# Engineering Adaptive Interfaces – Enhancement of Comprehension and Decision-Making

Zur Erlangung des akademischen Grades eines

Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

von der KIT-Fakultät für

Wirtschaftswissenschaften

am Karlsruher Institut für Technologie (KIT)

genehmigte

DISSERTATION

von

M.Sc. Hakan Pusmaz

Tag der mündlichen Prüfung: 10. Mai 2023

Referent: Prof. Dr. Christof Weinhardt

Korreferent: Prof. Dr. Verena Dorner

Karlsruhe, 2023

# Abstract

The role of information systems is growing steadily and permeating more and more all levels of our society. Meanwhile, information systems have to support different user groups in various decision situations simultaneously. Hence, the existing design approach to creating a unified user interface is reaching its limits. This work examines adaptive information system design by investigating user-adaptive information visualization and situation-aware nudging.

An exploratory eye-tracking study investigates participants' perception and comprehension of different financial visualizations and shows that none of them can be preferred across the board. Moreover, it reveals expertise knowledge as the research direction for visualization recommendations. Afterward, two empirical studies are conducted to relate different visualizations to participants' domain-specific knowledge. The first study, conducted with a broad sample of the population, shows that financial and graphical literacy increases participants' financial decision-making competency with certain visualizations. The second study, conducted with a more specific sample and an additional visualization, underlines a large part of the first study's results. Additionally, it identifies statistical literacy as an increasing factor in financial decision-making. Both studies are demonstrating that different visualizations cause different cognitive loads despite the same amount of information. After all, the results are used to derive visualization recommendations based on domain-specific knowledge and cognitive load.

This work also investigates the situation-aware effectiveness of nudging with the example of decision inertia. In a preliminary study, an experimental task is systematically transferred to different situational contexts by observing situational user characteristics. The identified contexts are examined in a subsequent large-scale empirical study with different nudges to reduce decision inertia. The results show gender-specific differences in decision inertia across the context. Hence, information system design has to adapt to gender and situational user characteristics to support users in their decision-making. Moreover, the study delivers empirical evidence for the contextual effectiveness of nudging. Future nudging research has to incorporate situational user characteristics to provide effective nudges in different situational contexts. Especially, further fundamental research is needed to understand the situational effectiveness of nudging. The study identifies individual situational preferences as one promising research stream.

# Contents

# List of Abbreviations

| | | |
|---|---|---|
| *IS* | Information System | 3 |
| *UI* | User Interface | 4 |
| *SAAP* | Situational Affordance and Adaptive Problems | 30 |
| *SIS* | Social Interdependence Scale | 30 |
| *RSQ* | Riverside Q-Sort | 31 |
| *AOI* | Area of Interest | 54 |
| *KIT* | Karlsruhe Institute of Technology | 89 |
| *KD*$^2$*Lab* | Karlsruhe Decision and Design Lab | 89 |
| *AME* | Average Marginal Effect | 96 |
| *Urn* | Urn Game | 116 |
| *Robo* | Robo-advisor | 117 |
| *Dating* | Dating Game | 117 |
| *Exam* | Exam Game | 117 |

# List of Figures

# List of Tables

# Part I

# Introduction

# Chapter 1

# Introduction

> “ The beginnings of all things are small”
>
> MARCUS TULLIUS CICERO, 1813

## 1.1 Motivation

THE digital transformation gained more and more momentum in the last two decades (Hanelt et al., 2021). Companies have been developing fundamentally new business models, products, and services (Châlons and Dufft, 2016). At the same time, more and more people shift common decisions, such as dating, shopping, or financial planning, into the digital sphere by using information systems (information system: IS).[1] As shown by digital government services during the Covid-19 pandemic, there are also cases in which analog services are substituted by digital counterparts so that users are forced to switch (Seifert, 2020; Burlacu et al., 2021). As a result, ISs have to support a growing diversity of users in an increasing variety of decisions.

Current ISs failing to live up to their growing responsibility to support a broad spectrum of our society in their decision-making (Miraz et al., 2021). In numerous cases, only a

[1] `https://www.dentsu.com/de/de/dsi-tech` last accessed: August 23, 2022.

particular group of users is targeted in the development process (Hussain et al., 2018; Miraz et al., 2021). The result is one unified user interface (UI) designed to meet the specific group's requirements, leaving the needs of other user groups unsatisfied (Hussain et al., 2018). For example, not customized robo advisors are not capable of meeting novice investors needs concerning the trustworthiness (Jung et al., 2018).

If ISs are not properly designed, comprehension of containing information is impaired, leading to trust issues, decreased user satisfaction, and lower acceptance (Jung et al., 2018; Belanche et al., 2019; Miller et al., 2005). Even though ISs typically allow users to manually set UI parameters to personalize them according to their needs, the opportunity is rarely used (Langley, 1999). Many users find this annoying, and some needs are reflected in their interactions but are not subject to introspection (Langley, 1999). In addition, the complexity for users is increased as they have to do fundamental work by learning the adaptation possibilities (Fischer, 2001). Paradoxically, inexperienced users would benefit most from personalization and still have to put in additional learning effort to adapt the UI to their needs (Mackay, 1991). Therefore, the question arises of how ISs needs to be designed to be simultaneously useful for various user groups in various decision situations.

Usefulness is defined as the degree to which an IS supports an individual in his or her decision-making (DeLone and McLean, 1992). IS design has to improve the decision-making competency of an individual to come up with a decision from an available set of decision alternatives against the chosen criteria for a given decision goal (Seddon et al., 1996). Two major aspects influence the competence in decision-making: (1) comprehension of the decision-relevant information and (2) rational and internally consistent information integration (Finucane et al., 2002).

The comprehension of decision-relevant information is influenced by its representation (Rudolph et al., 2009). Data visualization is an essential means of communicating decision-relevant information in ISs (Qin et al., 2018; Conati et al., 2015). As long as certain user characteristics are present, it provides more efficient information communication than verbalized data. Users' cognitive abilities influence their comprehension of particular visualizations (Steichen et al., 2013; Conati et al., 2015; Toker et al., 2012; Lee et al., 2019). These could be queried by the IS in order to recommend a visualization which improves users' comprehension (Toker and Conati, 2014).

However, there are reasons not to query cognitive abilities. The measurement tools developed to this date are pretty time-consuming. (Velez et al., 2005; Carenini et al., 2014; Toker et al., 2012). In addition, users might not want to disclose their cognitive abilities (Bansal et al., 2010). The cognitive state of an individual is particularly sensitive in terms of privacy. It allows conclusions to be drawn about the mental health, such as cognitive disabilities (Zhang et al., 2018; Jang and Yoo, 2009). In order to enhance comprehension, user characteristics are needed that have an acceptable measurement effort and are less sensitive to privacy concerns. Alternatively, research is needed to identify user characteristics by the interactions with the IS to customize the visualization during the use.

Besides comprehension, the rational and internally consistent information integration, which is the ability to weigh decision-relevant information in an internally consistent manner, affects decision-making competency (Finucane et al., 2002). An internal consistency ensures that an individual repeatedly makes good decisions in the same or similar decision situation and can be manipulated by the choice architecture (Malloy et al., 1992).

Choice architecture interventions, such as nudges, can change the expected utility of choice outcomes (Fischer et al., 2014). Nudges are well-established means in changing users' behavior in ISs (Sunstein and Thaler, 2008; Weinmann et al., 2016). A nudge is not always equally effective. There are cases in which a previously effective nudge has not changed behavior or even has an opposite effect than intended (Bolton et al., 2019). Most scholars investigated user characteristics in order to explain the differences in effectiveness (Ingendahl et al., 2021; Zhang and Xu, 2016; Beshears and Kosowsky, 2020).

However, there are meta-analyses suggesting situational dependency of nudging (Lehner et al., 2016; Hummel and Maedche, 2019; Mertens et al., 2022). The scholars identified large differences in nudge effectiveness across various situational contexts, such as health, environment, and privacy. The pending empirical evidence of the situational effectiveness of nudges leads to the application of situational information as a further design criterion for nudges.

Examining situational differences in nudge effectiveness requires a research subject that leads to systematic irrational behavior and occurs in different situational contexts. For this purpose, cognitive biases are suitable since they occur in different situational contexts and have already been well researched in combination with nudges (Sunstein, 2018). Decision inertia is one of these, resulting in systematic deviations from rationality. It describes the subconscious tendency to repeat a previously made choice, regardless of its outcome (Alós-Ferrer et al., 2016). The phenomenon is present in different situational contexts, such as finance, e-mobility, and ethical decision-making (Zhang et al., 2014; Jung et al., 2018; Stryja et al., 2017). The relevance for IS design, especially for choice architecture, could already be shown (Jung and Weinhardt, 2018; Jung et al., 2018). In financial decision situations, particular nudges improve decision-making competency by reducing decision inertia (Jung and Weinhardt, 2018). Research in other situational contexts is not available yet, making decision inertia the ideal candidate to explore the contextual effectiveness of nudges.

## 1.2 Research Outline

This thesis stresses the importance of personalized IS design to increase individuals' decision-making competency. For this purpose, user-specific information visualizations and context-aware nudging are investigated.

First of all, the overarching question is how to leverage a user-characteristics-based information visualization. An exploratory eye-tracking study is conducted in order to get a deeper understanding of how users perceive visualizations and which factors influence their information processing. Within the eye-tracking study, a recorded presentation with information about market bubbles and different financial visualizations, such as tables and line charts, is used as the experimental task. Participants' perceptual and cognitive processes during the presentation are examined with eye-movement analysis. In particular, the study addresses the following research question:

> **Research Question 1:** *How do users comprehend visualizations, and what factors influence their comprehension?*

The exploratory study showed differences in comprehension of the investigated visualizations and revealed experience as a potential influencing factor of visualization comprehension. Graphical and financial literacy are identified as user characteristics that reflect users' experience with two different visualizations, namely line charts and tabular representation. Two experimental studies are conducted to clarify the following research question:

**Research Question 2:** *How do user characteristics influence the comprehension of information with different visualizations?*

The cognitive bias decision inertia is chosen to analyze situational differences in nudges effectiveness. Situational differences could also exist in the case of decision inertia. For this purpose, a well-established experimental task for investigating decision inertia is systematically transferred to different situational contexts. These are used in a subsequent experimental study to determine the situational dependency of decision inertia addressing the following research question:

**Research Question 3:** *How do situational characteristics influence decision inertia?*

Finally, the situational dependency of nudging is examined with decision inertia as the research subject. Based on existing research, particular nudges are identified, which could reduce decision inertia. An experimental study examines their effectiveness in reducing decision inertia across situational contexts. The study addresses the following research question:

**Research Question 4:** *How do situational characteristics affect the effectiveness of nudges to reduce decision inertia?*

## 1.3 Structure of the Thesis

The structure of the thesis (see figure 1.1) is based on the research outline described in the previous section. The thesis comprises six parts. Part I motivates the thesis and introduces the research agenda and research questions.

Part II contains the foundations of the thesis. Chapter 2 shows the role of adaptive information visualization. Moreover, it defines adaptive information visualization and provides a taxonomy for visualization classification. Finally, it reviews relevant literature on adaptive visualization systems. Chapter 3 introduces situational research, provides situation terminology, introduces situational characteristics, and provides measurement tools.

Chapter 4 introduces decision inertia from a dual-processing perspective and reviews relevant literature considering the framing and applicational context of decision inertia. Chapter 5 explains nudging, especially in the digital context. It reviews relevant literature on nudges to reduce decision inertia or similar phenomena.

Part III investigates adaptive information visualization. Chapter 6 presents a qualitative and exploratory study to gain insights into perception and comprehension-related factors of different visualization in financial decision-making. The eye-tracking methodology measures participants' eye movements during a recorded presentation with different visualizations. None of the investigated visualizations is superior concerning comprehension and cognition. The line chart and tabular representation are identified as the most common visualizations in the financial context. Moreover, the user characteristics reflecting occupational experience or expertise could be used for visualization recommendations.

Chapter 7 consists of two studies to empirically investigate user-adaptive visualizations. The first study validates the experimental design and confirms the main assumptions that are derived in the exploratory study in chapter 6. In the second study, the experiment is replicated with extensions. Financial, graphical, and statistical literacy are identified as increasing factors of financial decision-making. Moreover, financial literacy can be used to recommend line charts and tabular representations to a broad sample of the population.

Part IV contains a large-scale study that incorporates perceived situational characteristics to investigate the situational dependencies of decision inertia and nudging. Using situational characteristics, an urn game, which is an established experimental task to investigate decision inertia, is transferred to different situational contexts. These are used to investigate decision inertia and nudges situationally. Results reveal situational drivers of decision inertia and gender-specific differences in decision inertia across situational contexts. Moreover, they confirm the situational effectiveness of nudges.

Part V summarizes and concludes the main contributions of this thesis and puts them into an overall picture. Moreover, it points out possible directions for future research.

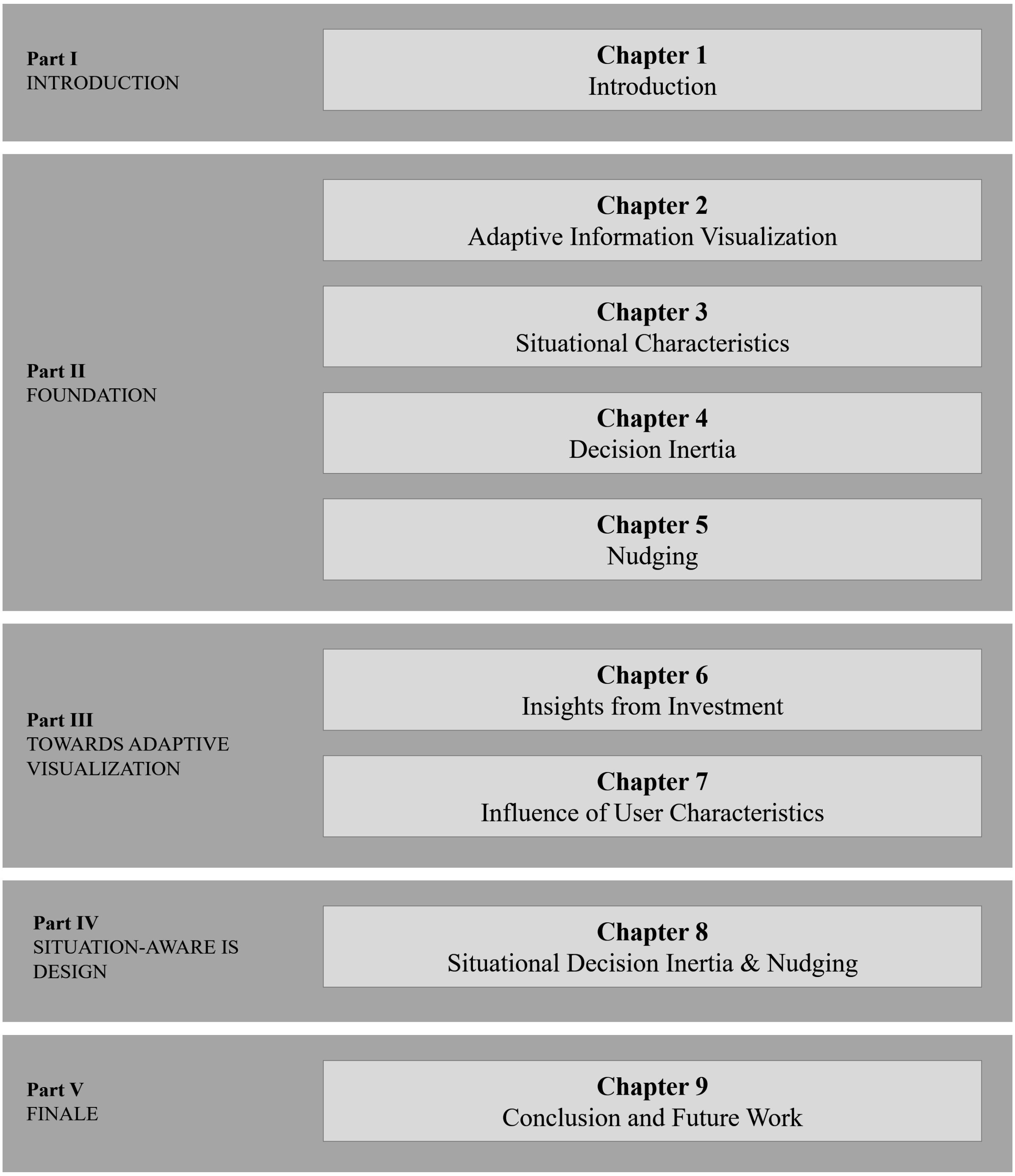



FIGURE 1.1: *Structure of this thesis.*

# Part II

# Foundations

# Chapter 2

# Adaptive Information Visualization

> “The ability to take data – to be able to understand it, process it, extract value from it, to communicate it – that's going to be a hugely important skill in the next decades”
>
> HAL VARIAN, 2009

THIS chapter introduces the foundations of adaptive information visualization. First of all, the relevance and the terms are explained. Moreover, relevant literature is reviewed.

## 2.1 The Role of Adaptive Information Visualization

Before the digital age, data and visualizations were predominantly used by specialists, such as statisticians, scientists, or engineers (Kirk, 2012). Nowadays, with ubiquitous access to powerful technologies, everybody is constantly recorded, creating a vast amount of data at a tremendous rate (Dennett and Roy, 2015). The available data describes different aspects of an individual, such as financial or health status, and thus, it is used to make vital decisions (Lupton, 2016; Saal et al., 2017).

In addition, innovative digital self-service products are emerging, in which users have to make decisions without human advice. It is expected that these will replace traditional advisory services sooner or later (Kretzschmar et al., 2019; Payne et al., 2021). A prominent example is robo-advisory (Bhattacharyya et al., 2019; Fein, 2015). To date, it is pretty obvious that robo-advisory has not replaced traditional bank services. However, this development shows where the trend is heading in the long term. More and more services will be transferred into the digital sphere so that users or customers are no longer advised by traditional human specialists who can communicate decision-relevant information according to the needs of their customers. This is where visualizations come into play. They are an efficient means of communicating relations of data. The following example shows their advantage (see figure 2.1) (Anscombe, 1973). The tabular representation is not well suited to discover the relation of the variables X and Y, whereas the scatter plot immediately conveys their relation.

| X | Y |
|---|---|
| 10 | 9.14 |
| 8 | 8.14 |
| 13 | 8.74 |
| 9 | 8.77 |
| 11 | 9.26 |
| 14 | 8.1 |
| 6 | 6.13 |
| 4 | 3.1 |
| 12 | 9.13 |
| 7 | 7.26 |
| 5 | 4.74 |

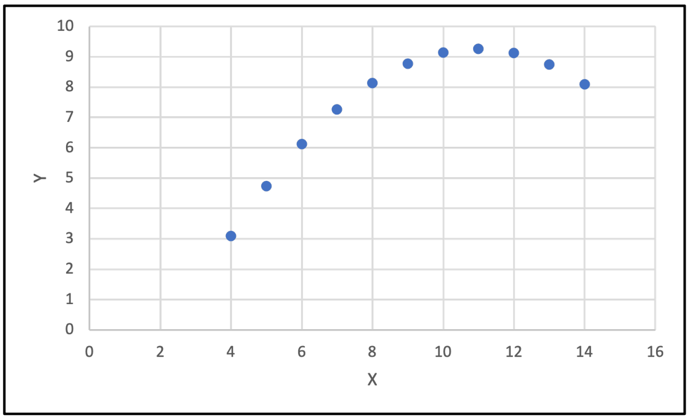



FIGURE 2.1: *Tabular representation and visualized data.*

Visualizations have become an integral part of today's ISs (Qin et al., 2018). Robo advisors use visualizations to explain the relationship between risk and return to their clients (Salo and Haapio, 2017). Information about Covid-19, such as the temporal and spatial progression of the infection rate, is communicated on the web or through smartphone apps predominantly with visualizations (Comba, 2020). Mobile health applications provide visualizations that enable users to monitor and predict their health status (Wang and Wang, 2021).

Recently, the federal government of Germany surveyed a broad sample of the population on data strategy[1]. 617 private individuals and 652 participants from different institutions (public administration, scientific institutions, companies, and organizations) participated in the online survey. The survey included 36 questions to assess the public's views on data literacy, data infrastructures, data ecosystem, data framework, and the role of government. Results show that data usage is a field of activity in which almost all participants are active. Moreover, data is frequently used for analysis and visualization purposes in order to interpret results and derive recommendations for action.

With the increasing diversity of visualization users, the predominantly used one-size-fits-all approach, resulting in one visualization for everyone, is no longer practicable. In simple cases, as already shown in the scatter plot example, the approach might be sufficient. Certain knowledge or user characteristics must be present for more complex visualizations. A ubiquitous example in science are boxplots. If there is no understanding of distributions and location parameters they are useless.

User characteristics, such as cognitive abilities, influence the comprehension of represented data (Steichen et al., 2013; Conati et al., 2015; Lee et al., 2019). Linguistic learners are adept at using words, whereas visual learners benefit through the use of images and graphics (Wright et al., 2007). In most cases, these abilities determine an individual's interests, educational, and professional development. Over time, individuals acquire knowledge or skills reflecting their strengths and weaknesses. These can be used to customize the representation of data according to their needs.

## 2.2 Definition and Taxonomy

Kirk (2012) used the information exchange model (see figure 2.2) to define information visualization. Accordingly, there are three main agents: the messenger, the receiver, and the message. The messenger, who is also the designer of the visualization, has to transmit information in the form of analysis, results, or stories. On the other side, there is the receiver, which is the user in the case of visualizations. The message, the visualization itself, is in the middle of the information exchange channel.

[1] `https://www.bundesregierung.de/resource/blob/974430/1761674/aec4dd81733f4bd4a7109bffc4914b37/2020-06-18-ergebnisse-der-oeffentlichen-konsultation-data.pdf` last accessed: August 23, 2022.

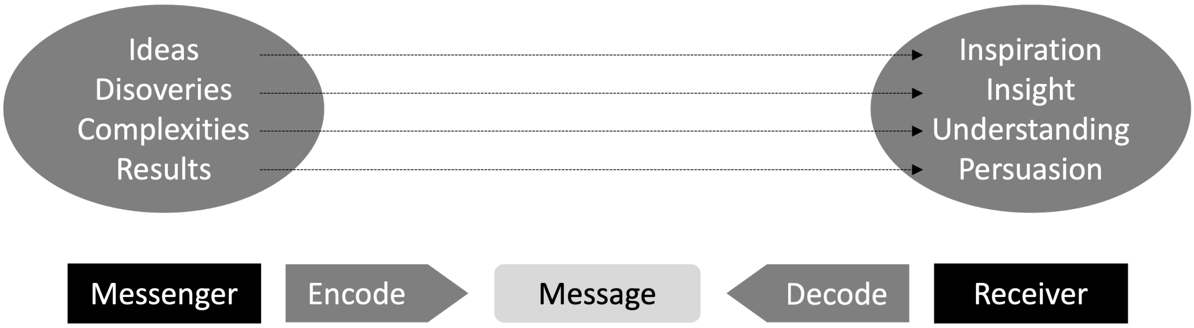



FIGURE 2.2: *Main agents of information exchange (Kirk, 2012)*
.

The main challenge of the designer is to know the receiver. More precisely, to anticipate and determine what the user wants to know and how the information can be conveyed in a visualization. Kirk (2012) denotes considering user needs as the best practice, which underlines the necessity of customized visualization. According to Kirk (2012), a visualization is "the **representation** and **presentation** of data that exploits our **visual perception abilities** in order to **amplify cognition**". The highlighted key words are further described:

- **Representation**: Representation describes how data is encoded into physical forms (visual variables). The most used visual variables are lines, bars, and circles.
- **Presentation**: The presentation defines how the visualized data is displayed and, thus, communicated. The representation is embedded into the presentation. The designer has to choose colors, annotations, and interactive features.
- **Visual perception abilities**: The exploitation of visual perception deals with how visual perception and cognitive processes are related to each other. In accordance with users' cognitive and perceptive abilities, the designer has to visualize the data in a way that enables spatial reasoning and pattern recognition.
- **Amplify cognition**: To amplify cognition describes the most effective and efficient processing of perceived information into thoughts, insights, and knowledge. Ideally, existing user knowledge and processes should be built upon in order to support the user as far as possible.

Visualizations can be classified according to the task. Thereby, the following taxonomy (see table 2.1) summarizes visualizations with similar characteristics, namely chart types (Kirk, 2012). It distinguishes five categories and lists each category's communication intention and the most common chart types.

| Task | Communication Intention | Common Charts |
|---|---|---|
| Comparing categories | To facilitate comparisons between relative and absolute values of categorical variables. | Bar chart, histogram, word cloud |
| Assessing hierarchical relations | Breakdown of categorical values in their relation to a population of values or as constituent elements of hierarchical structures. | Pie chart, tree map, bubble chart |
| Assessing temporal change | Use of temporal data to show the changing trends and patterns of values over a continuous or multiple discrete time periods. | Line chart, area chart, whiskers plot |
| Assessing connections and relationships | Using multivariate data sets to evaluate associations, patterns, and distributions among variables. | Scatter plot, heatmap, network diagram |
| Assessing geospatial relationships | Using datasets with geospatial properties to provide spatial relationships and patterns. | Choropleth map, cartogram, network connection map |

TABLE 2.1: *Visualization taxonomy containing five different categories.*

One research subject of this thesis is the personalization of visualization to the needs of the users. The existing research differentiates two variants of personalization, namely adaptability and adaptiveness (Findlater and McGrenere, 2004).

## 2.3 Adaptability, Adaptiveness, and User Model

Adaptability refers to the possibility of personalizing the UI and the visualizations it contains. Popular examples are Microsoft Excel and business intelligence tools such as Tableau or PowerBI. These applications recommend a diversity of visualizations for specific data depending on its dimensions and scales.

Usually, the personalization is explicit and has to be done in pre-runtime or before carrying out a task. In the worst case, an inappropriate personalization can lead to ineffective information representation. Moreover, the complexity for users is increased as they have to do fundamental work by learning the usefulness, pros and cons of different visualizations (Fischer, 2001). Paradoxically, inexperienced users have to reduce the complexity and benefit more from personalization than experienced users (Mackay, 1991). In their case, the adaptive visualization approach, is more purposeful.

Adaptiveness is characterized by automatic and constant alteration of IS characteristics in terms of visualization design (Miraz et al., 2021; Schneider-Hufschmidt et al., 1993; Oppermann, 1994). Thus, the personalization is implicit and done by the system during runtime. The goal of such a system is to steadily improve its usability (Miraz et al., 2021). A prominent example is the mobile usage of geographical information visualization. The IS adjusts the point of view, brightness, and even colors of the visualization during runtime (Reichenbacher et al., 2008).

The most compelling advantage of adaptiveness is that users do not have to learn adaptation possibilities. The IS asks maybe the user, or it reacts to the environment or interactions of the user automatically. Thus, the complexity for users is reduced. However, it could be difficult for users to build a coherent mental model of the system. This would result in the feeling of loss of control (Fischer, 2001). Thus, established usability principles such as predictability and controllability could be violated by adaptive systems (Story, 1998).

Adaptiveness is achieved in the background. The system incorporates incomplete prior experience with a particular user to form a user model. At the time of use, the user model is updated. Thus, building an adequate user model is the inevitable mechanism of adaptiveness (Langley, 1999; Benyon and Murray, 1993).

