Rosch and Vogel-Walcutt's (2013) review also showed that increased average fixation durations correspond with increases in cognitive load (Chang & Choi, 2014). Moreover, Bodala, Ke, Mir, Thakor, and Al-Nashash (2014) used peak saccade velocity as a measure of cognitive load and concluded that higher cognitive load requires faster saccades. To sum up, a

³ The sympathetic nervous system, which is also called the “fight-or-flight system,” represents the body's automatic response to external stimuli. Activation of the sympathetic nervous system can therefore lead to various responses of the body, such as an increase in breathing frequency, heart rate, as well as changes in pupil diameter.

variety of measures have been proposed to quantify cognitive load, but it remains an open question up to the present which of these measures is most reliable.

Perceptual fluency. Plumhoff and Schirillo (2009, p. 722) defined perceptual fluency as “a subjective experience reflecting the ease with which we are able to process an image.” More generally, perceptual fluency can be understood as the subjective experience or feeling of ease of processing new information (Schwarz, 2004). Reber, Fazendeiro, and Winkielman (2002) argued that fluency is an indicator of processing flow and as such is indicative of how efficiently a stimulus (or visual image) can be processed. Reber, Schwarz, and Winkielman (2004) suggested that the perception of beauty, for example, is affected by perceptual fluency such that increased fluency leads to a more positive aesthetic response.

Several ET measures have been proposed to quantify perceptual fluency. Early work by Stern et al. (1984) suggests that the blink rate can be informative for revealing transition points within the information processing flow. In other words, a high frequency of eye blinks may indicate that the flow of information processing has changed or was interrupted. Results by Wang, Reimer, Dobres, and Mehler (2014) in the context of driving suggest that more spatially concentrated fixations and smaller saccadic amplitudes might indicate higher fluency. Due to the limited empirical evidence, however, it remains unclear how perceptual fluency is best measured using ET.

Areas of Application

Across academic disciplines, ET has been broadly employed in various areas of application. In the following, we synthesize influential literature reviews that focus on the use of ET in specific areas of application. Overall, we identified four substantive areas of application, all of which are highly relevant to organizational research: information search and decision-making, learning, training systems, and expertise.

Information search and decision-making. Studying information search is of particular interest in marketing and information systems. In their review paper, Wedel and Pieters (2008b) outlined that ET has been used intensively to improve the usability of websites, search engines, or e-commerce shops. ET has been used to understand the information search and decision-making processes of consumers for many years (Wedel, 2015). More broadly, Orquin and Mueller Loose

(2013) reviewed ET research in the field of decision-making. The authors argued that one of the main contributions of ET is that it makes it possible to obtain a better understanding of the role attention plays in constructing decisions. Moreover, Rayner (2009) reviewed ET studies in reading, scene perception, and visual search and argued that some of the paradigms developed to study reading (i.e., the moving-window paradigm and boundary paradigm) can be adopted to study visual search, which makes his review paper even more relevant for organizational researchers interested in ET.

Learning. Lai et al. (2013) reviewed the use of ET in research on learning and identified an application spectrum comprising seven topics: patterns of information processing, effects of instructional strategies, reexamination of existing theories, individual differences, effects of learning strategies, patterns of decision-making, and social cultural effects. The number of identified studies across those seven themes is, however, found to be highly unbalanced. ET studies on learning were found to predominantly focus on the topics of information processing and instructional strategies. Studies that fell under the effects of instructional strategies topic focused on how the design of multimedia can support learning. In this context, ET has primarily been used to obtain feedback on how manipulation of the design of multimedia environments (e.g., cues, guidance, displays) influences learning. The interested reader is referred to a special issue in *Learning and Instructions* (van Gog & Scheiter, 2010).

Training Systems. Tien et al. (2014) reviewed the ET literature focusing on surgical training and the assessment of surgeon skills. Their review showed that task performance can be improved by training participants to maintain their gaze on important fixation points or in “certain expert eye-tracked benchmark patterns” (Tien et al., 2014, p. 176). The review makes evident that several ET studies in the fields of surgical training find differences regarding how attention is directed to stimuli by experts and novices. For instance, experienced radiologists were found to require fewer fixations to identify pathology, and expert surgeons showed significantly more fixation to task-relevant areas.

Rosch and Vogel-Walcutt (2013) reviewed applications of ET in the area of adaptive training systems. The authors focused on papers that measured cognitive load to adapt the effectiveness of training in real time. The rationale is that cognitive load can either be too high or too low, thus negatively impacting users’ learning performance. By measuring cognitive load through ET, training can be adapted and learning performance can potentially be increased, e.g.,

by increasing (decreasing) cognitive load if it is too low (high). The authors found that training processes were investigated in various different contexts and constructed the following categories: indicating cognitive load, reading comprehension, presentation design, distraction and attention guiding, and human-computer interaction. They emphasized that ET is still not used to its full potential because the technology is rarely built into adaptive training systems in order to adapt the system to the user's needs.

Expertise. Gegenfurtner et al. (2011) reviewed differences in expertise in the comprehension of visualizations. Building on a meta-analysis of ET research in professional domains, the authors concluded that three different theories provide important insights into visual expertise. In line with the theory of long-term working memory, experts process information more rapidly than novices, as indicated by shorter fixation durations. The finding that experts more frequently fixate task-relevant areas supports the information-reduction hypothesis. Experts thus outperform novices in terms of selective direction of attention to stimuli and in distinguishing relevant from irrelevant information. Expertise also extends the visual span through parafoveal processing. As suggested by the holistic model of image perception, experts had longer saccadic amplitudes and required shorter periods to first fixate task relevant information.

Summary

ET provides a wealth of behavioral data enabling scholars to investigate psychological constructs in detail. The interest of ET researchers usually goes beyond solely the measurement of these constructs and includes the study of bottom-up and top-down processes as well as working memory effects. Information search and decision-making, learning, training systems, and expertise represent substantive areas of application in which ET has been utilized across academic disciplines up to the present.

A Review of Eye-Tracking Use in Organizational Research

Our review incorporated the same 30 management-related journals as captured in previous works (e.g., Aguinis & Bradley, 2014; Podsakoff, MacKenzie, Podsakoff, & Bachrach, 2008), as these journals represent influential outlets across principal domains. To identify ET studies, we adopted a systematic and keyword-based search strategy using three variations of the term ET (*eye track**,

*eye-track**, and *eyetrack**). A combination of databases was employed for our search. We conducted a full text search using EBSCO Business Source Complete, which covers 28 out of the 30 selected journals. A search in titles, abstracts, and keywords was carried out through ScienceDirect for the *Journal of Vocational Behavior* and via SAGE Journals for the *Industrial & Labor Relations Review*. We constrained our search, starting with the year 1975, as ET-based research starting from the mid-1970s “has been marked by improvements in eye movement recording systems that have allowed measurements to be more accurate and more easily obtained” (Rayner, 1998, p. 372). Our initial search strategy yielded 29 articles, which we manually screened according to the following predefined criteria. First, articles had to be empirical ET studies. Second, articles had to report on ET data. We thus excluded articles in which the authors simply referred to (their own) ET research (e.g., Teixeira, 2012, 2013) as well as papers discussing the use of ET methodology (e.g., Patrick & James, 2004). Overall, 15 studies remained after the screening process. Each article was subsequently coded according to the following 10 dimensions: journal, year of publication, ET system and experimental environment, sample size, driver of attention investigated, ET measures, psychological constructs, area of application, research question, and key ET-based finding. The review’s findings are depicted in Appendix 1 and provide some general conclusions about the use of ET in organizational research.

With only 15 studies identified, the application of ET in leading management journals is not only surprisingly rare but also restricted to a very limited number of outlets. Only four out of the 30 selected journals published ET studies, namely, the *Journal of Applied Psychology*, the *Journal of Business Research*, *Management Science*, and *Organizational Behavior & Human Decision Processes*. The adoption of ET was further found to be a rather recent phenomenon as 12 of the studies appeared since the year 2012. Regarding the use of ET systems and the experimental environment, studies were predominantly built on desktop-based ET systems in the lab. With only three articles making use of mobile ET systems—either in the lab (Lindström, Berg, Nordfält, Roggeveen, & Grewal, 2016), the field (Wästlund, Otterbring, Gustafsson, & Shams, 2015), or using virtual reality (Bigné, Llinares, & Torrecilla, 2016)—technological advances appear to be underutilized as of the present. The sample size of studies ranged between

15 (Venkatraman, Payne, & Huettel, 2014) and 190 (Wästlund et al., 2015) participants, with a respective mean of 73 and a median of 66.⁴

Turning to the four dimensions of our proposed taxonomy, researchers were found to be particularly interested in studying top-down processes followed by bottom-up processes and working memory effects. Utility effects form the most frequently investigated top-down process, but learning effects, tasks instructions, and heuristics have also been explored. While visual saliency and position represent the most commonly studied bottom-up processes, research on working memory effects has been conducted on information complexity, decision difficulty, and presentation effects. A few papers deal with top-down as well as bottom-up processes (Menon, Sigurdsson, Larsen, Fagerstrøm, & Foxall, 2016; Shi, Wedel, & Pieters, 2013; Venkatraman et al., 2014).

Total dwell time and number of fixations represent the most commonly applied ET measures, which is indicative of the exclusive focus on the psychological constructs of attention directed to stimuli and attention patterns in the research reviewed. All 15 articles investigated how attention was directed to stimuli, while five articles additionally explored attention patterns (Fiedler, Glöckner, Nicklisch, & Dickert, 2013; Lohse & Johnson, 1996; Shi et al., 2013; Van Hooft & Born, 2012; Venkatraman et al., 2014). Much to our surprise, not a single study utilized ET to explore other psychological constructs, such as emotional arousal or cognitive load.

Among the 15 articles, the great majority of studies can be assigned to information search and decision-making as the area of application. Seven papers assigned to this area focus on marketing-related topics, such as advertising, branding, and pricing. However, studies unrelated to the field of marketing exist in this group as well. For example, Madera and Hebl (2012) researched the effects of facial stigmas in employee selection, while Raveendran, Puranam, and Warglien (2016) explored information search in a division of labor task. Noticeably, the group of information search and decision-making articles also captured one methodological paper, namely, one in which Lohse and Johnson (1996) compared the process-tracing technique Mouselab with ET.

⁴ Whenever provided, we used sample size information adjusted for participants without valid ET data. As Pieters, Wedel, and Zhang (2007, p. 1820) referred to a sample size of “about 100 adult consumers,” we used the number 100 for calculating the mean and median.

In sum, ET studies in leading management journals, although still rare, have been increasingly published in recent years. Technological advances are far from being fully tapped as researchers predominantly use desktop-based ET systems in the lab. Studies also show surprisingly little variation regarding the psychological constructs under investigation as all the identified studies focus on how attention is directed to stimuli. Up to the present, the capability of ET to measure further psychological constructs has thus remained unexploited. In the reviewed journals, ET is predominantly applied in the area of information search and decision-making to address marketing-related research questions.

Future Avenues for Eye Tracking in Organizational Research

Keeping the methodological potential as well as the current use of ET as reflected in leading management journals in mind, we now turn to future avenues for ET research in substantive domains of organizational science. Table 2 includes a selective set of illustrative research questions that could be addressed with ET. More specifically, we provide questions for research in strategic management, entrepreneurship, and human resources to exemplify the broad application potential of ET. We also offer key references as a starting point for interested readers seeking to delve deeper into the respective topics.

Table 2 *Illustrative research questions in OR domains that could be addressed with ET*

Research question	Building on previous work by
<i>Strategic management</i>	
How do investors direct their attention to CSR and financial information? Which affective reactions (e.g., changes in the level of arousal) can be observed?	Elliott, Jackson, Peecher, & White (2014)
How does the reporting format (e.g., integrated versus separate reporting) affect investors' information acquisition, evaluation, and weighting of financial and non-financial performance information?	Reimsbach, Hahn, & Gürtürk (2017)
How does the value orientation of investors influence the direction of attention to financial and non-financial performance information?	Oll, Hahn, Reimsbach, & Kotzian (2016)
How do visual characteristics (e.g., color) of the balanced-scorecard affect managerial attention during performance evaluations?	Chen, Jermias, & Panggabean (2016)
How do performance incentives shape managers' propensity to search for new information?	Lee & Meyer-Doyle (2017)
How do managers allocate their attention to internal and external search sources and how does this affect innovation outcomes?	Dahlander, O'Mahony, & Gann (2016)
<i>Entrepreneurship</i>	
How do visual characteristics (e.g., images) of entrepreneurs' business plans affect venture capitalists' screening decisions?	Chan & Park (2015)

How do visual symbols (e.g., displays of prototypes, entrepreneur dress, office furniture) used by entrepreneurs influence the resource acquisition process?	Clarke (2011)
How do novice and experienced (e.g., serial or portfolio) entrepreneurs differ with respect to their information search strategies?	Fiet, Norton, & Clouse (2012)
How does expert feedback change the information search processes of novice entrepreneurs?	Haynie, Shepherd, & Patzelt (2012)
How does the entrepreneur's level of arousal influence opportunity evaluations and exploitation decisions?	Welp, Spörrle, Grichnik, Michl, & Audretsch (2012); Foo (2011)
How do working memory effects (e.g., time pressure, information complexity) contribute to biased entrepreneurial decision-making? Which information search heuristics are used?	Shepherd, Williams, & Patzelt (2015)
<i>Human resources</i>	
How do organizational website characteristics influence job seekers' search behavior and organizational image perceptions?	Cober, Brown, Keeping, & Levy (2004); Walker, Feild, Giles, Bernerth, & Short (2011)
How does the employment situation (e.g., new entrant, job loser, employed job seeker) affect job search behavior?	Boswell, Zimmerman, & Swider (2012)
Does the integration of ET to Web 2.0 training applications increase learning performance?	London & Hall (2011)
How do nonverbal behavior cues (e.g., facial expressions, gestures) and physical characteristics of applicants (e.g., facial stigmas, gender) affect personnel selection decisions and discrimination?	Bonaccio, O'Reilly, O'Sullivan, & Chiochio (2016); Madera & Hebl (2012)
How do human resource professionals make use of social media to gather information about potential employees?	Roth, Bobko, Van Iddekinge, & Thatcher (2016)
Does ET allow the detection of fakers in actual selection contexts (e.g., job interviews)?	Van Hooft & Born (2012)

In the following, we illustrate in more detail two exemplary research areas that could benefit from utilizing ET. Our first example is situated in the field of strategic management and focuses on investors' decision-making. Given that financial market participants are increasingly supplied with non-financial information, such as corporate social responsibility (CSR) information, our proposed study resonates well with recent calls to engage more deeply with attention and decision-making in information-rich environments (van Knippenberg et al., 2015). We provide insights into the advantages of ET for measuring affective reactions as compared to survey research. Our second example suggests an ET study in the realm of virtual reality and seeks to encourage interested scholars to better leverage the rapid advances of ET technology. We elaborate on the role of ET in trade-off decisions between ecological validity and experimental control. Both examples thus provide general insights and information that are also relevant and applicable for setting up experiments in other research contexts. Building on our proposed taxonomy, Table 3 provides a summary of information on both exemplary avenues.

Table 3 *Taxonomy applied to suggested future avenues*

Organizational research domain /	Strategic management /	Human resources /
Based on previous work by	Elliott et al. (2014)	Madera and Hebl (2012)
Research question addressed in the referenced paper	How do CSR performance and explicit assessment of CSR performance influence investors' estimates of a firm's fundamental value?	Which factors bias the interviewer's assessment of applicant quality? How does visual attention directed to a facial stigma affect memory of interview facts and applicant ratings?
Drivers of attention	Top-down process: task instructions	Bottom-up process: saliency
Psychological constructs	1. Attention directed to stimuli 2. Emotional arousal	Attention directed to stimuli
ET measures	1. Number of fixations, total dwell time 2. Pupil diameter	Total dwell time
Area of application	Information search and decision-making	Information search and decision-making
Suggested ET technology	Desktop-based ET in the lab	Mobile ET in the virtual reality

Strategic Management

An area of growing importance for strategic management studies relates to the topic of CSR. An interesting and crucial property of CSR research and practice centers upon the distinct nature of CSR information (Orlitzky, 2013; Simnett, Vanstraelen, & Chua, 2009). In contrast to financial (traditional) reporting information, CSR information is predominantly non-financial, sometimes even qualitative, and based on voluntary disclosure in most countries. Further differences exist, as CSR information targets a broader spectrum of stakeholders inside and outside the organization, including employees, suppliers, government agencies, and investors. To date, CSR research is still highly fragmented at the individual level of analysis, although “individual actors are those who actually strategize, make decisions, and execute CSR initiatives” (Aguinis & Glavas, 2012, p. 953).⁵ As ET allows measuring psychological constructs at the individual level, the methodology has much to contribute toward the advancement of CSR literature. ET could, for instance, be adopted to investigate how investors direct their attention to CSR and standard financial information when making investment decisions. This is an important question due to mixed empirical evidence on the relationship between firms' CSR and financial performance (Orlitzky, 2013). In the following, we therefore illustrate the benefits of ET by suggesting an extension of the study by Elliott, Jackson, Peecher, and White (2014).

⁵ In the review by Aguinis and Glavas (2012), only 8 out of 181 articles on CSR were found to focus on the individual level of analysis.

In a 2 X 2 + control, between-subjects design, Elliott et al. (2014) investigated the influence of CSR performance (first factor) and the explicit assessment of CSR performance (second factor) on investors' estimates of a firm's fundamental value. CSR performance was manipulated by providing participants with analyst reports depicting positive, negative, or neutral (control) performance. The explicit assessment of CSR performance was manipulated by either prompting participants to explicitly assess the firm's CSR performance or not prompting participants. After reviewing the information about the firm's CSR performance, participants received summary financial statement information for the same company. Participants were then asked to assess the firm's financial performance and to estimate the fundamental value of the company's stock. At the end of the experiment, participants completed a post-test questionnaire including questions on their CSR awareness and affective reactions. The former part asked participants to retrospectively report on the extent to which CSR performance influenced their estimates, whereas the latter part asked participants to evaluate their affective reactions to the firm's CSR performance.

The results revealed an interesting interaction effect: positive (negative) CSR performance was found to significantly increase (decrease) investors' estimates of fundamental value, but only among participants who were not prompted to explicitly assess CSR performance. The authors also found that the affective reaction scores were significantly higher when CSR performance was positive than when it was negative, but there were no significant differences between participants who did or did not explicitly evaluate CSR performance. Regarding the results on investor awareness, the findings showed that investors believed that CSR performance influenced their estimates to a lesser extent in the condition of not having to explicitly evaluate CSR performance. The authors therefore concluded that the effect of CSR performance on investors' fundamental value estimates was unintentional. Based on the "affect-as-information" theory, Elliott et al. (2014) argued that investors unintentionally used their affective reactions to CSR performance to derive estimates of fundamental value, which they did not do when they were explicitly asked to assess CSR performance.

We propose to replicate Elliott et al.'s (2014) empirical study by monitoring participants' information processing using desktop-based ET. By either prompting or not prompting participants to explicitly assess CSR performance, the authors changed the task instructions in their study. The manipulation of the task instruction is expected to influence how attention is

directed to stimuli, i.e., participants in the explicit assessment condition are supposed to look more frequently at CSR information. From an information-processing perspective, the effect studied is a *top-down process*, as participants will direct attention to CSR information in line with their goals. Attention directed to CSR information could, for example, be quantified by the *number of fixations* or *total dwell time*. The process of CSR performance evaluation could be investigated even further by monitoring *which* social and environmental criteria participants gaze at most frequently. By examining attention patterns, new management implications could also be derived: for example, how to design CSR reports that facilitate information uptake.

Furthermore, Elliott et al. (2014) argued that, in line with “affect-as-information” theory, participants’ affective reactions were the main driver that subconsciously influenced investors’ judgments. The authors, however, only constructed survey research to measure affective reaction, i.e., by asking participants to retrospectively report their emotional reactions. The self-assessment of affective reactions makes it difficult to compare differences in the level of arousal between participants, as they might rate their level of arousal quite differently based on their subjective experience. ET provides the possibility of measuring affective reactions more objectively, by means of the pupillary response in particular (Bradley et al., 2008; Partala & Surakka, 2003). As outlined in the taxonomy chapter, emotional arousal can be measured by observing changes in pupil diameter. Thus, it can be hypothesized that participants’ pupils will dilate when looking at very positive or very negative CSR performance information. We therefore propose to assess the participants’ affective reactions by measuring pupil diameter and by testing how participants’ affective reactions influence the assessment of fundamental value.

Human Resources

Job and employee searches are, by their very nature, information-gathering activities for both sides of the interaction. In this context, the employment interview forms one of the most popular approaches to personnel selection. The purpose of the interview is to evaluate the quality of an applicant’s answers to interview questions. However, factors such as gender, race, religion, dress, and physical imperfections, such as facial stigmas, can bias the interviewer’s assessment of applicant quality. For example, previous research has shown that applicants’ physical attractiveness is an important determinant of personnel selection decisions (e.g., Hosoda, Stone-Romero, & Coats, 2003).

Madera and Hebl (2012) conducted two empirical studies to examine possible discrimination against applicants with facial stigmas and to investigate the underlying process involved. In Study 1, the authors used still images of applicants' faces, either with or without a facial stigma, and measured how participants' evaluation of applicants differed. The authors applied ET to measure the time participants looked at the stigma. ET enabled the researchers to unravel the process involved in discrimination against facially stigmatized applicants: visual attention directed to the stigma negatively affected memory of interview facts, which in turn resulted in lower applicant ratings. As acknowledged by Madera and Hebl (2012), the major limitations of Study 1 relate to its ecological validity due to the absence of a social context and the lack of interaction between interviewers and applicants.

To address the limitations of Study 1, the authors conducted an additional study (Study 2) in which they used face-to-face interviews to replicate their key findings but did not measure visual attention using ET. The use of face-to-face interviews in Study 2 led to higher ecological validity. The disadvantage of face-to-face interviews, however, is that other potential influence factors can hardly be controlled for. The applicants' body language and facial mimicry, for example, are factors that are supposed to influence the interviewers' quality assessments. A standardization of these factors in an experiment could potentially be achieved by using trained confederates as applicants, i.e., "actors that are instructed and trained to maintain the same verbal and non-verbal reactions across participants and across conditions" (Bombari, Mast, Canadas, & Bachmann, 2015, p. 2). However, it is very difficult to ensure that non-verbal behavior is indeed standardized.

An alternative solution to this trade-off decision between ecological validity and control over experimental factors (as further discussed by Loomis, Blascovich, & Beall, 1999) is to investigate social interaction in virtual reality (Bombari et al., 2015) and to monitor participant's eye movements in virtual reality settings (Bigné, Llinares, & Torrecilla, 2016; Meißner, Pfeiffer, Pfeiffer, & Oppewal, 2018). Applied to the employee-selection article discussed above by Madera and Hebl (2012), interaction with virtual humans would make the job interview situation much more realistic as compared to Study 1. At the same time, virtual environments allow manipulating factors that would be very difficult or impossible to test in real life (Bombari et al., 2015). For example, a participant's height can be easily changed in a virtual reality experiment. Furthermore, researchers could also change individual characteristics of the interaction partners,

such as facial stigmas, and control other subtle behavioral cues that are difficult to observe or measure.

Applying ET in a virtual environment would thus make replicating Madera and Hebl's (2012) results possible in an experimental setting with higher ecological validity than their Study 1 and more experimental control than Study 2. Moreover, ET in virtual environments allows researchers to analyze visual attention in real-time (as further explained in Meißner et al., 2018). This ability has radical consequences for conducting research because virtual humans could now react to a participant's gaze in real time. For example, if a participant frequently looked at the facial stigma of the virtual human, the virtual human could react by acknowledging the stigma. In Study 1, Madera and Hebl (2012) tested whether acknowledgment is effective as was supposed in reducing the attributions and preoccupation that one might have regarding the stigma. The authors did not find that acknowledgment influenced visual attention significantly. One could, however, argue that the interaction was simply too unrealistic. Yet, the virtual environment makes it possible to test whether directly addressing one's stigma during an interaction is more effective if the virtual human is responsive to the participant's gaze.

As noted by Bombari et al. (2015, p. 2), using "virtual humans as interaction partners to simulate interpersonal encounters is still in its infancy." We believe that the adaptation of social interaction based on a participant's visual attention is a fruitful future avenue for research in the human resource field. By using ET, researchers could, for instance, also monitor to what degree participants avoid eye contact when in a situation of stress in a job interview (Hartanto et al., 2014). Virtual humans could also react to a participant's gaze by moving to locations that are in the visual focus of the participant (Bombari et al. 2015) in order to test how this might influence the participant's behavior.

Methodological Guidance on Eye Tracking

The goal of this article is to introduce ET to the field of organizational science. Our intention is therefore not to provide a tutorial on how to conduct ET studies but to carve out the potential of ET for future organizational research. Interested scholars who are considering using ET in their research thus need to dig deeper into the methodological literature. Various sources are available, but sifting through the literature can be challenging. Two seminal books on ET are particularly

helpful for learning about and delving deeper into ET methodology: a book by Duchowski (2017) includes a detailed description of ET technology while also providing recommendations on designing ET experiments and empirical guidelines, and Holmqvist et al.'s (2011) book supports researchers by providing guidelines on how to plan ET experiments and provides valuable advice on how to record, process, analyze, and interpret ET data. We would like to make organizational researchers aware of important issues that, based on our own experience, are not extensively or not at all discussed in the abovementioned two books.

First, researchers need to be aware that the environment in which ET studies take place affects which ET devices are most useful. If it is the researcher's goal to observe participants' eye movements in everyday life situations, the researcher will most likely need mobile ET equipment, which implies less experimental control over external factors but higher ecological validity. When conducting a desktop-based ET study in the laboratory, experimental factors can be more easily controlled, but the situation may be less ecologically valid. Researchers frequently face that trade-off between experimental control and ecological validity (Loomis et al., 1999) and are encouraged to reflect on it before carrying out their own ET experiments. The reader is referred to Meißner et al. (2018), who developed a set of criteria researchers can use to make the decision whether to use desktop-based ET or mobile ET in natural or virtual environments. Mobile ET is best suited for research contexts in which the realism of the situation or the interaction investigated is central to the research question. For example, mobile ET could be used in one-to-one or group meetings or in training situations where the focus of the study is to investigate how participants interact in real-world natural settings and direct their attention to a particular object or person. In such an experiment, participants would be able to move around freely while wearing the mobile ET equipment. New mobile ET devices are lightweight so that after a while participants might even forget that their eye movements are being recorded. Using mobile ET equipment and gathering data is relatively easy to accomplish given the improvements in ET technology in recent years. The reader is referred to papers by Wästlund et al. (2015) and Hendrickson and Ailawadi (2014) that elaborate in greater detail important aspects that need to be considered when using mobile ET in research projects. However, preparing the gathered data is one of the major challenges when using mobile ET (Meißner et al., 2018). The head movements of participants change the position of objects in a head-mounted scene camera recording the environment. Every recorded video frame therefore needs to be analyzed separately to decide which areas of interest in the environment were fixated. This process is called manual

annotation and is extremely time-consuming. The high workload of preparing the gathered mobile ET data might be one of the reasons that relatively few papers have made use of mobile ET equipment so far (Meißner et al., 2018), as also further evidenced by the Review of ET in Organizational Research section. Another major disadvantage is that the experimental setting is often not reproducible. As real-world environments change continuously, it is very difficult to set up the exact same experimental situation at a later point in time.

Second, the design of the stimuli and the definition of the AOIs are closely related to each other. When designing stimuli to be presented on the computer screen in a desktop ET study, the researcher should carefully consider how the AOIs will be defined when analyzing the data. Orquin, Ashby, and Clarke (2016) tested how the definition of AOIs influenced the results of several previously published empirical studies. They concluded that the definition of the AOIs can significantly affect the results. We therefore recommend that researchers employ the guidelines for the use of AOIs in behavioral ET research developed by Orquin et al. (2016).

Third, it is important to enable researchers to determine the quality of ET data. Holmqvist, Nyström, and Mulvey's (2012) paper discusses how data quality can be measured, evaluated, and reported. We think that it is important to make organizational researchers aware of this issue, in particular as ET data quality can have a substantial impact on the validity of research results.

Conclusion

In recent years, rapid improvements have been made with respect to the ease of use, accuracy, unobtrusiveness, and costs of ET devices (Duchowski, 2017). Such evolving technological advances bring about promising opportunities for ET research. ET methodology will probably be adopted more frequently to support decision-makers in organizational settings, such as by highlighting important information when managers make decisions or by giving novices feedback on their search performance. As ET is to become increasingly widespread in the near future, and as all virtual reality devices will include ET by default, it is time to bring ET to the minds and hands of organizational scholars.

References

- Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods, 17*(4), 351–371.
- Aguinis, H., & Glavas, A. (2012). What we know and don't know about corporate social responsibility: A review and research agenda. *Journal of Management, 38*(4), 932–968.
- Aguinis, H., Pierce, C. A., Bosco, F. A., & Muslin, I. S. (2009). First decade of Organizational Research Methods: Trends in design, measurement, and data-analysis topics. *Organizational Research Methods, 12*(1), 69–112.
- Anderson, J. R., Bothell, D., & Douglass, S. (2004). Eye movements do not reflect retrieval processes: Limits of the eye-mind hypothesis. *Psychological Science, 15*(4), 225–231.
- Ashby, N. J. S., Johnson, J. G., Krajbich, I., & Wedel, M. (2016). Applications and innovations of eye-movement research in judgment and decision making. *Journal of Behavioral Decision Making, 29*(2–3), 96–102.
- Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience, 28*(1), 403–450.
- Ball, C. (1997). A comparison of single-step and multiple-step transition analyses of multiattribute decision strategies. *Organizational Behavior and Human Decision Processes, 69*(3), 195–204.
- Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7*(1), 66–80.
- Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science, 2*(4), 396–403.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91*(2), 276–292.
- Bigné, E., Llinares, C., & Torrecilla, C. (2016). Elapsed time on first buying triggers brand choices within a category: A virtual reality-based study. *Journal of Business Research, 69*(4), 1423–1427.
- Bodala, L. P., Ke, Y., Mir, H., Thakor, N. V., & Al-Nashash, H. (2014). Cognitive workload estimation due to vague visual stimuli using saccadic eye movements. In *Engineering in*

- Medicine and Biology Society (EMBC)*, 36th Annual International Conference of the IEEE, 2993–2996.
- Bombari, D., Mast, M. S., Canadas, E., & Bachmann, M. (2015). Studying social interactions through immersive virtual environment technology: Virtues, pitfalls, and future challenges. *Frontiers in Psychology*, 6, 1–11.
- Bonaccio, S., O'Reilly, J., O'Sullivan, S. L., & Chiocchio, F. (2016). Nonverbal behavior and communication in the workplace: A review and an agenda for research. *Journal of Management*, 42(5), 1044–1074.
- Bono, J. E., & McNamara, G. (2011). Publishing in AMJ—Part 2: Research design. *Academy of Management Journal*, 54(4), 657–660.
- Boswell, W. R., Zimmerman, R. D., & Swider, B. W. (2012). Employee job search: Toward an understanding of search context and search objectives. *Journal of Management*, 38(1), 129–163.
- Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 45(4), 602–607.
- Chaffin, D., Heidl, R., Hollenbeck, J. R., Howe, M., Yu, A., Voorhees, C., & Calantone, R. (2017). The promise and perils of wearable sensors in organizational research. *Organizational Research Methods*, 20(1), 3–31.
- Chan, C. S. R., & Park, H. D. (2015). How images and color in business plans influence venture investment screening decisions. *Journal of Business Venturing*, 30(5), 732–748.
- Charness, N., Reingold, E. M., Pomplun, M., & Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory & Cognition*, 29(8), 1146–1152.
- Chen, Y., Jermias, J., & Panggabean, T. (2016). The role of visual attention in the managerial judgment of balanced-scorecard performance evaluation: Insights from using an eyetracking device. *Journal of Accounting Research*, 54(1), 113–146.
- Claeys, A.-S., Cauberghe, V., & Pendelaere, M. (2016). Is old news no news? The impact of self-disclosure by organizations in crisis. *Journal of Business Research*, 69, 3963–3970.
- Clarke, J. (2011). Revitalizing entrepreneurship: How visual symbols are used in entrepreneurial performances. *Journal of Management Studies*, 48(6), 1365–1391.

- Cober, R. T., Brown, D. J., Keeping, L. M., & Levy, P. E. (2004). Recruitment on the net: How do organizational web site characteristics influence applicant attraction? *Journal of Management*, 30(5), 623–646.
- Çöltekin, A., Fabrikant, S. I., & Lacayo, M. (2010). Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings. *International Journal of Geographical Information Science*, 24(10), 1559–1575.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Review Neuroscience*, 3(3), 201–215.
- Dahlander, L., O'Mahony, S., & Gann, D. M. (2016). One foot in, one foot out: How does individuals' external search breadth affect innovation outcomes? *Strategic Management Journal*, 37(2), 280–302.
- Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology*, 17(2), 245–260.
- Duchowski, A. T. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods*, 34(4), 455–470.
- Duchowski, A. T. (2017). *Eye tracking methodology: Theory and practice* (2nd ed.). London, England: Springer.
- Elliott, W. B., Jackson, K. E., Peecher, M. E., & White, B. J. (2014). The unintended effect of corporate social responsibility performance on investors' estimates of fundamental value. *The Accounting Review*, 89(1), 275–302.
- Fiedler, S., Glöckner, A., Nicklisch, A., & Dickert, S. (2013). Social value orientation and information search in social dilemmas: An eye-tracking analysis. *Organizational Behavior and Human Decision Processes*, 120, 272–284.
- Fiet, J. O., Norton, W., & Clouse, V. (2012). Search and discovery by repeatedly successful entrepreneurs. *International Small Business Journal*, 31(8), 890–913.
- Foo, M.-D. (2011). Emotions and entrepreneurial opportunity evaluation. *Entrepreneurship Theory and Practice*, 35(2), 375–393.
- Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, 23(4), 523–552.

- Hartanto, D., Kampmann, I. L., Morina, N., Emmelkamp, P. G. M., Neerincx, M. A., Brinkman, W.-P. (2014). Controlling social stress in virtual reality environments. *PLoS ONE*, 9(3), 1–17.
- Hayes, T. R., & Petrov, A. A. (2016). Mapping and correcting the influence of gaze position on pupil size measurements. *Behavior Research Methods*, 48(2), 510–527.
- Haynie, J. M., Shepherd, D. A., & Patzelt, H. (2012). Cognitive adaptability and an entrepreneurial task: The role of metacognitive ability and feedback. *Entrepreneurship Theory and Practice*, 36(2), 237–265.
- Hendrickson, K., & Ailawadi, K. L. (2014). Six lessons for in-store marketing from six years of mobile eye tracking research. *Shopper marketing and the role of in-store marketing. Review of Marketing Research*, 11, 57–74.
- Holmqvist, K., Nyström, M., & Mulvey, F. (2012). Eye tracker data quality: What it is and how to measure it. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, 45–52.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. New York, NY: Oxford University Press.
- Hosoda, M., Stone-Romero, E. F., & Coats, G. (2003). The effects of physical attractiveness on job-related outcomes: A meta-analysis of experimental studies. *Personnel Psychology*, 56(2), 431–462.
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), 1254–1259.
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8(4), 441–480.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354.
- Kuvaas, B. (2002). An exploration of two competing perspectives on informational contexts in top management strategic issue interpretation. *Journal of Management Studies*, 39(7), 977–1001.

- Lai, M. L., Tsai, M. J., Yang, F. Y., Hsu, C. Y., Liu, T. C., Lee, S. W. Y., ... & Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, 10, 90–115.
- Lee, S., & Meyer-Doyle, P. (2017). How performance incentives shape individual exploration and exploitation: Evidence from microdata. *Organization Science*, 28(1), 19–38.
- Lindström, A., Berg, H., Nordfält, J., Roggeveen, A. L., & Grewal, D. (2016). Does the presence of a mannequin head change shopping behavior? *Journal of Business Research*, 69, 517–524.
- Lohse, G. L., & Johnson, E. J. (1996). A comparison of two process tracing methods for choice tasks. *Organizational Behavior and Human Decision Processes*, 68(1), 28–43.
- London, M., & Hall, M. J. (2011). Unlocking the value of Web 2.0 technologies for training and development: The shift from instructor-controlled, adaptive learning to learner-driven, generative learning. *Human Resource Management*, 50(6), 757–775.
- Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers*, 31(4), 557–564.
- Madera, J. M., & Hebl, M. R. (2012). Discrimination against facially stigmatized applicants in interviews: An eye-tracking and face-to-face investigation. *Journal of Applied Psychology*, 97(2), 317–330.
- Marshall, S. P. (2002). The index of cognitive activity: Measuring cognitive workload. In *Proceedings of the 2002 IEEE 7th Conference on Human Factors and Power Plants*. IEEE Computer Society Press, Los Alamitos, CA, 7.5–7.9.
- Meißner, M., Musalem, A., & Huber, J. (2016). Eye tracking reveals processes that enable conjoint choices to become increasingly efficient with practice. *Journal of Marketing Research*, 53(1), 1–17.
- Meißner, M., Pfeiffer, J., Pfeiffer, T., & Oppewal, H. (2018). Combining virtual reality and mobile eye tracking to provide a naturalistic experimental environment for shopper research. *Journal of Business Research*, forthcoming.
<https://doi.org/10.1016/j.jbusres.2017.09.028>
- Menon, R. G. V., Sigurdsson, V., Larsen, N. M., Fagerstrøm, A., & Foxall, G. R. (2016). Consumer attention to price in social commerce: Eye tracking patterns in retail clothing. *Journal of Business Research*, 69, 5008–5013.

- Moore, T., & Zirnsak, M. (2017). Neural mechanisms of selective visual attention. *Annual Review of Psychology*, 68, 47–72.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231–259.
- Ocasio, W. (2011). Attention to attention. *Organization Science*, 22(5), 1286–1296.
- Oll, J., Hahn, R., Reimsbach, D., & Kotzian, P. (2016). Tackling complexity in business and society research: The methodological and thematic potential of factorial surveys. *Business & Society*, 1–34. doi: 10.1177/0007650316645337
- Orlitzky, M. (2013). Corporate social responsibility, noise, and stock market volatility. *The Academy of Management Perspectives*, 27(3), 238–254.
- Orquin, J. L., Ashby, N. J., & Clarke, A. D. (2016). Areas of interest as a signal detection problem in behavioral eye-tracking research. *Journal of Behavioral Decision Making*, 29(5), 103–115.
- Orquin, J. L., Bagger, M. P., & Mueller Loose, S. (2013). Learning affects top down and bottom up modulation of eye movements in decision making. *Judgment and Decision Making*, 8(6), 700–716.
- Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, 144(1), 190–206.
- Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, 59(1), 185–198.
- Patrick, J., & James, N. (2004). Process tracing of complex cognitive work tasks. *Journal of Occupational and Organizational Psychology*, 77, 259–280.
- Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, 16(2), 366–387.
- Pieters, R., Warlop, L., & Wedel, M. (2002). Breaking through the clutter: Benefits of advertisement originality and familiarity for brand attention and memory. *Management Science*, 48(6), 765–781.
- Pieters, R., Wedel, M., & Zhang, J. (2007). Optimal feature advertising design under competitive clutter. *Management Science*, 53(11), 1815–1828.
- Plumhoff, J. E., & Schirillo, J. A. (2009). Mondrian, eye movements, and the oblique effect. *Perception*, 38(5), 719–731.