According to Maybury and Wahlster (1998), a user model helps to find relevant information by predicting user's behavior and adjusting the presented information and interface features to the user. The user model considers the user's goals, tasks, plans, and knowledge (Wahlster and Kobsa, 1989). Moreover, the context of use can be incorporated into the user model since each situation and, therefore, context is characterized and evaluated by each individual differently (Hussain et al., 2018; Rauthmann et al., 2014).

Rothrock et al. (2002) reviewed the extant literature and gathered the possible content of user models. The authors identified three relevant dimensions:

- Taxonomy of user abilities

- Taxonomy of tasks
- Taxonomy of environments

A simple example shows the necessity of all three dimensions: A visualization for three-dimensional spatial orientation, which the first two dimensions can classify, works well for indoor navigation in buildings. However, the same visualization would not work underwater for the user of a reconnaissance vessel.

Furthermore, an individual's personality can be incorporated into user models since it could be shown that certain traits impact information-seeking (Heinström et al., 2014; Alves et al., 2020). Users perform more effectively when the design matches their type of personality (Kostov and Fukuda, 2001).

## 2.4 Review of Adaptive Visualization Systems

In this section, an excerpt of the relevant literature on adaptable or adaptive visualization systems is reviewed. Identified research is described with the purpose of use, adapted properties of the visualization, and characteristics used for the adaptation.

Brusilovsky and Loboda (2006) proposed an adaptive system for the visualization of expression evaluation in the C programming language. The system is designed for the education of novice programmers. The main assumption is that individuals differ with respect to their level of knowledge about algorithms and or elements of a programming language. The level of visualization details corresponds to the level of knowledge. The lower the level, the more details are visualized. The potential of the systems in increasing the efficiency and effectiveness of learning programming language is theoretically discussed but not empirically shown.

Brusilovsky et al. (2006) designed a knowledge-based adaptive visualization system for retrieving educational resources. The main purpose of personalization is to support students in identifying learning materials that match their goals, interests, and knowledge. Their system combines spatial text-similarity visualizations with adaptive annotations. Thereby, the relevance of documents is determined and annotated depending on the current knowledge of students. The guidance of the prototype system, which was used in the programming course of the university, was appreciated by the students, and it encouraged them to explore significantly more documents than traditional information retrieval systems. The empirical evidence is not present, so the question of users' performance improvement is still pending.

Golemati et al. (2006) proposed a context-based adaptive visualization system for information retrieval in digital libraries. It contains several visualizations with predefined properties. Based on the system properties (mobile or stationary), user characteristics (demography, profession, and cognitive abilities), and document characteristics (metadata and document categories), the system matches possible visualizations and presents them to the user iteratively. The users can accept or deny the recommendation. The system stores the preferences of the users with a dynamic user model so that the representation is improved with respect to user needs in the course of the interactions. To date of publication, the proposed system was prototyped; therefore, evaluation is missing.

Nivala and Sarjakoski (2007) investigated adaptive cartographic maps in the mobile context. They argued that non-personalized maps and symbols on the map would lead to misinterpretation and frustration. They provided a system that adapts the visualization to particular situational contexts and personalize the symbols on the visualization according to user characteristics, such as age, nationality, and preferences. The design of the personalized cartographic maps were evaluated by qualitative expert interviews. Solely the intuitiveness of the personalized symbols was empirically evaluated with a relatively small sample (22 participants). The authors have shown that the personalization to specific user groups lowers the misinterpretation of the symbols and that the visualizations were perceived as more pleasant and intuitive than the non-adapted maps. However, a quantitative study with a larger number of participants, which objectively evaluates the performance in directing users towards their destination, would be desirable.

Shi et al. (2009) prototyped and investigated a visualization system for large-scale online social networks. The main goal of the system is to avoid visual cluttering (density overload of the connections), which is common in social network visualizations. For this purpose, the network visualization is summarized by hierarchical grouping based on the connectivity of the network participants. The users can interactively navigate through the visualization and, thus, individualize it according to their needs. The authors argued that the system is capable of rigidly controlling the visual density and provides a more effective exploration of network graphs. Thereby, with a self-selected use case, they calculated density metrics and showed that these are in the desired range. Therefore, the generalizability is not given and an empirical evaluation of the performance was not conducted.

Toker et al. (2012) examined the effectiveness, efficiency, and satisfaction of users with particular information visualizations. They investigated the influence of the cognitive abilities (perceptual speed, verbal working memory, and visual working memory) and self-reported expertise on the effectiveness of bar charts and radar plots. Both visualizations contain the same amount of information. The participants of their study, mainly students, had to answer a series of questions capturing the content of the visualizations. All cognitive abilities have a negative and significant influence on completion time. Furthermore, participants with high visual working memory preferred radar plots, and participants with low verbal working memory showed higher ease of use with bar graphs. In simple tasks, the bar graph outperformed the radar plot with respect completion time, whereas, in complex tasks, both visualizations delivered a similar performance. Both self-reported expertise measurements (radar plot and bar graph) negatively influenced participants' completion time in solving the experimental task. Remarkably, the authors did not investigate task performance (number of correct answers).

Steichen et al. (2013) built upon the study of Toker et al. (2012). They argued that the one-size-fits-all approach is inappropriate for satisfying different user groups' needs. They also used the bar chart and the radar plot to capture the knowledge of their participants. The main difference is that they recorded the eye movements with an eye-tracker during the experimental task. With the gaze patterns, the authors reliably predicted the task type, the visualizations, and the cognitive abilities of their participants. The long-term goal of their research is the gaze-based real-time adaptation of visualizations to user needs. However, it is unclear if and when the eye-tracking technology will gain acceptance in the consumer technology sector. Therefore, it is currently not a viable option for personalization.

Ahn and Brusilovsky (2013) deployed Adaptive VIBE. It is an adaptive visualization system for information retrieval. The authors claimed that with the ongoing rapid growth of volume and breadth of online information, the traditional search systems have become less efficient for users. They provide a system that spatially visualizes the similarity of retrieved documents based on the browsing history. For example, the browsing history contains the terms *Russia* and *Japan*. The user searches for the term *nuclear*, the system plots the similarity (based on frequency) of the retrieved documents with the previous search terms as cluster centers. Thereby, the user can adjust the visualization manually by adding or deleting cluster centers. The authors showed that Adaptive VIBE improves the precision of retrieved documents and encourages the user to engage in more exploration. Their approach is not actually personalization of visualization since only the current search result is clustered based on the previous search terms.

Bai et al. (2013) argued that generic business intelligence systems are not capable of efficiently visualizing the needs and requirements of companies. Therefore, the authors proposed an adaptive visualization framework, which considers the task, purpose of communication, time dependency, and the stakeholders. With the use cases *electricity network fault monitoring for operational level* and *KPI monitoring for managerial level*, they evaluated the usefulness of their framework. For this purpose, specific visualizations were added which are able to represent the corresponding use case. Moreover, stakeholder-specific visual elements were chosen in order to improve their efficiency in processing the visualized data. However, the chosen use cases are obviously pretty different and therefore, the use of different visualizations is self-evident. The scholars' goal is to evaluate their framework with further use cases.

Yelizarov and Gamayunov (2014) designed a visualization system that adapts the visualization to the cognitive resources to prevent information overload and improves decision speed and accuracy. The system derives metrics from the interaction with mouse and keyboard, such as reaction time, perception speed, and working memory capacity, that are able to make inferences about the cognitive state. If certain individual thresholds are exceeded, which are derived by initial measurements with a baseline stimulus, the granularity of the visualization is gradually adapted until the cognitive load is in an efficient state. In an experimental study, the system improved the decision-making efficiency of the participants by 42 %. It would be interesting to know if the metrics can also be used to classify users to recommend or adapt different chart types.

Inibhunu and Langevin (2016) designed an adaptive visualization system for dynamic computer networks. Traditionally, cyber administrators have to manually navigate through complex network structures to map incidents to mission impact. The authors argued that this procedure is inefficient, causing cognitive overload, and is prone to errors. The purpose of the system is to support the administrators in monitoring, identifying, planning, and conducting network operations in real-time. They propose an adaptive level of detail visualization system for hierarchical networks. It considers the network structure, users' tasks, and users' cognitive load to personalize the represented level of detail. The users can interact in real-time with the visualization by zooming in, zooming out, or viewing from different angles. The authors anticipated that user tasks and cognitive resources are appropriately managed with an adaptive level of detail network visualization. The next step of their project is the empirical evaluation of their system.

Álvarez et al. (2022) examined an adaptable visualization system for the unsupervised classification of astronomical objects in spectrophotometric data. The system's goal is to enable information processing in a reasonable time and to provide physical and statistical properties of the clustered objects. The authors used self-organizing maps (neural network-based clustering technique). The visualization of self-organizing map depends on the task type (analyzing density or distance of objects). Thus, the system does not live up to the claim of visualization personalization since the adaptation achieves no improvements for the user, but different tasks are processed with different visualizations.

The following table represents the summary of the reviewed literature. It differentiates the adaptation type, contains the content of the user models, and briefly explains the adapted content.

In most of the cases, the reviewed literature used the adaptive approach, which automatically personalize the representation to specific characteristics in the user model. Furthermore, task, cognitive factors, and users' experience is predominantly used for personalization. In the case of adapted content, chart types and granularity of visualization are most often used.

Table 2.2 compares the reviewed literature with this thesis. The purpose of the two studies in chapter 7 is to empirically investigate the automatic adaptation of chart types to the needs of the users with respect to task performance and cognitive load. Hence, the relevant factor are adaptiveness (user model), investigation of different chart types, and empirical evaluation of task performance and cognitive load. In none of the publications, all factors are present. The most similar study is that of Toker et al. (2012) since it empirically investigated different chart types. However, the authors did not evaluated the task performance in their study. In addition, they investigated cognitive abilities as user model

| Publication | Type | User Model | Content |
|---|---|---|---|
| Brusilovsky and Loboda (2006) | Adaptive | Knowledge-level | Highlighting of programming expressions and elements |
| Brusilovsky et al. (2006) | Adaptive | Knowledge-level | Annotation of visualizations |
| Golemati et al. (2006) | Adaptive | Context of use, demography, profession, cognitive abilities | Chart types |
| Nivala and Sarjakoski (2007) | Adaptive | Situational context, age, nationality, preference | Appearance and symbols of cartographic maps |
| Shi et al. (2009) | Adaptable | - | Granularity of social network visualization |
| Toker et al. (2012) | Adaptive | Cognitive abilities, self-reported expertise, preference | Chart types |
| Ahn and Brusilovsky (2013) | Adaptive | Key words of browsing history | Similarity based text visualization of retrieved documents |
| Steichen et al. (2013) | Adaptive | Eye gaze data | Chart types |
| Bai et al. (2013) | Adaptable | Task, purpose of communication, time dependency, stakeholders | Chart types and layout of representation |
| Yelizarov and Gamayunov (2014) | Adaptive | Cognitive state | Granularity of visualization |
| Inibhunu and Langevin (2016) | Adaptive | Task, cognitive resources | Level of detail of network structure visualization |
| Álvarez et al. (2022) | Adaptable | Task | Chart types |

characteristic instead of cognitive load as dependent variable. Cognitive abilities can be used to make inferences about cognitive load; therefore, this factor is partially given.

| **Puplication** | **User Model** | **Chart Types** | **Task Performance** | **Cognitive Load** |
|---|---|---|---|---|
| Brusilovsky and Loboda (2006) | + | - | - | - |
| Brusilovsky et al. (2006) | + | - | - | - |
| Golemati et al. (2006) | + | + | - | - |
| Nivala and Sarjakoski (2007) | + | - | - | - |
| Shi et al. (2009) | - | - | - | - |
| Toker et al. (2012) | + | + | - | +/- |
| Steichen et al. (2013) | + | + | - | +/- |
| Ahn and Brusilovsky (2013) | +/- | - | +/- | - |
| Bai et al. (2013) | - | + | + | - |
| Yelizarov and Gamayunov (2014) | + | - | + | + |
| Inibhunu and Langevin (2016) | + | - | - | - |
| Álvarez et al. (2022) | - | + | - | - |
| **Chapter 7 of this thesis** | + | + | + | + |

TABLE 2.2: *The comparison of this thesis with the reviewed literature.*

# Chapter 3

# Situational Characteristics

> "Everything begins in thought and the reality is only our interpretation of the situation."
>
> Marshall Sylver, 2018

THIS chapter introduces situational research. The existing literature on terminology and taxonomy of situations is described. Moreover measurement tools are presented which are used in the study in chapter 8.

## 3.1 The Triad: Situation, Person, and Behavior

It is common sense that behavior depends on the situation: People fulfill their different roles to act adequately. For example, parents sometimes act authoritarian to raise their children. The same behavior might not be appropriate for an employee. Lewin (2013) postulated with the well-known equation $B = f(P, E)$ that a function $f$ consisting of the environment ($E$) and P person ($P$) determines behavior ($B$). Later, the equation was reevaluated as the personality triad (Funder, 2009). To comprehend any member of the triad, one has to consider the other two members. Despite this fact, researchers primarily investigated aspects of persons to understand and predict their behavior (Rauthmann, 2016).

Recent studies have shown that incorporating situational information can lead to more accurate predictions of behavior (Rauthmann et al., 2014; Geukes et al., 2017).

## 3.2 Situation Terminology

This section describes the current understanding of situational research. There is no clear definition of a situation (Kenny et al., 2001; Saucier et al., 2007). According to Rauthmann (2016), existing definitions differ according to different psychological disciplines and theoretical perspectives. In order to create a shared view across several disciplines and theories, a terminology was proposed, which distinguishes three different basic kinds of situational information: Cues, characteristics, and classes – in ascending order of abstraction (see figure 3.1) (Rauthmann et al., 2015).

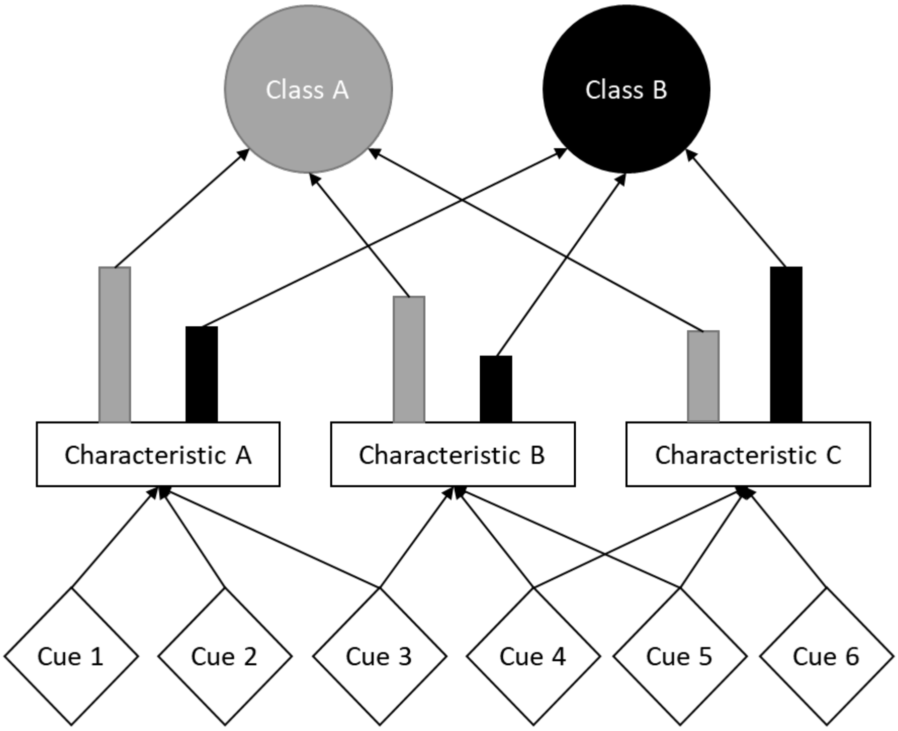



FIGURE 3.1: *Hierarchical relations of situational information (Rauthmann et al., 2015).*

Cues are objectively quantifiable or physically measurable situational information, such as place, time, and person. The purely use of cues for situation assessment might be insufficient because it is extremely difficult to assign a psychological meaning perceived by every person in the same situation (Horstmann et al., 2018).

The processing principle helps to overcome the shortcoming of cue-based situation assessment. In the course of perception and processing, cues feed into psychologically important situational characteristics (Horstmann et al., 2018). The principle is based on the assumption that people form psychologically active representations of stimuli (Rauthmann, 2012), which states that the subjective reality of situations solely matters (Rauthmann et al., 2015).

Based on the reality principle, each situation constitutes a physical, consensual, and idiosyncratic reality (Rauthmann et al., 2015; Horstmann et al., 2018). Situational cues cover physical reality. The consensual reality describes the shared view of several people in the same situation. Lastly, the idiosyncratic reality refers to the individual's unique perspective of a situation. The psychologically relevant situational characteristics can identify shared interpretations of situations across persons and individual differences between them (Rauthmann et al., 2014).

Finally, situational classes are the most abstract classifications of situations. They can be seen as types, domains, or contexts and summarize similar situations based on cues or situational characteristics (Rauthmann et al., 2015). It is assumed that the selection of a particular class is most likely driven by perceived characteristics (Horstmann et al., 2018; Rauthmann and Sherman, 2016). For example, going out for dinner is a situation of leisure for most people. In contrast, if the supervisor joins the dinner, the situation is more likely perceived as a duty. In this case, a person's role changes the perceived characteristic of the situation. Thus, assessing psychologically relevant situational characteristics is more convenient to investigate behavioral differences across individuals (Rauthmann et al., 2015).

## 3.3 Situational Characteristics Taxonomies

In the previous section, we get an understanding of situational information. The present work examines the situational influences on the behavior. Therefore, two questions arise, (1) which situations (classes) are perceived as different, such that the difference could lead to behavioral changes, and (2) how can we measure them. Horstmann et al. (2018) reviewed the extant situation taxonomies and identified two clusters of them. The first cluster was published between 1970 to 1984 and was developed assuming that no effective strategies for measuring situations exist. Consequently, there were no measurement tools. The second cluster that emerged after the millennium is described in the following.

For the present work, the situation research after the year 2014 is relevant because it has provided taxonomies and measurement tools and focuses on the psychologically important characteristics of situations. These taxonomies are DIAMONDS (Rauthmann et al., 2014), Situation 5 (Ziegler et al., 2019), CAPTION (Parrigon et al., 2017), Situational Affordance and Adaptive Problems (SAAP) (Brown et al., 2015), and Social Interdependence Scale (SIS) (Gerpott et al., 2017). There are differences between the taxonomies. These are outlined below.

The first differentiating factor is the intended use of the taxonomies. The SIS and SAAP are solely capturing social interactions, whereas Situation 5 assesses occupational situations. DIAMONDS and CAPTION are comprehensive concerning the intended use (Halevy et al., 2019). They are broader taxonomies for everyday life situations and measure additional characteristics besides professional and social ones. Another decisive criterion concerns the development of taxonomies. For example, DIAMONDS was constructed with samples across countries, whereas CAPTION was constructed with US samples. Both taxonomies rely on several dimensions to capture situational characteristics. They overlap considerably despite different derivation approaches (Halevy et al., 2019; Rauthmann and Sherman, 2018).

A comprehensive taxonomy that unites the different taxonomy models would be most purposeful in the long term (Rauthmann and Sherman, 2018). For this thesis, the DIAMONDS taxonomy is used for the following reasons. The thesis examines different situational contexts while the SIS and SAAP cover only social interaction (SIS) or occupational situations (SAAP). DIAMONDS and CAPTION are comprehensive with respect to content and context. They capturing cognitive, emotional, and social aspects and both are applicable for a wide range of contexts which are of interest in this thesis. It is assumed that DIAMONDS is evaluated more thoroughly since cross-cultural samples were used in the validation studies. In contrast, only a US sample was used for the evaluation of CAPTION. The last and practical reason is that there are validated German measurement tools for

DIAMONDS (Rauthmann and Sherman, 2015b,a).

## 3.4 The Situational Eight DIAMONDS

Rauthmann et al. (2014) pointed out the relevance of situations for behavior research. The authors argued that no purposeful taxonomy and no practicable measurement tool were available. In a multinational data set, they examined the structure of the Riverside Q-Sort (RSQ) – It consists of 89 items and was recognized as the most widely available measure for psychological important situational characteristics – and identified eight major dimensions on which people perceive, describe, and evaluate situations: Duty (does something need to be done?), intellect (is deep thinking required or desired?), adversity (are there external threats?), mating (is the situation sexually and/or romantically charged?), positivity (is the situation enjoyable?), negativity (does the situation elicit unpleasant feelings?), deception (is someone being untruthful or dishonest?), and sociality (are social interactions and relationship formations possible, required, or desired?) (Sherman et al., 2015; Rauthmann et al., 2014). The DIAMONDS model connects situational characteristics to personality and is suitable to predict behavior. It explains additional variance in behavior analysis (Sherman et al., 2015; Rauthmann et al., 2014).

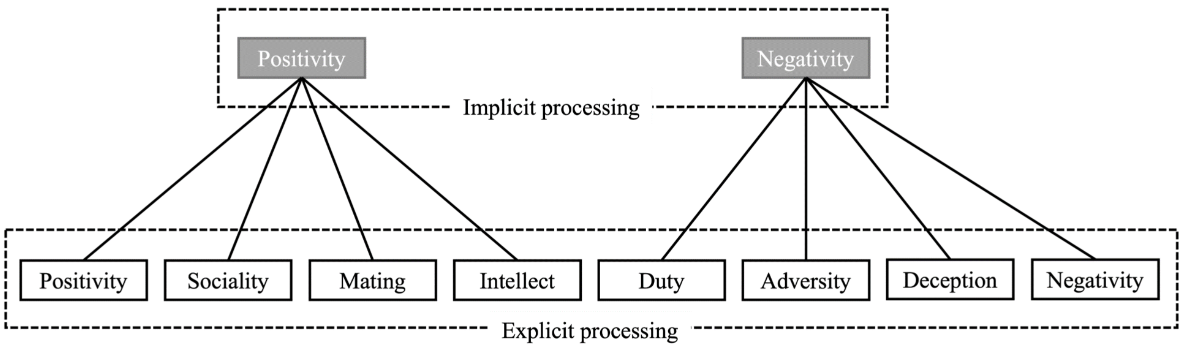



FIGURE 3.2: *DIAMONDS' implicitly and explicitly processed dimensions (Rauthmann, 2016).*

Emotions are crucial in situational experience (Halevy et al., 2019). Based on the level of processing, the eight dimensions form a two-layer hierarchy (see figure 3.2) (Rauthmann, 2016). At the top level, positivity and negativity are processed implicitly, which means no deliberation is needed to evaluate them. The bottom level represents explicit processing and is grouped within two clusters. Positivity and negativity are assumed to be evaluated on both processing levels, which means that all dimensions are evaluated by active deliberation. These give insight into the motivational content of the situation.

The first cluster differentiates the motivational content of implicitly evaluated positivity and is formed by explicitly evaluated positivity, sociality, mating, and intellect. Hence, if these dimensions are active, it is supposed that a person seizes rewards or opportunities concerning his motivational factors. The second cluster differentiates the content of implicitly evaluated negativity and is formed by explicitly evaluated negativity, duty, adversity, deception, and intellect. If these are active, it is supposed that a person's goal is hindered. Thus, they represent obstacles or threads in situations (Rauthmann, 2016; Rauthmann et al., 2014).

There are several possible applications for DIAMONDS. The taxonomy can be used to describe momentary situations or life spans. For example, "being in a club at midnight" represents a snapshot of an individual's life, whereas "being in a university course" is an enduring life space. Both can be characterized the DIAMONDS taxonomy (Rauthmann et al., 2014).

Moreover, the taxonomy can be used to compare different situations with each other. Differences or similarities can be analyzed in order to identify consistent or coherent behavior (Sherman et al., 2010). Additionally, situations can be clustered with clustering algorithms along the DIAMONDS dimensions. In the form of distinct profiles, the characteristics of clusters can be examined to identify corresponding descriptive classes or contexts (Rauthmann et al., 2014).

Finally, the taxonomy provides the possibility to investigate individual differences in situation perception. Thereby, the perception of one situation could be different across different individuals (interindividual differences). There could also be differences in an individual's perception across different situations (intraindividual differences). By measuring the situation at another point in time, DIAMONDS dimensions can be analyzed to answer how, when, and why a situation has changed (Rauthmann et al., 2014).

## 3.5 DIAMONDS Measurment Tools

As stated previously, there are several measurement tools (Rauthmann et al., 2014; Rauthmann and Sherman, 2015a,b, 2018). Initially, the RSQ-8 was created by analyzing the structure of the RSQ (Rauthmann et al., 2014). It contains four items for each of the eight dimensions. (Rauthmann and Sherman, 2015b) created the shorter version S8* (three items for each dimension) by selecting the maximal informative items of the RSQ-8. S8* represents a compromise between content coverage and item homogeneity to be more economical (Rauthmann et al., 2015). Thus, in an experimental setting, the S8* could be used to select and evaluate stimuli. To obtain even more economic assessment tools, the authors have provided four shorter questionnaires (S8-I, S8-II, S8-III-A, and S8-III-P) (Rauthmann and Sherman, 2015a). These have larger limitations with respect to the covered content and can be differentiated by their intended use. It is recommended to use the S8-I and S8-II for substantial research. Especially the S8-I for experience sampling, whereas the S8-II for stimuli validation. The S8-III-A and S8-III-P have not performed as well in terms of discriminant validity. Lastly, the authors developed and validated German translated versions of the S8*, S8-I, and S8-II questionnaires (Rauthmann and Sherman, 2015b,a).

In summary, despite prior existing insights, psychology research has, for a long time, failed to operationalize situations to understand human behavior. Meanwhile, research catching up this lack and is providing a deeper understanding of situational influences on behavior. Additionally, taxonomies and measurement tools were developed to capture situational differences.

# Chapter 4

# Decision Inertia

> “This is the essence of intuitive heuristics: when faced with a difficult question, we often answer an easier one instead, usually without noticing the substitution.”
>
> DANIEL KAHNEMAN, 2017

THIS chapter introduces the cognitive bias decision inertia. The existing literature on situational and applicational insights and drivers of decision inertia are described. It provides the theoretical foundation of the study presented in chapter 8.

## 4.1 Decision Inertia in IS

Inertia can lead to negative consequences or at least to missed opportunities. Prominent examples are mobile phone contracts or power supply contracts. A considerable amount of people do not change their old contract, although a change would result in better conditions or additional bonuses. In research, this behavior is known as decision inertia and has gained momentum in the last decade (Alós-Ferrer et al., 2016; Jung et al., 2019; Jung and Weinhardt, 2018; Power and Alison, 2017; Alison et al., 2015).

Initially, decision inertia was called "resistance to change" (Pitz, 1969). With variants of sequential probability updating tasks, it could be shown that participants tend towards their initial choices (Geller and Pitz, 1970). If a decision is followed after disconfirming information, the participants fail to update their confidence adequately (Geller and Pitz, 1968; Pitz, 1969; Grabitz, 1971; Brody, 1965; Grabitz and Grabitz-Gniech, 1972). Some researchers assumed motivational drivers and, inter alia, investigated the commitment effect (Brody, 1965; Geller and Pitz, 1968; Pitz, 1969; Grabitz, 1971).

The inertia effect is not present if the participants do not have to make an initial choice (the first decision is made without prior information choice) (Brody, 1965; Pitz, 1969). Also, the effect disappears if participants can not recall their previous selection (Pitz, 1969). Further findings have shown that disconfirming information increases decision speed (Geller and Pitz, 1968). It was concluded that contradictory information could lead to cognitive dissonance (Brehm and Cohen, 1962). In order to reduce the dissonance, decision-makers could rely on inertia and refuse rational deliberation (i.e., Bayesian updating) (Pitz, 1969). Further evidence for the commitment effect was provided by Grabitz (1971). Early disconfirming events in a sequence of samples caused much more cognitive dissonance (participants reduced their confidence significantly less) than later ones.

Nowadays, there are two distinct research streams, which use the term decision inertia. Naturalistic decision-making defines it as the "inability to reach a decision" due to redundant cognitive deliberation (Power and Alison, 2017; Alison et al., 2015). Thus, there is no deviation from rational behavior. The corresponding research investigates the cognitive processes in real-world scenarios so that the results are hardly reproducible or comparable.

In contrast, this thesis focuses on the design of ISs. Consequently, an improvement of the usefulness of ISs is targeted by identifying potential suboptimal behavior (i.e., decision inertia) and incorporating this knowledge into IS design. Hence, its understanding of decision inertia is covered by judgment and decision-making, representing the second research stream (Alós-Ferrer et al., 2016; Jung et al., 2019; Jung and Weinhardt, 2018; Jung and Dorner, 2018). In this context, decision inertia is defined as "the tendency to repeat a previous choice, regardless of its outcome, in a subsequent decision" (Alós-Ferrer et al., 2016). Thus, there are at least two subsequent decisions. Decision inertia occurs if the first decision leads to an undesired outcome but is repeated. Otherwise, it is a repetition of a previously successful choice.

## 4.2 Dual-Process Perspective of Decision Inertia

In 2016, decision inertia received a new impulse with the research from Alós-Ferrer et al. (2016), which sheds light from a dual-process view. Before that, the reflective-impulsive model of human decision-making was proposed (Strack and Deutsch, 2004). Accordingly, decision-making is a joint function of an impulsive and reflective information processing system. The impulsive system is associative, intuitive, fast, effortless, and uses decision heuristics (cognitive shortcuts) (Kahneman et al., 1982). While the reflective system is deliberative, conscious, slow, effortful, and represents rational thinking. Cognitive biases, such as decision inertia, occur if the impulsive and reflective system conflicts (Alós-Ferrer et al., 2016). The authors build upon these findings to model decision inertia as an automatic, unconscious, and effortless process, conflicting with rational, slow, and demanding rational processes like Bayesian updating. The authors investigated conflictive decision situations after a preliminary random choice, with a variant of an urn game (Charness and Levin, 2005). Decision inertia was assessed as choice repetition resulting in suboptimal decisions. Results indicate more errors and increased decision speed in conflictive decision situations. However, decision inertia disappears if the initial choice is not made autonomously.

The present work builds upon the definition of decision inertia as a bias resulting from conflictive cognitive processes (Alós-Ferrer et al., 2016). Other studies, such as Zhang et al. (2014) and Sautua (2017), used different definitions and hardly comparable experimental paradigms.

Achtziger and Alós-Ferrer (2014) delivered first insights but did not explicitly mention decision inertia in their study. They investigated response times when rational processes (i.e., Bayesian updating) conflict with intuitive processes. There were four variations of an urn game (Charness and Levin, 2005). One of these variants induced decision inertia in the same way as Alós-Ferrer et al. (2016). Results indicate a considerable amount of errors due to decision inertia (mean error rate: 31.5 %). Additionally, response times are longer when decision inertia conflicts with Bayesian updating.

Furthermore, personality traits as motivational drivers of decision inertia were investigated (Alós-Ferrer et al., 2016; Jung et al., 2019). In line with the commitment hypothesis (Pitz, 1969), it could be shown that an increasing preference for consistency (Cialdini et al., 1995) causes significantly more errors due to decision inertia (Alós-Ferrer et al., 2016). Despite the same experimental paradigm and measurement tool for preference for consistency, this finding could not be replicated by another study. Jung et al. (2019) could not find any association of preference for consistency with decision inertia. The authors concluded that motivational drivers might not be as stable as initially assumed.

Jung et al. (2019) investigated further motivational factors. An individual's action orientation, indecisiveness, and decision avoidance were measured and compared to suboptimal choice repetitions. Solely, a significant association with action orientation could be shown. The more action-oriented an individual is, the more he or she is prone to decision inertia. It was concluded that action-oriented individuals do not sufficiently consider initial suboptimal decisions. In line with previous research, if the first of two subsequent decisions is not made autonomously, errors due to decision inertia are lower (Alós-Ferrer et al., 2016). It is assumed that cognitive decision processes are not initiated if the initial decision is not required (Alós-Ferrer et al., 2016; Jung et al., 2019).

In contrast to prior research, Jung et al. (2019) additionally investigated the cognitive factors conservatism and evidence threshold as possible explanations for decision inertia. Conservatism refers to updating of probabilities in subsequent belief updating tasks. It hypothesizes, if newly incoming information is not sufficiently incorporated, it leads to suboptimal choice repetition. Consequently, the Bayesian updating skills of participants were measured and compared with decision inertia. The evidence threshold is based on the assumption that individuals differ in their level of evidence for accepting a hypothesis (Kozielecki, 1966). Jung et al. (2019) assumed that higher levels of evidence threshold might result in increased decision inertia. Both hypotheses are not confirmed, leaving the cognitive antecedents unclear.

## 4.3 Framing and Applicational Context of Decision Inertia

Decision inertia occurs in different domains or contexts. For instance, in financial decision-making, a considerable amount of people tend to repeat suboptimal investment decisions (Jung and Weinhardt, 2018; Jung et al., 2018). Another example is the resistance to innovative technologies. Some people avoid the adoption of electric cars and maintain the status quo, despite harmful consequences for the environment (Stryja et al., 2017). Since different framings of the decision situation could lead to different decision inertia, and the identified drivers are not stable across different situational contexts, the present research aims to answer whether the tendency to rely on decision inertia is a stable phenomenon across decision situations. Consequently, the following section reviews the existing research concerning the applicational context of decision inertia. Considering the lack of research on decision inertia, similar phenomena such as status-quo bias or indecisiveness are also reviewed. These can be summarized by inertial behavior. The different concepts have large overlaps and thus offer the possibility to investigate and transfer existing insights.

In most cases, decision inertia was examined in economic decision-making. In sequential (Brody, 1965; Geller and Pitz, 1968; Pitz, 1969; Grabitz, 1971) or subsequent (Achtziger and Alós-Ferrer, 2014; Alós-Ferrer et al., 2016; Sautua, 2017; Jung et al., 2019; Jung and Dorner, 2018; Jung et al., 2018) choice tasks, it could be shown that decision inertia manifests as a tendency towards prior commitments.

Alós-Ferrer et al. (2017) investigated the effect of loss versus win framing with a pretty similar decision task. The participants of the win-frame relied significantly less on suboptimal decision heuristics than participants of the loss-frame. The authors linked this to loss aversion – the asymmetry in the perception of gains and losses (Kahneman and Tversky, 2013). Jung et al. (2019) also examined the effect of win and loss framing on decision inertia. They showed a confirming tendency but without any significance.

Zhang et al. (2014) investigated decision inertia in ethical decision-making. They measured in subsequent decisions the cheating behavior of participants. The personality trait prevention-focus leads to significantly more choice repetitions. However, their findings suggest no association of preference for consistency with decision inertia, at least in an ethical context. Their results are hardly comparable because they did not investigate decision inertia as an outcome of conflicting cognitive processes.

Furthermore, decision inertia plays a role in accepting new software systems (Jermias, 2001; Polites and Karahanna, 2012; Li et al., 2016). Jermias (2001) showed in an experimental setting that commitment to a prior accounting system leads to higher valuation compared to non-commitment. Negative feedback leads to lower valuation by committed participants than non-committed. In the case of positive feedback, there is no difference. This insight indicates that the kind of feedback might have an impact on valuation and also on decision inertia.

Polites and Karahanna (2012) and Li et al. (2016) examined resistance towards new software systems via surveys. With a self-developed inertia scale, it was shown that the resistance is driven by increased usability and familiarization (Polites and Karahanna, 2012). Furthermore, it could be linked to loss aversion and contextual factors such as transition costs and social norms (Li et al., 2016).

Lastly, a study in the context of financial decision-making investigated the reduction of decision inertia with an appropriate choice architecture (Jung and Weinhardt, 2018; Jung et al., 2018). The authors incorporated two different digital nudges in a robo-advisor. The decision task within the robo-advisor is based on the dual-choice paradigm (Alós-Ferrer et al., 2016; Jung et al., 2019; Charness and Levin, 2005). Participants showed an inertia effect. This effect is significantly lower in the nudging treatments. Notably, compared to studies without an applicational context (Jung et al., 2019; Jung and Dorner, 2018), the mean error rate is remarkably higher (27.3 % vs. 42.13 %). Furthermore, a higher level of domain expertise (i.e. financial literacy - Lusardi and Mitchell, 2007) was linked to reduced decision inertia.

In summary, the phenomenon of intuitive suboptimal choice repetition is called decision inertia and emerges from conflicting decision processes. It occurs in different contexts. Insights suggest that the antecedents of decision inertia differ across contexts. Consequently, the tendency to rely on decision inertia might vary depending on situational contexts.

# Chapter 5

# Nudging

> “If you want people to comply with some norm or rule, it is a good strategy to inform them that most other people comply.”
>
> RICHARD H. THALER, 2008

THIS chapter introduces nudging. The existing literature on nudging for reducing decision inertia and its situational dependencies are described. The content of this chapter is the theoretical foundation for the study in chapter 8.

## 5.1 Definition of Nudging

Traditional economic theory suggests that individuals behave fundamentally rational. Accordingly, we have complete information, stable preferences, and maximize our utility (homo economicus). Behavioral economics, for example in the field of finance, contradicts this view (Tseng, 2006). In addition, psychological research helps economists to understand this phenomenon: People are bounded-rational – decision-making is not always fully rational due to cognitive limitations or disproportion between the utility of a decision and the costs of information gathering (Simon, 1959). Based on the dual-process theory, people tend to use automatic and intuitive decision processes such as heuristics, especially in uncertain situations (Kahneman and Tversky, 1979).

Thaler and Sunstein (2003) built upon these findings and introduced the concept of libertarian paternalism. It legitimates private and public institutions to help individuals in making better decisions in their interest while preserving their freedom of choice. Decisions can be altered with an appropriate choice architecture (Thaler and Sunstein, 2003). Simple interventions can increase the likelihood of a specific behavior. A nudge is "any aspect of the choice architecture that alters people's behavior predictably without forbidding any option or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid" (Thaler and Sunstein, 2003). Nudges improve decision-making by addressing biases. For instance, it could be shown that a simple change from opt-in to opt-out doubles the number of people as organ donors (Johnson and Goldstein, 2003). In this case, the use of a default option addresses the status-quo bias and changes the behavior towards an altruistic goal.

## 5.2 Digital Nudging

Nowadays, with the ongoing adoption of digital technologies, many decisions are made in the online environment, ranging from life partner search to purchase of goods, investments, or insurances. Thereby, nudging is becoming more and more an essential part of the digital sphere. For instance, Amazon nudges its customers during the buying process with social references by showing products that other customers typically purchased along with the initially selected product. Another prominent example is Booking.com. On the result page, the information for temporary discounts or counts of remaining offers is prominently presented. This nudge uses the psychological principle of loss aversion. Losses have a stronger influence on preferences than gains.

Weinmann et al. (2016) introduced the term digital nudging. They defined it as "the use of user interface design elements to guide people's choices or influence user's inputs in online decision environments". The online environment entails some advantages. Digital nudges are easier, faster, and cheaper to implement. Additionally, there are web services such as clickstream analysis, which provide detailed usage metrics. In usability studies, the usage metrics can be used to compare different nudges and identify the most effective ones (Weinmann et al., 2016). In contrast to Thaler and Sunstein (2003), they explicitly mentioned that the focus of digital nudging is beyond libertarian paternalism. It can be used regardless of morality or virtue of the goal. Nevertheless, they discouraged the use of digital nudges for unethical behavior.

## 5.3 Digital Nudging Process

Weinmann et al. (2016) proposed a five steps process for developing nudges in online environments (see figure 5.1) (Weinmann et al., 2016).

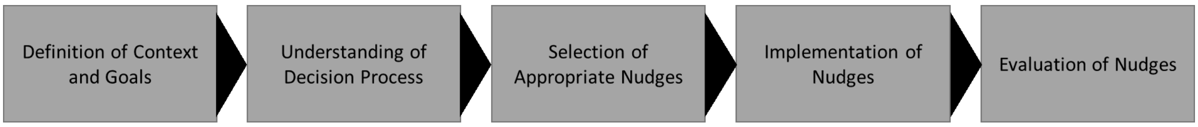



FIGURE 5.1: *Five steps process of digital nudge development (Weinmann et al., 2016).*

The process starts with the (1) definition of context and goals of the intervention. The usage behavior alters with the context of the IS. For instance, it could be shown that customers' motivational and cognitive factors change depending on goal-directed or experiential consumption behavior (Novak et al., 2003). In experiential settings, purchase decisions are more impulsive, which makes them more prone to biases. Additionally, Hummel and Maedche (2019) indicate that the effectiveness of nudges might vary with the context. The second step is concerned with the (2) comprehension of the decision process. Here, potential biases should be identified. The following third step is the (3) selection of nudges. The decision should be based on the corresponding bias, which is planned to use. For example, if the anchoring effect ought to be tackled for product ratings, the presentation of other customers' ratings should be avoided until the user rates the product. Afterward, the nudges are (4) implemented, and the effectiveness is (5) evaluated.

## 5.4 Nudges to Reduce Decision Inertia

Several kinds of nudges have been established in the last decades. Sunstein (2018) identified ten of the most important nudge categories. The three most used in descending order are default settings, warning messages, and social references (Hummel and Maedche, 2019). Nudges can be grouped into two clusters (Johnson et al., 2012; Hummel and Maedche, 2019). The first cluster addresses nudges to structuring the choice task, whereas the second one contains nudges to describe choice options. Elements for structuring the choice task are concerning the question "what to present", while the latter deals with "how to present" (Johnson et al., 2012).

In the case of decision inertia, both clusters seem to be relevant. For instance, decision inertia could be reduced in the context of robo-advisory with warning and default nudges (Jung and Weinhardt, 2018; Jung et al., 2018). A warning message was shown when participants decided on a suboptimal choice in the subsequent decision. In contrast, the optimal choice was already set for the subsequent decision in the default nudge treatment. Default nudge belongs to the first cluster, while warning messages to the second one.

Only a few nudging publications explicitly handle decision inertia (Jung et al., 2018; Jung and Weinhardt, 2018; Jung, 2019). There is much more research on similar constructs, such as status-quo bias or inertia in general. Thirty-four publications were identified and reviewed to look for relevant nudges.

In the literature, four different nudges have been noted. The default nudges are used most often (26). This finding is not surprising because default nudges were recommended for inertia-related problems (Johnson et al., 2012; Thaler and Sunstein, 2009). For instance, Stryja et al. (2017) examined innovation resistance. In a decision support system for rental cars, simple default options generated significantly more bookings of electric vehicles.

The second and third most used nudges are warnings (8) and framings (8). Warnings effectively reduce decision inertia in financial decision support systems (Jung et al., 2018; Jung and Weinhardt, 2018). Framing is concerned with the presentation and orientation of information (Thaler and Sunstein, 2009). It is assumed that these alter the perceived meaning of choice alternatives such as attractiveness. For example, the proximity of snacks to beverage stations was examined in a field study at Google's canteen (Baskin et al., 2016). The likelihood of snacking was nearly doubled, if the snack station was closer to the beverage station. Another example in the digital context was provided by Forwood et al. (2015). In an experimental online supermarket, they examined the effects of within-category food swap offerings to alter purchase decisions towards healthier substitutes. Thereby, the energy density of the shopping basket could be reduced significantly.

The last and most rarely mentioned nudge is social-norm. Social cues are used to emphasize what most people do (Sunstein, 2018). Czajkowski et al. (2019) experimentally investigated the effects of various levels of social nudges on household recycling behavior. In general, the sorting behavior could be altered positively. However, the effects were not stable. Higher levels of social-norm information backfired. Also, households who already sort a lot showed adverse reactions. This is also a segue to the next section, as particular nudges are not equally effective, and situational differences may motivate these differences.

## 5.5 Situational Effectiveness of Nudging

Nudges are not equivalent effective. There are considerable differences between them (Hummel and Maedche, 2019). Hummel and Maedche (2019) reviewed the extant nudging literature and calculated the nudge categories' average and median effect sizes. Warnings are the most effective (average relative effect size: 107 %), whereas pre-commitments are the worst effective (average relative effect size: 7 %). Notably, the authors did not find any differences between conventional and digital nudges.

In addition, sometimes, nudges do not work (Huber et al., 2019) and even backfire (Czajkowski et al., 2019). Sunstein (2018) reported individual differences and counter-nudges as potential reasons for failure. For instance, in the case of marital names, default rules vary in their effect on gender (Emens, 2007). In contrast to men, women change their surnames quite more often. A possible explanation is that strong social norms influence women. These could counteract the effects of nudges (Huh et al., 2014).

Furthermore, personality traits could affect the effectiveness of nudges. Stutzer et al. (2011) investigated the blood donation behavior using nudges. In their experiment, conscientiousness (Big Five - John et al. 1991) explained behavioral differences in the default nudge treatment. Inline, Hummel and Maedche (2019) proposed that individual differences, such as gender or personality, should be taken into account when designing choice architectures.

Beyond individual differences, the decision context could influence the outcome of a decision and thus the usefulness of nudges. For instance, Lehner et al. (2016) reviewed the effectiveness of Swedish policy interventions in the domains of energy, food, and mobility. The actual outcomes of the interventions vary across the contexts. Hummel and Maedche (2019) examined the effectiveness of nudges across several contexts. Besides the initial domains of health and finance (Thaler and Sunstein, 2009), they identified energy, environment, policy-making, and privacy as relevant. Results show that nudges are most effective in privacy-related decisions (average relative effect size: 259 %). In contrast, they are least effective in policy-making (average relative effect size: 8 %). Furthermore, the authors noticed a relationship between category and context. For instance, the majority of nudges in the privacy domain are warnings, whereas the most applied in finance are reminders and defaults.

There are quite large differences within the same context and same nudge. In the health context, the effectiveness of default nudges is very diverse. For instance, parents' nutritional choices for their children have been successfully altered towards healthier foods (Loeb et al., 2017). Parents were about four times (relative effect size 444 %) more likely to choose a healthier breakfast with a default nudge. In comparison, in an experimental study for influenza vaccination, changing defaults resulted in 70 % more appointments for vaccination. Possibly, this could indicate that the distinction at the domain level is not sufficient. Hence, a more detailed classification might be needed that takes into account the differentiation between one's health and the health of children.

In summary, nudges are appropriate interventions to regulate biased decision-making. In online environments, simple UI elements guide people without restricting their freedom of choice. Nudges can be used to reduce decision inertia or inertia-related phenomena. There are indications that their effectiveness might be different across contexts. However, the empirical evidence is pending.

# Part III

# Towards Adaptive Visualization

# Chapter 6

# Insights from Investment

> “Above all others, the eye was the organ with which I grasped the world.”
>
> JOHANN WOLFGANG GOETHE, 1811

THIS chapter reports a qualitative and exploratory eye-tracking study, which is conducted to get a deeper understanding of how users perceive visualizations and which factors influence their information processing. Thereby, a small but in-depth study design is used to examine the usability of a recorded presentation that contains financial visualizations. The visualizations are investigated with respect to perception, comprehension, and comprehension-related factors. The insights of this study are used to explore user-adaptive visualizations in chapter 7.

## 6.1 Introduction

Today’s UIs contain various information representation formats such as graphs, tables, and texts (Kim et al., 2014; Majooni et al., 2015, 2018). These have to be comprehended adequately in order to foster well-grounded decision-making. There is evidence that user characteristics are relevant for information processing with different visualizations (Ziemkiewicz et al., 2011; Toker et al., 2012), which implies that a particular visualization is not ideal to comprehend for all users. Hence, research is needed for the personalization of visualization based on user characteristics.

This study's purpose is to gain a deeper understanding of visualization comprehension and its influencing factors for further research. It is an exploratory study in which a professionally created presentation with different financial visualizations is investigated. The presentation was made for internal use of an investment company. It provides information about herding behavior and possibilities to identify and overcome it. Task completion, subjective experience, perception, and cognitive processes, which occurred during the presentation, are investigated by eye movement analysis and questionnaires.

## 6.2 Methodology

This section introduces the methodology of the study. It contains the experimental design, experimental task, measures, experimental setup, and participants.

### Experimental Design

The study's primary objective is to generate insights for hypothesis development with respect to the personalization of visualization. The secondary objective is to evaluate the usability of the presentation slides. The analysis of eye movements is combined with questionnaires. Participants' perceptual and cognitive processes during the presentation are assessed with the eye-tracking methodology. Subsequent questionnaires capture the comprehension and subjective experience of the participants. The insights of this study answer the following research question:

> **Research Question 1:** *How do users comprehend visualizations, and what factors influence their comprehension?*

Exploratory studies use a modest number of samples to determine the direction of research in the early stages (Stebbins, 2001). It is also pretty common to use small sample sizes in usability studies. Nielsen and Landauer (1993) reviewed a large number of usability studies and formulated and evaluated that with seven users, 93 % of usability issues can be found. A further increase in the sample size does not significantly increase the revealed usability issues' ratio.

## Experimental Task

The content for the study is a presentation containing information about certain market situations and corresponding investment strategies. It contains bar graphs, line charts, and tabulars as different formats of information visualization. Experts in presentation design and moderation created the slides and held the presentation. In order to ensure consistent quality for all participants, the presentation was recorded. Due to a non-disclosure agreement, it is not allowed to publish the original slides, the audio track of the presentation and the content of the knowledge questionnaire. A modified version of the slides, excluding sensitive information, can be found in the appendix A.

In order to evaluate the usefulness of the eye-tracking methodology, three usability issues were knowingly placed on the slides. On slide 9, slide 10, and slide 12, a title is used that does not match the remaining content of the slide. If the eye-tracking methodology identifies them, it is assumed that the methodology is appropriately chosen for this study.

The presentation lasts about 20 minutes and contains thirteen slides. The first slide is a cover and therefore, not investigated. The remaining presentation is divided into four thematically distinct sections (see table 6.1). The participants were instructed to follow the presentation as usual. After each section, a break was made and a questionnaire was conducted.

| **Part 1** | **Part 2** | **Part 3** | **Part 4** |
|---|---|---|---|
| Slide 2 | Slide 5 | Slide 8 | Slide 11 |
| Slide 3 | Slide 6 | Slide 9 | Slide 12 |
| Slide 4 | Slide 7 | Slide 10 | Slide 13 |

TABLE 6.1: *The slides of each presentation part.*

## Eye-Tracking Methodology

Unlike traditional questionnaires, eye-tracking delivers quantitative metrics to assess perception and information processing. Nowadays, highly technical devices based on infrared technology are used to measure eye movements accurately. These are either head-mounted or stationary devices. While head-mounted systems are designed as helmets or glasses, external devices, such as infrared bars, are positioned in front of the subject. The advantage of the latter is that no equipment is attached to the body or head of the subject, creating a comfortable condition during the measurement and no disturbance of natural behavior (Yarbus, 2013).

The following two findings are cornerstones of eye-tracking research and allow conclusions to be drawn about cognitive processes occurring during the perception of visual stimuli: First, the eye-mind hypothesis assumes that fixation of information and the cognitive processing of the fixated information are closely linked (Just and Carpenter, 1980). Second, the immediacy-of-interpretation hypothesis assumes that processing in the brain happens immediately after the fixation of the information (Just and Carpenter, 1980).

According to the functioning of the human eye, there are two basic eye movements, namely fixations and saccades. A fixation is characterized by focusing a visual object for a certain duration to perceive information – typically 250 to 300 ms. A saccade is a rapid eye movement that serves to focus on an object. Its duration is 10 to 150 ms. It is assumed that no signals are transmitted to the brain during a saccade (Yarbus, 2013).

A stationary eye-tracking device, such as the one used in this study, examines the eye movements on a monitor. Thereby, it provides the location of the eyes as X and Y coordinates in a time series. The eye-tracker software identifies the fixations and saccades from the time series and links them to corresponding visual elements in the stimulus.

### Eye-Tracking Metrics

Based on fixations and saccades, more complex metrics can be computed to evaluate perceptional and cognitive processes. The following metrics are used in this study (Sharafi et al., 2015)

- **Fixated:** The fixation indicates whether an object is perceived. If an object is not fixated, it should be verified whether the object is necessary or not.
- **Duration of fixation (duration):** The duration of fixation refers to the total time that an object is fixated. It is assumed that objects are fixated several times over the course of time. The metric is used to evaluate the impairment of cognitive processes. If the fixation duration of an object is disproportionally higher than that of others, this is an indication of impaired cognitive processes leading to poorer comprehension. A good example of this is the repeated reading of a word or text passage that is difficult to comprehend.
- **Frequency of fixation (frequency):** The frequency of fixation indicates how often an object has been fixated over time. It is assumed that the number of fixations allows statements about the complexity and memorability of objects. The more often an object is fixated, the more complex its information content is and, accordingly, the more difficult it is to recall it. Therefore, disproportionally higher values indicate intensive cognitive processes.
- **Sequence of perception (sequence):** The sequence of perception is a complex metric that refers to all fixations of a stimulus. Here, a stimulus contains multiple visual objects. The metric indicates the order in which the different objects are perceived. In this study, the sequence of perception is compared with the ideal sequence of perception to identify divergences that indicate comprehension issues or inappropriateness of visual stimuli.

### Procedure of Eye-Tracking Analysis

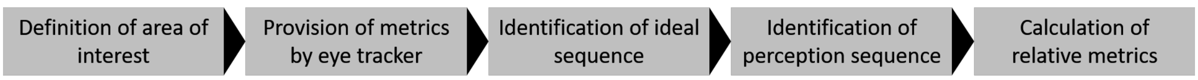



FIGURE 6.1: *Procedure of the eye-tracking analysis.*

In order to provide an understanding of how the presentation slides are analyzed, the analysis procedure (see figure 6.1) is described in the following with an illustrative slide (see figure 6.2).