- Podsakoff, P. M., & Dalton, D. R. (1987). Research methodology in organizational studies. *Journal of Management*, 13(2), 419–441.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *The Journal of Applied Psychology*, 88(5), 879–903.
- Podsakoff, P. M., MacKenzie, S. B., Podsakoff, N. P., & Bachrach, D. G. (2008). Scholarly influence in the field of management: A bibliometric analysis of the determinants of university and author impact in the management literature in the past quarter century. *Journal of Management*, 34(4), 641–720.
- Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, 12(4), 531–544.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32(1), 3–25.
- Preuschoff, K., Marius't Hart, B., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience*, 5.
- Raveendran, M., Puranam, P., & Warglien, M. (2016). Object salience in the division of labor: Experimental evidence. *Management Science*, 62(7), 2110–2128.
- Ray, J. L., & Smith, A. D. (2012). Using photographs to research organizations: Evidence, considerations, and application in a field study. *Organizational Research Methods*, 15(2), 288–315.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422.
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, 62(8), 1457–1506.
- Reber, R., Fazendeiro, T. A., & Winkielman, P. (2002). Processing fluency as the source of experiences at the fringe of consciousness. *Psyche*, 8(10), 1–21.
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8(4), 364–382.
- Reimsbach, D., Hahn, R., & Gürtürk, A. (2017). Integrated reporting and assurance of sustainability information: An experimental study on professional investors' information processing. *European Accounting Review*, 1–23. doi: 10.1080/09638180.2016.1273787

- Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science*, 12(1), 48–55.
- Rosch, J. L., & Vogel-Walcutt, J. J. (2013). A review of eye-tracking applications as tools for training. *Cognition, Technology & Work*, 15(3), 313–327.
- Roth, P. L., Bobko, P., Van Iddekinge, C. H., & Thatcher, J. B. (2016). Social media in employee-selection-related decisions: A research agenda for uncharted territory. *Journal of Management*, 42(1), 269–298.
- Samuels, E. R., & Szabadi, E. (2008). Functional neuroanatomy of the noradrenergic locus coeruleus: Its roles in the regulation of arousal and autonomic function part I: principles of functional organisation. *Current Neuropharmacology*, 6(3), 235–253.
- Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, 14(4), 332–348.
- Shepherd, D. A., Williams, T. A., & Patzelt, H. (2015). Thinking about entrepreneurial decision making: Review and research agenda. *Journal of Management*, 41(1), 11–46.
- Shi, S. W., Wedel, M., & Pieters, F. G. M. (2013). Information acquisition during online decision making: A model-based exploration using eye-tracking data. *Management Science*, 59(5), 1009–1026.
- Simnett, R., Vanstraelen, A., & Chua, W. F. (2009). Assurance on sustainability reports: An international comparison. *The Accounting Review*, 84(3), 937–967.
- Stern, J. A., Walrath, L. C., & Goldstein, R. (1984). The endogenous eyeblink. *Psychophysiology*, 21(1), 22–33.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4(4), 295–312.
- Teixeira, T. (2012). The new science viral ads. *Harvard Business Review*, 90(3), 25–27.
- Teixeira, T. (2013). How to profit from “Lean advertising”. *Harvard Business Review*, 91(6), 23–25.
- Tien, T., Pucher, P. H., Sodergren, M. H., Sriskandarajah, K., Yang, G. Z., & Darzi, A. (2014). Eye tracking for skills assessment and training: A systematic review. *Journal of Surgical Research*, 191(1), 169–178.
- Uhlmann, E. L., Leavitt, K., Menges, J. I., Koopman, J., Howe, M., & Johnson, R. E. (2012). Getting explicit about the implicit: A taxonomy of implicit measures and guide for their use in organizational research. *Organizational Research Methods*, 15(4), 553–601.

- van der Lans, R., Pieters, R., & Wedel, M. (2008). Eye-movement analysis of search effectiveness. *Journal of the American Statistical Association*, 103(482), 452–461.
- van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning & Instruction*, 20, 95–99.
- van Hooft, E. A. J., & Born, M. Ph. (2012). Intentional response distortion on personality tests: Using eye-tracking to understand response processes when faking. *Journal of Applied Psychology*, 97(2), 301–316.
- van Knippenberg, D., Dahlander, L., Haas, M., & George, G. (2015). Information, attention, and decision making. *Academy of Management Journal*, 58(3), 649–657.
- Velichkovsky, B. M. (1999). From levels of processing to stratification of cognition. In B. H. Challis, & B. M. Velichkovsky (eds.), *Stratification in cognition and consciousness* (pp. 203–235). Amsterdam/Philadelphia: John Benjamins
- Velichkovsky, B. M., Rothert, A., Kopf, M., Dornhöfer, S. M., & Joos, M. (2002). Towards an express-diagnostics for level of processing and hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour*, 5(2), 145–156.
- Venkatraman, V., Payne, J. W., & Huettel, S. A. (2014). An overall probability of winning heuristic for complex risky decisions: Choice and eye fixation evidence. *Organizational Behavior and Human Decision Processes*, 125(2), 73–87.
- Vila, J., & Gomez, Y. (2016). Extracting business information from graphs: An eye tracking experiment. *Journal of Business Research*, 69(5), 1741–1746.
- Waldman, D. A., Wang, D., & Fenters, V. (2016). The added value of neuroscience methods in organizational research. *Organizational Research Methods*, 1–27. doi: 10.1177/1094428116642013
- Walker, H. J., Feild, H. S., Giles, W. F., Bernerth, J. B., & Short, J. C. (2011). So what do you think of the organization? A contextual priming explanation for recruitment web site characteristics as antecedents of job seekers' organizational image perceptions. *Organizational Behavior and Human Decision Processes*, 114(2), 165–178.
- Wang, J. T. Y., Spezio, M., & Camerer, C. F. (2010). Pinocchio's pupil: Using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. *American Economic Review*, 100(3), 984–1007.
- Wang, Y., Reimer, B., Dobres, J., & Mehler, B. (2014). The sensitivity of different methodologies for characterizing drivers' gaze concentration under increased cognitive

- demand. *Transportation Research Part F: Traffic Psychology and Behaviour*, 26, 227–237.
- Wästlund, E., Otterbring, T., Gustafsson, A., & Shams, P. (2015). Heuristics and resource depletion: Eye-tracking customers' in situ gaze behavior in the field. *Journal of Business Research*, 68(1), 95–101.
- Wedel, M. (2015). Attention research in marketing: A review of eye tracking studies. In J. Fawcett, E. F. Risko, & A. Kingstone (eds.), *The handbook of attention* (pp. 569–588). Cambridge, MA: MIT Press.
- Wedel, M., & Pieters, R. (2008a). Eye tracking for visual marketing. *Foundations and Trends in Marketing*, 1(4), 231–320.
- Wedel, M., & Pieters, R. (2008b). A review of eye-tracking research in marketing. In N. K. Malhotra (ed.), *Review of Marketing Research*, Vol. 4 (pp. 123–147). Armonk, NY: M.E. Sharpe.
- Wedel, M., Pieters, R., & Liechty, J. (2008). Attention switching during scene perception: how goals influence the time course of eye movements across advertisements. *Journal of Experimental Psychology: Applied*, 14(2), 129.
- Welpel, I. M., Spörrle, M., Grichnik, D., Michl, T., & Audretsch, D. B. (2012). Emotions and opportunities: The interplay of opportunity evaluation, fear, joy, and anger as antecedent of entrepreneurial exploitation. *Entrepreneurship Theory and Practice*, 36(1), 69–96.

Appendix 1

Summary of published ET studies

Study by Journal	ET System & Experimental Environment	Sample Size	Drivers of Attention	Psychological Constructs	ET Measures	Area of Application	Research Question	Key ET-based Finding
<i>Journal of Applied Psychology</i>								
Madera & Hebl (2012)	Desktop-based ET in the lab	171	Bottom-up process: saliency	Attention directed to stimuli	Total dwell time	Information search and decision-making	Are facially stigmatized applicants discriminated against and what are the involved attentional processes?	Participants with facially stigmatized applicants attended more to the stigma, which led participants to recall fewer interview facts, which in turn led to lower applicant ratings.
Van Hoof & Born (2012)	Desktop-based ET in the lab	129	Top-down process: task instructions	Attention directed to stimuli, attention patterns	Number of fixations, fixation path analysis	Other	Does ET increase our understanding of the response processes when faking on personality and integrity items and can faking behavior be identified?	Faking on personality tests is characterized by a faster and less cognitively demanding response pattern. ET was demonstrated to be potentially useful in detecting faking behavior.
<i>Journal of Business Research</i>								
Bigné, Llinares, & Torrecilla (2016)	Mobile ET in virtual reality	41	Top-down process: utility effect	Attention directed to stimuli	Total dwell time, number of fixations per second	Information search and decision-making	How does the time people spend on examining a brand influence subsequent purchase decisions and brand choices within the same product category?	The key driver of additional brand choices is the time buyers spend on the first choice, showing that the allocation of less for the first choice triggers additional purchases within the product category and, therefore, increases sales.
Claeys, Cauberghe, & Pandelaere (2016)	Desktop-based ET in the lab	Study 1: 66; Study 2: 86	Top-down processes: utility effects	Attention directed to stimuli	Total dwell time	Other	Does organizational self-disclosure reduce crisis damage?	When an organization does not self-disclose crisis information, participants pay more attention to subsequent negative publicity from a third party than when the organization self-discloses the crisis.

Study by Journal	ET System & Experimental Environment	Sample Size	Drivers of Attention	Psychological Constructs	ET Measures	Area of Application	Research Question	Key ET-based Finding
Lindström, Berg, Nordfält, Roggeveen, & Grewal (2016)	Mobile ET in the lab	70	Bottom-up process: saliency	Attention directed to stimuli	Total dwell time	Information search and decision-making	Does the impact of mannequin style on purchase intentions depend on the level of consumers' fashion knowledge?	The eye-tracking results provide further evidence that the role of mannequin style on attention paid to the merchandise depends on the customers' level of fashion knowledge.
Menon, Sigurdsson, Larsen, Fagerstrøm, & Foxall (2016)	Desktop-based ET in the lab	31	Bottom-up processes: saliency, position; Top-down process: utility effect	Attention directed to stimuli	Total dwell time	Information search and decision-making	How do price points, price position, and price saliency influence attention to price in social commerce?	Besides utility effects, stimulus driven processes substantially affect the attention to price.
Wästlund, Otterbring, Gustafsson, & Shams (2015)	Mobile ET in the field	Study 1: 190; Study 2: 98; Study 3: 66	Top-down processes: task instructions	Attention directed to stimuli	Number of fixations, number of different AOIs fixated	Information search and decision-making	How does the specificity of a customer's shopping tasks influence how much visual attention is directed to task-relevant in-store stimuli?	Differences do exist in viewing behavior based on whether shopping goals are planned or unplanned as more complex and unplanned shopping goals lead to increased observations of in-store stimulus.
Vila & Gomez (2016)	Desktop-based ET in the lab	30	Effect related to working memory: presentation effects	Attention directed to stimuli	Total dwell time	Other	How do people extract information from statistical charts and how do they discriminate between relevant and irrelevant pieces of information?	Successful participant focused their attention in a small subset of relevant and informative AOIs, whereas the others spare their attention in a wider set of AOIs.
Management Science								
Pieters, Warlop, & Wedel (2002)	Desktop-based ET in the lab	119	Top-down process: learning effects	Attention directed to stimuli	Number of fixations, total dwell time	Information search and decision-making	What is the influence of ad originality on attention to advertising and memory for brands?	Ad originality enhanced information storage about the advertised brand in memory by increasing the attention devoted to it. The positive impact of ad originality on information storage was found to further increase for familiar, original advertisements.

Study by Journal	ET System & Experimental Environment	Sample Size	Drivers of Attention	Psychological Constructs	ET Measures	Area of Application	Research Question	Key ET-based Finding
Pieters, Wedel, & Zhang (2007)	Desktop-based ET in the lab	About 100	Bottom up-processes: surface size, visual clutter	Attention directed to stimuli	Fixation likelihood (attention selection), total dwell time (attention engagement)	Information search and decision-making	How do design characteristics of feature ads affect consumers' visual attention to them?	Significant improvements in attention to feature advertising can be achieved without increase in costs. The resultant optimal feature ad designs create win-win opportunities for manufacturers and retailers.
Shi, Wedel, & Pieters (2013)	Desktop-based ET in the lab	108	Bottom up-process: position; Top-down processes: utility effect, heuristics	Attention directed to stimuli, attention patterns	Number of fixations, number of saccades	Information search and decision-making	How do consumers gather product and attribute information over time?	Consumers switch frequently between acquisition strategies, and they obtain information on only two or three attributes or products in a particular acquisition strategy before switching.
Raveendran, Puranam, & Warglien (2016)	Desktop-based ET in the lab	16	Effect related to working memory: information complexity	Attention directed to stimuli	Total dwell time	Information search and decision-making	What aspects of the problem capture participants' attention as they engage in choosing the division of labor?	Individuals displayed greater fixation time on the pictorial representations of the object-based than on the activity-based instructions.
Organizational Behavior & Human Decision Processes								
Lohse & Johnson (1996)	Desktop-based ET in the lab	36	Effect related to working memory: information complexity; top-down process: learning effects	Attention directed to stimuli, attention patterns	Total dwell time, search index	Information search and decision-making	Are Mouselab and eye-tracking equally suited to investigate decision processes?	Mouselab increases the amount of time needed to acquire information and induces more systematic information acquisition behavior compared to eye tracking.

[illegible]

5 Fourth Article

Oll, J. (2022). The Concept of Values in Socially Responsible Investment Research: A Critical Evaluation and Recommendations for Future Work. *Unpublished Working Paper*.

The Concept of Values in Socially Responsible Investment Research: A Critical Evaluation and Recommendations for Future Work

Abstract

Values form one of the most fundamental factors governing human decision-making, and prior research confirms the importance of values in various sustainability-related domains. In this vein, the study of investors' values could provide crucial insights into a key question in socially responsible investment (SRI) research: why do some investors adopt SRI practices, whereas others refrain from doing so? Despite the prominence of investors' values in the SRI literature, this paper argues that the field's values perspective is not yet well developed, a situation which impedes scientific progress and the accumulation of a coherent body of knowledge. To substantiate this assertion, we review the conceptualization of values in SRI research and provide a critique aimed at three major problem areas. First, our analysis indicates that values represent, to a considerable degree, a buzzword rather than a substantive concept in the SRI literature. Second, we document a substantial lack of conceptual clarity as well as conceptual confusion. Third, we identify methodological issues regarding the measurement of values. On this basis, we provide specific recommendations to address the identified conceptual and methodological issues. To advance the understanding of values as a motive for SRI, we further offer promising future research avenues. These contributions provide the foundation for more rigorous values inquiry and a stronger values-based perspective in the field of SRI research.

Keywords: Socially responsible investment, sustainability, values, concept, review

Introduction

In recent years, private and institutional investors have increasingly incorporated sustainability considerations through a broad spectrum of investment practices (Eurosif, 2016; GSIA, 2016; US SIF, 2018). These practices are commonly described by overlapping and supplementary terms, including socially responsible investment (SRI), sustainable investment, ethical investment, and others (Busch et al., 2016; Cadman, 2011; Eccles and Viviers, 2011; Schueth, 2003). In general terms, SRI refers to a shift away from purely financial objectives toward the integration of environmental, social, and governance (ESG) factors into the investment process (Capelle-Blancard and Monjon, 2012). Although the SRI market has gained considerable traction, mirrored by increased academic interest (Busch et al., 2016; Capelle-Blancard and Monjon, 2012), relatively little is known about the underlying drivers, motives, and determinants (Diouf et al., 2016; Glac, 2009; Hong and Kostovetsky, 2012; McLachlan and Gardner, 2004; Scholtens and Sievänen, 2013). The key question of why some investors adopt SRI practices, whereas others refrain from doing so, thus remains a largely unresolved issue. We argue that a promising explanation may rest on a profound study of investors' values.

Values are widely accepted as one of the most fundamental factors governing human decision-making and behavior (Homer and Kahle, 1988; Rokeach, 1973; Weber, 2015) and "have long been considered important to explaining action in and around organizations" (Gehman et al., 2013, p. 84). As such, values can be seen as "prime drivers of personal, social, and professional choices" (Suar and Khuntia, 2010, p. 443). In the field of management scholarship, the importance of values has also long been acknowledged, both in theoretical (Hemingway and MacLagan, 2004; Unsworth et al., 2013) and empirical work on sustainability (Marcus et al., 2015). For instance, empirical evidence indicates that certain values significantly relate to ethical decision-making (Fritzsche and Oz, 2007), unethical practices and work behavior (Suar and Khuntia, 2010), justice perception (Schminke et al., 2015), attitudes toward sustainable business practices (Ng and Burke, 2010), corporate sustainability outcomes (Marcus et al., 2015), and pro-environmental behavior (Andersson et al., 2005; Gatersleben et al., 2014).

The notion of investors' values is also quite common in the literature on SRI, which itself has been referred to as values-based or values-driven investing (Derwall et al., 2011; Fehrenbacher, 2001; Schueth, 2003). However, although values research may provide important insights into the underlying motives for SRI, the respective research is diverse and not well integrated, implying that no coherent picture has yet evolved. Prior research has associated SRI with an extensive list of different types of values, including personal, organizational, societal,

sustainability, counter-cultural, political, religious, ethical, social, environmental (ecological), protected, and materialistic values, among others (Bengtsson, 2008; Diouf et al., 2016; Hofmann et al., 2009; Hong and Kostovetsky, 2012; Juravle and Lewis, 2009; Kashyap and Iyer, 2009; Pasewark and Riley, 2010). Furthermore, in some cases, inconclusive findings circulate in the SRI literature. For instance, whereas some scholars identify self-transcendent (environmental and social) values as an important motive for SRI (e.g., Brodback et al., 2018; Jansson and Biel, 2011), other researchers (e.g., Iyer and Kashyap, 2009; Kashyap and Iyer, 2009) make the case for materialistic values. Signs also point toward a lack of conceptual clarity as distinct concepts, such as attitudes and norms, are frequently subsumed under review sections on investors' values (e.g., Diouf et al., 2016, p. 48; Pasewark and Riley, 2010, p. 240).

One possible explanation for the present state of the SRI literature builds upon the peculiarities and pitfalls of values research, which constitutes a complex field of inquiry: values exist and (inter-)act at various levels (e.g., personal/individual, organizational, societal values), and a plethora of different values theories and measurement approaches are at researchers' disposal (Agle and Caldwell, 1999; Braithwaite and Scott, 1991; Watkins, 2010). In addition, scholars across disciplines often "use the word values in Humpty Dumpty fashion: They make it mean just what they choose it to mean" (Rohan, 2000, p. 255), despite considerable progress toward conceptual agreement and clarity (Agle and Caldwell, 1999; Connor and Becker, 1994; Schwartz, 1994; Schwartz and Bilsky, 1987). Values, therefore, continue to be regularly confused with distinct socio-psychological concepts, such as attitudes, norms, and traits (Connor and Becker, 1994; Hitlin and Piliavin, 2004; Rohan, 2000). This lack of conceptual clarity, however, comes with severe downsides that can eventually retard scientific progress. A lack of conceptual clarity makes it difficult to interpret the results of different (values) studies and to develop consistent findings across studies (Dose, 1997; Meglino and Ravlin, 1998), thereby impeding communication between researchers and the accumulation of a coherent body of knowledge (Connor and Becker, 1994). At the operational level, poorly defined concepts induce measurement problems by increasing the likelihood of deficient concept operationalizations and "a mismatch between the concept and measures or manipulations of it" (Podsakoff et al., 2016, p. 166). Clearly defined concepts further serve as the essential building blocks of theory (Bacharach, 1989; Klein and Zedeck, 2004; Newstead, 2018), whereas a lack thereof can lead to an unjustified "acceptance (or rejection) of theoretical statements" and the "unwarranted modification of theories" (Podsakoff et al., 2016, p. 166).