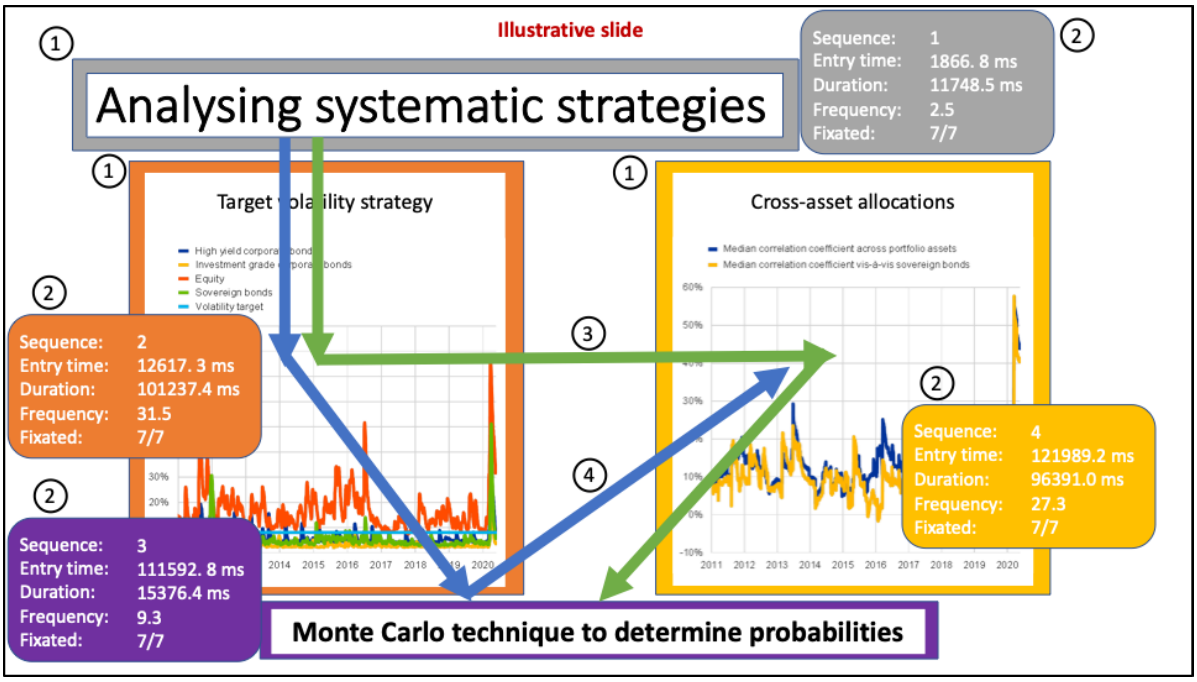



FIGURE 6.2: *The processed illustrative slide with all eye-tracking metrics*

1. **Definition of area of interest (AOI):** AOIs are important parts of a stimulus. They are typically defined by the semantics of the information. For each visual element on the slide, a corresponding AOI is defined so that the eye-tracking metrics can be aggregated in this visual area (see figure 6.2 – colored frames around the visual elements).

2. **Provision of metrics by eye-tracker:** After defining AOIs, the eye-tracker (see section 6.2) provides the aggregated eye-tracking metrics (see figure 6.2 – boxes with metrics in the same color as the AOIs).

3. **Identification of ideal sequence:** The ideal perception sequence of the visual elements is identified with the audio track and confirmed by the creator of the presentation. The number of the names of the visual elements is in line with the ideal perception sequence, which means element 1 is the first visual element mentioned on the audio track. If a visual element is not mentioned on the audio track, the corresponding name consists of a letter, for example, element X. The ideal sequence is reflected by the green arrows (see figure 6.2).

4. **Identification of perception sequence:** In step 2, the eye-tracker provides the

sequence metric, which is calculated by the averaged entry time off all participants. For example, if a visual element has the lowest entry time, it was perceived at first. The sequence of all visual elements is identified (see figure 6.2 – blue arrows) in order to compare it with the ideal perception sequence.

5. **Calculation of relative metrics:** Finally, for each visual element, the relative duration and frequency (see table 6.3) is calculated for further analysis.

The following table contains the processed eye-tracking metrics, which are used to examine the usability of visual elements.

| Area of Interest | Fixated | Relative Duration | Relative Frequency | Sequence |
|---|---|---|---|---|
| Element 1 | 7 | 0.05 | 0.04 | 1 |
| Element 2 | 7 | 0.45 | 0.45 | 2 |
| Element 3 | 7 | 0.43 | 0.39 | 4 |
| Element 4 | 7 | 0.07 | 0.13 | 3 |

TABLE 6.2: *The eye-tracking metrics of the illustrative slide.*

## Measures

Besides the eye-tracking metrics (see section 6.2), the following questionnaires are used. For each part of the presentation, the experts, who designed the presentation, provided knowledge questions to query the containing information. The questionnaire contains one item for each visualization capturing the comprehension of its information.

After each part, the cognitive load is measured with the NASA-TLX questionnaire (Hart, 2006). It is an established questionnaire to gather subjective experience concerning task load and completion. It contains six sub-scales. The subjective task load is measured by the sub-scales *mental, physical*, and *temporal demand* — the remaining sub-scales *performance, effort*, and *frustration-level* capture the subjective task completion.

Finally, age, gender, occupation, department in the company, and eye-tracking confounding variables, such as vision, are queried.

### Experimental Setup

The eye-tracking setup includes a SMI RED250 eye-tracker. It is a screen-based solution that measures eye movements at 250 Hz. The eye-tracker unit is mounted on a 24 inch monitor with a resolution of 1980 x 1024 pixels. The participants are placed at a distance of around 60 – 70 cm and are calibrated with a 7-point SMI calibration. In advance, the participants are examined for their appropriateness with a vision and color blindness test (Dain, 2004).

### Participants

As mentioned in section 6.2, a small but in depth study design is used. Thus, ten experts of various fields were invited to the study. In total, seven subjects followed the invitations, three portfolio managers, two accountants, and two marketing experts. The eye movements of one participant were not properly trackable, resulting in the exclusion of the corresponding eye-tracking data. All participants had sufficient vision and non colour blindness.

The experiment took place on-site in a sufficiently darkened room of the investment company. Thus, the participants joined the experiment during their work. Their identity was disclosed, and none of the results allow conclusions to be drawn about them. The participants took part individually, and the experiment took approximately 60 minutes.

## 6.3 Results

This section reports the analyses. First, the results of the questionnaires are presented, followed by the eye movement analysis. Only the eye-tracking metrics relevant to the analysis are mentioned in the following. The table with the detailed eye-tracking metrics can be found in the appendix A.1.

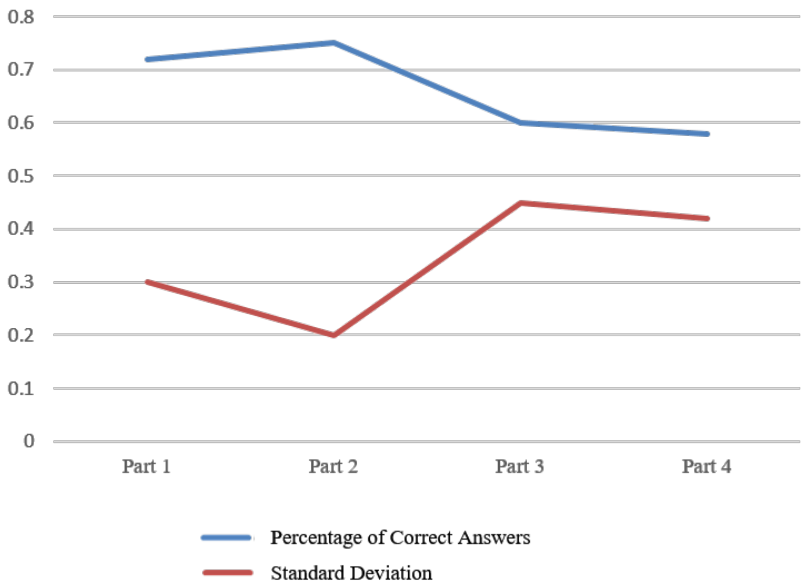



FIGURE 6.3: *The results of the comprehension questions of each part of the presentation.*

## Questionnaires

As depicted in figure 6.3, the comprehension questions were answered most correctly in part 2 (mean = 0.75, sd = 0.2), followed by part 1 (mean = 0.72, sd = 0.3), part 3 (mean = 0.6, sd = 0.45), and part 4 (mean = 0.58, sd = 0.42). The relatively high standard deviation of part 3 and part 4 indicates that some participants have correctly processed the containing information, whereas others had comprehension issues. The following table shows which participants answered the visualization comprehension questions correctly.

| **Participant** | **Tabular** | **Line Chart** | **Bar Graph** |
|---|---|---|---|
| Portfolio Manager 1 | - | + | + |
| Portfolio Manager 2 | + | + | + |
| Portfolio Manager 3 | + | - | + |
| Accountant 1 | + | + | - |
| Accountant 2 | + | - | - |
| Marketing Expert 1 | + | + | - |
| Marketing Expert 2 | - | - | - |

TABLE 6.3: *Correctness of the participants' answers regarding the questions about the comprehension of the visualizations.*

The tabular representations seems to be well known by all participants since most participants (5 of 7) answered the corresponding comprehension question correctly. Four participants with varying profession answered the knowledge question of the line chart correctly. The content of the bar graph was the most poorly comprehended, as only three participants answered the related question correctly – only the portfolio managers. It is assumed that none of the investigated visualizations can be preferred across the board. Moreover, there could be a relation between profession and visualization type. Precisely, during their education or training, the participants may have developed their ability to comprehend a specific visualization.

In general, the results of the comprehension questions are reflected by the results of the NASA-TLX. The overall task load is highest in part 4 (mean = 0.4), followed by part 3 (mean = 0.39), part 2 (mean = 0.32), and part 1 (mean = 0.3). The mental demand is relatively high compared to the remaining sub-scales. In contrast, the physical demand is relatively low. 1 – Performance indicates the inverted subjective performance of the subjects. The lower the better the self-rated performance. The increase of mental demand and effort could be linked to fatigue symptoms. However, after each part, the participants were instructed to take a break in order to prevent fatigue.

Interestingly, the frustration level increases by jumps after part 2. The bar chart and tabular could cause this since they are placed in part 3. As mentioned above, the content of the bar chart was the worst understood. In the case of tabular, research has shown that it causes higher cognitive load than visualized counterparts (Rudolph et al., 2009; Anderson et al., 2011). The cognitive effort of the visualizations is further evaluated by eye-tracking analysis.

It is also noticeable that the temporal demand of part 3 was perceived to be very high compared to the other parts. Especially considering the fact that the presentation duration of all parts is about the same. This could also be due to the bar chart and the tabular, as the rest of part 3 does not contain sophisticated information.

## Eye-Tracking

Next, the eye-tracking analysis of the visual elements of each slide is reported. The detailed table of eye-tracking metrics can be found in the appendix (see table A.1). In total, eight out of perception issues could be identified on the slide 2, slide 4, slide 5, slide 6, slide 7,

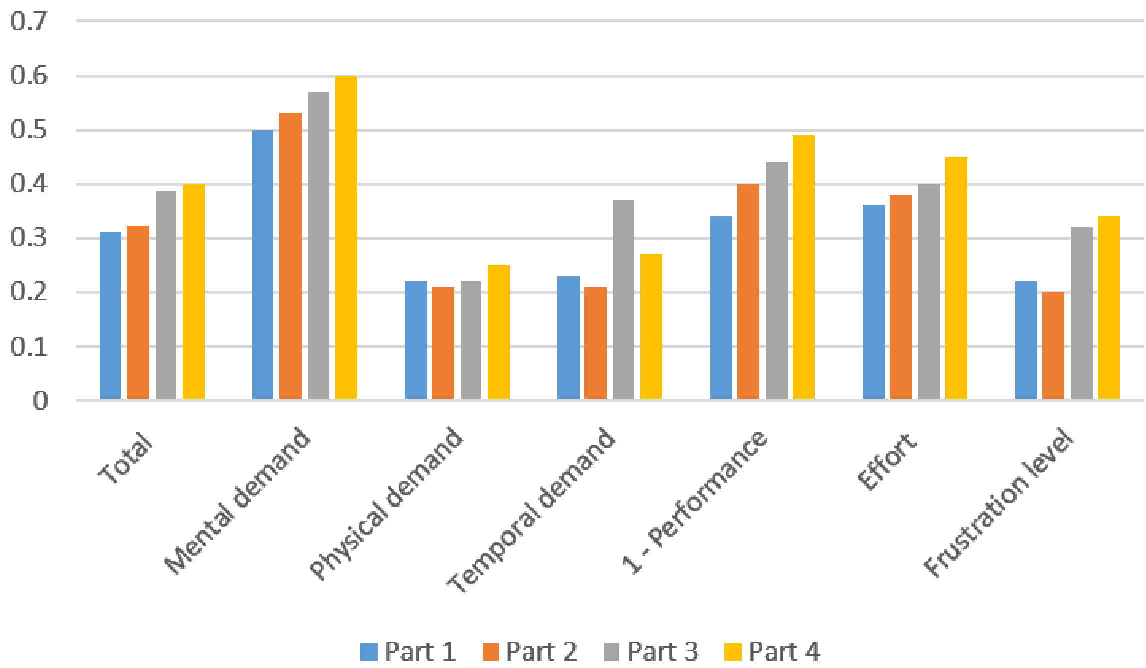



FIGURE 6.4: *The results of NASA-TLX of each part of the presentation.*

slide 9, slide 10, and slide 13. The corresponding visual elements were not perceived by all participants. In order to ensure the perception of the visual elements, it is recommended to assess their necessity. If the corresponding visual element is necessary, it should be highlighted, or the visual elements on the slide should be animated piece by piece with sufficient reading time to make sure that they are perceived and processed appropriately.

Furthermore, on slide 1, slide 3, slide 6, and slide 13, the sequence of perception deviates more than two deviations from the ideal sequence. Memorability is increased if different modalities present the same information at the same time (Penney, 1989), which means that the perception sequence should match the audio sequence (ideal sequence). Thus, in order to ensure that modalities deliver the same content, the visual elements should be presented or highlighted sequentially and parallel to the audio sequence of the presentation.

Next, the relative fixation duration and the relative fixation frequency are combined to analyze comprehension issues. If the relative fixation duration deviates from the relative fixation frequency, it may signify comprehension issues. Therefore, the difference (delta) is calculated, and visual elements with a difference larger than ten percent are further investigated.

In three cases (slide 9: delta = 19 %, slide 10: delta = 20 %, slide 12: delta = 11 %), the participants refixated the title of the slide (element 1) too often. As mentioned in section 6.2, the title of the slide is manipulated knowingly in order to evaluate the usefulness of the eye-tracking methodology. Since the three issues are identified successfully, it is assumed that the eye-tracking methodology is suitable for the evaluation of presentation slides in terms of perception and comprehension.

In four cases (slide 2: delta = 14 %, slide 9: delta = 18 %, slide 6: delta = 14 %, slide 7: delta = 19 %) the participants refixated images relatively often. Images are exceptions, as they have aesthetic properties, which can attract the attention of users (Massaro et al., 2012). Thus, it is refrained from making any conclusions for them. However, the use of images should be chosen carefully in order to prevent unnecessary distractions.

The remaining, including the largest deltas, occurred in the context of different visualizations. The largest deviation showed the tabular (slide 8 element 4: delta = 29 %). This insight underlines the assumption that the tabular causes a higher cognitive load than its visual counterparts. The second highest deviation is caused by the bar graph (slide 8 element 3: delta = 27 %). This insight is plausible since the corresponding comprehension question was not answered correctly in most cases (3 of 7 participants). The lowest deviation is produced by the line chart indicating the comprehension of the containing information is less cognitively demanding than the other visualizations.

In summary, it is assumed that some participants had problems comprehending the containing information. In addition, the line chart was the easiest to comprehend, while the tabular caused the highest cognitive load.

## 6.4 Conclusion

The secondary goal of the study is the identification of potential usability issues of presentation slides. Remarkably, the slides were created by an expert in presentation design. Despite the high quality of the slides, 19 potential usability issues could be identified. This study does not aim to decide whether or not the issues should be solved by adjusting the presentation. Rather, it serves as a list of elements with improvement potential with respect to visual perception and processing.

Multimedia design, which includes presentation design, is a well-researched field (Issa et al., 2011; Moreno and Mayer, 2000; Mayer, 2017). The contributions concerning usability of the slides are already researched as design principles. The existing design principles are identified for each type of usability issue in the following.

Roughly half of the issues are due to not perceived visual elements. In all cases, the information is redundant as it is mentioned on the audio track of the presentation. According to the dual-channel theory of multimedia learning (Issa et al., 2011), there are two separate cognitive systems: verbal and non-verbal. Learning is impaired if the two systems contain contradictory information (split-attention principle, Moreno and Mayer, 2000). Here, the question arises whether the not perceived information is essential. If yes, the creator has to highlight the corresponding visual stimulus (signaling principle - Issa et al., 2011). If not, the visual stimulus can be seen as a distraction and has to be excluded (coherence principle - Issa et al., 2011).

In four slides, the perception sequence of visual elements does not match the audio sequence. In this case, learning and memorability are impaired as the verbal and non-verbal cognitive systems contain different information. The temporal contiguity principle proposes to place corresponding visual information and audio information at the same time (Mayer, 2005). Additionally, the segmentation principle comes to bear, which recommends presenting in learner-paced segments (Issa et al., 2011).

Further potential usability issues were identified by the combined analysis of the relative fixation duration and relative fixation frequency. The fixation duration gives insight into the complexity or length of a visual stimulus. In contrast, fixation frequency gives insight into comprehension processes (Poole and Ball, 2006). An increasing frequency indicates a lack of meaningfulness. Consequently, the combination of the metrics indicates whether the frequency is in line with the length or complexity of a stimulus.

This study identified three cases in which the frequency deviates from fixation duration. The first one, which served as evaluation for the eye-tracking methodology, occurred in the context of non matching slide titles. For example, the title of slide 9 is "Analysing discretionary investors' goals to start with expectations." In contrast, the slide does not contain any information on discretionary investors or goals. The slide contains only three images — one image with bull and bear, representing the corresponding market situations — the remaining two images showing the manager and the fans of a football team. The spatial contiguity principle recommends placing a title that does not contradict the corresponding images (Mayer, 2005). The titles are knowingly inappropriate. Since all of them are identified successfully, it is assumed that the eye-tracking methodology is suitable for these analysis.

Further deviations could be identified in the case of images. Images can distract from essential information as they additionally contain aesthetic properties (Massaro et al., 2012). Therefore, it is recommended to use images carefully, resulting in weighing whether the image is essential to comprehend the content of the slide.

The largest deviations occurred in the context of visualizations. This result is underlined by the comprehension questions and the cognitive load in the corresponding presentation part. Therefore, it is recommended to use visualizations that can be processed by the users adequately, which is one primary goal of this thesis and will be discussed in the following.

The deviation of relative fixation duration and relative fixation frequency is highest in the case of tabular, which could be attributed to the higher cognitive demand (as a reminder, the higher the deviation the higher the probability of comprehension issues). Anderson et al. (2011) have also shown that tabular representations are most cognitively demanding. The line chart produced the lowest deviation. It is concluded that this representation demands lower cognitive resources than the other visualizations.

The questions concerning the comprehension of visualizations were answered most correctly in the case of the line chart and the tabular. Most comprehension issues occurred with the bar chart. It is assumed that the line charts and tabular are more established in financial domain than bar charts. Nevertheless, none of the investigated visualizations can be preferred across the board since the participants' comprehension of the line chart is between tabular and bar chart. However, the cognitive load of the line chart is lower than the other visualizations.

There could be a relation between profession and comprehension of particular visualization. In this study, only the portfolio managers answered the comprehension question of the bar chart correctly. In addition, the review of adaptive visualization systems (see section 2.4) revealed that users' experience and knowledge could be used to recommend different visualizations. Consequently, there must be latent variables in the form of user characteristics reflecting users' experience or knowledge of particular visualizations.

Only seven persons participated the study, which is an established number in usability studies (Nielsen and Landauer, 1993). The insights are not generalizable since the participants were employees of an investment company but are promising with regard to the personalization of visualizations.

After all, this study aims to explore possible comprehension issues of different visualizations to generate insights for subsequent studies. None of the used visualizations could be adequately processed by all participants. It could be identified that line charts and tabular are most common in financial decision-making. Moreover, user characteristics related to users' profession might explain the users' performance with particular visualizations. Further research is needed to evaluate the generated assumptions.

# Chapter 7

# Influence of User Characteristics

> "Human beings are the center of the universe from only one perspective, and that is our own."
>
> JAMES BERNARD MACKINNON, 2013

THIS chapter reports two studies that investigate the influence of user characteristics on different visualizations. A financial decision situation is used to examine participants' cognitive load and decision-making. Furthermore, the feasibility of representation recommendations based on user characteristics is investigated.

## 7.1 Introduction

Visualizations have become a crucial part of today's ISs. With the ongoing digital transformation, more and more data is available. This has given rise to a completely new business sector, ranging from new business models for breaking down data silos to business intelligence software for analyzing and visualizing data (Iansiti and Lakhani, 2014). Visualizations have also found their way to the consumers. For example, robo-advisors are using financial visualizations to communicate risk and return of investment alternatives

(Salo and Haapio, 2017) or fitness applications represent users' progress with visualizations (Langner et al., 2015).

When developing visual representation of data, a one-size-fits-all approach is used predominantly, resulting in one visualization that each user should comprehend (Toker et al., 2012; Hussain et al., 2018). Although visualizations enable more efficient information transfer, certain user characteristics are presupposed to process the information adequately (Toker and Conati, 2014). Differences in user characteristics, such as cognitive abilities, prior experience, and demographics, influence comprehension of visualizations (Steichen et al., 2013; Toker and Conati, 2014; Lee et al., 2019).

There are also exceptions in the form of visualization tools, such as Tableau or Power BI, which allow users to customize the visualizations according to their needs. These applications are mostly used by experts in the context of business intelligence in order to analyze data and represent actionable information that helps executives and managers to make informed business decisions (Gowthami and Kumar, 2017). The use of visualization tools requires a lot of training (Milligan, 2019). Accordingly, inexperienced users already have a high complexity and have to gather additional knowledge to learn adaptation possibilities (Mackay, 1991; Milligan, 2019). Moreover, users are not aware of their needs and the visualization they prefer (Langley, 1999).

In order to get an understanding of how users comprehend visualizations and to generate insights for further research, an exploratory eye-tracking study was conducted (see chapter 6). The results underline the assumption that there is no one visualization for all users. The study identified the line chart and tabular as established visualizations in the financial context.

Furthermore, there could be a relation between users' familiarity and decision quality with a visualization. Participants with the same profession had no issues comprehending a particular visualization's information. In addition, the review of adaptive visualization systems (see section 2.4) revealed that users' knowledge and experience might influence decision-making with different visualizations (Toker et al., 2012; Steichen et al., 2013). It is assumed that latent user characteristics reflect users' expertise with particular visualizations. In contrast to prior research (Toker et al., 2012; Steichen et al., 2013), this chapter's studies investigated participants' objective expertise measurements and empirically evaluate task completion and cognitive load in order to answer the following research question:

> **Research Question 2:** *How do user characteristics influence the comprehension of information with different visualizations?*

Toker et al. (2012) have already shown that cognitive abilities influence task completion time and ease of use with different visualizations. Participants with low perceptual speed and low verbal working memory completed the radar plot task faster. In contrast, participants with high perceptual speed and high verbal working memory were more efficient in processing bar charts (Toker et al., 2012).

The assessment of cognitive abilities in order to adapt visualization is not practicable in ISs. The corresponding questionnaires are pretty time-consuming (approximately 30 minutes) (Velez et al., 2005; Carenini et al., 2014; Toker et al., 2012). In addition, the disclosure of cognitive abilities is sensitive in terms of data privacy, as they allow statements to be made about an individual's mental health (Zhang et al., 2018; Jang and Yoo, 2009). More appropriate user characteristics are needed.

A promising alternative in research is the analysis of eye movements to predict cognitive abilities and thus adapt visualizations (Steichen et al., 2013). To date, eye-tracking technology is not standard in the consumer technology sector, and it is unclear if and when the technology will gain acceptance. Therefore, it is currently not a viable option for personalization.

Fortunately, cognitive abilities are reflected in education and expertise knowledge acquisition. Cognitive abilities are linked to corresponding literacy (Lee et al., 2019; Hindal et al., 2009). Linguistic learners are adept at using words, whereas visual learners benefit through the use of images and graphics (Wright et al., 2007).

Many people prefer to avoid financial planning despite the fact that it is a major factor in achieving life goals like buying a house or obtaining a comfortable retirement. Two reasons for this avoidance are low financial literacy (of terms like risk, return, and compound interest) and the high cognitive demand of evaluating investment options (Van Rooij et al., 2011; Lusardi and Mitchell, 2007; Alessie et al., 2007; Lusardi et al., 2010).

Comprehension of financial information, especially probabilities of risk and return, can be boosted by appropriate visualizations (Cleveland and McGill, 1984). Most robo-advisors do indeed use visualizations to communicate portfolio risk and return of forecasts (Salo and Haapio, 2017). As previously mentioned, uniform visualizations can not fulfill the needs of a wide range of users.

In the following, two studies are reported that investigate the influence of identified user characteristics on financial decision-making and cognitive load with different visualizations.

## 7.2 Study 1

The main purpose of this study is to evaluate the experimental design and to investigate the assumptions derived from the exploratory study in chapter 6 and the adaptive visualization review in chapter 2. The exploratory study identified the line chart and tabular as common visualizations in the financial context. These are highly contrasted representations and, therefore, capable of obtaining meaningful insights with relatively small sample size. This study examines the influence of objective expertise measurements that could reflect users' knowledge with the line chart and the tabular.

### 7.2.1 Research Model

Chapter 6 suggests a relation between certain visualizations and experience due to profession. Moreover, Toker et al. (2012) investigated the self-rated experience of different chart types with respect to completion time with promising results. In this study, domain expertise is investigated. Based on the results of knowledge-based domain expertise questionnaires, highly contrasted subgroups are identified with a median split and compared with each other (Iacobucci et al., 2015).

Financial literacy is a relevant factor in individuals' financial decision-making (Lusardi and Mitchell, 2007). Financially educated individuals are more likely to plan for retirement (Lusardi and Mitchell, 2007). They also make more rational decisions in the case of robo-advisory (Jung et al., 2018). Klapper et al. (2012) used the median split to examine

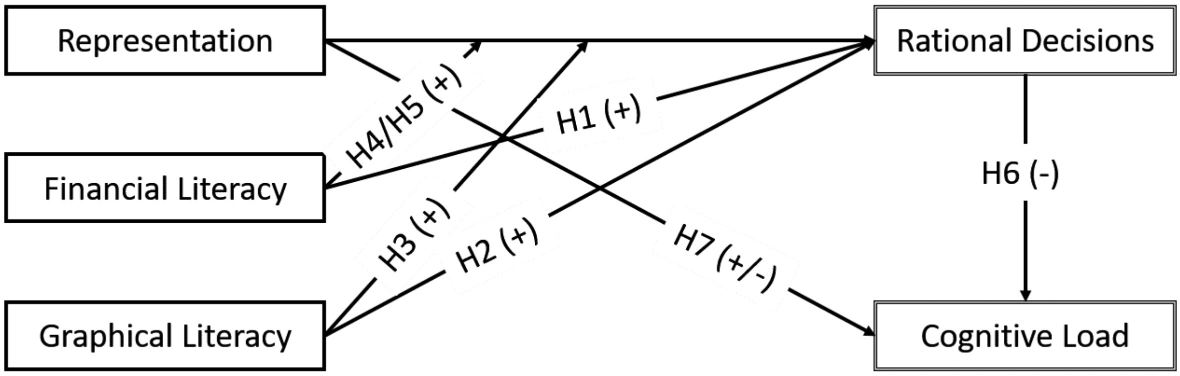



FIGURE 7.1: *Research model of the study.*

the effects of literate and non-literate subgroups on investment behavior. They showed that financially literate individuals make more rational investment decisions than uneducated individuals during the financial crisis. In this study, rational investment decisions are equivalent to maximizing expected return (see section 7.2.2). The median split is adopted to generate subgroups and to form the following hypothesize (Iacobucci et al., 2015):

> ***H1**: Financially literate individuals make more rational investment decisions than non-literate individuals.*

Graphical literacy is a crucial factor in making rational inferences. Okan et al. (2012) showed that graphically literate individuals make more reasoned inferences than non-literate individuals. Literate individuals can comprehend information with higher complexity (Okan et al., 2012). Individuals who have to make financial decisions might be overwhelmed by the complexity of presented information. In the case of graphical literacy, the median split is also used to identify two subgroups and to hypothesize (Iacobucci et al., 2015):

> ***H2**: Graphically literate individuals make more rational investment decisions than non-literate individuals.*

The exploratory study (see chapter 6) suggests that individuals with varying professions process presented information differently. It also suggests the tabular and line chart as common representations in financial decision-making. This study investigates these two highly contrasted representations (see section 7.2.2).

Galesic and Garcia-Retamero (2011) developed the graphical literacy questionnaire to assess individuals' competence in using various visualizations, including line charts. Shah (2002) have shown that graphically literate individuals perform better with line charts than non-literate individuals. Hence, it is expected that the line chart increases the financial decision making competency of graphically literate participants. In contrast, graphical literacy does not indicate experience with tabular representations. Therefore, it is refrained from making any assumption for the tabular.

> ***H3**: Graphically literate make more rational investment decisions with line charts than non-literate individuals.*

Financially literate individuals are more risk-tolerant and experienced with investment decisions (Awais et al., 2016). Visualizations, especially line charts, are established means in financial decision-making to depict market situations or communicate risk and return (Lux, 1997; Zhang et al., 2020). It is assumed that in the course of acquiring expertise, financially literate individuals have already prior experience with line charts.

> ***H4**: Financially literate individuals make more rational investment decisions with financial line charts than non-literate individuals.*

Besides line charts, tabular representations are common in financial decision-making (Zhu et al., 2021; Li et al., 2020). They provide an efficient method for presenting structured data in financial reports (Li et al., 2020). In this study, risk and return of investment alternatives have to be considered to come up with an investment decision. A financially literate individual can process risk and return adequately (Mahdzan et al., 2017). Thus, it is assumed that financially literate individuals will make more rational decision with tabular representations.

> ***H5**: Financially literate individuals make more rational investment decisions with tabular representations than non-literate individuals.*

The review of adaptive information systems in chapter 2 revealed cognitive load as a further relevant dependent variable for designing user visualizations. A higher cognitive load impairs the numeracy and recall ability of individuals (Rose et al., 2004). Scholars suggest that individuals with a high cognitive load are more likely to use intuitive decision processes that can lead to irrational decisions, such as the anchoring effect (Deck and Jahedi, 2015; Raoelison et al., 2020). In this study, the subjective cognitive load is measured after decision-making. Due to the measurement after the experimental task, it can not be used to predict rational investment decisions. Nevertheless, it is assumed that rational investment decisions negatively correlates with subjective cognitive load.

> ***H6***: *Rational investment decisions negatively correlates with subjective cognitive load.*

Rudolph et al. (2009) compared a tabular with an area diagram in an interactive software tool and showed that graphical representations help to overcome cognitive limitations. In another study, it could be demonstrated that boxplot representations cause significantly less cognitive load than a tabular representation of data (Anderson et al., 2011). This study assumes that the financial line chart causes lower subjective cognitive load than the tabular representation.

> ***H7***: *The financial line charts cause a lower cognitive load than tabular representations.*

### 7.2.2 Methodology

This section introduces the methodology of the study, containing the experimental design, the procedure of the study, experimental task, and measures.

#### Experimental Design

Following the procedures of experimental economics (Friedman et al., 1994), an online experiment with a between-subject design is used to investigate the working hypotheses. Participants are randomly assigned to one of two representations of investment scenarios.

### Procedure

The following experimental procedure was used (see figure 7.2). First, participants were informed about the general experimental rules and then are introduced to the experimental task (Instructions) (see section 7.2.2). After the briefing, they performed a quiz with a series of questions to ensure that they correctly comprehend the task (Quiz). Then, following a between-subject design, the participants were randomly assigned to either the line chart representation or tabular representation of the investment scenarios (Experimental Task). There are five different investment scenarios that were represented in random order. Right after the experimental task, participants worked on several questionnaires (Questionnaires). First, they rated their cognitive load. Then, they answered the questionnaires for financial literacy, graphical literacy, and demography.

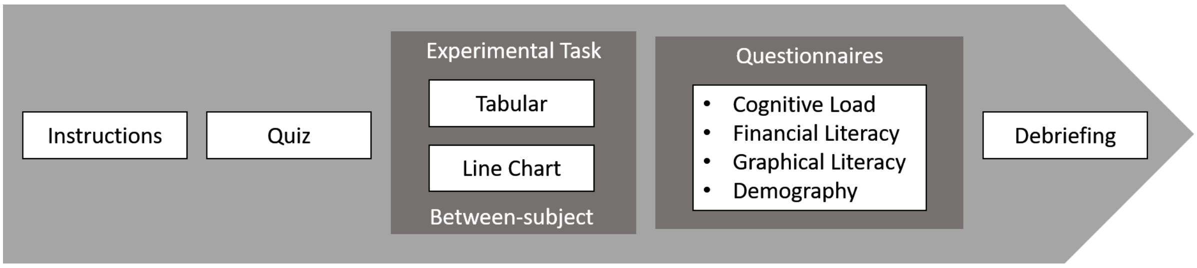



FIGURE 7.2: *Procedure of the study.*

### Experimental Task

The experimental task consists of five different investment scenarios, based on Rudolph et al. (2009). In each scenario, the participants have an investment capital in the amount of 1000 euro. Participants have to choose one of three investments with an investment duration of three years. The investments are represented either in tabular representation (see table 7.1) or line chart (see figure 7.3). Each investment differs concerning the investment amount, risk, and return. If the investment capital is not completely exhausted, the remaining amount is invested in an interest-free account. After the three-year investment period, the earned investment capital and the interest-free investment are paid out. The participants are instructed to identify the investment with the highest expected total return for each investment scenario. Due to the three-year investment period, participants have to calculate compound interests to identify the investment with the expected highest return.

| Investment Option | Investment Amount | Risk ($\sigma$) | Return (%) |
|---|---|---|---|
| A | 600 euro | +/-0.07 | 3 |
| B | 1000 euro | +/-0.12 | 7 |
| C | 600 euro | +/-0.06 | 5 |

TABLE 7.1: *Tabular representation of an investment scenario.*

The investment scenario presented in the table 7.1 is used in the following in order to illustrate the calculation process. First, for each investment option the expected return has to be calculated:

- **Investment option A:** $600\ euro\ x\ 1,03^3 + 400\ euro = 1055.64\ euro$
- **Investment option B:** $1000\ euro\ x\ 1,07^3 + 0\ euro = 1225.04\ euro$
- **Investment option C:** $600\ euro\ x\ 1,05^3 + 400\ euro = 1094.58\ euro$

After the calculation of expected returns, the maximizing investment option can be identified easily. It is option B in the above example.

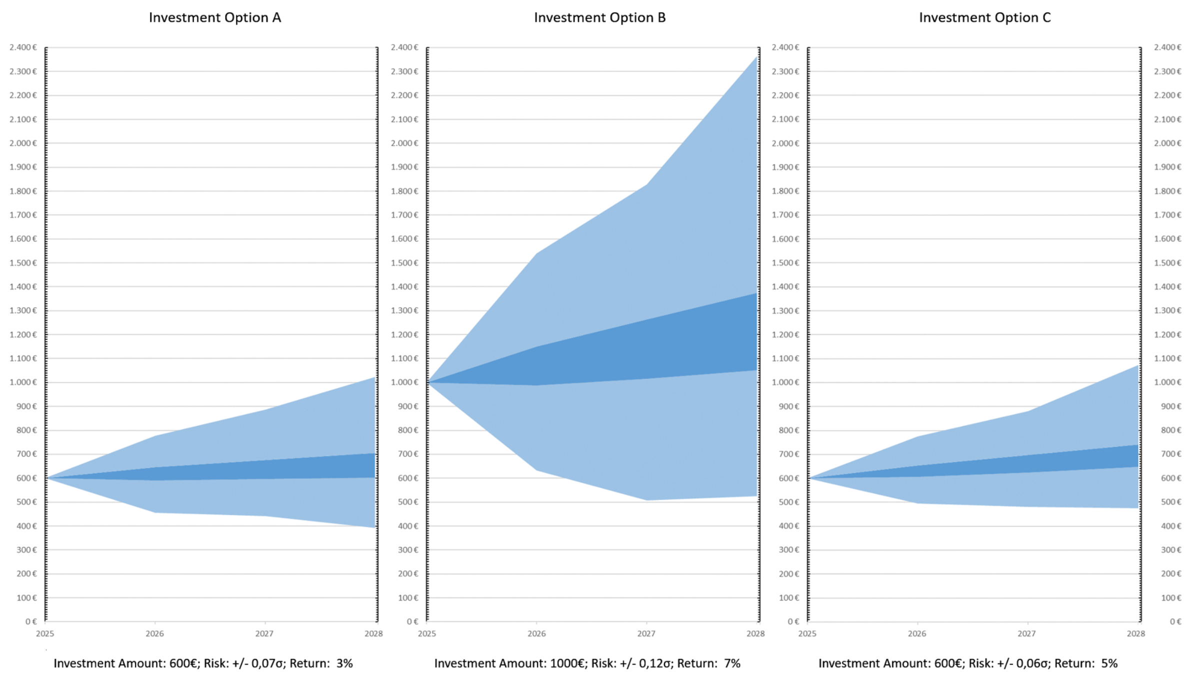



FIGURE 7.3: *Line chart representation of an investment scenario.*

### Measures

The dependent variables of the study are *rational investment decisions* and *cognitive load*. There are five different investment scenarios. There is a rational decision based on the maximized expected return in each scenario (see section 7.2.2). Thus, *rational investment decisions* can vary from zero to five correct answers.

The *cognitive load* of a participant is the second dependent variable. This study used the subjective *cognitive load* after the processing of the experimental task. For the measurement, the NASA-TLX is used (Hoonakker et al., 2011). The questionnaire measures the task-dependent workload on the six subjective subscales *mental demand, physical demand, temporal demand, effort,* and *frustration*. The participants rated each subscale on a 10-point Likert scale.

The study's independent variables are *representation format, financial literacy,* and *graphical literacy*. There are two different representations as treatments, namely *line chart* (see figure 7.3) and *tabular* (see table 7.1). These will be coded as dummy variables for the analysis.

*Financial literacy* is queried by the questionnaire from Lusardi and Mitchell (2007). It consists of three questions concerning compound interest and funds comprehension and two questions for financial retirement planning. Thus, the *financial literacy* score ranges from zero to five correct answers.

*Graphical literacy* is also measured by a knowledge-based questionnaire adopted from Galesic and Garcia-Retamero (2011). It contains 13 items. For each item, a specific visualization is used to query the interrelationship of containing information.

The demographic data, such as *student, age,* and *gender*, are collected as control variables.

### Participants

In order to obtain a broad sample of the population, the participants are acquired via mailing lists and social media. Regarding participants' payoff, it is shown that incentives increase participants' performance in easy effort-responsive tasks (Camerer and Hogarth, 1999). In contrast, incentives in choice tasks do not change participants' mean performance. In many cases, incentives solely reduce the variance of participants' responses (Camerer and Hogarth, 1999). This study refrained from incentivizing the participants to obtain high variance with relatively small sample size.

### 7.2.3 Results

In total, 40 persons participated the study. The study lasted approximately 25 minutes (sd = 11.7). Two subjects who failed the comprehension quiz are excluded from the analysis, resulting in 18 complete observations in the treatment *tabular* and 20 in the treatment *line chart*. On average, the participants were 37.5 (sd = 12.3) years old. 50 % were male, whereas 44.7 % of the participants were students.

#### Analysis of Treatment Homogeneity

First of all, the homogeneity of the treatments is analyzed. For this purpose, the differences between the treatments *tabular* and *line chart* are examined with the variables *age, gender, student, graphical literacy*, and *financial literacy*. Thereby, the following procedure is used. In the case of continuous variables, the distribution is tested with a Shapiro-Wilk normality test. Depending on the outcome, parametric or non-parametric statistical methods are used to examine the treatment differences. In the case of nominal variables, the frequencies are plotted and the treatment differences are tested with Pearson's chi-squared tests. The significance levels of the analysis are described in the footnote[1].

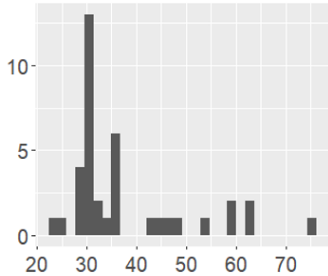
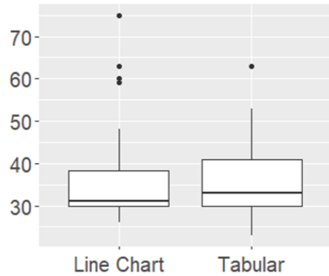
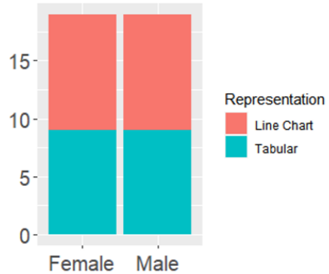
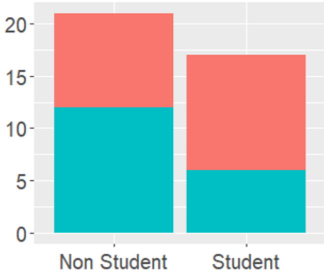
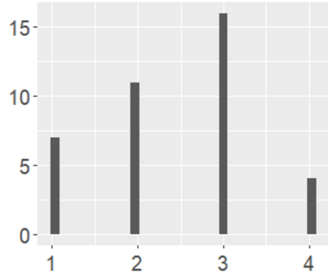
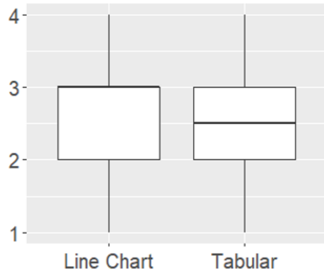
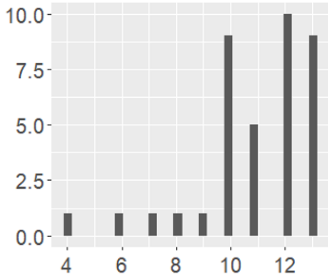
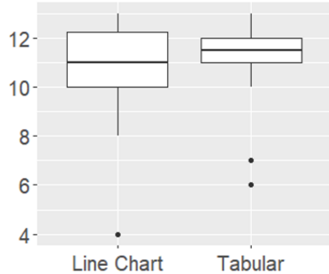
Table 7.2 summarizes the treatment difference analysis. The variable *age* is not normally distributed[2] (Shapiro-Wilk normality test: W = 0.77625, p-value < 0.001). Hence, a Wilcoxon rank-sum test indicates (W = 178, p-value = 0.964) that there is no significant difference between the treatments.

The variables *gender* and *student* are nominal variables. Hence, Pearson's chi-squared tests are used to examine treatment differences. The results show no significant difference

[1] Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

[2] Age is typically normally distributed (Kronthaler, 2021). With a larger sample size, a normal distribution would be expected (central limit theorem).

TABLE 7.2: *Treatment difference analysis of age, gender, student, graphical literacy, and financial literacy.*

| Variable | Overall Distribution | Treatment Distribution | Normally Distributed | Treatment Difference |
|---|---|---|---|---|
| Age |  |  | No (Shapiro-Wilk normality test: W = 0.77625, p-value < 0.001***) | No (two-sided Wilcoxon rank-sum test: W = 178, p-value = 0.964) |
| Gender & Student |  |  | | Gender: No (Pearson's chi-squared test: X-squared = 0, df = 1, p-value = 1); Student: No (X-squared = 1.0292, df = 1, p-value = 0.310) |
| Financial Literacy |  |  | No (Shapiro-Wilk normality test: W = 0.86996, p-value < 0.001***) | No (two-sided Wilcoxon rank-sum test: W = 182.5, p-value = 0.950) |
| Graphical Literacy |  |  | No (Shapiro-Wilk normality test: W = 0.83618, p-value < 0.001***) | No (two-sided Wilcoxon rank-sum test: W = 162.5, p-value = 0.610) |

in *gender* (X-squared = 0, df = 1, p-value = 1) and *student* (X-squared = 1.0292, df = 1, p-value = 0.310).

In the case of *financial literacy*, the participants answered a mean of 2.4 (sd = 0.9) items correctly. A Shapiro-Wilk normality test shows (W = 0.86996, p-value $< 0.001$) that the variable is not normally distributed. Hence, a two-sided Wilcoxon rank-sum test (W = 182.5, p-value = 0.950) indicates no significant difference across the treatments.

Finally, the participants answered a mean of 10.9 (sd = 2.1) items of the *graphical literacy* questionnaire correctly. The variable is not normmally distributed (Shapiro-Wilk normality test: W = 0.83618, p-value $< 0.001$). Therefore, the result of a two-sided Wilcoxon rank-sum test (W = 162.5, p-value = 0.610) shows no significant treatment difference.

In summary, as expected from randomized treatment allocation, no significant treatment differences could be observed (see table 7.2).

**Analysis of Rational Investment Decisions**

In the following, the dependent variable *rational investment decisions* is analyzed. The hypotheses *H1, H2, H3, H4,* and *H5* refer to *rational investment decisions*. As mentioned in section 7.2.2, *rational investment decisions* correspond to the investment options that maximize the expected returns.

Participants made on average 2.4 (sd = 1.3) *rational investment decisions* (5 investment scenarios). As the histogram suggests (see figure 7.4) and a Shapiro-Wilk normality test (W = 0.85596, p-value $< 0.001$) shows, *rational investment decisions* is not normally distributed. Hence, non-parametric statistical tests have to be used for further analysis. In the treatment *line chart*, participants made an average of 2.6 (sd = 1.5) *rational investment decisions*, whereas, in the treatment *tabular*, participants reached a mean of 2.2 (sd = 1.1) *rational investment decisions*. The treatment difference is examined with a two-sided Wilcoxon rank-sum test (W = 227, p-value = 0.156) showing no significant difference. This finding underlines the necessity of this study, since none of the representation formats can be preferred across the board.

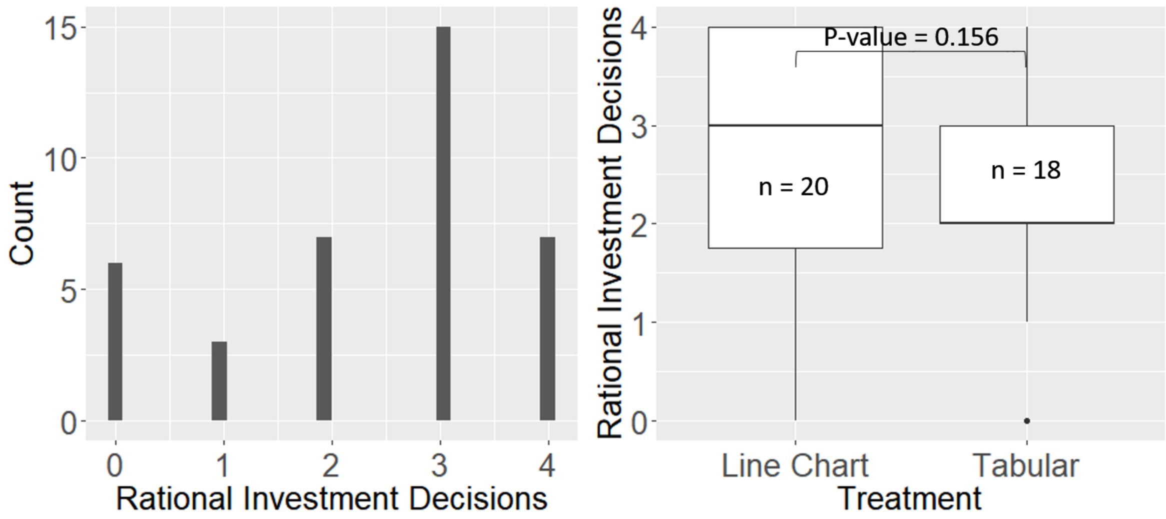



FIGURE 7.4: *Histogram and boxplots of rational investment decisions*

Next, the working hypotheses are investigated. A median split is used for the analysis to compare the resulting highly contrasted subsamples (Iacobucci et al., 2015; Klapper et al., 2012). The number of participants in the subsamples are depicted in the corresponding box plots.

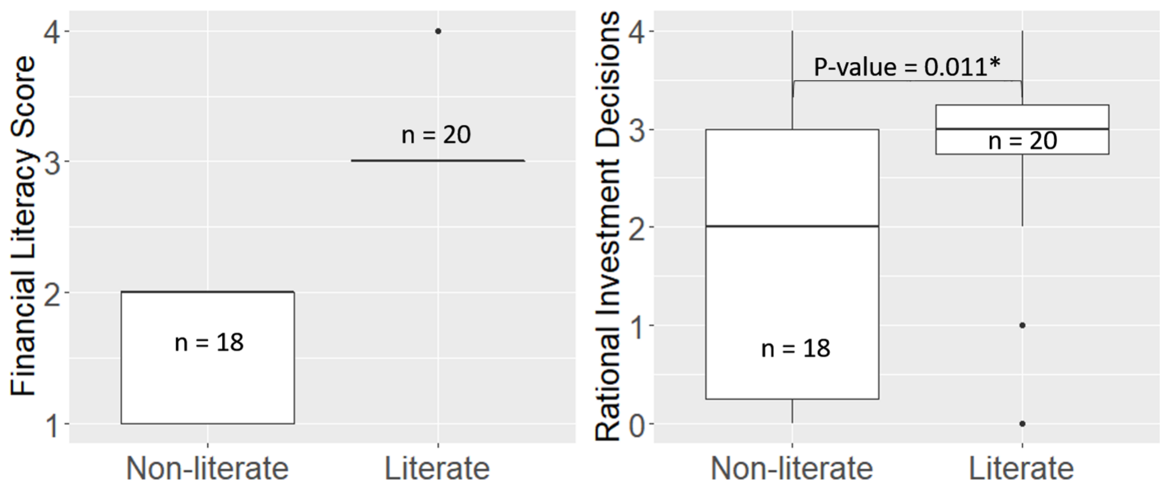



FIGURE 7.5: *Boxplots of financial literacy and rational investment decisions.*

*H1* assumes that *financially literate* individuals make more *rational investment decisions* than *non-literate* individuals. Participants who achieved less than three (median) correct answers on the financial literacy questionnaire form the *non-literate* subgroup (see figure

7.5), while participants who answered three or more questions correctly form the *literate* subgroup. *Financially literate* individuals answered 3.2 (sd = 0.4) items correctly and made an average of 2.9 (sd = 1.0) *rational investment decisions*, whereas *non-literate* individuals answered 1.6 (sd = 0.5) items of the financial literacy questionnaire correctly and made an average of 1.8 (sd = 1.4) *rational investment decisions*. Since an increase of *rational investment decisions* is expected, the difference is investigated with a one-sided Wilcoxon rank-sum test. The result (W = 255.5, p-value = 0.011) shows a significant increase of *rational investment decisions* in the *financially literate* subgroup. Hence, the alternative hypothesis of *H1* accepted.

> ***Result 1****: Financially literate individuals make more rational investment decisions than non-literate individuals.*

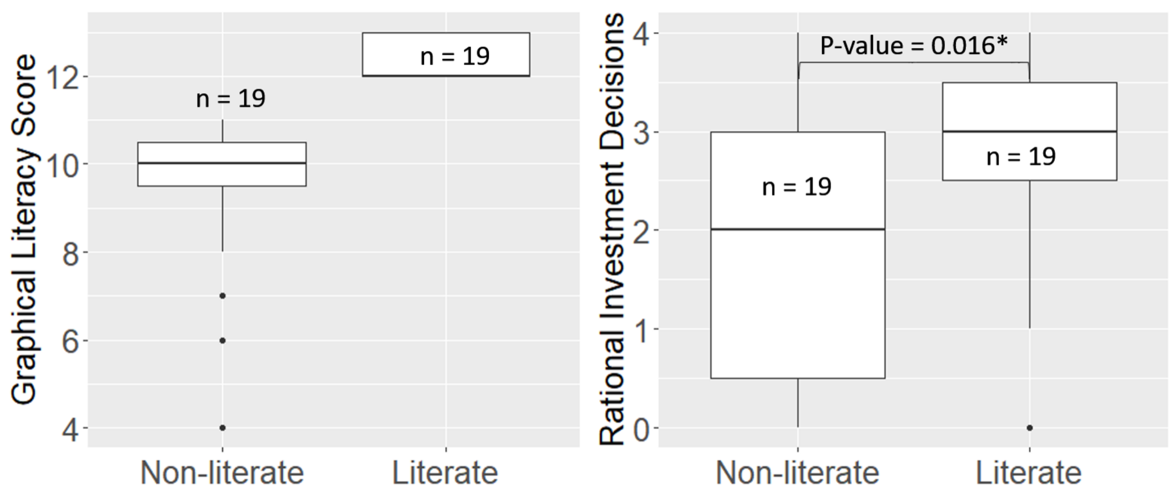



FIGURE 7.6: *Boxplots of graphical literacy and rational investment decisions.*

*H2* assumes that *graphically literate* individuals make more *rational investment decisions* than *non-literate* individuals. The relation of *financial literacy* with *graphical literacy* is investigated with a Pearson's product-moment correlation test. The result (cor = 0.310, t = 1.9632, df = 36, p-value = 0.057) shows a significantly moderate positive relationship, which could be attributed to the indirect measurement of the participants' intelligence with both questionnaires. The participants are allocated into two subgroups (see figure 7.6) with a median split (median = 11.5). *Graphically literate* participants answered 12.5 items (sd = 0.5) of the corresponding questionnaire correctly and made an average of 2.8

(sd = 1.1) *rational investment decisions*, whereas *non-literate* participants answered 9.4 (sd = 1.9) items of the literacy questionnaire correctly and reached a mean of 1.9 (sd = 1.4) *rational investment decisions*. The result of a one-sided Wilcoxon rank-sum test (W = 251.5, p-value = 0.016) shows an significant increase of *rational investment decisions*. Consequently, the alternative hypothesis of *H2* is accepted.

> ***Result 2**: Graphically literate individuals make more rational investment decisions than non-literate individuals.*

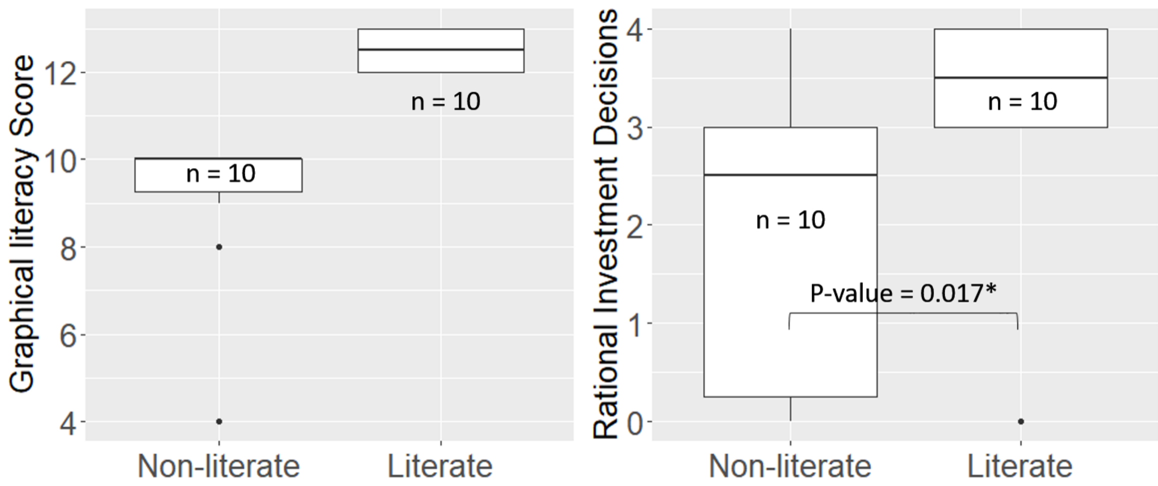



FIGURE 7.7: *Boxplots of graphical literacy and rational investment decisions in the line chart treatment.*

*H3* assumes that *graphically literate* individuals make more *rational investment decisions* with financial *line charts* than *non-literate* individuals. Thus, the subgroups *graphically literate* and *non-literate* are further divided by the treatment variable *representation format*, resulting in ten observations in the *graphically literate* subgroup and ten observations in the *graphically non-literate* subgroup. In fact, such sample sizes can only be used to determine tendencies. Nevertheless, statistical tests are carried out with the awareness that the significance could not be achieved due to small sample sizes. The subgroup *graphically literate* answered an average of 12.5 (sd = 0.5) items of the literacy questionnaire correctly and made a mean of 3.2 (sd = 1.2) *rational investment decisions*, whereas the *non-literate* subgroup answered an average of 9.1 (sd = 1.9) items correctly and made a mean of 1.9 (sd = 1.5) *rational investment decisions*. A one-sided Wilcoxon rank-sum test (W = 77, p-value = 0.017) shows a significant increase of *rational investment decisions*. Hence, the alternative hypothesis of *H3* is accepted.