Despite the prominence of investors' values in the SRI literature, this article argues that the field's values perspective is not yet well developed and that the potential explanatory power

of values is thus far from being fully tapped. Against this backdrop, the present article seeks to substantiate this critical allegation, to cast light on the underlying reasons, and to offer promising paths to solutions. We first explicate central and commonly shared features in conceptual definitions of values and delineate related concepts that are often confused or conflated with values. Following this, we review the conceptualization of values in SRI research and provide a critique aimed at three major problem areas. First, our analysis indicates that values represent, to a considerable degree, a buzzword rather than a substantive concept in the SRI literature. Second, we document a substantial lack of conceptual clarity and conceptual confusion. In the great majority of reviewed articles, the concept of values is neither defined nor investigated on grounds of established values theories and/or measurement approaches. Third, we identify methodological issues regarding the measurement of values. Even SRI studies that build on established values instruments are associated with the incomprehensive application of values measures. Against this backdrop, we provide specific recommendations to address the identified conceptual and methodological issues. Furthermore, to advance the understanding of values as a motive for SRI, we offer promising future research avenues before closing with concluding remarks. We contribute to the field of SRI research by providing the foundation for more rigorous values inquiry and a stronger values-based perspective.

A primer on the concept of values

The study of values has a long tradition in a wide range of disciplines (Hitlin and Piliavin, 2004; Rohan, 2000), and values have been proclaimed as the central and integrative concept “able to unify the apparently diverse interests of all the sciences concerned with human behavior” (Rokeach, 1973, p. 3). Although values represent a prominent topic in the business world as well as in academic and public debates, two distinctively different meanings of the term need to be separated (Finegan, 2000; Meglino and Ravlin, 1998; Rokeach, 1973; Sánchez-Fernández and Iniesta-Bonillo, 2007; Schwartz and Bilsky, 1987). On the one hand, the word can be used in relation to the value that a subject places on an object, such as the monetary value assigned to a product or service. According to this understanding, the concept describes “the ‘value’ of an object” (Finegan, 2000, p. 150) and reflects the “outcome of an evaluative judgment” (Sánchez-Fernández and Iniesta-Bonillo, 2007, p. 429). On the other hand, the term can also connote a meaning that “is more likely to be used to describe a person as opposed to an object” (Meglino and Ravlin, 1998, p. 353). From this perspective, also known as the “values-as-crite-

tion” view (Finegan, 2000, p. 150; Williams, 1968, p. 283), values refer to the principles, standards, or criteria of social entities that serve as the basis for how they select or evaluate actions, people, and events (Sánchez-Fernández and Iniesta-Bonillo, 2007; Schwartz and Bilsky, 1987). By focusing on investors’ internalized values, this paper draws on the latter understanding of the term.

Multiple strands of literature have contributed to the understanding of values in the past. Given the dissemination of values research across all social science disciplines, it is no surprise that values have been conceptualized in different ways (Van Quaquebeke et al., 2014). Nevertheless, considerable progress toward conceptual clarity and agreement has been accomplished over the past decades (Agle and Caldwell, 1999; Connor and Becker, 1994; Schwartz, 1994; Schwartz and Bilsky, 1987). As most definitions of values tend to have several key features in common (Rindova and Martins, 2018; Schwartz, 1994; Van Quaquebeke et al., 2014), the concept has been assessed as “fairly well defined” (Agle and Caldwell, 1999, p. 359), and an emergence of “a consensual definition of values” (Connor and Becker, 1994, p. 68) has been asserted.

One of the most influential conceptions of values was provided by the American anthropologist Clyde Kluckhohn (1951, p. 395), who defined values as “a conception, explicit or implicit, distinctive of an individual or characteristic of a group, of the desirable, which influences the selection from available modes, means, and ends of action.” According to Hitlin and Piliavin (2004), this definition was particularly influential, as it considered values as an individual- and group-level phenomenon. Elaborating on Kluckhohn’s (1951) definition, Rokeach (1968, p. 124) conceptualized values as “abstract ideals, not tied to any specific object or situation, representing a person’s belief about modes of conduct and ideal terminal modes.” Reviewing and integrating the writings of seminal values theorists (e.g., Allport, 1961; Kluckhohn, 1951; Rokeach, 1973; Williams, 1968), Schwartz (1994, p. 21) provides the following summary definition of values: “desirable transsituational goals, varying in importance, that serve as guiding principles in the life of a person or other social entity.”

These definitions point toward several fundamental (defining) characteristics that are common to most conceptions of values and that make it possible to distinguish values from related concepts. First, values are transsituational as they represent abstract goals that apply across specific domains and contexts (Bardi and Schwartz, 2003; Connor and Becker, 1994; Hitlin and Piliavin, 2004; Rindova and Martins, 2018; Rokeach, 1973; Schwartz and Bilsky,

1987). Second, people hold numerous values, which are ordered in a personal hierarchy of importance (Bardi and Schwartz, 2003; Verplanken and Holland, 2002). Stated differently, values can be “ordered by importance relative to other values to form a system of value priorities” (Schwartz, 1994, p. 20). Third, values refer to ideals and the desirable and therefore carry an inherently positive valence (Hitlin and Piliavin, 2004; Rindova and Martins, 2018; Rokeach, 1973). The conceptualization of values as desirable goals also implies that values express motivational concerns (Bardi and Schwartz, 2003; Rindova and Martins, 2018; Van Quaquebeke et al., 2014; Verplanken and Holland, 2002). People who regard specific values as important are motivated to attain these goals by taking appropriate action. Values thus serve as guiding principles, i.e., they “guide selection or evaluation of behavior and events” (Schwartz and Bilsky, 1987, p. 551). Fourth, values not only exist at the individual level but also act as guiding principles for other social entities (e.g., groups, organizations, and countries), implying possible multilevel interaction among values (Agle and Caldwell, 1999; Kluckhohn, 1951).

According to Podsakoff et al. (2016, p. 165), “[c]lear conceptual definitions not only identify the nature of the concept of interest, but they also help distinguish the focal concept from other, seemingly similar concepts in the field.” In the following, the concept of values is therefore differentiated from related concepts that are often confused or conflated with values, namely, attitudes, norms, and traits. *Attitudes* are generally defined as favorable or unfavorable evaluations of concrete objects, situations, or entities (Ajzen, 1991; Hitlin and Piliavin, 2004; Rohan, 2000), whereas the abstract nature of values transcends specific situations (Hitlin and Piliavin, 2004; Ng and Burke, 2010; Rohan, 2004). Values form a structured system comprising different values that vary in importance (Schwartz, 1992, 1994). In contrast, “the core characteristic of an attitude is its variation on an evaluative dimension (favorable–unfavorable)” (Verplanken and Holland, 2002, p. 435). Unlike attitudes, which are relatively peripheral, values are more central to individuals’ personhood and, in turn, more durable (stable) over time (Hitlin, 2003; Hitlin and Piliavin, 2004; Meglino and Ravlin, 1998). Values and attitudes both represent antecedents of behavior, but values underlie attitudinal processes (Connor and Becker, 1994; Homer and Kahle, 1988; Rokeach, 1973), i.e., “[a]n attitude results from the application of a general value to concrete objects or situations” (Theodorson and Theodorson, 1969, p. 19).

Norms represent situation-specific expectations about appropriate or inappropriate behaviors, whereas values refer to desirable abstract goals that transcend specific situations (Hitlin and Piliavin, 2004; Rindova and Martins, 2018). Norms “specify how things should be done” (Scott, 2013, p. 64) and therefore “capture an ‘ought’ sense” (Hitlin and Piliavin, 2004, p. 361). People are usually aware of the fact that committing norm violations can lead to penalties and

sanctions. Thus, “[p]eople acting in accordance with values do not feel pushed as they do when acting under normative pressure” (Hitlin and Piliavin, 2004, p. 361).

Traits and values are both relatively stable across time and context (Roccas et al., 2002). However, traits refer to “dimensions of individual differences in tendencies to show consistent patterns of thoughts, feelings and actions” (McCrae and Costa, 1990, p. 23) and can therefore be understood as enduring dispositions, whereas values refer to enduring abstract goals that people consider desirable to pursue (Hitlin and Piliavin, 2004; Roccas et al., 2002). Furthermore, “[t]raits vary in the frequency and intensity of their occurrence” (Roccas et al., 2002, p. 790) and may be seen as positive or negative. Values, in contrast, are mainly considered positive, as they refer to the desirable and vary “in their importance as guiding principle[s]” (Roccas et al., 2002, p. 790). Traits form descriptive variables capturing “individual’s responses and behaviors,” whereas values are a motivational variable expressing “a person’s motivations that may or may not be reflected in behavior” (Parks-Leduc et al., 2015, p. 5). This distinction is important as specific values and traits can refer to the same term (e.g., aggression, creativity, ambition, competence), but with different meanings (Hitlin and Piliavin, 2004; Parks-Leduc et al., 2015; Roccas et al., 2002). For instance, as illustrated by Parks-Leduc et al. (2015), an individual needs to exhibit a certain degree of creative thinking, acting, or behaving to be described as a person possessing the trait of creativity. Creativity as a value, in contrast, refers to the importance that an individual assigns to creativity as a guiding principle in life, but this does not inevitably imply or lead to creative behavior. Stated differently, creative people (trait) may or may not be guided by the goal of creativity (value), and people guided by creativity (value) are not necessarily creative (trait).

As pointed out by Verplanken and Holland (2002, p. 434), “[i]t might be flattering to think of values as a fundamental characteristic of human beings.” To be informative, however, the study of values has to provide meaningful insights into behavioral domains (Bardi and Schwartz, 2003; Schwartz, 2005; Verplanken and Holland, 2002). As regards the relation between values and behavior, researchers emphasize humans’ desire for self-consistency (Bardi and Schwartz, 2003; Hitlin, 2003; Miles 2015; Rokeach, 1973). Substantial evidence indeed documents individuals’ motivation to act in a self-consistent manner (Robinson, 2007), which they can accomplish “if they see their different behaviors as manifestations of a common set of values” (Miles, 2015, p. 683). Furthermore, value-consistent behavior is usually deemed beneficial and rewarding as it supports individuals in achieving their desired goals (Bardi and Schwartz, 2003; Rindova and Martins, 2018). Research on the link between values and behavioral intentions in hypothetical situations indicates that people seek to live up to their important

values (Feather, 1995; Sagiv and Schwartz, 1995). In real-life situations, however, values represent one of many influencing variables that may shape behavior, implying possible inconsistencies in the value-behavior relationship (Bardi and Schwartz, 2003; Verplanken and Holland, 2002).

According to Schwartz (2005), four processes help to explain how values can influence behavior. The first linking process relates to *value activation*, a necessary condition for values to have an effect on behavior (Verplanken and Holland, 2002). In this context, the likelihood of activation is partly determined by the accessibility of a value, i.e., how easily it comes to mind. Values that are important (central) to a person tend to relate more strongly to behavior than values of lower importance. In a given situation, value-relevant aspects and clues can also reinforce the activation of high-priority values. The second linking process refers to *values as a source of motivation*. Considering a specific value of high importance to an individual (e.g., self-direction), respective actions that promote the attainment of the desired goal (self-direction) become more attractive to the individual than interfering actions. In some areas of life, behavior may result from carefully weighing up the advantages and disadvantages of alternatives. In these domains, people are likely to consciously consider their value priorities, which then affect decisions. Importantly, however, effects of values on behavior can also occur with little or no conscious awareness. For instance, values can “influence behavior through mechanisms, such as habits, that do not require conscious decisions” (Bardi and Schwartz, 2003, p. 1209). The third linking process builds upon the *influence of values on attention, perception, and interpretation in situations*. Following Verplanken and Holland (2002, p. 435), high priority (central) values represent desirable goals that influence how individuals “define and interpret a situation,” increase attention to value-relevant information, “enhance the weight of such information,” and “elicit a motivation to act” in value-consistent ways. The *influence of values on the planning of action* forms the fourth and last linking process suggested by Schwartz (2005). This process builds upon the argument by Gollwitzer (1996) that the motivation for thorough planning increases with the importance of goals. The higher the importance of a specific value, the higher therefore the probability that people make action plans that can lead to value-expressive behavior.

Scope and method of the literature review

To review the concept of values in extant SRI research, we applied the following systematic, multistep approach for selecting publications and organizing the review. First, based on prior

literature (e.g., Busch et al., 2016; Cadman, 2011; Capelle-Blancard and Monjon, 2012; Schueth, 2003), we developed a preliminary list of search terms that captured common synonyms for SRI. This list was subsequently discussed with experienced SRI scholars and slightly adapted.⁶

In a second step, we applied our search string to titles, abstracts, and keywords in the Social Sciences Citation Index (SSCI) database. The SSCI database covers all social sciences journals with an impact factor, which presumably represent the field's leading outlets. To obtain an even broader coverage of literature, the search was complemented with the following three major research databases: EBSCO Business Source Complete, EBSCO EconLit, and ScienceDirect. We additionally applied several limiters to only include English-language, peer-reviewed work published in academic journals. After the deletion of duplicates, this approach yielded an overall body of 1,098 articles.

Step three involved the identification of relevant articles based on a screening of each article's title, abstract, and keywords. In the first screening round, we eliminated all articles not broadly situated in the field of SRI. In the subsequent round, all remaining publications were analyzed regarding the (implied) meaning of the term "value*". In line with the goal of this review, articles were only included when the term described or referred to the internal characteristics of investors. In contrast to this subject-related understanding, articles were excluded if they related to an object or the outcome of an evaluative judgment (e.g., firm value, market value, or stock value). We further excluded those articles employing the term "value*" in a different context of meaning, for example, when stressing the value (originality) of the research conducted. Overall, this process resulted in the identification of 47 relevant articles.

Fourth, we conducted the actual review and synthesized our findings. To organize the review, we grounded our work on deductively derived categories. According to Connor and Becker (1994), three major problem areas related to the issues of neglect, level, and method

⁶ Target publications were required to include the word "value*" and have at least one match from the following list of synonyms: "sustainab* invest*", "ESG invest*", "eco* invest*", "green invest*", "environment* invest*", "social* invest*", "socially conscious invest*", "socially-conscious invest*", "socially aware invest*", "socially-aware invest*", "governance invest*", "SRI", "responsible invest*", "ethical invest*", "values invest*", "values based invest*", "values-based invest*", "values driven invest*", "values-driven invest*", "mission based invest*", "mission-based invest*", "mission related invest*", "mission-related invest*", and "impact invest*". We utilized the asterisk as a truncation symbol to allow for different endings of our search terms (e.g., invest, investing, investment, investments). Regarding the synonyms for SRI, we explicitly excluded variations of "value invest*", as this terminology refers to the investment strategy of selecting securities that appear undervalued by some form of fundamental analysis (Graham, 1949; Graham and Dodd, 1934).

impede the progress of values research. The problem of neglect involves the issue that researchers often neglect values as a possible variable in their investigations. Regarding the issue of level, Connor and Becker (1994) criticize that values are often treated as situation-specific concepts, contrary to the established conception of values at an abstract level. The issue of level thus deals with attitudes being mistaken for values and, in a more general sense, with a lack of conceptual clarity and conceptual confusion. With respect to the issue of method, Connor and Becker (1994) draw on two problematic observations. First, it is sometimes the case that a researcher “simply generates, ad hoc, an instrument to be used for the project at hand” and that subsequently “[a] list of terms are asserted to be values” (Connor and Becker, 1994, p. 71). A second methodological issue stems from researchers modifying “a traditionally valid and reliable instrument to fit the expedient needs of a project” (Connor and Becker, 1994, p. 71).

Building on Connor and Becker’s (1994) problem areas of neglect, level, and method, we derived three guiding questions to organize the review. First, do values form a substantive concept in SRI research, or are they more of a buzzword that is simply mentioned in relation to the topic of SRI (Section 4.1)? Second, to what extent is the SRI literature associated with a lack of conceptual clarity and conceptual confusion (Section 4.2)? Third, to what extent is the measurement of values in SRI research associated with methodological issues (Section 4.3)?⁷ The following chapter presents the synthesis of our findings.

The concept of values in SRI research: Three central problem areas

Buzzword or substantive concept?

All 47 articles included in the present review have two central aspects in common: they are all situated in the field of SRI research *and* refer to investors’ values in the title, abstract, or keywords. Yet, the identified body of literature differs considerably regarding the content-wise importance of the values concept. For instance, while some studies empirically investigate investors’ values, others simply mention the term in a rather superficial way. To reflect this het-

⁷ We fully acknowledge that the aforementioned questions are not completely independent from each other. In particular, one may argue that studies measuring values through self-generated instruments are more likely to suffer from conceptual confusion than studies utilizing established measures. In the following, methodological aspects will therefore also be taken into consideration when exploring the prevalence of a lack of conceptual clarity and conceptual confusion (Section 4.2). In contrast, methodological issues discussed in Section 4.3 exclusively relate to those SRI studies that build on established values theories and/or measurement approaches.

erogeneity and to classify the collected material accordingly, we formed the following two mutually exclusive groups. Articles assigned to Group 1 refer to investors' values but do not focus on the underlying motives for SRI. These studies, therefore, do not empirically study or elaborate on values in any depth. In contrast, articles assigned to Group 2 relate to values but, in addition, place the primary focus on investors' motives or the underlying factors that contributed to the development and growth of the SRI market. Allocation decisions were based on articles' stated research objective(s) and/or question(s). When no research objective or question was explicitly provided, a decision was made based on our own reading of the entire article.

Overall, we assigned 24 articles to Group 1. Although these articles do not focus on the underlying motives for SRI, they include a (very) limited number of references to investors' values, serving four major functions. One central function relates to the definition of SRI. In this context, SRI is usually defined as an investment approach allowing investors to integrate their personal values into investment decisions and to select securities that are consistent with their value system (e.g., Henningsen, 2002; Derwall et al., 2011; Sauer, 1997; Schueth, 2003; Shank et al., 2005). A second function is apparent in financial performance studies that do not provide evidence for any underperformance of SRI as compared to conventional forms of investing. Here, the respective authors conclude that socially responsible investors can align their values with their portfolios without sacrificing financial performance (e.g., BinMahfouz and Hassan, 2013; Hassan et al., 2005; Ortas et al., 2012, 2014). Other performance studies refer to investors' values as they investigate the financial performance of the KLD Catholic Values Index (e.g., Carosella et al., 2012; Kurtz and diBartolomeo, 2005) or conduct financial modeling to build portfolios that take investors' values into account (e.g., Bilbao-Terol et al., 2015; Simister and Whittle, 2013). A third function of the term *values* is associated with the historical emergence of SRI. In this regard, some authors briefly point out that modern SRI market growth can be explained by SRI activists' social and environmental values as well as profound societal value shifts (e.g., Vivo and Franch, 2009; Welker and Wood, 2011). The fourth and last function broadly relates investor values to the practical implementation of different SRI strategies (e.g., De Colle and York, 2009; Dembinski et al., 2003; Derwall et al., 2011; Henningsen, 2002; Jennings and Martin, 2007; Muñoz et al., 2014). In quantitative financial performance studies, for instance, values-based terminology is employed to distinguish between "values-driven" (using negative screens) and "profit-driven" (using positive screens) socially responsible investors (Derwall et al., 2011; Muñoz et al., 2014).

Based on the aforementioned criteria for categorization, we assigned 23 articles to Group 2 (see Appendix 1). This stream of literature relates to values and provides important

insights into the underlying drivers of private (retail) investors, professional investors, and/or, at the aggregate level, SRI market growth. In the context of *private* investors, environmental, social, societal, and health-related values have been found to positively relate to SRI in Austria, Canada, and the United States (Diouf et al., 2016; Hofmann et al., 2009; Pasewark and Riley, 2010). Based on research in Germany, Brodback et al. (2018) identify a positive (negative) relation between altruistic (egoistic) values and the decision to invest responsibly. Focusing on private Indian investors, Nair and Ladha (2014) identify collectivism, but not materialism, to significantly influence investors' pursuit of non-economic investment goals. On the contrary, in the study by Iyer and Kashyap (2009, p. 230), materialism relates positively to U.S. private investors' non-economic goals and social investing efficacy, whereas collectivism is not found to have a significant effect.⁸ Kashyap and Iyer (2009) provide a more nuanced picture revealing two different types of U.S. private investors attracted to SRI. Both types assign high importance to social goals but differ in terms of their economic goal priorities. The so-called "Sustainers" attach great importance to economic and social goals and "are likely driven by materialistic values" (Kashyap and Iyer, 2009, p. 127). In contrast, "Bleeding Hearts" regard social goals as central but assign low importance to economic goals, although they also seem to possess materialistic values. However, Bleeding Hearts are, more than any other investor group, guided by their protected values, implying a strong unwillingness "to compromise ideals to achieve financial gains" (Kashyap and Iyer, 2009, p. 127).