> ***Result 3**: Graphically literate individuals make more rational investment decisions with line charts than non-literate individuals.*

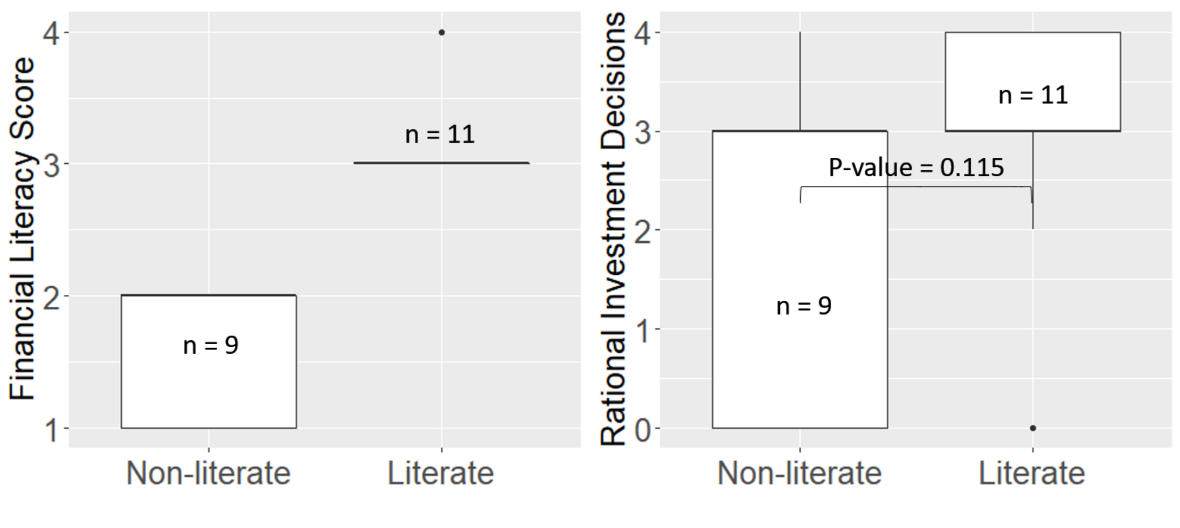



FIGURE 7.8: *Boxplots of financial literacy and rational investment decisions in the line chart treatment.*

Next, the hypothesis *H4* is investigated, which assumes that *financially literate* individuals make more *rational investment decisions* with financial *line charts* than *non-literate individuals*. The subgroups of *financially non-literate* (n = 9) answered a mean of 1.6 (sd = 0.5) items of the corresponding questionnaire correctly and made an average 2.0 (sd = 1.7) *rational investment decisions*, whereas, *financially literate* participants (n = 11) answered a mean of 3.2 (sd = 0.4) items correctly and made a mean of 3.0 (sd = 1.2) *rational investment decisions*. A one-sided Wilcoxon rank-sum test is performed without a significant result (W = 65, p-value = 0.115). However, it is assumed that the non-significant result is caused by the small sample size. The null hypothesis of *H4* can not be rejected, but a promising tendency is present.

> ***Result 4**: Financial literacy does not significantly influence rational investment decisions with line charts.*

*H5* hypothesize that *financially literate* individuals make more *rational investment decisions* with *tabular* representation than *non-literate* individuals. The *financially literate*

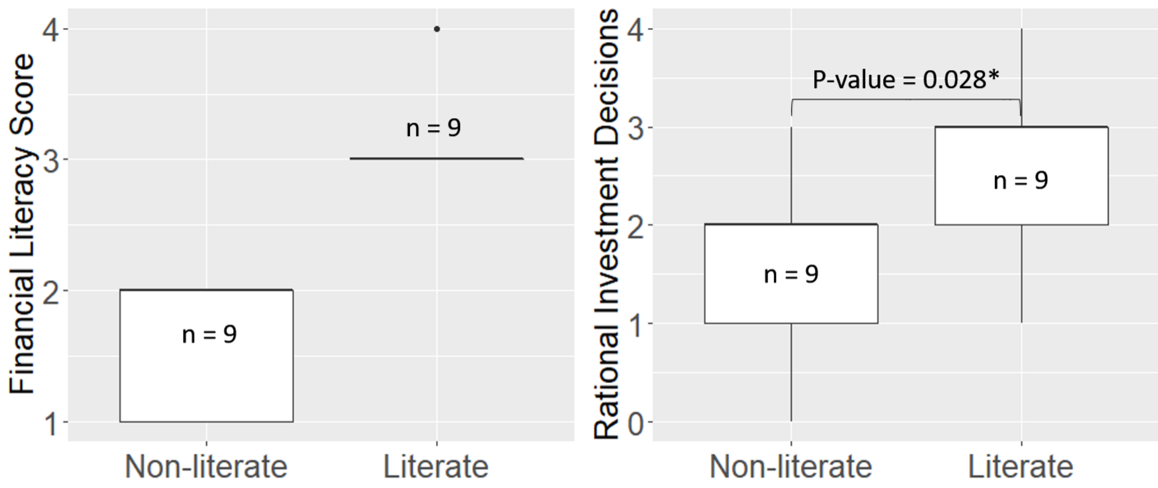



FIGURE 7.9: *Boxplots of financial literacy and rational investment decisions in the tabular treatment.*

subgroup (n = 13) answered a mean of 3.3 (sd = 0.6) items correctly and made an average of 2.5 (sd = 0.9) *rational investment decisions*. The *financially non-literate* subgroup correctly answered a mean of 1.5 (sd = 0.5) items and made an average of 2.0 (sd = 1.5) *rational investment decisions*. The result of the one-sided Wilcoxon rank-sum test (W = 61.5, p-value = 0.028) is significant. Hence, the alternative hypothesis is accepted.

> ***Result 5****: Financially literate individuals make more rational investment decisions with tabular than non-literate individuals.*

| # | Sample 1 | Size | Sample 2 | Size | Acceptance |
|---|---|---|---|---|---|
| H1 | Financially Literate | 18 | Financially Non-literate | 20 | True (0.011*) |
| H2 | Graphically Literate | 19 | Graphically Non-literate | 19 | True (0.016*) |
| H3 | Graphically Literate with Line Chart | 10 | Graphically Non-literate with Line Chart | 10 | True (0.017*) |
| H4 | Financially Literate with Line Chart | 9 | Financially Non-literate with Line Chart | 11 | False (0.115) |
| H5 | Financially Literate with Tabular | 9 | Financially Literate with Tabular | 9 | True (0.028*) |

TABLE 7.3: *Summary of results with respect to rational investment decisions.*

In summary, the results confirm four hypotheses (see table 7.3 – *H1, H2, H3*, and *H5*).

### Analysis of Cognitive Load

In this section, the dependent variable *cognitive load* is investigated. The participants reported a mean of 5.2 points (sd = 1.6) in a possible range from zero to ten. According to a Shapiro-Wilk test (W = 0.98318, p-value = 0.826), *cognitive load* is normally distributed (see figure 7.10). Hence, parametric tests can be used for further analysis. Since it is hypothesized that financial line charts cause lower cognitive load, it makes no sense to check for treatment difference.

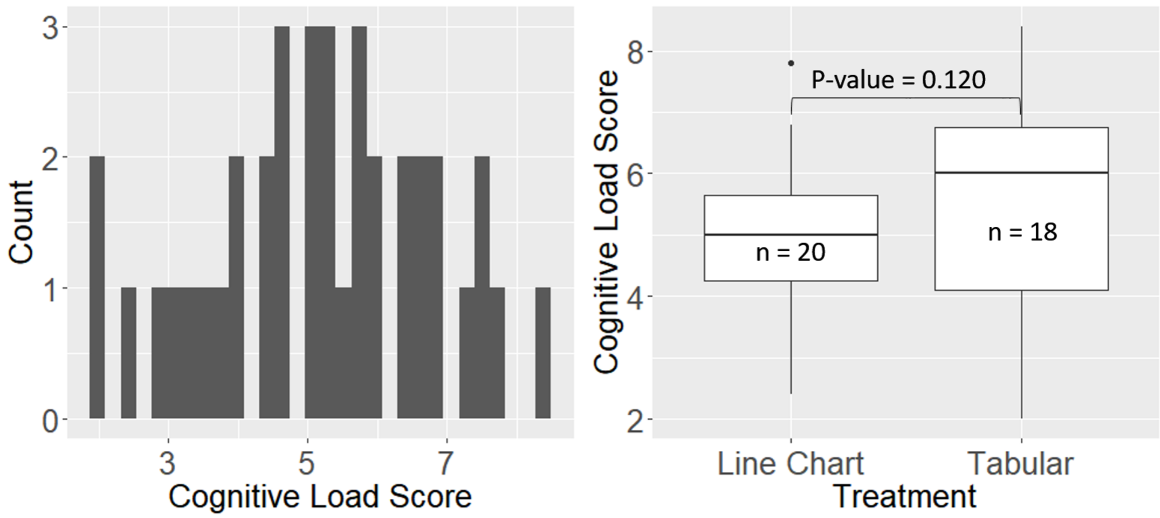



FIGURE 7.10: *Histogram and boxplots of cognitive load*

*H6* assumes that *rational investment decisions* negatively correlate with *cognitive load*. A Pearson's product-moment correlation test is used to investigate the linear relationship. The result (cor = -0.281, t = -1.7605, df = 36, p-value = 0.086) shows that *cognitive load* decreases significantly with increasing *rational investment decisions*. Hence, the alternative hypothesis of *H6* is accepted.

> ***Result 6**: Rational investment decisions negatively correlate with subjective cognitive load.*

Finally, *H7* assumes that the *financial line charts* cause a lower *cognitive load* than the *tabular* representations. The mean *cognitive load* in the *tabular* treatment amounts to 5.5 points (sd = 1.9), whereas, in the *line chart* treatment, the mean is 4.9 points (sd = 1.3). Due to the inhomogeneity of the standard deviations, a one-sided Wilcoxon rank-sum test is performed to test the difference. The result indicates no significant ($W = 220.5$, p-value $= 0.120$) difference. Thus, the alternative hypothesis of *H7* is rejected.

> ***Result 7**: Financial line charts do not cause a lower cognitive load than tabular representations.*

In essence, there is a significant negative linear relationship between subjective *cognitive load* and *rational investment decisions*. Moreover, it was expected that *line charts* cause less *cognitive load* than *tabulars*. The descriptive analysis confirms the expectation, but the difference is not significant. In the following section, the results are discussed.

### 7.2.4 Discussion

The main goal of this study is the evaluation of the experimental design. Since five of seven hypotheses are accepted, and for the remaining ones, tendencies towards the expectations are observed, it is assumed that the experimental design is appropriate to examine the research question.

Remember, the research question deals with how user characteristics influence comprehension with different representations. First, it is shown that financial literacy and graphical literacy influence decision-making in the financial context. In both cases, rational investment decisions increases with the literacy. Accordingly, financial ISs could check the literacy of their users and, if necessary, improve it in order to enable them to make better investment decisions. The corresponding questionnaires do not require a long time to complete, nor is the information disclosed sensitive in terms of data privacy.

Turning to the influence of user characteristics on different representations, graphically literate participants make better investment decisions with financial line charts than non-literate individuals. Moreover, financially literate individuals make better investment decisions with tabular representations than non-literates. The study also assumed that financially literate individuals perform better with line charts than non-literate individuals. The descriptive statistics confirm the expectation, but the difference is not significant. However, the non-significant result could be caused by the small sample sizes. In the subsequent study, all hypotheses will be investigated with a larger sample.

The study revealed a significant and moderately positive relation between graphical and financial literacy. Since both questionnaires are knowledge-based, it is assumed that an indirect measurement of participants' intelligence took place. This is a common phenomenon in experimental economics and is negligible since the constructs have not been analyzed in combination. Moreover, it is much more important to what extent the questionnaires explain additional variance besides the intelligence measurement. Both questionnaires show a moderate correlation with education level, a robust predictor of intelligence. However, discriminant validity is present since both questionnaire are able to predict behavior independent of education level (Galesic and Garcia-Retamero, 2011; Lusardi and Mitchell, 2007).

Participants of the study were recruited via mailing lists and social media, resulting in a broad sample of the population ($mean_{age} = 37.5$, $sd_{age} = 11.7$, 44.7 % student). It could be argued that students are not the ideal subjects for investment decisions, as they do not have the necessary financial resources. However, high school students already receive instruction on topics related to household financial decision-making. Bernheim et al. (2001) have shown that exposure to financial curricula explains the asset accumulation in adulthood.

A larger and specific student sample for the second study will be used for more control. Moreover, the line chart conveys more information than the tabular representation since the probability of the risk (25 % and 75 % quantile) have two different colors. This information does not play a role in the decision-making since the goal is to maximize the expected return. Nevertheless, the two quantiles will be added to the tabular representation.

## 7.3 Study 2

The first study (see section 7.2) confirmed the insights of the exploratory eye-tracking study (see chapter 6) and the review of adaptive visualization systems (see chapter 2). The primary goal of this research is to identify user characteristics that influence decision-making with different visualizations. The experimental design is approved by the first study so that the research model of this study is extended with further user characteristics.

Moreover, an additional representation is implemented to evaluate the effectiveness of visualizations more thoroughly.

### 7.3.1 Research Model

The research model of the first study (see section 7.2.1) is extended with further working hypotheses. In the following, only the extensions are reported.

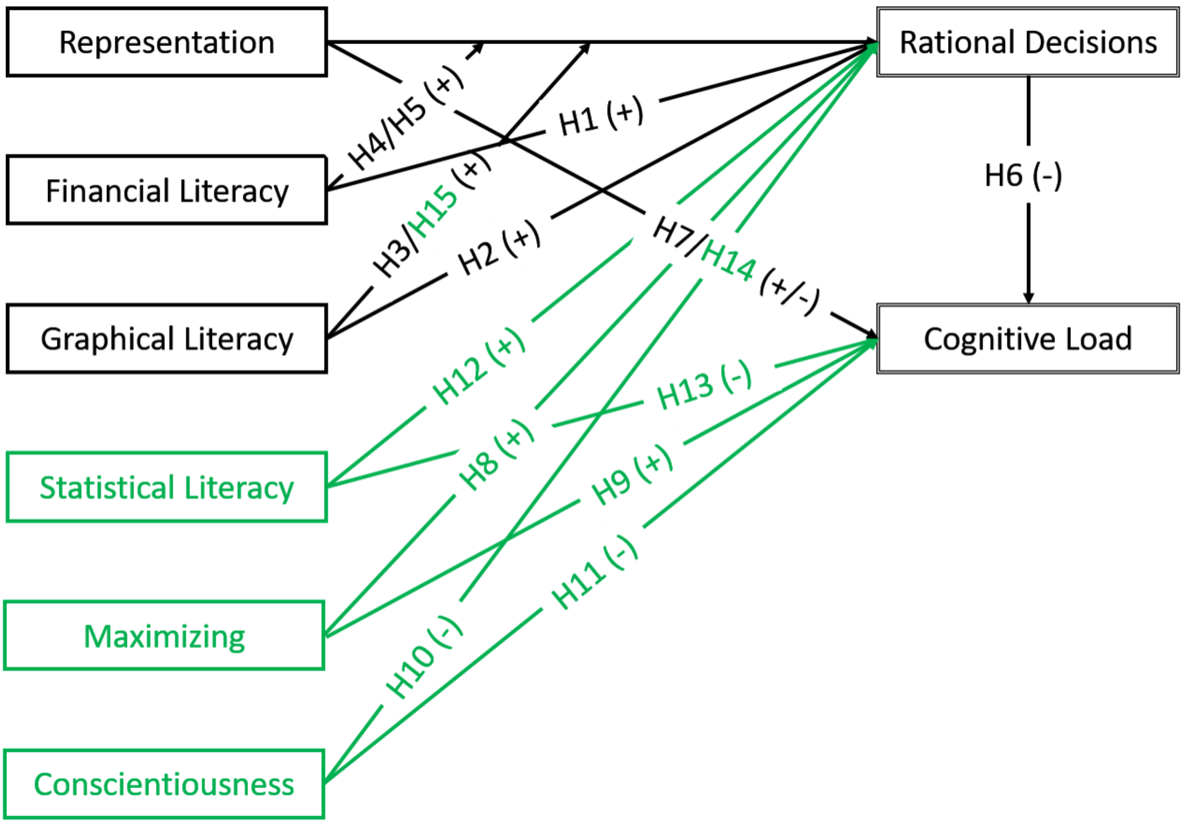



FIGURE 7.11: *Research model of the study.*

There are three further user characteristics, namely *maximizing, conscientiousness,* and *statistical literacy*, which are investigated in this study. These are added because they potentially influence the quality of financial decision-making, and therefore, could be used in ISs to evaluate users' decision-making competency.

Behavioral decision-making research investigated the decision-making styles of individuals with respect to the axiom of utility-maximizing (Hastie and Dawes, 2009; Von Neumann and Morgenstern, 2007; Yates, 1990). As a result, individuals can be classified on a spectrum ranging from *satisficer* to *maximizer*. *Satisficing* describes the choice of a decision alternative, which is good enough. In contrast, *maximizing* describes the strategy of choosing the alternative with the highest utility (Parker et al., 2007).

*Maximizing* is related to problematic decision-making since individuals are less behavioral coping, show more dependence on others' opinions when deciding, and are more prone to experience regret, and therefore, tend to avoid decisions. However, when they are forced to make a decision, as it is also the case in this study, *maximizers* are more likely to make spontaneous decision (Parker et al., 2007). Hence, this study assumes that *maximizers* make less *rational investment decisions*[3] than *satisficers*.

> ***H8****: Maximizers make less rational investment decisions than satisficers.*

Furthermore, due to their claim of maximizing and the additional cognitive processing caused by deliberation on others' opinions, *maximizers* are expected to show a higher cognitive load than *satisficers* (Parker et al., 2007).

> ***H9****: Maximizers have a higher cognitive load than satisficers.*

In the case of the personality trait *conscientiousness*, meta-analyses show that it influences the outcomes of different types of tasks (Barrick et al., 2002; Dudley et al., 2006). Especially, it could be shown that a lower *conscientiousness* leads to better decisions in a multiple cue probability learning task (Hollenbeck et al., 1995; LePine et al., 2000). LePine et al. (2000) related this behavior to dutifulness and additional deliberation. Inline, this study assumes that a higher level of *conscientiousness* leads to less *rational investment decisions* and more *cognitive load* than a lower level.

> ***H10****: A higher level of conscientiousness leads to less rational investment decisions than a lower level of conscientiousness.*

> ***H11****: A higher level of conscientiousness leads to a higher cognitive load than a lower level of conscientiousness.*

[3] A Rational investment decision is equivalent to the investment option that maximizes the expected return (see section 7.2.2).

*Statistical literacy* assesses the statistical numeracy and risk literacy of an individual (Cokely et al., 2012; Ghazal et al., 2014). Ghazal et al. (2014) have shown a strong correlation for task performance in the medical context. Moreover, *statistical literacy* is a robust predictor for meta-cognitive abilities (Ghazal et al., 2014). Consequently, it is expected that a higher level of *statistical literacy* leads to more *rational investment decisions* and a lower *cognitive load* than a lower level.

> ***H12**: Statistically literate individuals make more rational investment decisions than non-literate individuals.*

> ***H13**: A higher level of statistical literacy leads to lower cognitive load than a lower level.*

In order to evaluate the effectiveness of visualizations more thoroughly, the *boxplot* visualization is added as a further representation format of the investment decision situation. Due to its simplistic representation, it is an omnipresent visualization in science (Streit and Gehlenborg, 2014). *Boxplots* lead to lower *cognitive load* than *tabular* representation in distribution interpretation tasks (Anderson et al., 2011). In line, this study also assumes that *boxplots* lead to a lower *cognitive load* in financial decision-making.

> ***H14**: Boxplots cause lower cognitive load than the tabular representations.*

In the case *graphical literacy*, the relationship with *boxplots* is intuitive. It is assumed that *graphically literate* individuals make more *rational investment decisions* with *boxplots* than *non-literate* individuals.

> ***H15**: Graphically literate individuals make more rational investment decisions with boxplots than non-literate individuals.*

*Boxplots* are not very common in the financial domain (Li et al., 2016; Benjamini, 1988). They are mainly used by statistically well-grounded financial experts (Chan, 2004). Hence, it is refrained from making a hypothesis for financial literacy.

### 7.3.2 Methodology

This section introduces the methodological extensions for the second study. The experimental design, the procedure, the experimental task, and the measures are adapted from the first study and reported in section 7.2.2.

This study was conducted with Karlsruhe Decision and Design Lab (KD$^2$Lab) sample. It mainly consists of German students from the Karlsruhe Institute of Technology (KIT). The participants were acquired via Hroot (Bock et al., 2014). This study refrains from performance-based incentives in order to prevent participants from choosing risky investment options instead of maximizing expected returns. Participants were incentivized to choose the investment that maximizes the expected return by raffling 650 euro (20 x 20 euro, 25 x 10 euro) under participants who processed the experiment adequately (processing time $\geq$ 10 minutes, correct answers in quiz $\geq$ 3) and identified the expected return maximizing investment option in all investment scenarios.

The experimental task is the same as in the first study. Remember, the identification of the investment option that maximizes the expected return (see section 7.2.2). The line chart and the tabular representation are adapted from the first study. The quantile information is added to the tabular representation to convey the same amount of information in all treatments. Boxplots as a further graphical representation (see figure 7.12) and a further investment scenario is added to get more variance, resulting in six investment decisions.

| Investment | 1 |
|---|---|
| Investment Amount | 450.00 Euro |
| Risk | +/- 0.07 |
| Return | 7.00 % |
| Minimum | 349.97 Euro |
| $Q_{0,25}$ | 508.51 Euro |
| $Q_{0,75}$ | 594.80 Euro |
| Maximum | 845.23 Euro |

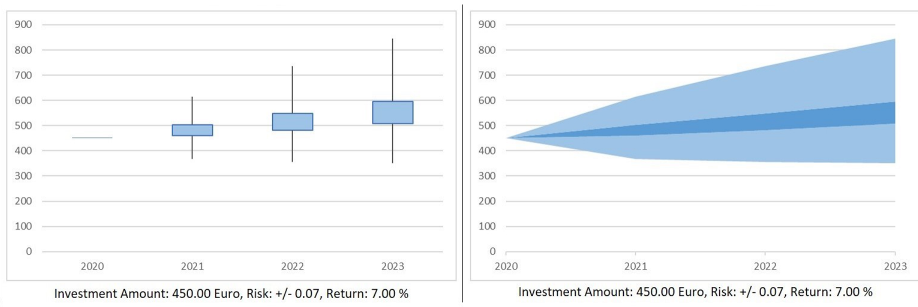



FIGURE 7.12: *The treatments of the study, namely tabular representation, boxplot, and line chart.*

The measures are extended by the questionnaires for *maximizing*, *conscientiousness*, and *statistical numeracy*. For the assessment of *maximizing*, the maximizing scale with

12 items is used (Schwartz et al., 2002). Each item describes a behavior in an everyday situation. The participants responded to each item using a 7-point Likert scale ranging from "completely disagree" to "completely agree".

In the case of *conscientiousness*, the big five questionnaire is used (Donnellan et al., 2006). The subscale of *conscientiousness* contains four items. Participants rated each item on a 5-point Likert scale ranging from "complete disagree" to "complete agree".

For the measurement of *statistical literacy*, the berlin numeracy questionnaire with four items is used (Cokely et al., 2012). Each item contains a computational task. Thus, the *statistical numeracy* score ranges from zero to four correct answers.

In the case of *financial literacy*, the sophisticated financial literacy questionnaire is used (Lusardi and Mitchell, 2007). It contains three more items than the basic version, which was used in the first study. Hence, the *financial literacy* score ranges from zero to eight correct answers.

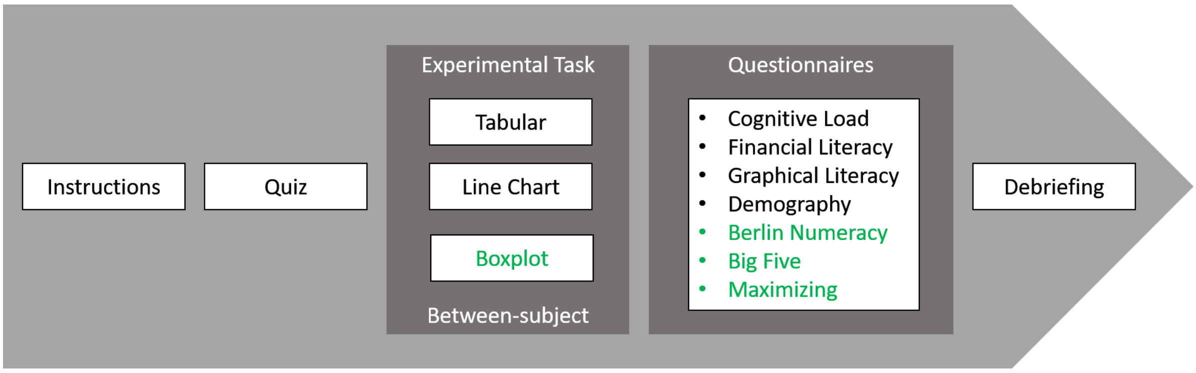



FIGURE 7.13: *The extended experimental procedure of the study. The extensions are colored green.*

Following the experimental procedures of Friedman et al. (1994), a between-subject design is chosen for the study (see figure 7.13).

### 7.3.3 Results

In total, 159 persons participated the study. After data cleansing with quiz result (minimum 3 of 4 correct answers) and processing time ($\geq$ 10 minutes), 20 observations are excluded from the analysis, resulting in 40 female and 89 male participants. The experiment took 39.1 (sd = 15.5) minutes on average, and the participants are on average 23.7 (sd = 4.6) years old. Forty-eight observations are collected in the treatment *tabular*, 45 in the treatment *line chart*, and 36 in the treatment *boxplot*. The most common educational qualification is abitur (48.0 %), followed by the bachelor's degree (42.6 %), and the master's degree (7.0 %).

**Analysis of Treatment Homogeneity**

First of all, the homogeneity of the treatments is examined. For this purpose, the differences between the treatments *tabular*, *boxplot* and *line chart* are examined with the variables *age, gender, education, maximizing, conscientiousness, statistical literacy, graphical literacy*, and *financial literacy*.