Work by Jansson and Biel (2011) investigates the underlying motives of Swedish investors and provides insights into the *private and professional* investment context. SRI among private and institutional investors is found to be guided by self-transcendent values (environmental and social values); yet, fund managers working in investment institutions are primarily affected by their own beliefs about the long-term returns of SRI and not by self-transcendent values.⁹ Focusing on major Swedish investment institutions (e.g., public and private pension fund companies and mutual fund companies), Jansson and Biel (2014) document that conventional (non-SRI) and socially responsible investors do not differ in their endorsement of self-enhancement and self-transcendent values: both groups regard self-enhancement values as of higher importance to their organization than self-transcendent values. In the SRI group, but not

⁸ Iyer and Kashyap (2009, p. 230) define social investing efficacy as "the assessment an individual makes of the likelihood that his/her investment will lead to the desired outcome or increase the probability of the desired outcome."

⁹ In the study by Jansson and Biel (2011), the term *institutional investor* captures entities that mainly invest their own capital (e.g., companies, municipalities, and universities), whereas fund managers working in investment institutions primarily manage the capital of other investors (e.g., investment banks and pension funds).

in the non-SRI group, values further influenced the intention to increase SRI in the future. However, the authors reveal a relationship in opposition to their own expectations, finding a positive (negative) association between self-enhancement values (self-transcendent values) and the intention to increase SRI. Hong and Kostovetsky (2012) explore the influence of U.S. mutual fund managers' political values on investment decisions, finding that Democrats underweight socially irresponsible companies. Research by Juravle and Lewis (2009) and Lewis and Juravle (2010, p. 486) reveals that sustainable investment champions in the UK "do not speak with one voice" as they are driven by a range of different values. The authors further provide evidence for the importance of values at different levels. At the organizational and institutional level, counter-cultural and materialistic values are found to impede SRI, whereas moral and sustainability values are identified as an important facilitator at the organizational level. At the individual level, sustainability values of sustainable investment champions formed a critical driver, particularly in the early days of SRI (Juravle and Lewis, 2009). A further stream of literature includes research on the link between SRI and the values of professional investors with a religious background. Building on an international sample, Louche et al. (2012) conclude that religious organizations do not consider investing to be at odds with their religious values. Quite the contrary, organizations' religious values form a key driver of responsible investment. According to Kreander et al. (2004), theological values, such as creationism, stewardship, agapism, engagement, and witness, guide the ethical investment programs of the Methodist Church and the Church of England. However, both churches are faced with "the challenge that the values at the core of a truly Christian ethic are ultimately incompatible with the values of capitalism" (Kreander et al., 2004, p. 433). Research by Kreander et al. (2009) provides additional evidence on a disconnection between the underlying values basis and actual investment practices. Despite the strong values-based mission of many charities, Kreander et al. (2009, p. 166) conclude that "the degree of values alignment at the operational level was often significantly lower than that espoused at the policy level."

Besides research with a focus on private and/or professional investors, Group 2 also includes several articles relating values to the *aggregate market level* and historic market growth in particular. For instance, Lewis and Cullis's (1990) hypothesis that the growing ethical investment segment can be seen as a reflection of societal value shifts (post-industrial values) finds support in Park's (2009, p. 208) assertion that the origins of the modern SRI movement can be traced back to "a radical change, in the way society viewed faith, values, and commerce." Research by Majoch et al. (2017) indicates that the first signatories (2006–2007) to the Principles for Responsible Investment (PRI) explain their signing decision by stating that the values

of the PRI reflect the values of their organization. However, reference to management values dropped dramatically in the subsequent years, suggesting a shift “from activist and socially driven to profit[-]driven” investors (Majoch et al., 2017, p. 734). Scholars also link the notion of values to SRI markets in specific geographic regions, such as France (Déjean et al., 2013), Norway (Reiche, 2010), Scandinavia (Bengtsson, 2008), and Nordic countries (Scholtens and Sievänen, 2013). For instance, covering Denmark, Finland, Norway, and Sweden, Scholtens and Sievänen (2013) relate Hofstede’s (1980, 1991) cultural value dimensions to the size and composition of SRI. Their findings suggest that “[f]emininity connects with more SRI in general, uncertainty avoidance with more core SRI” (Scholtens and Sievänen, 2013, p. 612).

With 47 identified articles, investors’ values represent a prominent and recurrent topic in the SRI literature. However, considering the distribution of articles between Groups 1 (24) and 2 (23) lends support to the assertion that values represent, to a considerable degree, a buzzword rather than a substantive concept. The following sections elaborate on the question of whether, or to what extent, Group 2 articles suffer from a lack of conceptual clarity and conceptual confusion (4.2) as well as problematic methodological practices (4.3).

The prevalence of a lack of conceptual clarity and conceptual confusion

Values are frequently conflated or confused with distinct socio-psychological concepts despite considerable progress toward conceptual agreement and clarity (Agle and Caldwell, 1999; Connor and Becker, 1994; Schwartz, 1994; Schwartz and Bilsky, 1987). In this context, the provision of proper concept definitions forms an essential step to foster conceptual clarity and mitigate the possibility of conceptual confusion (Podsakoff et al., 2016). Yet, considering the sample of Group 2 articles, only two studies provide an explicit definition of the term *values*. Drawing on Schwartz (1992), Jansson and Biel (2014, p. 35) define values “as guiding principles for individuals or organizations that go beyond specific situations.” Similarly, Brodback et al. (2018, p. 120) refer to Schwartz and Bilsky (1987, 1990) and “define values as concepts that relate to desired end states, which go beyond specific situations.” Our observation resonates with the assertion by Van der Wal et al. (2006, p. 317) that the term *values* “is often not defined at all, even in studies in which values themselves are the object of research.”

Other articles assigned to Group 2 do not provide a general definition of values but define specific values, such as materialism (Iyer and Kashyap, 2009; Kashyap and Iyer, 2009), collectivism (Iyer and Kashyap, 2009), protected values (Kashyap and Iyer, 2009), and management values (Majoch et al., 2017) as well as power distance, individualism, masculinity

(femininity), and uncertainty avoidance (Scholtens and Sievänen, 2013). However, taking a closer look at the respective measurement of values, the latter approach (i.e., defining specific values) does not entirely preclude the possibility of conceptual confusion. Building on Baron and Spranca (1997), Kashyap and Iyer (2009, p. 122) define protected values as “those that resist trade-offs with other values, particularly economic values.” To measure investors’ protected values, the authors confront participants with the following statement: “No matter how large the benefits, my personal values *would not allow me to ever invest* in a company that...” (Kashyap and Iyer, 2009, p. 133, italics in the original). Based on common SRI screening criteria, eight companies (firms that produce firearms, alcoholic beverages, tobacco, etc.) are subsequently presented to investors who respond on a strongly disagree to strongly agree scale. In the presented case, investors’ protected values are clearly not abstract but instead highly situation specific. Furthermore, the constructed statement and response format leads to answers reflecting investors’ favorable or unfavorable evaluations of specific entities (companies). As such, we argue that Kashyap and Iyer (2009) collect investors’ (protected) attitudes rather than protected values.

About half of all articles assigned to Group 2 explicitly seek to measure/collect values and therefore utilize specific values measurement instruments (see Appendix 1). In some instances, these instruments are grounded in established theories and approaches, such as Schwartz’s (1992) theory of human values or Hofstede’s (1980, 1991) cultural values dimensions. Other studies, in contrast, collect/measure investors’ values by means of self-generated approaches. Without providing a definition of the term *values* and by departing from established instruments, this stream of literature is deemed to be particularly vulnerable to conceptual confusion.

Diouf et al. (2016), for example, combine interviews and survey research to explore the underlying factors associated with SRI decisions. Their interview questions relate to the characteristics of those investors that choose SRI products, including “attitudinal (values based) characteristics” (Diouf et al., 2016, p. 51). The provided synthesis of the interview material, however, fails to distinguish between attitudinal and values-based characteristics. Considering the survey component, Diouf et al. (2016) conflate the concept of values with awareness by measuring investors’ social values through questions on investors’ awareness of ESG issues. Pasewark and Riley (2010) seek to explore the role of personal values in investment decision-making. However, the measurement of personal values is closely linked to the topic of tobacco use. More specifically, the authors utilize a tobacco survey comprising “eight items relating to medical, legal, and societal aspects of tobacco use” (Pasewark and Riley, 2010, p. 242). For

instance, participants are asked to indicate their level of (dis-)agreement with statements such as: “Tobacco use should be eliminated in public” (Pasewark and Riley, 2010, p. 252). The above statement captures the (un)favorable evaluation of a specific object (tobacco), suggesting that attitudes are again mistaken for values. Hong and Kostovetsky (2012) investigate the influence of mutual fund managers’ political values on investment decisions. To measure political values, the authors draw on the level of fund managers’ donations (contributions) to Democratic and Republican candidates. However, operationalizing values through behavior is problematic as values express motivational concerns “that may or may not be reflected in behavior” (Parks-Leduc et al., 2015, p. 5). Furthermore, in the case of value-ambivalent behaviors, the same behaviors can, for different individuals, be motivated by distinct (opposing) values (Lönnqvist et al., 2013).¹⁰ As such, it remains debatable whether, or to what extent, Hong and Kostovetsky (2012) actually measure fund managers’ political values.

About half of all articles assigned to Group 2 do *not* explicitly seek to measure/collect values. This stream of literature, for instance, elaborates on the underlying drivers for SRI and substantiates values-related claims by previous research. Other studies explore the motives of investors (e.g., via interviews) and derive values-related findings from the collected and analyzed material. The fact that these articles refrain from specific values measurement instruments makes it difficult, if not impossible, to precisely assess the prevalence of conceptual confusion. Yet, some indication of, at least, blurry conceptual boundaries is apparent in studies that draw values-related conclusions, although concepts other than values are stated to be the object of study. For instance, Hofmann et al. (2009) apply the theory of planned behavior (TPB) to explain ethical investment behavior. In the TPB, an individual’s behavioral intentions and behaviors are shaped by three determinants: attitude toward the behavior, subjective norms, and perceived behavioral control (Ajzen, 1991). With respect to the specific attitudinal factors influencing investment decisions, Hofmann et al. (2009, p. 112) point out that respondents’ “descriptions of these attitudes were relatively vague (for example, I hold certain values and certain moral standards, and I want to live by them—an ethical investor).” However, instead of extending the TPB, the respondents’ references to values and guiding principles are subsumed under the determinant of attitude. Furthermore, although Hofmann et al. (2009, p. 102) situate their research in “[a]ttitude theory,” the authors’ conclude as follows: “Our findings demonstrate how essential it is for ethical investors to lead a certain lifestyle driven by particular ecological

¹⁰ Lönnqvist et al. (2013) provide several examples of value-ambivalent behaviors. Publicly giving to charity may, for instance, “be driven by the intrinsic desire to help others” (universalism values), “or it could be a means of signalling to others one’s wealth and social status” (power values) (Lönnqvist et al., 2013, p. 556).

and social values.” Similarly, Louche et al. (2012, p. 302) claim to “study the opinions and practices of religious organisations towards investing in general and more especially towards responsible investing.” Opinions and practices are further specified as “expressions of attitudes” (Louche et al., 2012, p. 302). Yet, from the study’s findings, Louche et al. (2012) draw conclusions such as “investing is not perceived as being in contradiction with religious values, religious values are important drivers” (p. 301), and “[t]he relationship between religious values and investment strategy needs to be better understood” (p. 317).

Overall, our analysis lends support to the assertion that the concept of values lacks clarity in the SRI literature. Our analysis further points toward a considerable degree of conceptual confusion. In particular, blurry boundaries between the concepts of values and attitudes appear to be widespread.

Methodological issues in the measurement of values

Regarding the measurement of values, only a limited number of Group 2 articles (appear to) utilize established values instruments. Three studies (Brodback et al., 2018; Jansson and Biel, 2011, 2014) draw on the Schwartz (1992) theory of basic human values, which encompasses two fundamental aspects relating to the content and structure of values. First, Schwartz (1992) found that people recognize ten basic values that differ with respect to the underlying motivational goal expressed. Second, people vary in terms of their specific value priorities, although their value systems are structured in the same way. More specifically, the theory postulates that the ten values form a circular motivational continuum structured along two orthogonal dimensions: self-enhancement versus self-transcendence and conservation versus openness-to-change. The former dimension captures the conflict between the pursuit of self-interest and the concern for the interest and well-being of others. The latter dimension refers to the conflict between values emphasizing one’s own independent thought and action as well as readiness for change and values emphasizing self-restriction, preservation of traditional practices, and resistance to change (Schwartz, 1992, 1994). These relations of conflict and congruity are reflected in the circular motivational continuum, i.e., the “closer any two values in either direction around the circle, the more similar their underlying motivations; and the more distant any two values, the more antagonistic their underlying motivations” (Schwartz and Boehnke, 2004, p. 231).

Whereas Jansson and Biel (2011) only measure investors’ self-transcendent orientation, Brodback et al. (2018) and Jansson and Biel (2014) restrict the measurement of values to the

self-enhancement versus self-transcendence dimension. However, not measuring all values of the circular motivational continuum is problematic on various grounds. First, dropping values upsets the balance of value scales and can therefore bias subsequent responses, for instance, due to framing effects or social desirability. Furthermore, different motivations than expected may underlie a particular behavior (Lönngqvist et al., 2013), and these can be missed when studying only part of the circular continuum. For instance, “[t]raditional financial metrics integrate ESG factors with great difficulty as these intangibles are often more difficult to quantify” (Juravle and Lewis, 2009, p. 82). SRI may thus particularly resonate with investors that welcome novelty and challenge, i.e., give priority to openness-to-change values. On the other hand, religious investors have represented a central group in SRI ever since its inception (Eurosif, 2016; Kreander et al., 2004), and religiosity correlates positively with values favoring conservation (Saroglou et al., 2004; Schwartz and Huismans, 1995). As such, leaving out the openness-to-change versus conservation dimension may reduce the explanatory power of investors’ values in the SRI context. Recent sustainability-related research (Schaefer et al., 2018) provides evidence for the usefulness of studying the complete range of values when applying the theory of basic human values.

Scholtens and Sievänen (2013) build on Hofstede’s (1980, 1991) original cultural value dimensions (individualism-collectivism, power distance, uncertainty avoidance, masculinity-femininity) from the “four-dimension period” (Minkov and Hofstede, 2011, p. 13) but exclude the two dimensions that have been identified since. Research by Hofstede and Bond (1988) led to the adoption of a fifth universal dimension, “Confucian dynamism,” which captures a culture’s long-term versus short-term orientation. The sixth universal dimension, labeled “indulgence versus restraint,” was later added based on research by Minkov (2007, 2009). This dimension “reflects the degree to which it is culturally acceptable to indulge in leisurely and fun-oriented activities, either with family and friends or alone, and spend one’s money, at one’s own discretion” (Minkov, 2009, p. 174). Particularly as SRI is frequently associated with a “long-term paradigm” (Busch et al., 2016, p. 310), moving beyond Hofstede’s (1980, 1991) four dimensions promises deeper insights into cultural values as determinants for differences in SRI across countries.

Three SRI studies measure materialism (Iyer and Kashyap, 2009; Kashyap and Iyer, 2009; Nair and Ladha, 2014), with two of them (Iyer and Kashyap, 2009; Nair and Ladha, 2014) additionally collecting data on collectivism. Regarding the measurement of materialism, all three studies report the items used, but the respective source of items is not clearly/directly specified. The fact that all three studies refer to work by Marsha L. Richins (e.g., Richins, 1994;

Richins, 2004; Richins and Dawson, 1992) suggests that the Material Values Scale (MVS) provided the underlying measurement basis. In the original version, the MVS contained 18 items (Richins and Dawson, 1992). In subsequent years, a shorter version (a nine-item scale) was found to possess “acceptable levels of reliability and validity for measuring overall materialism” (Richins, 2004, p. 216). In the respective SRI studies, however, the number of items varies between three (Nair and Ladha, 2014), seven (Kashyap and Iyer, 2009), and eight (Iyer and Kashyap, 2009), and no rationale for dropping the item(s) is provided. Regarding the measurement of collectivism, Iyer and Kashyap (2009) employ five items without specifying their source. However, the “CVSCALE, a 26-item five-dimensional scale of individual cultural values that assesses Hofstede’s cultural dimensions at the individual level” can be identified as the underlying basis (Yoo et al., 2011, p. 193). Whereas the CVSCALE measures collectivism through six items, Iyer and Kashyap (2009) drop one item without explanation. The source of the four items used by Nair and Ladha (2014) to measure collectivism remains unspecified as well. Furthermore, the respective items do not represent standard measures.¹¹

From the reviewed material, only a limited number of studies (appear to) draw on established values instruments, and even these studies contain methodological issues regarding the measurement of values, especially concerning the incomprehensive application of measures.

Recommendations for future work

Building on Connor and Becker’s (1994) problem areas of neglect, level, and method, we have substantiated our critical assertion that the values perspective in SRI research is not yet well developed. Considering the number of articles assigned to Group 2 and, in particular, the limited number of studies explicitly seeking to measure investors’ values, we agree with Hong and Kostovetsky (2009) in affirming that the influence of values in investment decision-making is under-researched. As has been pointed out before in the values literature (e.g., Agle and Caldwell, 1999; Connor and Becker, 1994; Dose, 1997; Kirkman et al., 2006; Van der Wal et al., 2006), a lack of conceptual clarity coupled with the constant application of different instruments makes it almost impossible to develop a coherent body of knowledge. As such, we are cautious about the idea that “simply” more research will suffice and instead call for future research that

¹¹ A phrase search in Google Scholar for each of the four items did not result in a hit, apart from the study by Nair and Ladha (2014).

is more coherent. Against this backdrop, we first provide recommendations to address the identified conceptual and methodological issues. We subsequently specify future research avenues to advance the understanding of values as a motive for SRI.

Recommendations to address conceptual and methodological issues

We agree with Rohan (2000) in that much of the confusion surrounding the values concept could be reduced if scientists would clearly specify what they mean when they refer to values. To foster conceptual clarity and to prevent even more confusion, we thus encourage SRI scholars to provide an explicit definition of their conceptualization of values. With respect to the use of terminology, we further recommend employing values-related wording (e.g., values investing, values-driven investing, values-based investing, etc.) with care and deliberation. More specifically, we recommend against using the abovementioned terms to separate SRI from conventional investing or to distinguish between different SRI strategies, such as negative (values-driven) and positive (profit-driven) screening (see, e.g., Derwall et al., 2011; Muñoz et al., 2014). The reason for our reservation is twofold. First, such verbalizations do not explicate the specific values (e.g., self-transcendence) that may underlie SRI or a particular investment strategy. The meaningfulness of such terminology is thus limited, but it contributes to the manifestation of values as a buzzword. Second, the respective wordings preclude the conceivability that conventional and/or profit-driven investors may also be guided by specific values (e.g., self-enhancement or materialistic values), thereby fostering a simplistic view of the role of values in investment decisions.

To generate reliable and valid results and to enhance the comparability of findings, we additionally advise SRI scholars to comprehensively apply established values theories and measurement approaches. In the selection process, paying close attention to the level of analysis is imperative as values operate at various levels (Agle and Caldwell, 1999). For example, the Schwartz (1992, 1994) theory of basic human values and its revised version (Schwartz et al., 2012) have been validated in a multitude of cross-cultural samples (e.g., Bilsky et al., 2011; Lee et al., 2017; Schwartz and Boehnke, 2004; Schwartz and Rubel-Lifschitz, 2009; Schwartz and Sagiv, 1995) and therefore provide an excellent basis for values research at the individual level. Scholars seeking to explore the underlying drivers for SRI across countries can build on validated instruments at the cultural level, such as Hofstede et al.'s (2010) cultural value di-

mensions or Schwartz's (2006) theory of cultural value orientations. In this context, the interested reader is referred to Agle and Caldwell (1999), who provide a comprehensive review of values research in management, with a particular emphasis on the level of analysis.