For the analysis, the following procedure is used. In the case of continuous variables (*age, maximizing, conscientiousness, statistical literacy*, and *graphical literacy*), the mean value and standard deviation is reported, and the distribution is tested with a Shapiro-Wilk normality test. If the normal distribution is present, parametric statistical methods are used for further analysis, whereas non-parametric statistical methods are used in the other case. For nominal variables (*gender* and *education*), the frequencies are plotted, and Pearson's chi-squared tests are used to examine treatment differences.

The table 7.4 summarizes the analysis of the variables *age, gender, education*, and *maximizing*. *Age* (Shapiro-Wilk normality test: W = 0.69998, p-value < 0.001) is not normally distributed. Hence, a Kruskal-Wallis rank-sum test (chi-squared = 0.7975, df = 2, p-value = 0.671) shows no significant treatment difference.

*Gender* and *education* are nominal variables. Consequently, Pearson's chi-squared tests show that there is no difference in *gender* (X-squared = 1.0437, df = 2, p-value = 0.593) and *education* (X-squared = 3.9652, df = 6, p-value = 0.681) between the treatment.

In the case of *maximizing*, participants have a mean of 52.9 (sd = 11.8). *Maximizing* is normally distributed (Shapiro-Wilk normality test: W = 0.99424, p-value = 0.883). Hence, an ANOVA shows (F = 1.473, df = 2, p-value = 0.233) no significant difference across the treatments.

TABLE 7.4: *Treatment difference analysis of age, gender, education, and maximizing.*

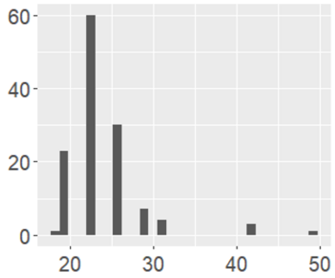
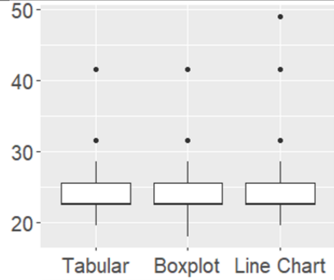
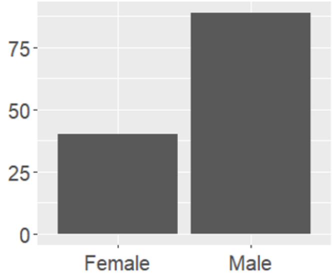
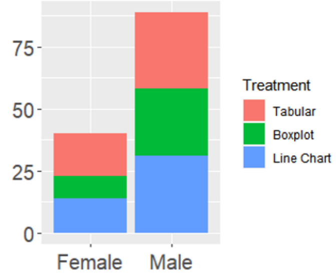
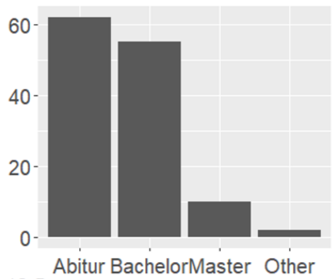
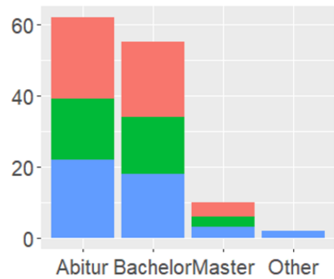
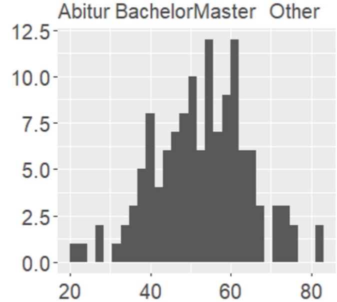
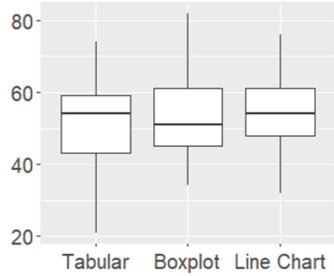
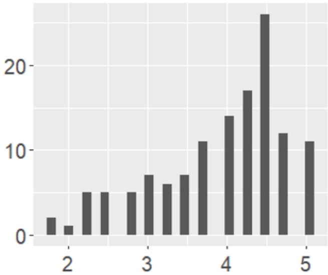
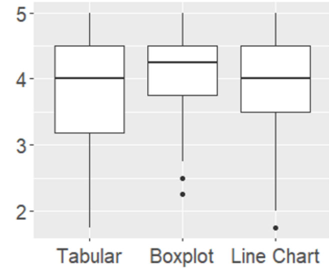
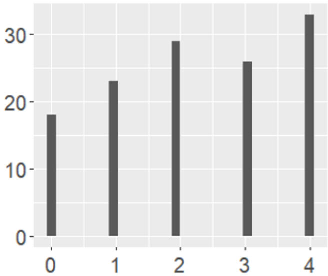
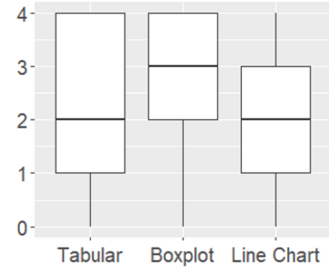
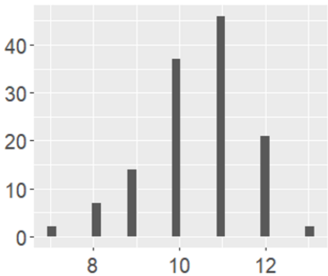
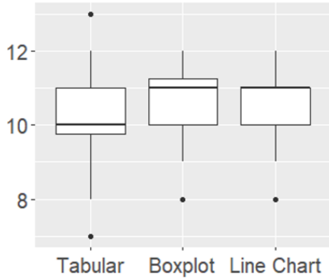
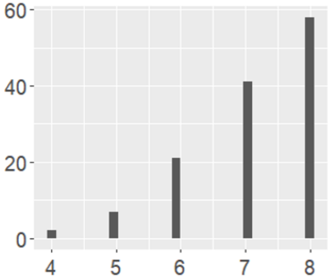
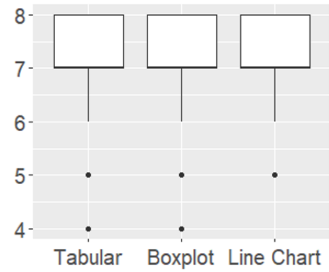
| Variable | Overall Distribution | Treatment Distribution | Normally Distributed | Treatment Difference |
|---|---|---|---|---|
| Age |  |  | No (W = 0.69998, p-value < 0.001***) | No (Kruskal-Wallis rank-sum test: chi-squared = 0.7975, df = 2, p-value = 0.671) |
| Gender |  |  | | No (Pearson's chi-squared test: X-squared = 1.0437, df = 2, p-value = 0.593) |
| Education |  |  | | No (Pearson's chi-squared test: X-squared = 3.9652, df = 6, p-value = 0.681) |
| Maximizing |  |  | Yes (W = 0.99424, p-value = 0.883) | No (ANOVA: F = 1.473, df = 2, p-value = 0.233) |

TABLE 7.5: *Treatment difference analysis of conscientiousness, statistical literacy, graphical literacy, and financial literacy.*

| Variable | Overall Distribution | Treatment Distribution | Normally Distributed | Treatment Difference |
|---|---|---|---|---|
| Conscientiousness |  |  | No (W = 0.91501, p-value < 0.001***) | No (Kruskal-Wallis ranks-sum test: chi-squared = 1.6545, df = 2, p-value = 0.437) |
| Statistical Literacy |  |  | No (W = 0.88691, p-value < 0.001***) | No (Kruskal-Wallis ranks-sum test: chi-squared = 4.0166, df = 2, p-value = 0.134) |
| Graphical Literacy |  |  | No (W = 0.91754, p-value < 0.001***) | No (Kruskal-Wallis ranks-sum test: chi-squared = 1.1543, df = 2, p-value = 0.561) |
| Financial Literacy |  |  | No (W = 0.80379, p-value < 0.001***) | No (Kruskal-Wallis ranks-sum test: chi-squared = 0.02547, df = 2, p-value = 0.987) |

The table 7.5 summarizes the analysis of the variables *conscientiousness, statistical literacy, graphical literacy*, and *financial literacy*. The participants' *conscientiousness* amounts to a mean of 3.9 (sd = 0.8) points. *Conscientiousness* is not normally distributed (Shapiro-Wilk normality test: W = 0.91501, p-value < 0.001). Thus, a Kruskal-Wallis rank-sum test shows (chi-squared = 1.6545, df = 2, p-value = 0.437) no significant difference.

In the case of *statistical literacy*, participants reached a mean of 2.3 (sd = 1.4) correct answers. Since *statistical literacy* is not normally distributed (Shapiro-Wilk normality test: W = 0.88691, p-value < 0.001). The result of the Kruskal-Wallis rank-sum test indicates (chi-squared = 4.0166, df = 2, p-value = 0.134) no significant treatment differences.

The participants answered a mean of 10.4 (sd = 1.2) items of the *graphical literacy* questionnaire correctly. *Graphical literacy* is not normally distributed (Shapiro-Wilk normality test: W = 0.91754, p-value < 0.001). Furthermore, no significant treatment differences are present (Kruskal-Wallis rank-sum test: chi-squared = 1.1543, df = 2, p-value = 0.561).

Finally, the participants' *financial literacy* amounts to a mean of 7.1 (sd = 1.0) correct answers. A Shapiro-Wilk normality test (W = 0.80379, p-value < 0.001) shows that *financial literacy* is not normally distributed. There are no significant differences between the treatments (Kruskal-Wallis ranks-sum test: chi-squared = 0.02547, df = 2, p-value = 0.987).

The *financial literacy* questionnaire is too simple for the current sample (see histogram of financial literacy in table 7.4). This result was not expected, since the sophisticated version of the questionnaire is used (see section 7.3.2). With this insight, the median-split based analysis's weaknesses come to bear. The generated subgroups would be highly unbalanced ($n_{low}$ = 99, $n_{high}$ = 30), and the variance, especially in the literate subgroup ($sd_{high}$ = 0.4), would be too small to conduct further meaningful analyses. Consequently, regression analysis are used in this study, since they are more robust with respect to skewed distributions.

In essence, no treatment differences are observed since none of the test results are significant. As expected from random treatment allocation, the treatments are balanced with respect to the variables *age, gender, education, statistical literacy, maximizing, conscientiousness, graphical literacy*, and *financial literacy*.

**Analysis of Rational Investment Decisions**

In this section, the dependent variable *rational investment decisions* is analyzed. *rational investment decisions* are equivalent to the investment options that maximize the expected returns (see section 7.2.2). There are six different investment scenarios. Hence, *rational investment decisions* ranges from zero to six correct answers.

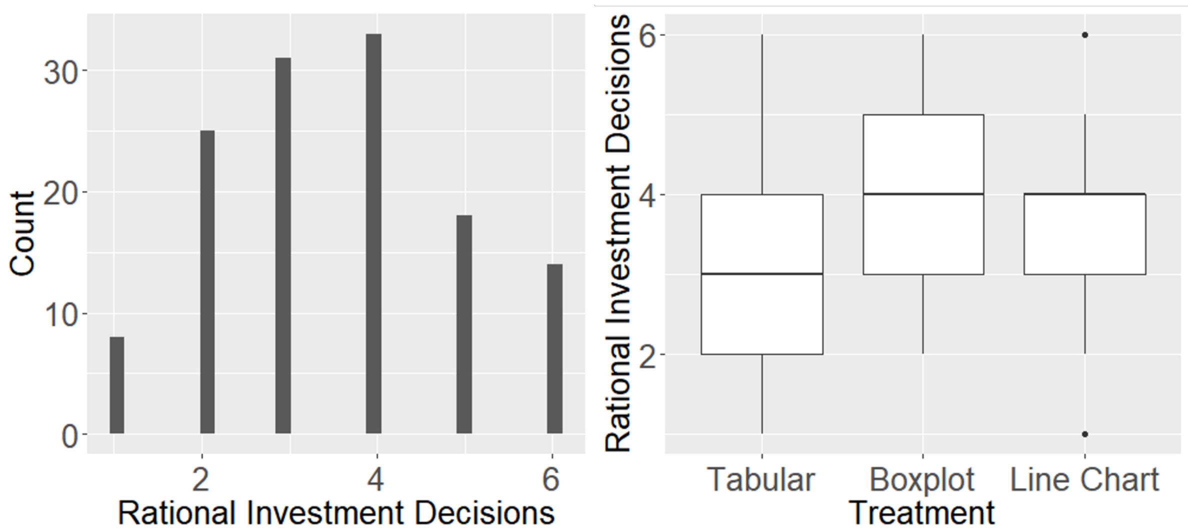



FIGURE 7.14: *Histogram and boxplots of rational investment decisions.*

The mean of *rational investment decisions* amounts to 3.5 (sd = 1.4). A Shapiro-Wilk normality test indicates (W = 0.93388, p-value $< 0.001$) that the variable is not normally distributed (see figure 7.14). Hence, non-parametric tests are used for further analysis. The mean values in the treatments are (descending order): $\text{mean}_{boxplot} = 3.9$ ($\text{sd}_{boxplot}$ = 1.4), $\text{mean}_{linechart} = 3.6$ ($\text{sd}_{linechart}$) = 1.4, and $\text{mean}_{tabular} = 3.3$ ($\text{sd}_{tabular} = 1.4$). A Kruskal-Wallis rank-sum test shows (chi-squared = 4.3539, df = 2, p-value = 0.113) no significant differences across the treatments. As in the previous study, none of the representation alternatives is generally preferable, which emphasizes the need for personalized visualizations.

In the following, the working hypotheses with respect to *rational investment decisions* are investigated. Since participants made six investment decisions, learning or fatigue effects have to be examined, and if present, controlled. Hence, it is opted for logistic regression analyses. The dependent variable is binary, representing a rational investment decision (1 = rational investment decision) for each of the six investment scenarios (round number). The participants are considered as random effects and all other factors as fixed effects.

In the case of logistic regressions and regressions with interactions, the effects and the significances of the effects are conditional. Hence, a variable's effect and significance vary for each value of the remaining variables. The standard regression table reflects the effects of a particular condition. Consequently, the standard regression table must be ignored and the average marginal effects (AME) must be used to evaluate the hypotheses (Schunck and Nisic, 2020; Brambor et al., 2006).

| **Variable** | **AME** | **SE** | **P-value** |
|---|---|---|---|
| Round Number | -0.035 | 0.010 | < 0.001 *** |
| Financial Literacy | 0.064 | 0.020 | 0.001 ** |

TABLE 7.6: *Average marginal effects of financial literacy (H1).*

*H1* hypothesizes that an increasing *financial literacy* leads to more *rational investment decisions*. Hence, the independent variable of the regression (see appendix B.3) is *financial literacy*. The average marginal effects (see table 7.6) indicates a significant and negative effect of *round number* (AME = -0.035, p-value $< 0.001$) on *rational investment decision*. This shows that the participants performed worse with each round of the experimental task; therefore, a fatigue effect could be present. *Financial literacy* also significantly influences (AME = 0.064, p-value = 0.001) *rational investment decision*. With increasing *financial literacy* the probability for a *rational investment decision* increases. Accordingly, the alternative hypothesis of *H1* is accepted.

> ***Result 1****: Financially literate individuals make more rational investment decisions than non-literate individuals.*

| **Variable** | **AME** | **SE** | **P-value** |
|---|---|---|---|
| Round Number | -0.035 | 0.010 | < 0.001 *** |
| Graphical Literacy | 0.030 | 0.017 | 0.081 + |

TABLE 7.7: *Average marginal effects of graphical literacy (H2).*

*H2* hypothesizes that an increasing *graphical literacy* leads to more *rational investment decisions*. Hence, the independent variable of the regression is *graphical literacy* (see appendix B.3). The corresponding average marginal effects (see table 7.7) reveals a significant and positive influence (AME = 0.030, p-value = 0.081). In line with the result of the first study, an increasing *graphical literacy* increases *rational investment decisions*.

> ***Result 2**: Graphically literate individuals make more rational investment decisions than non-literate individuals.*

| **Variable** | **Line Chart** | **AME** | **SE** | **P-value** |
|---|---|---|---|---|
| Round Number | 0 | -0.035 | 0.010 | < 0.001 *** |
| Round Number | 1 | -0.035 | 0.010 | < 0.001 *** |
| Graphical Literacy | 0 | 0.032 | 0.018 | 0.087 + |
| Graphical Literacy | 1 | 0.023 | 0.018 | 0.214 |

TABLE 7.8: *Average marginal effects of graphical literacy and line charts (H3).*

*H3* assumes that *graphically literate* individuals make more *rational investment decisions* with *line charts* than *non-literate* individuals. Hence, the interaction of the treatment variable *line chart* with *graphical literacy* is examined with a regression (see appendix B.4) and the average marginal effects are calculated (see table 7.8). The interaction of *graphical literacy* with *line chart* (line chart = 1) is not significant (AME = 0.023, p-value = 0.214). In contrast to the first study, *graphical literacy* does not influence decision-making with *line charts*. Thus, the null hypothesis of *H3* can not be rejected. At this point, the question arises why the studies produced different results. One obvious and possible answer is the sample difference. A broader sample was used for the first study, while a specific student sample was used for this study. However, in the first study, the largest part of the participants were students. Moreover, in contrast to the first study, this study's participants were incentivized. Both factors can influence decision-making and thus be responsible for the difference in results. This insight is further discussed in section 7.3.4.

> ***Result 3**: Graphical literacy does not influence rational investment decisions with line charts.*

| **Variable** | **Line Chart** | **AME** | **SE** | **P-value** |
|---|---|---|---|---|
| Round Number | 0 | -0.035 | 0.010 | < 0.001 *** |
| Round Number | 1 | -0.035 | 0.010 | < 0.001 *** |
| Financial Literacy | 0 | 0.053 | 0.019 | 0.007 ** |
| Financial Literacy | 1 | 0.088 | 0.018 | < 0.001 *** |

TABLE 7.9: *Average marginal effects of financial literacy and line charts (H4).*

Next, the hypothesis *H4* is investigated, which assumes that *financially literate* individuals make more *rational investment decisions* with financial *line charts* than *non-literate*. A regression with the interaction of *financial literacy* and the dummy variable *line chart* is performed (see appendix B.4). The corresponding average marginal effects (see table 7.9) shows that the interaction of *line chart* with *financial literacy* (AME = 0.088, p-value < 0.001) significantly influences *rational investment decisions*. With increasing *financial literacy*, *rational investment decisions* with *line charts* increases.

***Result 4****: Financial literate individuals make more rational investment decisions with line charts.*

| **Variable** | **Tabular** | **AME** | **SE** | **P-value** |
|---|---|---|---|---|
| Round Number | 0 | -0.034 | 0.010 | < 0.001 *** |
| Round Number | 1 | -0.036 | 0.010 | < 0.001 *** |
| Financial Literacy | 0 | 0.071 | 0.019 | < 0.001 *** |
| Financial Literacy | 1 | 0.052 | 0.018 | 0.006 ** |

TABLE 7.10: *Average marginal effects of financial literacy and tabular (H5).*

*H5* assumes that *financially literate* individuals make more *rational investment decisions* with *tabular* representations than *non-literate* individuals. A logistic regression with the interaction of *financially literacy* with *tabular* is performed (see appendix B.5) and the average marginal effects are calculated (see table 7.10). Financial literacy (AME = 0.052, p-value = 0.006) significantly and positively influences decision-making with tabular representations.

***Result 5****: Financial literate individuals make more rational investment decisions with tabular representations.*

| **Variable** | **AME** | **SE** | **P-value** |
|---|---|---|---|
| Round Number | -0.035 | 0.010 | < 0.001 *** |
| Maximizing | 0.001 | 0.021 | 0.947 |

TABLE 7.11: *Average marginal effects of maximizing (H8).*

Next, *H8*, which assumes that *satisficers* make more *rational investment decisions* than *maximizers*, is investigated with a regression (see appendix B.6). In contrast to the expectation, the average marginal effect (see table 7.11) indicates that *maximizing* (AME =

0.021, p-value = 0.947) does not influence *rational investment decisions*. The null hypothesis of *H8* is not rejected.

***Result 8**: Maximizing does not influence rational investment decisions.*

| **Variable** | **AME** | **SE** | **P-value** |
|---|---|---|---|
| Round Number | -0.035 | 0.010 | < 0.001 *** |
| Conscientiousness | 0.007 | 0.025 | 0.785 |

TABLE 7.12: *Average marginal effects of conscientiousness (H10).*

*H10* hypothesizes that individuals with a *high level* of *conscientiousness* make less *rational investment decisions* than those with *low level*. A regression with the independent variable *conscientiousness* is performed (see appendix B.7) and the average marginal effects are calculated (see table 7.12). *Conscientiousness* (AME = 0.007, p-value = 0.785) does not influence *rational investment decisions*.

***Result 10**: Conscientiousness does not influence rational investment decisions.*

| **Variable** | **AME** | **SE** | **P-value** |
|---|---|---|---|
| Round Number | -0.035 | 0.010 | < 0.001 *** |
| Statistical Literacy | 0.052 | 0.010 | < 0.001 *** |

TABLE 7.13: *Average marginal effects of statistical literacy (H12).*

In the case of *H12*, which assumes that an increasing *statistical literacy* increases *rational investment decisions*, the average marginal effects are calculated (see table 7.13) with the corresponding logistic regression (see appendix B.8). *Rational investment decisions* significantly increase with increasing *statistical literacy* (AME = 0.052, p-value < 0.001).

***Result 12**: An increasing statistical literacy increases rational investment decisions.*

| **Variable** | **Boxplot** | **AME** | **SE** | **P-value** |
|---|---|---|---|---|
| Round Number | 0 | -0.035 | 0.010 | < 0.001 *** |
| Round Number | 1 | -0.033 | 0.010 | < 0.001 *** |
| Graphical Literacy | 0 | 0.037 | 0.017 | 0.027 * |
| Graphical Literacy | 1 | 0.002 | 0.017 | 0.874 |

TABLE 7.14: *Average marginal effects of graphical literacy and boxplot (H5).*

Finally, *H15* assumes that *graphically literate* individuals make more *rational investment decisions* with *boxplots* than *non-literate* individuals. Hence, a logistic regression is performed with the interaction of *boxplot* with *graphical literacy* (see appendix B.9). The average marginal effect (see table 7.14) is not significant (AME = 0.002, p-value = 0.874). Thus, the null hypothesis of *H15* is not rejected.

> ***Result 15****: Graphical literacy does not influence rational investment decisions with boxplots.*

In summary, financial, graphical, and statistical literacy increase rational investment decisions. Moreover, financial literacy improves decision-making with line charts and tabular representations. In contrast to the first study, graphical literacy does not improve decision-making with line charts.

### Analysis of Cognitive Load

In this section, the dependent variable *cognitive load* is investigated. Compared with the first study (mean = 5.0, sd = 1.6), the participants reported a lower cognitive load (mean = 4.2, sd = 1.4). The histogram (see figure 7.15) resembles a normal distribution. The result of a Shapiro-Wilk normality test (W = 0.97924, p-value = 0.045) confirms this.

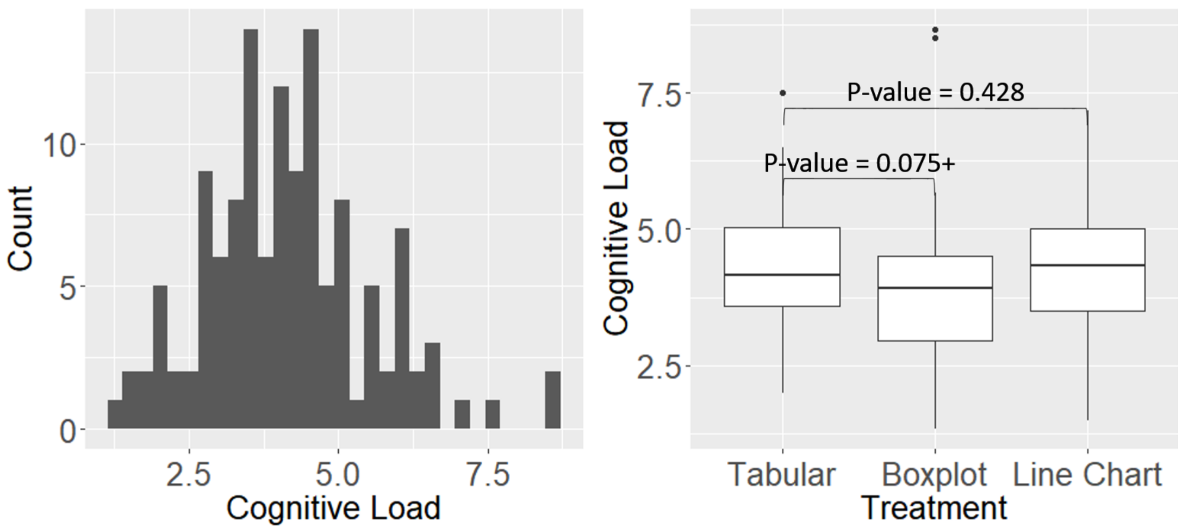



FIGURE 7.15: *Histogram and boxplots of cognitive load.*

*H6* hypothesizes a negative correlation of *cognitive load* with increasing *rational investment decisions*. Therefore, a Pearson's product-moment correlation test is performed (cor = -0.080, t = -0.90919, df = 127, p-value = 0.365) indicating a negative but not significant correlation. Thus, the null hypothesis of *H6* is not rejected.

> ***Result 6**: There is no linear relationship between rational investment decisions and cognitive load.*

*H7* assumes that *line charts* cause a lower *cognitive load* than the *tabular* representations. In the treatment *line chart*, participants reported a mean of 4.2 (sd = 1.3) points, whereas, in the treatment *tabular*, a mean of 4.3 (sd = 1.3) points. As depicted in the boxplots (see figure 7.15), the difference is relatively low. A one-sided Wilcoxon rank-sum test underlines this with a non-significant result (W = 1104, p-value = 0.428). Thus, there is no difference between the representations concerning *cognitive load*.

> ***Result 7**: Line charts do not cause a lower cognitive load than tabular representations.*

| | | **H9** | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | 3.413 | 0.563 | < 0.001 *** |
| Maximizing | 0.014 | 0.010 | 0.171 |
| **$R^2$** | | **0.0147** | |

Number of Participants/Observations: 129; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE 7.15: *Linear regression on cognitive load with maximizing as the independent variable (H9).*

*H9* assumes that *maximizers* have a higher *cognitive load* than *satisficers*. A linear regression is performed with the independent variable *maximizing* (see table 7.15). The result ($\beta$ = 0.014, p-value = 0.171) is not significant..

> ***Result 9**: Maximizing does not influence cognitive load.*

| | H11 | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | 5.117 | 0.598 | < 0.001 *** |
| Conscientiousness | -0.240 | 0.149 | 0.109 |
| **$R^2$** | **0.020** | | |
| Number of Participants/Observations: 129; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 | | | |

TABLE 7.16: *Linear regression on cognitive load with conscientiousness as the independent variable (H11).*

*H11* assumes a positive influence of *conscientiousness* on *cognitive load*. The corresponding linear regression (see table 7.16) shows that there is no significant relationship ($\beta$ = -0.240, p-value = 0.109).