Most of the reviewed articles that utilize specific instruments to explicitly measure values draw on surveys (see Appendix 1). Problematically, measures of the independent (values) and dependent (SRI) variable tend to be provided by the same person in a common measurement and item context. Furthermore, dependent SRI variables, derived from traditional survey research (e.g., "I will increase my investments in SRI assets during the next 2/5 years" (Jansson and Biel, 2011, p. 138)), are usually broadly phrased and draw on participants' implicit understanding of the term *SRI*. As how SRI is understood varies considerably (Cheah et al., 2011; Eurosif, 2016), the concept remains ambiguous to the participants in these studies. Overall, the observed relationships between an investor's values and SRI are thus exposed to various sources of common method biases (Podsakoff et al., 2003). One potential remedy for the above-mentioned issues lies in obtaining predictor and criterion variables from different methodological sources. Furthermore, as values instruments predominantly build on self-reported survey responses, we call for greater imagination regarding the measurement of the dependent (SRI) variable.

SRI scholars could, for instance, utilize quasi-experimental designs, such as the factorial survey (vignette) method (Oll et al., 2018), to collect criterion variables. Factorial surveys offer the opportunity to provide participants (e.g., investors) with multidimensional scenario descriptions (e.g., SRI products) to systematically manipulate those dimensions (e.g., high/low financial/sustainability performance, small/large company) and to determine the importance of dimensions in the decision-making process (e.g., likelihood to invest). In combination with survey-based values instruments, factorial surveys not only mediate common method biases but also enable a more granular operationalization of SRI. To enhance external validity, we further recommend more frequent utilization of administrative data, such as investors' monthly (SRI) fund holdings and linking to (quasi-)experimental and administrative data (Riedl and Smeets, 2017). We refer SRI scholars to the seminal article by Podsakoff et al. (2003) to identify potential sources of as well as remedies for common method biases.

Recommended future avenues to advance the understanding of values as a motive for SRI

Values exist and act at various levels, which implies the possibility of multilevel interaction (Agle and Caldwell, 1999; Kluckhohn, 1951). In the reviewed SRI literature, reference to investors' values is made at different levels, such as the personal/individual, organizational, or societal level. However, studies on the interaction of values at different levels are rare and do not build on specific multilevel values instruments (e.g., Juravle and Lewis, 2009). Future contributions are therefore likely to come from research adopting a multilevel perspective to the study of values that is based on established and validated measurement approaches. For example, in the majority of countries, SRI is predominantly undertaken by institutional investors (Eurosif, 2016; GSIA, 2016). As these professional investors are embedded within an organizational context, future research could examine the congruence between individual and organizational values. For instance, is there an alignment between the personal values of employed investors and the organization's values? If there is a "value fit," what are the specific values that underlie SRI? If individual and organizational values are not congruent, whose values are associated with SRI? Do organizational values rule or do the personal values of investors take precedence over organizational values? In the context of these sample research questions, we refer interested readers to the rich body of literature on person–organization (value) fit (e.g., Chatman, 1989; Edwards and Cable, 2009; Goodman and Svyantek, 1999; Kristof, 1996; Liedtka, 1989; O'Reilly et al., 1991; Vveinhardt and Gulbovaite, 2017).

SRI markets vary considerably across countries with respect to market size, growth rates, institutional and retail investors, asset classes, and investment strategies (Eurosif, 2016; GSIA, 2016; US SIF, 2018). Despite these substantial differences, the majority of the reviewed articles focus on SRI in a single country (see Appendix 1). Furthermore, (shifts in) values are often referred to as a central driver in the historical development of SRI markets. When considering the time dimension of Group 2 articles, however, longitudinal studies are surprisingly rare. Instead, researchers frequently use retrospective case analyses "where data were collected at one point in time, but events and histories were reconstructed" (Chandler and Lyon, 2001, p. 104). Against the backdrop of these observations, future contributions are likely to come from research conducting cross-country and longitudinal values studies. For instance, future research could merge longitudinal SRI market data with longitudinal values data to study associations between value shifts and changes in SRI markets. Collecting respective data for different countries would provide further insights into the relationship between societal values and the heterogeneity of SRI markets. Longitudinal SRI market data for different geographic regions are

available from the periodic trend reports published by Eurosif or the Global Sustainable Investment Alliance (GSIA). With respect to longitudinal, cross-country values data, scholars can directly draw on several major initiatives, such as the World Values Survey (WVS), the European Values Survey (EVS), and the European Social Survey (ESS). We refer interested readers to Jordaan et al. (2016) as an excellent example of a longitudinal values study in financial market research.

In the reviewed material, several studies provide empirical evidence for a significant association between investors' values and SRI. Although the employed reasoning and wording often suggest a causal relationship (e.g., values guide, drive, cause, influence SRI), the respective studies are essentially correlational.¹² Future contributions are therefore likely to come from research demonstrating causality in the value-behavior (SRI) relationship. A fruitful approach to establishing causality involves the technique of priming, "the activation of mental concepts through subtle situational cues" (Cohn et al., 2017). In an experimental setup, future research could prime investors' self-enhancement and/or self-transcendent values to explore whether the activation of specific personal values facilitates or impedes SRI. If the primed group is found to invest differently than the control (non-primed) group, evidence would be provided that the activation of values indeed causes behavior (Schwartz, 2005). We refer the interested reader to Verplanken and Holland (2002) as an insightful example of how environmental values can be primed and how priming can promote environmentally friendly choices. Besides the priming of personal/individual values, future research could also investigate how priming organizational values influences the propensity to engage in SRI practices. In this context, interested readers may find inspiration in research into the effects of priming bankers' professional identity on bankers' dishonesty (Cohn et al., 2014) and willingness to take risks (Cohn et al., 2017).

Conventional and socially responsible investors' decision-making does not come about in a vacuum but evolves from investors' processing of investment-relevant (financial and/or non-financial) information (Gödker and Mertins, 2017; Reimsbach et al., 2018). According to Maines and McDaniel's (2000) influential framework, investors' judgements develop in three stages comprising information acquisition, information evaluation, and information weighting. Considering the reviewed material, however, the link between investors' values and information processing is still unexplored. This void might come as a surprise as high-priority values influence how individuals attend to value-relevant information, weight such information, and

¹² This issue is not unique to SRI studies but relates to values research in general (Connor and Becker, 1994; Hitlin, 2003; Schwartz, 2005).

act upon it (Schwartz, 2005; Verplanken and Holland, 2002). Future contributions are therefore likely to come from research exploring the influence of investors' values in the "black box" between informational inputs (e.g., financial/non-financial information) and the decision outcomes (e.g., conventional versus SRI). Regarding the collection of data during the decision process, we recommend using process-tracing methods, such as verbal protocols, information boards (search displays), and eye-tracking methodology (Ford et al., 1989; Meißner and Oll, 2018). As eye-tracking methodology provides several advantages over other process-tracing methods (Ashby et al., 2016; Lohse and Johnson, 1996), we refer interested readers to recent eye-tracking studies in the investment context (Duclos, 2015; Hellmann et al., 2017; Rubaltelli et al., 2016; Shavit et al., 2010).

Conclusion

The present review lends support to the assertion that values represent, to a considerable degree, a buzzword rather than a substantive concept in the SRI literature. As pointed out by Thornton and Ocasio (2008, p. 99), buzzwords "are over used; as a result their meanings often get distorted and overextended and they burn-out of existence." Reviewing the findings of prior literature, we also compile empirical evidence indicating a significant relation between specific values and SRI. As such, we provide support for the idea that values form a promising concept to explain why some investors adopt SRI practices, whereas others refrain from doing so. However, no coherent picture has yet evolved in SRI research. We attribute this fact, at least partially, to the lack of conceptual clarity, the prevalence of conceptual confusion, and the constant application of different values measurement approaches. An alternative, although not mutually exclusive, explanation could be that SRI represents a value-ambivalent behavior, i.e., SRI may, for different investors, be motivated by the distinct values of each. However, disentangling the two explanations is not possible as long as the aforementioned conceptual and methodological issues prevail. Going beyond the review of prior literature, we thus provide recommendations to address the identified issues and to foster coherence in future research. In this regard, we recommend paying close attention to terminological and conceptual clarity, comprehensively applying established values theories and measurement approaches, and considering sources of as well as remedies for common method biases. To advance the understanding of values as a motive for SRI, we further specify promising future research avenues. Future contributions are likely to come from SRI research adopting a multilevel approach to the study of values, con-

ducting cross-country and longitudinal values studies, demonstrating causality in the value-behavior (SRI) relationship, and exploring the influence of values in the “black box” between information provision and SRI decision-making. Our hope is that this review of the current state of literature and recommendations for future inquiry will help to set the stage for the next phase of values studies in SRI research.

References¹³

- Agle, B. R. and Caldwell, C. B. (1999). Understanding research on values in business: A level of analysis framework. *Business and Society* 38(3), 326-387.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 50(2), 179-211.
- Allport, G. W. (1961). *Pattern and growth in personality*. Holt, Rinehart & Winston, New York.
- Andersson, L., Shivarajan, S. and Blau, G. (2005). Enacting ecological sustainability in the MNC: A test of an adapted value-belief-norm framework. *Journal of Business Ethics* 59(3), 295-305.
- Ashby, N. J., Johnson, J. G., Krajbich, I. and Wedel, M. (2016). Applications and innovations of eye-movement research in judgment and decision making. *Journal of Behavioral Decision Making* 29(2-3), 96-102.
- Bacharach, S. B. (1989). Organizational theories: Some criteria for evaluation. *Academy of Management Review* 14(4), 496-515.
- Bardi, A. and Schwartz, S. H. (2003). Values and behavior: Strength and structure of relations. *Personality and Social Psychology Bulletin* 29(10), 1207-1220.
- Baron, J. and Spranca, M. (1997). Protected values. *Organizational Behavior and Human Decision Processes* 70(1), 1-16.
- *Bengtsson, E. (2008). A history of Scandinavian socially responsible investing. *Journal of Business Ethics* 82(4), 969-983.
- *Bilbao-Terol, A., Arenas-Parra, M. and Cañal-Fernández, V. (2016). A model based on Copula Theory for sustainable and social responsible investments. *Revista de Contabilidad (Spanish Accounting Review)* 19(1), 55-76.
- Bilsky, W., Janik, M. and Schwartz, S. H. (2011). The structural organization of human values: Evidence from three rounds of the European Social Survey (ESS). *Journal of Cross-Cultural Psychology* 42(5), 759-776.
- *BinMahfouz, S. and Hassan, M. K. (2013). Sustainable and socially responsible investing: Does Islamic investing make a difference? *Humanomics* 29(3), 164-186.
- Braithwaite, V. A. and Scott, W. A. (1991). Values. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (eds.), *Measures of personality and social psychological attitudes*, (pp. 661-753). Academic Press, San Diego, California.

¹³ All 47 articles included in our review are marked with an asterisk (*).

- *Brodback, D., Guenster, N. and Mezger, D. (2018). Altruism and egoism in investment decisions. *Review of Financial Economics* 37(1), 118-148.
- Busch, T., Bauer, R. and Orlitzky, M. (2016). Sustainable development and financial markets: Old paths and new avenues. *Business & Society* 55(3), 303-329.
- Cadman, T. (2011). Evaluating the governance of responsible investment institutions: An environmental and social perspective. *Journal of Sustainable Finance & Investment* 1(1), 20-29.
- Capelle-Blancard, G. and Monjon, S. (2012). Trends in the literature on socially responsible investment: Looking for the keys under the lamppost. *Business Ethics: A European Review* 21(3), 239-250.
- *Carosella, N., Rodriguez, J., Williams, S., Nawrocki, D. and Doh, J. P. (2012). The return-to-risk performance of socially responsible investing according to Catholic values. *The Journal of Investing* 21(4), 47-58.
- Chandler, G. N. and Lyon, D. W. (2001). Issues of research design and construct measurement in entrepreneurship research: The past decade. *Entrepreneurship Theory and Practice* 25(4), 101-113.
- Chatman, J. A. (1989). Improving interactional organizational research: A model of person-organization fit. *Academy of Management Review* 14(3), 333-349.
- Cheah, E. T., Jamali, D., Johnson, J. E. and Sung, M. C. (2011). Drivers of corporate social responsibility attitudes: The demography of socially responsible investors. *British Journal of Management* 22(2), 305-323.
- Cohn, A., Fehr, E. and Maréchal, M. A. (2014). Business culture and dishonesty in the banking industry. *Nature* 516, 86-89.
- Cohn, A., Fehr, E. and Maréchal, M. A. (2017). Do professional norms in the banking industry favor risk-taking? *The Review of Financial Studies* 30(11), 3801-3823.
- Connor, P. E. and Becker, B. W. (1994). Personal values and management: What do we know and why don't we know more? *Journal of Management Inquiry* 3(1), 67-73.
- *De Colle, S. and York, J. G. (2009). Why wine is not glue? The unresolved problem of negative screening in socially responsible investing. *Journal of Business Ethics* 85(1), 83-95.
- *Déjean, F., Giamporcaro, S., Gond, J. P., Leca, B. and Penalva-Icher, E. (2013). Mistaking an emerging market for a social movement? A comment on Arjaliès' social-movement perspective on socially responsible investment in France. *Journal of Business Ethics* 112(2), 205-212.

- *Dembinski, P. H., Bonvin, J. M., Dommen, E. and Monnet, F. M. (2003). The ethical foundations of responsible investment. *Journal of Business Ethics* 48(2), 203-213.
- *Derwall, J., Koedijk, K. and Ter Horst, J. (2011). A tale of values-driven and profit-seeking social investors. *Journal of Banking & Finance* 35(8), 2137-2147.
- *Diouf, D. and Hebb, T. (2016). Exploring factors that influence social retail investors' decisions: Evidence from Desjardins fund. *Journal of Business Ethics* 134(1), 45-67.
- Dose, J. (1997). Work values: An integrative framework and illustrative application to organizational socialization. *Journal of Occupational and Organizational Psychology* 70(3), 219-241.
- Duclos, R. (2015). The psychology of investment behavior: (De)biasing financial decision-making one graph at a time. *Journal of Consumer Psychology* 25(2), 317-325.
- Eccles, N. S. and Viviers, S. (2011). The origins and meanings of names describing investment practices that integrate a consideration of ESG issues in the academic literature. *Journal of Business Ethics* 104(3), 389-402.
- Edwards, J. R. and Cable, D. M. (2009). The value of value congruence. *Journal of Applied Psychology* 94(3), 654-677.
- Eurosif. (2016). European SRI study 2016. EUROSIF A.I.S.B.L., Brussels.
- Feather, N. T. (1995). Values, valences, and choice: The influence of values on the perceived attractiveness and choice of alternatives. *Journal of Personality and Social Psychology* 68(6), 1135-1151.
- Fehrenbacher, S. (2001). Put your money where your morals are: A guide to values-based investing. Broadman and Holdman Publishers, Nashville, TN.
- Finegan, J. E. (2000). The impact of person and organizational values on organizational commitment. *Journal of Occupational and Organizational Psychology* 73(2), 149-169.
- Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. M. and Doherty, M. L. (1989). Process tracing methods: Contributions, problems, and neglected research questions. *Organizational Behavior and Human Decision Processes* 43(1), 75-117.
- Fritzsche, D. J. and Oz, E. (2007). Personal values influence on the ethical dimension of decision making. *Journal of Business Ethics* 75(4), 335-343.
- Gatersleben, B., Murtagh, N. and Abrahamse, W. (2014). Values, identity and pro-environmental behaviour. *Contemporary Social Science: Journal of the Academy of Social Sciences* 9(4), 374-392.

- Gehman, J., Trevino, L. K. and Garud, R. (2013). Values work: A process study of the emergence and performance of organizational values practices. *Academy of Management Journal* 56(1), 84-112.
- Glac, C. (2009). Understanding socially responsible investing: The effect of decision frames and trade-off options. *Journal of Business Ethics* 87(1), 41-55.
- Gödker, K. and Mertins, L. (2017). CSR disclosure and investor behavior: A proposed framework and research agenda. *Behavioral Research in Accounting* 30(2), 37-53.
- Gollwitzer, P. M. (1996). The volitional benefits of planning. In P. M. Gollwitzer & J. A. Bargh (eds.), *The psychology of action*, (pp. 287-312). Guilford, New York.
- Goodman, S. A. and Svyantek, D. J. (1999). Person–organization fit and contextual performance: Do shared values matter. *Journal of Vocational Behavior* 55(2), 254-275.
- Graham, B. (1949). *The intelligent investor: A book of practical counsel*. Harper and Brothers, New York.
- Graham, B. and Dodd, D. L. (1934). *Security analysis*. McGraw-Hill, New York.
- GSIA. (2016). Global sustainable investment review 2016. www.gsi-alliance.org/wp-content/uploads/2017/03/GSIR_Review2016.F.pdf
- *Haigh, M. (2006). Social investment: Subjectivism, sublation and the moral elevation of success. *Critical Perspectives on Accounting* 17(8), 989-1005.
- *Hassan, A., Antoniou, A. and Paudyal, D. K. (2005). Impact of ethical screening on investment performance: The case of the Dow Jones Islamic Index. *Islamic Economics Studies* 12-13(2-1), 67-97.
- Hellmann, A., Yeow, C. and De Mello, L. (2017). The influence of textual presentation order and graphical presentation on the judgements of non-professional investors. *Accounting and Business Research* 47(4), 455-470.
- *Hellsten, S. and Mallin, C. (2006). Are ‘ethical’ or ‘socially responsible’ investments socially responsible? *Journal of Business Ethics* 66(4), 393-406.
- Hemingway, C. A. and MacLagan, P. W. (2004). Managers’ personal values as drivers of corporate social responsibility. *Journal of Business Ethics* 50(1), 33-44.
- *Henningsen, C. (2002). Investing as if the world really mattered. *Corporate Environmental Strategy* 9(2), 163-171.
- Hitlin, S. (2003). Values as the core of personal identity: Drawing links between two theories of self. *Social Psychology Quarterly* 66(2), 118-137.
- Hitlin, S. and Piliavin, J. A. (2004). Values: Reviving a dormant concept. *Annual Review of Sociology* 30, 359-93.

- *Hofmann, E., Penz, E. and Kirchler, E. (2009). The ‘whys’ and ‘hows’ of ethical investment: Understanding an early-stage market through an explorative approach. *Journal of Financial Services Marketing*, 14(2), 102-117.
- Hofstede, G. (1980). *Culture’s consequences: International differences in work-related values*. Sage Publications, Beverly Hills.
- Hofstede, G. (1991). *Culture and organizations: Intercultural cooperation and its importance for survival - Software of the mind*. McGraw-Hill, London.
- Hofstede, G. (2001). *Culture’s consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage, Thousand Oaks, CA.
- Hofstede, G. and Bond, M. H. (1988). The Confucius connection: From cultural roots to economic growth. *Organizational Dynamics* 16(4), 5-21.
- Hofstede, G., Hofstede, G. J. and Minkov, M. (2010) *Cultures and organizations: Software of the mind*, 3rd edition. McGraw-Hill, New York.
- Homer, P. M. and Kahle, L. R. (1988). A structural equation test of the value-attitude-behavior hierarchy. *Journal of Personality and Social Psychology* 54(4), 638-646.
- *Hong, H. and Kostovetsky, L. (2012). Red and blue investing: Values and finance. *Journal of Financial Economics* 103(1), 1-19.
- *Iyer, E. S. and Kashyap, R. K. (2009). Noneconomic goals of investors. *Journal of Consumer Behaviour* 8(5), 225-237.
- *Jansson, M., and Biel, A. (2011). Motives to engage in sustainable investment: A comparison between institutional and private investors. *Sustainable Development* 19(2), 135-142.
- *Jansson, M. and Biel, A. (2014). Investment institutions’ beliefs about and attitudes toward socially responsible investment (SRI): A comparison between SRI and non-SRI management. *Sustainable Development* 22(1), 33-41.
- *Jennings, W. W. and Martin, G. W. (2007). Socially enhanced indexing: Applying enhanced indexing techniques to socially responsible investment. *The Journal of Investing* 16(2), 18-31.
- Jordaan, J. A., Dima, B. and Goleţ, I. (2016). Do societal values influence financial development? New evidence on the effects of post materialism and institutions on stock markets. *Journal of Economic Behavior & Organization* 132, 197-216.
- *Juravle, C. and Lewis, A. (2009). The role of championship in the mainstreaming of sustainable investment (SI): What can we learn from SI pioneers in the United Kingdom? *Organization & Environment* 22(1), 75-98.

- *Kashyap, R. and Iyer, E. S. (2009). Not everybody wants to save the world. *Journal of Financial Services Marketing* 14(2), 118-134.
- Kirkman, B. L., Lowe, K. B. and Gibson, C. B. (2006). A quarter century of “Culture’s Consequences”: A review of empirical research incorporating Hofstede’s cultural values framework. *Journal of International Business Studies* 37(3), 285-320.
- Klein, K. and Zedeck, S. (2004). Theory in applied psychology: Lessons (re)learned. *Journal of Applied Psychology* 89, 931-933.
- Kluckhohn, C. (1951). Values and value-orientations in the theory of action: An exploration in definition and classification. In T. Parsons & E. Shils (eds.), *Toward a general theory of action*, (pp. 388-433). Harvard University Press, Cambridge, MA.
- *Kreander, N., Beattie, V. and McPhail, K. (2009). Putting our money where their mouth is: Alignment of charitable aims with charity investments - Tensions in policy and practice. *The British Accounting Review* 41(3), 154-168.
- *Kreander, N., McPhail, K. and Molyneaux, D. (2004). God’s fund managers. *Accounting, Auditing & Accountability Journal* 17(3), 408-441.
- Kristof, A. L. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology* 49, 1-49.
- *Kurtz, L. and diBartolomeo, D. (2005). The KLD Catholic Values 400 Index. *The Journal of Investing* 14(3), 101-104.
- Lee, J. A., Sneddon, J. N., Daly, T. M., Schwartz, S. H., Soutar, G. N. and Louviere, J. J. (2017). Testing and extending Schwartz refined value theory using a best–worst scaling approach. *Assessment*, forthcoming.
- *Lewis, A. and Cullis, J. (1990). Ethical investments: Preferences and morality. *Journal of Behavioral Economics* 19(4), 395-411.
- *Lewis, A. and Juravle, C. (2010). Morals, markets and sustainable investments: A qualitative study of ‘Champions’. *Journal of Business Ethics* 93(3), 483-494.
- Liedtka, J. M. (1989). Value congruence: The interplay of individual and organizational value systems. *Journal of Business Ethics* 8(10), 805-815.
- Lohse, G. L. and Johnson, E. J. (1996). A comparison of two process tracing methods for choice tasks. *Organizational Behavior and Human Decision Processes* 68(1), 28-43.
- Lönnqvist, J.-E., Verkasalo, M., Wichardt, P. C. and Walkowitz, G. (2013). Personal values and prosocial behaviour in strategic interactions: Distinguishing value-expressive from value-ambivalent behaviours. *European Journal of Social Psychology* 43(6), 554-569.

- *Louche, C., Arenas, D. and Van Cranenburgh, K. C. (2012). From preaching to investing: Attitudes of religious organisations towards responsible investment. *Journal of Business Ethics* 110(3), 301-320.
- Maines, L. A. and McDaniel, L. S. (2000). Effects of comprehensive-income characteristics on nonprofessional investors' judgments: The role of financial-statement presentation format. *The Accounting Review* 75(2), 179-207.
- *Majoch, A. A., Hoepner, A. G. and Hebb, T. (2017). Sources of stakeholder salience in the responsible investment movement: Why do investors sign the Principles for Responsible Investment? *Journal of Business Ethics* 140(4), 723-741.
- Marcus, J., MacDonald, H. A. and Sulsky, L. M. (2015). Do personal values influence the propensity for sustainability actions? A policy-capturing study. *Journal of Business Ethics* 127(2), 459-478.
- McCrae R. R. and Costa, P. T. (1990). *Personality in adulthood*. Guilford, New York.
- McLachlan, J. and Gardner, J. (2004). A comparison of socially responsible and conventional investors. *Journal of Business Ethics* 52(1), 11-25.
- Meglino, B. M. and Ravlin, E. C. (1998). Individual values in organizations: Concepts, controversies, and research. *Journal of Management* 24(3), 351-389.
- Meißner, M. and Oll, J. (2018). The promise of eye-tracking methodology in organizational research: A taxonomy, review, and future avenues. *Organizational Research Methods*, forthcoming.
- Miles, A. (2015). The (re)genesis of values: Examining the importance of values for action. *American Sociological Review* 80(4), 680-704.
- Minkov, M. (2007). *What makes us different and similar: A new interpretation of the World Values Survey and other cross-cultural data*. Klasika y Stil Publishing House, Sofia, Bulgaria.
- Minkov, M. (2009). Predictors of differences in subjective well-being across 97 nations. *Cross-Cultural Research* 43(2), 152-179.
- Minkov, M. and Hofstede, G. (2011). The evolution of Hofstede's doctrine. *Cross Cultural Management: An International Journal* 18(1), 10-20.
- *Muñoz, F., Vargas, M. and Vicente, R. (2014). Fund flow bias in market timing skill. Evidence of the clientele effect. *International Review of Economics & Finance* 33, 257-269.
- *Nair, A. S., and Ladha, R. (2014). Determinants of non-economic investment goals among Indian investors. *Corporate Governance* 14(5), 714-727.

- Newstead, T. P., Macklin, R., Dawkins, S. and Martin, A. (2018). What is virtue? Advancing the conceptualization of virtue to inform positive organizational inquiry. *Academy of Management Perspectives* 32(4), 443-457.
- Ng, E. S. and Burke, R. J. (2010). Predictor of business students' attitudes toward sustainable business practices. *Journal of Business Ethics* 95(4), 603-615.
- Oll, J., Hahn, R., Reimsbach, D. and Kotzian, P. (2018). Tackling complexity in business and society research: The methodological and thematic potential of factorial surveys. *Business & Society* 57(1), 26-59.
- O'Reilly, C. A., Chatman, J. A. and Caldwell, D. F. (1991). People and organizational culture: A profile comparison approach to assessing person-organization fit. *Academy of Management Journal* 34(3), 487-516.
- *Ortas, E., Moneva, J. M. and Salvador, M. (2012). Does socially responsible investment equity indexes in emerging markets pay off? Evidence from Brazil. *Emerging Markets Review* 13(4), 581-597.
- *Ortas, E., Moneva, J. M., Burritt, R. and Tingey-Holyoak, J. (2014). Does sustainability investment provide adaptive resilience to ethical investors? Evidence from Spain. *Journal of Business Ethics* 124(2), 297-309.
- *Park, J. (2009). Sustainable consumption and the financial sector: Analysing the markets for responsible investment in Hong Kong and Japan. *International Journal of Consumer Studies* 33(2), 206-214.
- Parks-Leduc, L., Feldman, G. and Bardi, A. (2015). Personality traits and personal values: A meta-analysis. *Personality and Social Psychology Review* 19(1), 3-29.
- *Pasewark, W. R. and Riley, M. E. (2010). It's a matter of principle: The role of personal values in investment decisions. *Journal of Business Ethics* 93(2), 237-253.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y. and Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology* 88(5), 879-903.
- Podsakoff, P. M., MacKenzie, S. B. and Podsakoff, N. P. (2016). Recommendations for creating better concept definitions in the organizational, behavioral, and social sciences. *Organizational Research Methods* 19(2), 159-203.
- *Reiche, D. (2010). Sovereign wealth funds as a new instrument of climate protection policy? A case study of Norway as a pioneer of ethical guidelines for investment policy. *Energy* 35(9), 3569-3577.

- Reimsbach, D., Hahn, R. and Gürtürk, A. (2018). Integrated reporting and assurance of sustainability information: An experimental study on professional investors' information processing. *European Accounting Review* 27(3), 559-581.
- Richins, M. L. (1994). Special possessions and the expression of material values. *Journal of Consumer Research* 21(3), 522-533.
- Richins, M. L. (2004). The materialism values scale: Measurement properties and development of the short form. *Journal of Consumer Research* 31(1), 209-219.
- Richins, M. L. and Dawson, S. (1992). A consumer values orientation for materialism and its measurement: Scale development and validation. *Journal of Consumer Research* 19(3), 303-316.
- Riedl, A. and Smeets, P. (2017). Why do investors hold socially responsible mutual funds? *The Journal of Finance* 72(6), 2505-2550.
- Rindova, V. P. and Martins, L. L. (2018). From values to value: Value rationality and the creation of great strategies. *Strategy Science* 3(1), 323-334.
- Robinson, D. T. (2007). Control theories in sociology. *Annual Review of Sociology*, 33(1), 157-74.
- Roccas, S., Sagiv, L., Schwartz, S. H. and Knafo, A. (2002). The big five personality factors and personal values. *Personality and Social Psychology Bulletin* 28(6), 789-801.
- Rohan, M. J. (2000). A rose by any name? The values construct. *Personality and Social Psychology Review* 4(3), 255-277.
- Rokeach, M. (1968). *Beliefs, attitudes and values*. Jossey-Bass, San Francisco.
- Rokeach, M. (1973). *The nature of human values*. Free Press, New York.
- Rubaltelli, E., Agnoli, S. and Franchin, L. (2016). Sensitivity to affective information and investors' evaluation of past performance: An eye-tracking study. *Journal of Behavioral Decision Making* 29(2-3), 295-306.
- Sagiv, L. and Schwartz, S. H. (1995). Value priorities and readiness for out-group social contact. *Journal of Personality and Social Psychology* 69(3), 437-448.
- Sánchez-Fernández, R. and Iniesta-Bonillo, M. Á. (2007). The concept of perceived value: A systematic review of the research. *Marketing Theory* 7(4), 427-451.
- *Sandberg, J., Juravle, C., Hedesström, T. M. and Hamilton, I. (2009). The heterogeneity of socially responsible investment. *Journal of Business Ethics*, 87(4), 519-533.
- Saroglou, V., Delpierre, V. and Dernelle, R. (2004). Values and religiosity: A meta-analysis of studies using Schwartz's model. *Personality and Individual Differences* 37(4), 721-734.

- *Sauer, D. A. (1997). The impact of social-responsibility screens on investment performance: Evidence from the Domini 400 Social Index and Domini Equity Mutual Fund. *Review of Financial Economics* 6(2), 137-149.
- Schaefer, A., Williams, S. and Blundel, R. (2018). Individual values and SME environmental engagement. *Business & Society*, forthcoming.
- Schminke, M., Arnaud, A. and Taylor, R. (2015). Ethics, values, and organizational justice: Individuals, organizations, and beyond. *Journal of Business Ethics* 130(3), 727-736.
- *Scholtens, B. and Sievänen, R. (2013). Drivers of socially responsible investing: A case study of four Nordic countries. *Journal of Business Ethics* 115(3), 605-616.
- *Schueth, S. (2003). Socially responsible investing in the United States. *Journal of Business Ethics* 43(3), 189-194.
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. P. Zanna (ed.), *Advances in experimental social psychology*, vol. 25, (pp. 1-65). Academic Press, New York.
- Schwartz, S. H. (1994). Are there universal aspects in the structure and content of human values? *Journal of Social Issues* 50(4), 19-45.
- Schwartz, S. H. (2005). Robustness and fruitfulness of a theory of universals in individual human values. In A. Tamayo and J. Porto (eds.), *Valores e comportamento nas organizações* [Values and behavior in organizations], (pp. 56-95). Vozes, Petrópolis, Brazil.
- Schwartz, S. H. (2006). A theory of cultural value orientations: Explication and applications. *Comparative Sociology* 5(2), 137-182.
- Schwartz, S. H. and Bilsky, W. (1987). Toward a universal psychological structure of human values. *Journal of Personality and Social Psychology* 53(3), 550-562.
- Schwartz, S. H. and Bilsky, W. (1990). Toward a theory of the universal content and structure of values: Extensions and cross-cultural replications. *Journal of Personality and Social Psychology* 58(5), 878-891.
- Schwartz, S. H. and Boehnke, K. (2004). Evaluating the structure of human values with confirmatory factor analysis. *Journal of Research in Personality* 38(3), 230-255.
- Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., ... and Dirilen-Gumus, O. (2012). Refining the theory of basic individual values. *Journal of Personality and Social Psychology* 103(4), 663-688.
- Schwartz, S. H. and Huismans, S. (1995). Value priorities and religiosity in four Western religions. *Social Psychology Quarterly* 58(2), 88-107.

- Schwartz, S. and Rubel-Lifschitz, T. (2009). Cross-national variation in the size of sex differences in values: Effects of gender equality. *Journal of Personality and Social Psychology* 97(1), 171-185.
- Schwartz, S. H. and Sagiv, L. (1995). Identifying culture-specifics in the content and structure of values. *Journal of Cross-Cultural Psychology* 26(1), 92-116.
- Scott, W. R. (2013). *Institutions and organizations: Ideas, interests, and identities*, 4th edition. Sage, Thousand Oaks.
- *Seele, P., Jensen, O. A. and Dalva, C. D. (2015). Ethical free riding? The double entendre of “dirty hands” in finance, exemplified by ethical investment guidelines. *Studies in Communication Sciences* 15(1), 61-67.
- *Shank, T. M., Manullang, D. K. and Hill, R. P. (2005). Is it better to be naughty or nice? The *Journal of Investing* 14(3), 82-88.
- Shavit, T., Giorgetta, C., Shani, Y. and Ferlazzo, F. (2010). Using an eye tracker to examine behavioral biases in investment tasks: An experimental study. *Journal of Behavioral Finance* 11(4), 185-194.
- *Signori, S. (2009). Ethical (SRI) funds in Italy: A review. *Business Ethics: A European Review* 18(2), 145-164.
- *Simister, J. and Whittle, R. (2013). Ethical investment and portfolio theory: Using factor analysis to select a portfolio. *Journal of Mathematical Finance* 3, 145-152.
- Suar, D. and Khuntia, R. (2010). Influence of personal values and value congruence on unethical practices and work behavior. *Journal of Business Ethics* 97(3), 443-460.
- Theodorson, G. A. and Theodorson, A. G. (1969). *A modern dictionary of sociology*. Cromwell, New York.
- Thornton, P. H. and Ocasio, W. (2008). Institutional logics. In R. Greenwood, C. Oliver, R. Suddaby & K. Sahlin (eds.), *The Sage handbook of organizational institutionalism*, (pp. 99-129). Sage Publications, London.
- Unsworth, K. L., Dmitrieva, A. and Adriasola, E. (2013). Changing behaviour: Increasing the effectiveness of workplace interventions in creating pro-environmental behaviour change. *Journal of Organizational Behavior* 34(2), 211-229.
- US SIF. (2018). *Report on US sustainable, responsible and impact investing trends 2018*. US SIF, Washington, D.C.
- *Vandekerckhove, W., Leys, J. and Van Braeckel, D. (2008). A speech-act model for talking to management. Building a framework for evaluating communication within the SRI engagement process. *Journal of Business Ethics* 82(1), 77-91.

- Van der Wal, Z., Huberts, L., Van den Heuvel, H. and Kolthoff, E. (2006). Central values of government and business: Differences, similarities and conflicts. *Public Administration Quarterly* 30(3), 313-364.
- Van Quaquebeke, N., Graf, M. M., Kerschreiter, R., Schuh, S. C. and van Dick, R. (2014). Ideal values and counter-ideal values as two distinct forces: Exploring a gap in organizational value research. *International Journal of Management Reviews*, 16(2), 211-225.
- Verplanken, B. and Holland, R. W. (2002). Motivated decision making: Effects of activation and self-centrality of values on choices and behavior. *Journal of Personality and Social Psychology* 82(3), 434-447.
- *Vivo, L. A. and Franch, M. R. B. (2009). The challenges of socially responsible investment among institutional investors: Exploring the links between corporate pension funds and corporate governance. *Business and Society Review* 114(1), 31-57.
- Vveinhardt, J. and Gulbovaite, E. (2017). Models of congruence of personal and organizational values: How many points of contact are there between science and practice? *Journal of Business Ethics* 145(1), 111-131.
- Watkins, L. (2010). The cross-cultural appropriateness of survey-based value(s) research. *International Marketing Review* 27(6), 694-716.
- Weber, J. (2015). Identifying and assessing managerial value orientations: A cross-generational replication study of key organizational decision-makers' values. *Journal of Business Ethics* 132(3), 493-504.
- *Welker, M. and Wood, D. (2011). Shareholder activism and alienation. *Current Anthropology*, 52(S3), 57-69.
- Williams, R. M., Jr. (1968). Values. In E. Sills (ed.), *International encyclopedia of the social sciences*, (pp. 283-287). Macmillan, New York.
- Yoo, B., Donthu, N. and Lenartowicz, T. (2011). Measuring Hofstede's five dimensions of cultural values at the individual level: Development and validation of CVSCALE. *Journal of International Consumer Marketing* 23(3-4), 193-210.

Appendix 1: Summary of articles assigned to Group 2

References	Research methodology / data collection method	Data source / sample	Geographic focus	Time dimension	Definition of the term <i>values</i>	Values measurement instrument	Key values-related claims or findings
Bengtsson, 2008	Empirical: case study (interviews, document analysis)	Interviewees (not specified), annual reports, press releases, newspapers, etc.	Scandinavia	RCA	-	-	Changes in societal values shaped the development of Scandinavian SRI.
Brodback et al., 2018	Empirical: survey	306 individuals approached at a local citizen center	Germany	CS	✓	Five items to measure self-enhancement (SE) values and four items to measure self-transcendent (ST) values	There is a positive link between altruistic values and the relative importance of social responsibility. Egoistic values are negatively associated with the decision to invest responsibly, unless individuals associate responsible investing with higher returns.
Déjean et al., 2013	Non-empirical: commentary	-	France	RCA	-	-	The emergence of SRI in France was driven by loosely coordinated actors with conflicting values.
Diouf et al., 2016	Empirical: case study (secondary survey data, interviews)	893 retail investors (survey); 10 managers, analysts, and advisors (interviews)	Canada	CS	-	Questions about awareness of ESG issues (survey); questions about investors' attitudinal (values-based) characteristics (interviews)	Social values related to environmental, social, and governance issues represent one important factor that guides people's choices in SRI.
Hofmann et al., 2009	Empirical: interviews, focus groups	Seven financial experts (interviews); nine ethical and 12 conventional investors (focus groups)	Austria	CS	-	-	Ethical investment behavior is driven by ecological and social values.
Hong & Kostovetsky, 2012	Empirical: archival study	Morningstar Principia disks, CRSP mutual fund database, Thomson Reuters/CDA Spectrum mutual fund holdings database, FEC database, KLD database	US	L	-	Mutual fund managers campaign donations to Democratic and Republican candidates	Political values influence the investment decisions of mutual fund managers: managers who donate to Democrats underweight (relative to non-donors or Republican donors) stocks that are deemed socially irresponsible.

Iyer & Kashyap, 2009	Empirical: interviews, survey	15 interviewees (not specified), 216 MBA students	US	CS	-	Survey-based: Eight items to measure materialism, five items to measure collectivism	Materialism relates positively to investors' non-economic goals and social investing efficacy. Collectivism is not found to have a significant effect.
Jansson & Biel, 2011	Empirical: survey	60 employees from 19 investment institutions; 453 private investors; 71 institutional investors	Sweden	CS	-	Five items to measure self-transcendent (ST) values	ST values (environmental and social values) guide SRI among private and institutional investors but not among fund managers working in investment institutions.
Jansson & Biel, 2014	Empirical: survey	58 respondents from 17 different investment institutions (31 conventional and 27 responsible investors)	Sweden	CS	✓	Three items each to measure self-enhancement (SE) and ST values	Conventional and responsible investors both regard SE values to be more important in their own organization than ST values. In the non-SRI group, values do not influence the intention to increase SRI in the future. In the SRI group, self-enhancement values (self-transcendent values) are positively (negatively) associated with the intention to increase SRI.
Juravle & Lewis, 2009	Empirical: interviews	15 sustainable investing professionals	UK	RCA	-	-	At the institutional level, materialistic values impede SRI. At the organizational level, counter-cultural values impede, whereas moral and sustainability values facilitate SRI. At the individual level, sustainability values of SRI champions formed an important driver—particularly in the early days of SRI.
Kashyap & Iyer, 2009	Empirical: survey	348 retail investors	US	CS	-	Seven items to measure materialism, six items to measure protected values	Differences in materialistic and protected values explains disparities in investors' economic and social investing goals.
Kreander et al., 2004	Empirical: interviews, document analysis	Five church investment professionals, annual reports and policy documents from the Methodist Church and the Church of England	UK	RCA	-	-	Theological values played a significant role in the development of the ethical investment movement in the UK. Creationism, stewardship, agapism, engagement, and witness form the value basis of both churches ethical investment programs. Both churches face the challenge that values at the core of a truly Christian ethic

							are ultimately incompatible with the values of capitalism.
Kreander et al., 2009	Empirical: survey, interviews	88 charities (survey); 12 people with significant experience in charity investments	UK	CS	-	-	Given the values-based mission of charities, charities' investment behavior should exhibit the same values-based nature. Yet, the degree of values alignment at the operational level is often significantly lower than that espoused at the policy level.
Lewis & Cullis, 1990	Non-empirical: untested theoretical model	-	-	RCA	-	-	The growth of ethical investments might be viewed as the reflection of a societal value shift (post-industrial values).
Lewis & Juravle, 2010	Empirical: interviews	14 sustainable investing professionals	UK	RCA	-	-	Sustainable investment champions are driven by a range of values (e.g., environmental, religious), but do not speak with one voice.
Louche et al., 2012	Empirical: focus groups, survey	25 representatives from different religious organizations (focus groups); 103 religious organizations (survey)	North America, Europe, Africa, Asia, Oceania	CS	-	-	Religious organizations do not perceive investing as being in contradiction with religious values, but religious values represent an important driver of responsible investment.
Majoch et al., 2017	Empirical: survey	Signatories to the PRI (asset owners, asset managers, insurers) between 2007 (97 responses) and 2011 (464 responses)	Global	L	-	Analysis of signatories' responses to the question of why they signed the principles	The first signatories from 2006 to 2007 explain their decision to sign the PRI by stating that the values of the PRI reflect the values of their organization. Reference to management values dropped dramatically in the subsequent years.
Nair & Ladha, 2014	Empirical: survey	342 executives students with investment experience	India	CS	-	Four items to measure collectivism, three items to measure materialism	Collectivism, but not materialism, significantly influences Indian investors' pursuit of non-economic investment goals.
Park, 2009	Empirical: case study	Data from the Association for Sustainable & Responsible Investment in Asia	Japan, Hong Kong	CS, RCA	-	-	The origins of the modern SRI movement can be traced to the turbulent period in the 1960s when powerful social undercurrents changed the way society viewed faith, values, and commerce.
Pasewark & Riley, 2010	Empirical: experiment (decision task), survey	216 business students	US	CS	-	Eight items relating to medical, legal, and societal aspects of tobacco use.	Investors consider personal values in addition to financial factors in choosing in-

							vestments. Personal values regarding tobacco use interact with expected rates of return to determine an investment choice.
Reiche, 2010	Empirical: case study (interviews, document analysis)	Documents, policy papers; 12 representative actors from the Norwegian State Bank, the Ethics Council, and the Ministry of Finance	Norway	RCA	-	-	At the end of the 1990s, the voice of Norwegian civil society insisted that the sovereign wealth fund should not only ensure intergenerational justice but should also contribute to the implementation of values and norms of the present country.
Sandberg et al., 2009	Empirical: interviews, document analysis	14 SRI professionals; websites of signatories to the UN Principles for Responsible Investment	UK	RCA	-	-	The heterogeneity of the SRI movement can be explained by differences in values, norms, and ideology among the different actors which influence SRI or the SRI stakeholders.
Scholtens & Sievänen, 2013	Empirical: archival study	Data on economics, finance, culture, and institutions from Eurosif, OECD, UNDP, etc.	Denmark, Finland, Norway, Sweden	CS	-	Cultural value dimensions: individualism-collectivism, power distance, uncertainty avoidance, masculinity-femininity	MAS (femininity) and uncertainty avoidance can be associated with both size and composition of SRI in the four Nordic countries: Femininity connects with more SRI in general, uncertainty avoidance with more core SRI.

Note: L = longitudinal; CS = cross-sectional; RCA = retrospective case analyses; PRI = Principles for Responsible Investment

6 Appendices

6.1 Short Summary of Articles

Gemäß § 6 Abs. 5 der Promotionsordnung vom 24.08.2010 muss die Dissertation Kurzfassungen der Ergebnisse in deutscher und englischer Sprache enthalten. Diese werden in den folgenden Tabellen dargelegt.

Article 1
<p>Oll, J., Hahn, R., Reimsbach, D. & Kotzian, P. (2018). Tackling Complexity in Business and Society Research: The Methodological and Thematic Potential of Factorial Surveys. <i>Business & Society</i>, 57(1), 26-59. [VHB-JOURQUAL: B]</p>
<p><i>English summary:</i> Factorial surveys (FSs) integrate elements of survey research and classical experiments. Using a large number of respondents in a controlled setting, FSs approximate complex and realistic judgment situations through so-called vignettes—that is, carefully designed descriptions of hypothetical people, social situations, or scenarios. Despite being rooted, and predominantly applied, in sociology, FSs are particularly promising for business and society (B&S) scholars. Given the multiplicity, inherent complexity, and sometimes fuzziness of B&S research objects, conventional research methods inevitably reach their limits. This article, therefore, systematically presents methodological and thematic opportunities for FS studies in B&S research. It is argued that FSs are well suited to dealing with the complex interplay of societal-, organizational-, and individual-level factors in B&S research and to studying the principles underlying human perceptions, attitudes, values, social norms, and (anticipated) behavior. The application of the FS method is illustrated based on a showcase example in the realm of socially responsible investments (SRIs). As the literature on the conceptualization of FSs is limited, methodological challenges are addressed to guide B&S researchers past the common methodological pitfalls.</p>
<p><i>Deutsche Zusammenfassung:</i> Der faktorielle Survey (FS) integriert Bestandteile der Umfrageforschung mit Elementen klassischer Experimente. Aus diesem Grund können FSs sowohl von einer hohen Teilnehmerzahl profitieren als auch die Vorzüge einer kontrollierten Umgebung nutzen. Herzstück des FSs bilden die sogenannten Vignetten, die sorgfältig entworfene Beschreibungen von hypothetischen Menschen, sozialen Situationen oder Szenarien darstellen. Durch die Verwendung von Vignetten können sich FSs komplexen und realistischen Urteilssituationen annähern. Der FS entstammt ursprünglich der Soziologie und findet dort auch seine vorwiegende Anwendung. Gleichwohl ist der FS aber auch für Untersuchungen im Forschungsfeld Wirtschaft und Gesellschaft hervorragend geeignet.</p>

Grund hierfür ist, dass sich das Forschungsfeld durch einen hohen Grad an Komplexität und teilweise auch durch Unschärfe hinsichtlich zentraler Konzepte und Untersuchungsobjekte kennzeichnet. Herkömmliche Forschungsmethoden stoßen hierbei unweigerlich an ihre Grenzen. Dieser Artikel arbeitet daher systematisch das methodische sowie thematische Potenzial von FS-Studien für den Forschungsbereich Wirtschaft und Gesellschaft heraus. In diesem Zusammenhang wird argumentiert, dass der FS insbesondere gut geeignet ist, um das komplexe Zusammenspiel von Faktoren auf der gesellschaftlichen, organisationalen und individuellen Ebene abzubilden und zu beleuchten. Die Anwendung des FSs wird dabei anhand eines Vignetten-Beispiels im Kontext nachhaltiger Investitionen veranschaulicht. Da die methodische Literatur zur konkreten Umsetzung von FSs noch sehr begrenzt ist, werden ferner methodische Herausforderungen adressiert. Ziel ist es hierbei, Forscher des Forschungsfelds Wirtschaft und Gesellschaft an den gängigen methodischen Fallstricken vorbeizuführen.

Article 2

Bassen, A., Gödker, K., Lüdeke-Freund, F. & Oll, J. (2019). Climate Information in Retail Investors' Decision-Making: Evidence from a Choice Experiment. *Organization & Environment*, 32(1), 62-82. [VHB-JOURQUAL: B]

English summary: Financial markets play a decisive role in the transition to a low-carbon economy. This study investigates the role of climate information presentation for climate-friendly investing among retail investors. We conduct a choice experiment in which we vary the presentation format of climate information by means of three label designs to test their influence on investment practices. We provide empirical evidence for the effectiveness of climate labeling as a potential nudge for climate-friendly investing. Furthermore, we find heterogeneity in the influence of climate information across different label designs and cognitive characteristics of investors. Intuitive (reflective) decision makers tend to place significantly more (less) weight on funds' climate performance compared with financial performance—irrespective of a participant's environmental preference.

Deutsche Zusammenfassung: Finanzmärkte spielen hinsichtlich der Förderung einer nachhaltigen Entwicklung sowie des Übergangs zu einer kohlenstoffarmen Wirtschaft eine entscheidende Rolle. Vor diesem Hintergrund untersucht die Studie, ob bzw. inwieweit sich Veränderungen in der Präsentation von Klimainformationen auf das Investitionsverhalten

von privaten Anlegern auswirken. Basierend auf einem Auswahl-Experiment wird die Darstellung bzw. Präsentation von Klimainformationen entlang dreier Label-Designs variiert und der Einfluss auf das Investitionsverhalten beleuchtet. Die empirischen Studienergebnisse untermauern die Wirksamkeit von Klima-Labels als potenzielle Stellschraube bezüglich der Förderung klimafreundlicher Investitionen. Darüber hinaus zeigt die Studie auf, dass der Einfluss von Klimainformationen heterogen ist. Unterschiede existieren hierbei sowohl mit Blick auf die verschiedenen Label-Designs aber auch hinsichtlich der Charakteristika der Investoren. Im Vergleich zur finanziellen Performance der Aktienfonds, legen intuitive (reflektierende) Entscheider deutlich mehr (weniger) Gewicht auf die Klimaperformance. Dieser Zusammenhang vollzieht sich unabhängig und losgelöst von den Umweltpräferenzen der Studienteilnehmer.

Article 3

Meißner, M. & Oll, J. (2019). The Promise of Eye-Tracking Methodology in Organizational Research: A Taxonomy, Review, and Future Avenues. *Organizational Research Methods*, 22(2), 590-617. [VHB-JOURQUAL: A]

English summary: Technological advances in recent years have greatly lowered the barriers for using eye tracking (ET) as a research tool in laboratory and field settings. However, despite its potential and widespread application in other disciplines, the use of ET in organizational research remains sparse. This article therefore aims to introduce ET, and thus a new mode of behavioral data, to the field of organizational research. Based on a synthesis of prior literature, we propose an integrative taxonomy that unravels the methodological potential of ET as well as its scope of application. Building on our proposed taxonomy, we systematically review the use of ET in leading management journals and reflect on the current state of research. We further illustrate future avenues for ET in the domains of strategic management, entrepreneurship, and human resources to contribute to the method's future dissemination and to the advancement of organizational science as well.

Deutsche Zusammenfassung: In den letzten Jahren haben technologische Fortschritte bestehende Barrieren hinsichtlich der Anwendung von Eye-Tracking (ET) im erheblichen Maße abgebaut. Dies gilt für die Nutzung von ET als Forschungsinstrument in Labor- und Feldumgebungen gleichermaßen. Aus wissenschaftlicher Sicht bietet ET großes Potenzial und so überrascht es nicht, dass sich ET in vielen Disziplinen etabliert hat. Im Gegensatz

hierzu ist die Nutzung von ET in der Organisationsforschung erstaunlicherweise noch stark begrenzt. Der vorliegende Beitrag zielt daher darauf ab, ET in den Bereich der Organisationsforschung einzuführen. Basierend auf einer Synthese der bisherigen Literatur, wird eine integrative Taxonomie zur ET-Forschung entwickelt. Diese deckt sowohl das methodische Potenzial als auch das breite Anwendungsspektrum von ET auf. Auf der entwickelten Taxonomie aufbauend, wird im Anschluss die Verwendung von ET in führenden Management-Zeitschriften systematisch aufgearbeitet. Im Folgenden werden konkrete Forschungsfragen für ET in den drei Bereichen strategisches Management, Unternehmertum und Personalwesen herausgearbeitet. Diese dienen als Impuls für die zukünftige Verbreitung von ET und der Weiterentwicklung der Organisationswissenschaft gleichermaßen.

Article 4

Oll, J. (2022). The Concept of Values in Socially Responsible Investment Research: A Critical Evaluation and Recommendations for Future Work. *Unpublished Working Paper*.

English summary: Values form one of the most fundamental factors governing human decision-making, and prior research confirms the importance of values in various sustainability-related domains. In this vein, the study of investors' values could provide crucial insights into a key question in socially responsible investment (SRI) research: why do some investors adopt SRI practices, whereas others refrain from doing so? Despite the prominence of investors' values in the SRI literature, this paper argues that the field's values perspective is not yet well developed, a situation which impedes scientific progress and the accumulation of a coherent body of knowledge. To substantiate this assertion, we review the conceptualization of values in SRI research and provide a critique aimed at three major problem areas. First, our analysis indicates that values represent, to a considerable degree, a buzzword rather than a substantive concept in the SRI literature. Second, we document a substantial lack of conceptual clarity as well as conceptual confusion. Third, we identify methodological issues regarding the measurement of values. On this basis, we provide specific recommendations to address the identified conceptual and methodological issues. To advance the understanding of values as a motive for SRI, we further offer promising future research avenues. These contributions provide the foundation for more rigorous values inquiry and a stronger values-based perspective in the field of SRI research.

Deutsche Zusammenfassung: Im Kontext menschlichen Verhaltens stellen Werte einen zentralen Einflussfaktor dar. Darüber hinaus bestätigt die bisherige Forschung die hohe Bedeutung von Werten in verschiedenen nachhaltigkeitsrelevanten Entscheidungsbereichen. Vor diesem Hintergrund können Untersuchungen zu den Werten von Investoren entscheidende Einblicke in das Forschungsfeld nachhaltiger Investitionen bieten. Dies gilt insbesondere im Hinblick auf eine der Schlüsselfragen des Forschungsfelds, nämlich warum einige Anleger nachhaltig investieren und andere nicht. Die Forschung im Bereich nachhaltiger Investitionen nimmt regelmäßig auf die Werte von Anlegern Bezug. Dennoch argumentiert die vorliegende Studie, dass die Werteperspektive des Forschungsfelds bisher noch unterentwickelt ist. Dieser Zustand behindert den wissenschaftlichen Fortschritt und Erkenntnisgewinn im erheblichen Maße. Um die Aussage, dass die Werteperspektive des Forschungsfelds nachhaltiger Investitionen noch unterentwickelt ist, zu untermauern, setzt sich diese Untersuchung intensiv mit der bisherigen Konzeptualisierung von Werten auseinander. Hinsichtlich der Werteforschung im Bereich nachhaltiger Investitionen wird dabei eine Kritik entlang von drei zentralen Problemfeldern entwickelt. Zum einen weisen die Ergebnisse der Analyse darauf hin, dass Werte häufig eher als Schlagwort denn als substanzielles Konzept verwendet werden. Zweitens werden ein erheblicher Mangel konzeptioneller Klarheit sowie Fälle von konzeptioneller Verwechslung dokumentiert. Drittens identifiziert die Studie methodische Probleme hinsichtlich der Messung von Werten. Vor diesem Hintergrund werden spezifische Empfehlungen entwickelt, um die identifizierten konzeptionellen und methodischen Probleme zu beheben. Ferner werden zünftige Forschungsmöglichkeiten aufgezeigt, die zu einem besseren Verständnis von Werten als Motiv für nachhaltige Investitionen beitragen.

6.2 Statutory Declaration

Erklärung

Hiermit erkläre ich, Josua Erik Ottfried Oll, dass ich keine kommerzielle Promotionsberatung in Anspruch genommen habe. Die Arbeit wurde nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

Ort/Datum

Unterschrift Doktorand

Eidesstattliche Versicherung

Ich, Josua Erik Ottfried Oll, versichere an Eides statt, dass ich die Dissertation mit dem Titel: „Essays on the Advancement of Socially Responsible Investment Research: A Conceptual and Methodological Perspective“ selbst und bei einer Zusammenarbeit mit anderen Wissenschaftlerinnen oder Wissenschaftlern gemäß den beigefügten Darlegungen nach § 6 Abs. 3 der Promotionsordnung der Fakultät für Wirtschafts- und Sozialwissenschaften vom 24. August 2010 verfasst habe. Andere als die angegebenen Hilfsmittel habe ich nicht benutzt.

Ort/Datum

Unterschrift Doktorand

Unterschrift Verwaltung

6.3 Self-Declaration

Selbstdeklaration

Gemäß § 6 Abs. 2 b) der Promotionsordnung vom 24.08.2010 muss bei einer kumulativen Dissertation wie der vorliegenden, eine Liste der Titel der Einzelarbeiten vorgelegt werden sowie gemäß § 6 Abs. 5 eine Liste der aus dieser Dissertation bereits hervorgegangenen Veröffentlichungen. Entsprechend hierzu enthält die untenstehende Tabelle eine Liste der Titel der Einzelarbeiten dieser Dissertation sowie die Angabe des aktuellen Publikationsstandes.

Zusätzlich ist gemäß § 6 Abs. 3 der Promotionsordnung bei schriftlichen Promotionsleistungen, die in Zusammenarbeit mit anderen Wissenschaftlerinnen oder Wissenschaftlern entstanden sind, der Anteil der Doktorandin oder des Doktoranden eindeutig abzugrenzen. In der untenstehenden Tabelle lege ich deshalb dar, in welchem Umfang ich zu den in dieser Dissertation enthaltenen Artikeln beigetragen habe. Die Einschätzung teilt sich dabei gemäß § 6 Abs. 3 der Promotionsordnung in die drei Teilbereiche Konzeption, Durchführung und Berichtsabfassung auf und diese wurden wie folgt bewertet:

A Ich leistete einen Beitrag (0-33 Prozent)

B Ich leistete einen substanziellen Beitrag (34-66 Prozent)

C Ich leistete die Mehrheit der Arbeit eigenständig (67-100 Prozent)

Artikel	Status der Publikation	Selbstdeklaration	
Artikel 1: Oll, J., Hahn, R., Reimsbach, D. & Kotzian, P. (2018). Tackling Complexity in Business and Society Research: The Methodological and Thematic Potential of Factorial Surveys. <i>Business & Society</i> , 57(1), 26-59.	Publiziert [VHB-JOUR-QUAL: B]	Konzeption	B
		Durchführung	B
		Berichtsabfassung	B
Artikel 2: Bassen, A., Gödker, K., Lüdeke-Freund, F. & Oll, J. (2019). Climate Information in Retail Investors' Decision-Making: Evidence from a Choice Experiment. <i>Organization & Environment</i> , 32(1), 62-82.	Publiziert [VHB-JOUR-QUAL: B]	Konzeption	B
		Durchführung	B
		Berichtsabfassung	B
Artikel 3: Meißner, M. & Oll, J. (2019). The Promise of Eye-Tracking Methodology in Organizational Research: A Taxonomy, Review, and Future Avenues. <i>Organizational Research Methods</i> , 22(2), 590-617.	Publiziert [VHB-JOUR-QUAL: A]	Konzeption	C
		Durchführung	C
		Berichtsabfassung	C
Artikel 4: Oll, J. (2022). The Concept of Values in Socially Responsible Investment Research: A Critical Evaluation and Recommendations for Future Work.	Unpublished Working Paper	Allein-Autorenschaft	C