***Result 11**: Conscientiousness does not influence cognitive load.*

| | H13 | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | 4.763 | 0.228 | < 0.001 *** |
| Statistical Literacy | -0.262 | 0.086 | 0.002 ** |
| **$R^2$** | **0.067** | | |
| Number of Participants/Observations: 129; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 | | | |

TABLE 7.17: *Linear regression on cognitive load with statistical literacy as the independent variable (H13).*

*H13* hypothesizes that an increasing *statistical literacy* decreases *cognitive load*. The linear regression analysis (see table 7.17) indicates a significant and negative influence on *cognitive load* ($\beta$ = -0.262, p-value = 0.002). Hence, *cognitive load* decreases with increasing *statistical literacy*.

***Result 13**: An increasing statistical literacy decreases cognitive load.*

The last hypothesis *H14* assumes that *boxplots* cause less *cognitive load* than *tabular* representations. In the treatment *boxplot*, participants reported a mean of 3.9 (sd = 1.6) points, whereas, in the treatment *tabular*, the mean is 4.3 (sd = 1.3) points. Since an increase is expected, a one-sided Wilcoxon rank-sum test is performed. The result (W = 1023, p-value = 0.075) confirms the alternative hypothesis.

***Result 14****: Boxplots cause less cognitive load than the tabular representations.*

In summary, statistical literacy influences cognitive load negatively. Moreover, boxplots cause less cognitive load than tabular representations. The remaining hypotheses are not confirmed.

### 7.3.4 Discussion

Both studies investigated rational investment decisions, and cognitive load as dependent variables and have shown that financial literacy and graphical literacy are relevant factors in financial decision-making. Compared with existing literature, the participants showed a relatively high score on both questionnaires (Lusardi et al., 2010; Okan et al., 2012).

Statistical literacy also increases rational investment decisions. Moreover, it is negatively related to cognitive load. The analysis of the correlations with financial literacy and graphical literacy suggests that the measurements partially reflect general intelligence. Scholars reported that statistical literacy predicts meta-cognitive abilities, underlining this insight (Ghazal et al., 2014; Cokely et al., 2012). Nevertheless, it is shown that the questionnaires can explain additional variance besides intelligence-related factors, such as education (Galesic and Garcia-Retamero, 2011; Lusardi and Mitchell, 2007; Ghazal et al., 2014).

Especially, the statistical literacy questionnaire is relatively short (4 items). Therefore, it could be used in ISs to check whether users are competent in financial decision-making and, if necessary, build their knowledge to ensure well-grounded decisions. Furthermore, potential cognitive limitations could be overcome with the help of adaptive IS interventions such as nudges (see chapter 5).

In the case of maximizing, the working hypotheses with respect to rational investment decisions and cognitive load are not confirmed. In the existing literature, the negative influence of maximizing is shown in the context of consumer decisions (Parker et al., 2007). Maximizing could have a weaker influence in the financial context. Another possible explanation is the goal of the experimental task. Scholars argued that the effects of maximizing occur in association with utility maximizing. In contrast, participants in this study had to maximize the expected return of the investments, which is not necessarily the same thing.

The study does also not confirm the working hypotheses of conscientiousness. The negative effect on decision-making was shown in unforeseen tasks. Consequently, the participants had no routines in decision-making and had to react by developing new ways of doing the task (LePine et al., 2000). In this study, the participants showed a relatively high financial literacy, which indicates that they were already experienced in financial decision-making. Consequently, they might used their existing decision routines, resulting in a weaker or absent effect.

This study investigated visualization recommendations based on user characteristics. It has shown that financial literacy improves decision-making with tabular representations and line charts. This insight confirms the assumption that both representations are common in the financial domain. Therefore, future ISs could recommend the aforementioned visualizations as soon as a certain level of financial literacy is present. Moreover, users' financial literacy could be improved in front of the decision-making with line charts or tabular representations.

There are discrepancies between the results of the conducted studies. In the first study, rational decisions are negatively related to cognitive load. This could not be replicated in the main study. Moreover, in contrast to the first study, graphical literacy does not improve decision-making with line charts. One possible explanation is the change in the sample. A broad sample of the population participated in the first study, whereas a specific student sample was used for this study. Nevertheless, the deviations were not expected since the largest part of the first study's sample was also students. Hence, a further study with a sample of non-students is needed to finally clarify the question.

Another possible reason for the discrepancy of the results could be the incentive difference. In the first study, the participants were not incentivized in order to obtain high variance with a relatively small sample size. In this study, if all investment scenarios were answered correctly, the participants took part in a cash raffle. Performance-based payouts were deliberately omitted in order not to encourage participants to make risky investments. Participants in this study may have made more effort to enter the raffle. Camerer and Hogarth (1999) reviewed the extant literature to investigate the effects of incentives in experimental studies. The authors found out that incentives increase performance in effort-related tasks. In contrast, incentives predominantly reduced the variance of participants' answers in choice tasks (Camerer and Hogarth, 1999).

Both studies show that none of the investigated representation variants can be preferred across the board, which underlines the necessity of this research. Moreover, despite different samples and different incentivization, the results of both studies regarding financial literacy are as expected. This insight shows that financial literacy is a robust user characteristic and can be used by a wide range of the population to assess decision-making competency and to recommend particular visualizations.

In general, researchers explained behavioral differences mainly with user characteristics such as personality and experience (Rauthmann et al., 2015). However, as in this thesis, studies with the same research subject and hypotheses came up with deviating results that could not be sufficiently explained (Zhang et al., 2014; Alós-Ferrer et al., 2016; Jung et al., 2019). This raises the question of whether situational aspects should be taken into account in IS design research as a matter of principle.

Consideration of context and situation is not an innovation in IS design. With the upcoming of mobile devices and in the course of pervasive computing, scholars have already investigated situation-aware ISs (Selker and Burleson, 2000; Rothrock et al., 2002). However, this has been limited to objective situational cues such as device (size or type) or environmental aspects (temperature, light, or location).

Rauthmann et al. (2014) have done pioneering work with their research by finding out that people perceive situations on the basis of specific psychologically relevant situational characteristics (see chapter 3). These are able to capture inter- and intra-individual differences of situations. Moreover, personality traits can be predicted, and unique variance in behavior prediction can be explained.

In view of inconsistent study results, the situational characteristics could be used to capture perceived situational experience. These are capable of identifying differences in sample and incentive. For example, when people voluntarily participate in a study, they are not paid for it and may not feel their participation as mandatory. The situational characteristics duty and positivity can capture their situational perception since duty measures whether something needs to be done, and positivity indicates whether a situation is financially rewarding (see chapter 3.4). Hence, the measurement of situational characteristics can make a valuable contribution by recording under which situational circumstances the generated findings have their validity. There are already questionnaires that have been constructed exactly for this task and involve a relatively low measurement effort (Rauthmann and Sherman, 2015a,b). In chapter 8, a study is presented that evaluates the measurement of situational characteristics to situationally classify the perception of experimental stimuli.

# Part IV

# Situation-Aware IS Design

# Chapter 8

# Situational Decision Inertia & Nudging

> "Probably 90 percent of our life decisions are powered by the twin engines of inertia and laziness."
>
> ARNOLD STEPHEN JACOBS JR., 2009

THIS chapter reports a study which investigates decision inertia and its situational dependencies across situational contexts. Furthermore, using the example of decision inertia, the study examines whether the effectiveness of nudging depends on the situational context.

## 8.1 Introduction

Decision-makers tend to rely on their intuition in decision situations with overwhelming complexity or uncertainty. They make use of mainly unconscious cognitive shortcuts (heuristics) to come up with a decision (Kahneman et al., 1982). However, in some cases, heuristics can lead to non-rational choices. These are systematic deviations from rationality and are called cognitive biases (Kahneman, 2003).

Decision inertia is one of these biases. It describes the inability or unwillingness to change a previously made decision despite negative consequences (Jung et al., 2019; Alós-Ferrer et al., 2016). It is a ubiquitous phenomenon and relevant to the design of ISs (Jung et al., 2018). For instance, a considerable amount of people repeated suboptimal financial decisions. This behavior can be tackled with the use of an appropriate IS design (Jung et al., 2018; Jung and Weinhardt, 2018).

Behavior and thus, intuitive decision-making can vary in different situations (Funder, 2009). Also, the self-esteem and affective state of an individual depends on the situation and have an impact on behavior (Geukes et al., 2017). Regarding decision inertia, yet it is unclear whether the tendency to rely on it is stable across situational contexts. Antecedents of decision inertia might be influenced by surrounding context resulting in differences in its extent (Jung et al., 2019; Jung, 2019).

IS research deals with the challenge of designing systems, which support their users as best as possible in their decision-making process. One aspect of it is recognizing and reducing systematic errors resulting from intuitive decision heuristics (Barber and Odean, 2000; Jung et al., 2018). The choice architecture provides tools in the form of nudges as interventions to alter the behavior in biased decisions (Weinmann et al., 2016). However, findings indicate that their effectiveness might vary across situational contexts (Hummel and Maedche, 2019; Lehner et al., 2016).

Psychological research has long postulated that behavior is determined by the situation and individual's characteristics (Lewin, 1936). However, relatively recently, there has been a paradigm shift from using objective cues of situations to describe them to looking at perceived characteristics of a situation (Rauthmann et al., 2014; Ziegler et al., 2019; Parrigon et al., 2017). This makes sense because not all objective cues are perceived equally by all individuals and they interpret perceived cues differently, depending on their subjective relevance. In the process, taxonomies of psychologically relevant situational characteristics have emerged that facilitate the measurement of situations and thus, the identification of inter- and intra-individual differences (see chapter 3). These are a essential part of this research, enabling this study to be realized.

The contribution of this study is twofold. Firstly, the situational dependency of decision inertia has not been empirically shown so far. Thus, it is unclear whether the tendency to rely on decision inertia varies across situational contexts.

**Research Question 3:** *How do situational characteristics influence decision inertia?*

Secondly, nudges are an effective way to tackle biases, such as decision inertia. However, their effectiveness might also depend on the situational context (Hummel and Maedche, 2019). So far, the combination of decision inertia and nudging has only been studied in the context of financial decision-making – Particular nudges reduce decision inertia (Jung et al., 2018; Jung and Weinhardt, 2018; Jung, 2019). There are also studies which investigated nudging with similar phenomena. Handel (2013) reduced general inertia, which occurred in health insurance choices. Stryja et al. (2017) have shown that status quo bias can be changed in favor of electric cars in the context of rental car decisions. Huber et al. (2019) altered the charging behavior of electric car owners to to avoid load peaks in the energy infrastructure. However, the mentioned studies are hardly comparable because different experimental tasks were used due to the different phenomena. Consequently, the context-dependent effectiveness of nudging is not empirically shown yet.

**Research Question 4:** *How do situational characteristics influence the effectiveness of nudges to reduce decision inertia?*

The remainder of this chapter proceeds as follows: Following the research questions, section 8.2 presents the investigated research model. Section 8.3 introduces the methodology of the study. In section 8.4, the results of the study are analyzed. Finally, section 8.5 discusses the contributions and limitations.

## 8.2 Research Model

Decision inertia is defined as "the tendency to repeat a previous choice, regardless of its outcome, in a subsequent decision" (Alós-Ferrer et al., 2016). The dual-choice paradigm, which is an established way of measuring decision inertia (Alós-Ferrer et al., 2016; Jung et al., 2019; Charness and Levin, 2005), provides a binary decision that is repeated in a subsequent decision. If the first choice resulted in a suboptimal outcome and it is repeated in the subsequent decision, decision inertia occurs. Thus, the dependent variable of the study is the suboptimal choice repetition.

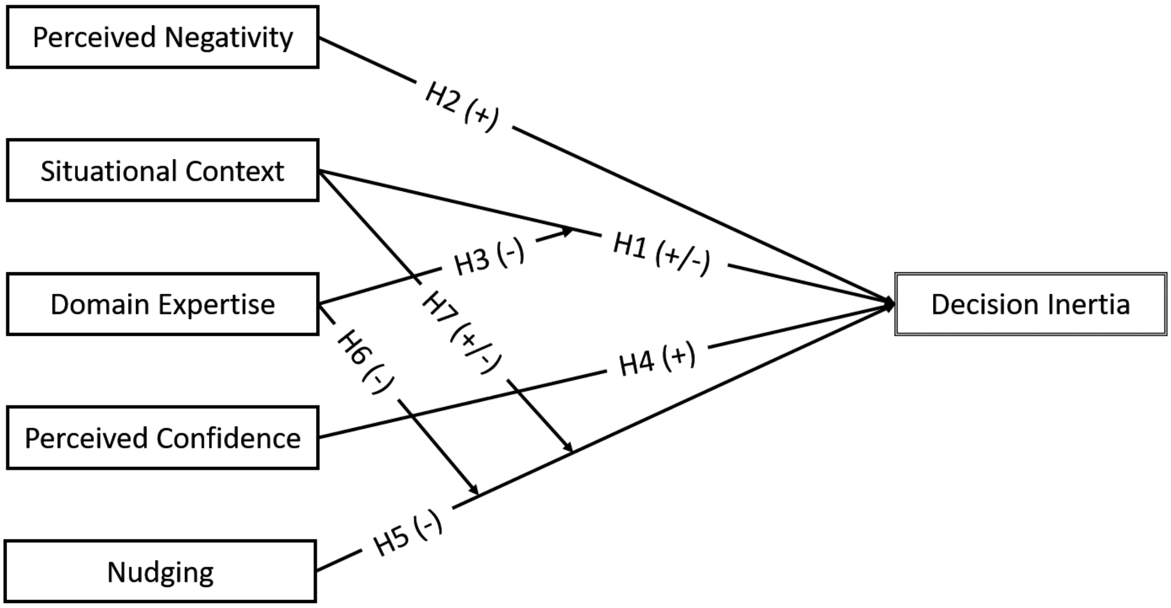



FIGURE 8.1: *Research model of the study*

This study investigates the influence of different situational contexts on decision inertia. As mentioned in chapter 3, situations can be measured with situational characteristics (Rauthmann et al., 2014). Situations with similar characteristics profiles are summarized by contexts (Rauthmann and Sherman, 2015a). In order to measure decision inertia across contexts, the dual-choice paradigm is transferred to different contexts, which are differing significantly in their situational characteristics profiles (see appendix C). In the following, the research questions are taken up to derive the working hypotheses.

**Research Question 3:** *How do situational characteristics influence decision inertia?*

Different framings of decision situations influence the valuation of outcomes (Kahneman and Tversky, 2013) and the tendency to rely on intuitive processes (Alós-Ferrer et al., 2016). Preferences, affective state, and self-esteem vary across situational contexts and crucially influence behavior (Rauthmann et al., 2014; Geukes et al., 2017; Halevy et al., 2019; Lichtenstein and Slovic, 2006). However, the influence of situational characteristics on decision inertia is empirically unexplored. This study assumes that situations with significant differences in their perceived characteristics profiles (contexts) lead to different decision inertia.

***H1**: Decision inertia varies across situational contexts.*

Valence is a crucial part of situational experience (Halevy et al., 2019) and determine the quality of decision-making (Bechara et al., 1997; Spence, 1995). The situational characteristics *positivity* and *negativity* capture the positive and negative aspects of situations (Rauthmann and Sherman, 2018). It is already shown that negative feedback lowers the valuation of previously committed choices (Jermias, 2001). It is also assumed that situationally experienced negativity reduces the flexibility in decision-making (Baumann and Kuhl, 2005) and thus, it reduces the change of a previously committed choice. In the case of new software adoption, scholars shown that the reduction of decision flexibility leads to loss aversion and therefore, resistance towards advantageous new software (Li et al., 2016). Consequently, this study proposes that situational contexts inducing high levels of perceived *negativity* increase decision inertia.

***H2**: Situations inducing high levels of perceived negativity increase decision inertia.*

The following hypothesis deals with context-specific knowledge, namely *domain expertise*. In general, high levels of *domain expertise* lead to more effective decision-making (Kahneman and Klein, 2009; Dane et al., 2012). Jung (2019) investigated decision inertia in the context of robo-advisory and showed that financially literate individuals are less prone to it.

Additional evidence comes from resilience research. Emotional intelligence, which means the ability to handle emotional situations, increases with situational experience (Shipley et al., 2010). Conveyed to decision inertia, experienced individuals could better deal with perceived negativity. Thus, this study proposes that higher levels of *domain expertise* improve the understanding of the decision situations and reduce decision inertia.

***H3**: Decision inertia decreases with increasing domain expertise.*

The next hypothesis deals with the confidence of an individual. The literature distinguishes between three basic types of confidence. Overconfident individuals judge their abilities reliably greater than the objective reality (Moore and Schatz, 2017). In contrast, underconfident individuals judge their abilities significantly lower than the objective reality (Moore and Schatz, 2017). Finally, realistic individuals tend to rate their abilities according to objective reality (Moore and Schatz, 2017). An individual's confidence varies across situational contexts (Sahin and Yilmaz, 2014). The variation emerges due to the kind of feedback experienced in situations. If an individual gets positive feedback because of successfully made decisions, he or she will be more confident in his or her next decision. However, if the individual experiences regret, his or her confidence will be lower in the next decision.

The effect of confidence on decision inertia could be twofold. *Underconfident* people tend to experience negative feelings because they fear being worse than others (Moore and Schatz, 2017). Also, *overconfident* people tend to experience negative feelings due to their unattainable goals and claims (Moore and Schatz, 2017). Both *underconfidence* and *overconfidence* can lead to emotional overreactions. Therefore, this study proposes that these overreactions increase decision inertia by triggering intuitive reactions and consequently, the use of the repetition heuristic instead of rational thinking (Kahneman et al., 1982).

> ***H4O****: Overconfidence increases decision inertia.*

> ***H4U****: Underconfidence increases decision inertia.*

Besides the investigation of decision inertia across contexts, this study aims to situation-aware reduce decision inertia with nudging:

> **Research Question 4:** *How do situational characteristics influence the effectiveness of nudges to reduce decision inertia?*

Nudging is used to improve decision-making by addressing cognitive biases like decision inertia (Thaler and Sunstein, 2003). Appropriate nudges improve inertia-related behavior, such as general inertia or status-quo bias (Handel, 2013; Stryja et al., 2017; Huber et al., 2019).

So far, only one study explicitly investigated decision inertia (Jung et al., 2018; Jung and Weinhardt, 2018; Jung, 2019). The study's authors have shown that defaults and warnings reduce decision inertia in the financial context. It is unclear whether the nudges reduce decision inertia in other situational contexts. This study aims to shed light by adopting the hypotheses.

In order to avoid cognitive effort, people tend to rely on the status-quo bias (Johnson and Goldstein, 2003). Pre-selected choice alternatives are usually chosen more often (Johnson and Goldstein, 2003). Consequently, setting optimal defaults is a promising way to alter inertia-related problems' behavior. Thereby, default options serve as a counter-bias. Hence, it is proposed that defaults reduce decision inertia.

> ***H5D****: Defaults reduce decision inertia.*

Based on the cognitive feedback theory (Balzer et al., 1989), warnings can be used to alter irrational behavior. They provide information about the current decision-making to discourage suboptimal decisions by reconsidering the decision situation (Bhandari et al., 2008; Jung et al., 2018; Jung and Weinhardt, 2018). In the case of decision inertia, warnings provide insight into the unconscious repetition. This gives decision-makers the opportunity to consciously deliberate their intuitive decision and change their choice.

> ***H5W****: Warnings reduce decision inertia.*

Moreover, the moderation effect of domain expertise on the effectiveness of nudging is investigated. In the context of social media, it could be shown that nudging of experts are more successful than the nudging of novice decision-makers (Nekmat, 2020). Nekmat (2020) assumed that domain experts are more likely to trust the nudges. Therefore, this study hypothesizes that an increasing level of domain expertise increases the effectiveness of defaults and warnings to reduce decision inertia.

> ***H6D****: A higher level of domain expertise increases the effectiveness of defaults to reduce decision inertia.*

> ***H6W****: A higher level of domain expertise increases the effectiveness of warnings to reduce decision inertia.*

Next, the context-dependent effectiveness of nudging is discussed. There are indications in the literature that the effectiveness of nudges varies across situational contexts. Hummel and Maedche (2019) reviewed the extant literature on nudging and found large domain-specific differences in the mean effect sizes of nudges. Furthermore, Lehner et al. (2016) investigated the effectiveness of Swedish policy interventions and noticed that the outcomes of the same interventions vary across the domains of energy, food, and mobility.

There is no research on whether and how nudges could be influenced by perceived situational characteristics. Nevertheless, there are investigations of emotional reactions triggered by specific nudges (Stryja et al., 2017; Zhang and Xu, 2016). However, with mixed results: In the context of novel technologies, Stryja et al. (2017) could not find any association, whereas Zhang and Xu (2016) could show that specific nudges influence anxiety and fear.

Due to limited research on nudging and situational characteristics, it is refrained from making predictions for perceived situational characteristics. Thus, it is proposed that the effectiveness of warnings and defaults in reducing decision inertia varies across situational contexts. It is a good starting point for further research.

> ***H7D****: The effectiveness of defaults in reducing decision inertia varies across situational contexts.*

> ***H7W****: The effectiveness of warnings in reducing decision inertia varies across situational contexts.*

## 8.3 Methodology

This section introduces the methodology of the study, containing the experimental design, decision task, selection of nudges, measures, and the procedure of the study.

### Experimental Design

The study investigates two treatment factors, namely *situational context* and *nudge*. In the run-up to this study, different situational contexts of the urn game, an established experimental task for investigating decision inertia (see section 8.3), were designed and identified. In the following, the situational contexts are briefly explained. The detailed description is reported in the appendix C.

- **Urn Game (Urn):** There are two different urns with different proportion of white and black balls. Participants have to draw black balls.

- **Robo-advisor (Robo):** There are two different portfolios with different proportion of stocks and bonds. Based on the market situation, which favors either stocks or bonds, participants have to choose the portfolio that leads to profit.
- **Dating Game (Dating):** There are two different chat partners with different preferences. Participants have to choose between two different mix of topics. If the chat partner likes the selected topic, the participants get a reward.
- **Exam Game (Exam):** There are two different types of exams. Participants have to decide between two learning strategies. If the selected strategy leads to the pass of the exam, the participants get a reward.

Besides the situational context *urn* as baseline, three further situational contexts, namely *robo, dating,* and *exam*, are used. The *robo* is adapted from Jung et al. (2018) in order to replicate previous results. The remaining situational contexts were designed and identified for the purpose of this study (see appendix C). In line with the literature (Rauthmann and Sherman, 2015a), this study's situational contexts significantly differ with respect to the profile of situational characteristics.

In the case of the treatment factor *nudge*, there is a *baseline* without nudging. In addition, *warnings* and *defaults* are used to reduce decision inertia.

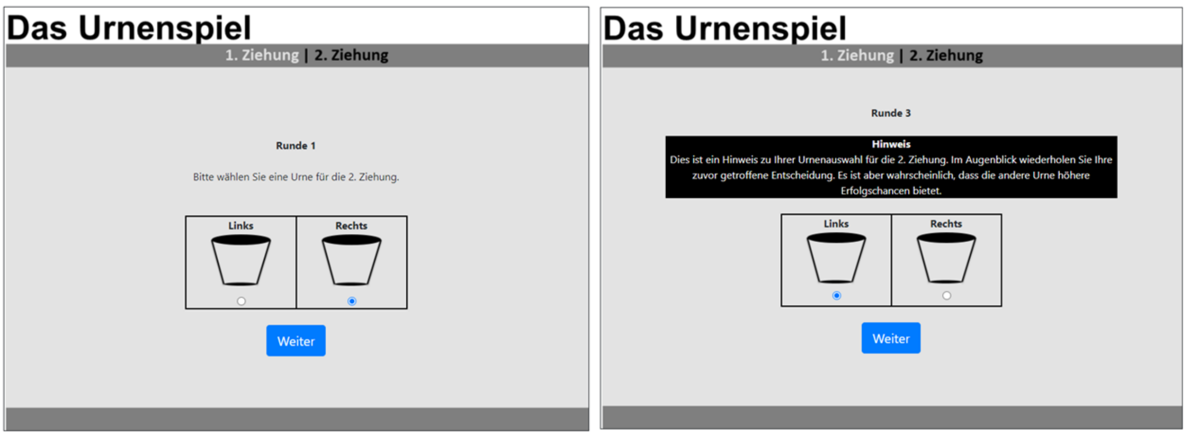



FIGURE 8.2: *The implemented defaults and warnings.*

The wording of the warning nudge is adopted from Jung et al. (2018). When decision inertia occurs, the decision-maker can revise his or her decision with the following displayed message (see figure 8.2): "This is a hint to your choice of urn for the 2nd draw. At the moment, you repeat the decision you made before. However, it is likely that the other urn offers a higher chance of success."[1]

In the case of default nudge, the optimal second choice is pre-selected (see figure 8.2). The optimal second choice maximizes the expected probability for a black ball and is calculated based on Bayesian updating.

| | Urn | Robo | Dating | Exam |
|---|---|---|---|---|
| Baseline | T1 | T2 | T3 | T4 |
| Warning | T5 | T6 | T7 | T8 |
| Default | T9 | T10 | T11 | T12 |

TABLE 8.1: *Experimental design of the study with the between-subject treatments nudge and situational context.*

An online experiment with a 4 x 3 between-subject design (see table 8.1) is chosen following experimental economics (Friedman et al., 1994). There is no single measurement tool for the domain expertise of the situational contexts. Hence, appropriate questionnaires are identified, and corresponding sub-hypotheses are formulated:

> ***H3U****: Higher levels of probabilistic reasoning decrease decision inertia in the urn context.*
>
> ***H3R****: Higher levels of financial literacy decrease decision inertia in the robo context.*
>
> ***H3D****: Higher levels of interpersonal competence decrease decision inertia in the dating context.*
>
> ***H3E****: Higher levels of learning competence decrease decision inertia in the exam context.*
>
> ***H6DU****: A higher level of probabilistic reasoning ability increases the effectiveness of defaults in the urn context.*
>
> ***H6DR****: A higher level of financial literacy increases the effectiveness of defaults in the robo context.*

[1]Translated from the original german version: "Dies ist ein Hinweis zu ihrer Urnenauswahl für die 2. Ziehung. Im Augenblick wiederholen sie ihre zuvor getroffene Entscheidung. Es ist aber wahrscheinlich, dass die andere Urne höhere Erfolgschancen bietet."

***H6DD***: *A higher level of interpersonal competence increases the effectiveness of defaults in the dating context.*

***H6DE***: *A higher levels of learning competence increases the effectiveness of defaults in the exam context.*

***H6WU***: *A higher level of probabilistic reasoning ability increases the effectiveness of warnings in the urn context.*

***H6WR***: *A higher level of financial literacy increases the effectiveness of warnings in the robo context.*

***H6WD***: *A higher level of interpersonal competence increases the effectiveness of warnings in the dating context.*

***H6WE***: *A higher level of learning competence increases the effectiveness of warnings in the exam context.*

## Decision Task

| State | Left urn | Right urn |
|---|---|---|
| Up (50%) | ●●●●○○ | ●●○○○○ |
| Down (50%) | ●●○○○○ | ●●●●○○ |

| Case | First decision | Inference | Bayesian updating | Decision Inertia |
|---|---|---|---|---|
| 1 | Left urn: Black | Up | Stay with left urn | - |
| 2 | Left urn: White | Down | Shift to right urn | Stay with left urn |
| 3 | Right urn: Black | Down | Stay with right urn | - |
| 4 | Right urn: White | Up | Shift to left urn | Stay with right urn |

FIGURE 8.3: *Urn composition and decision cases of the urn game.*

The decision task is the urn game, which follows the dual-choice paradigm (Charness and Levin, 2005; Alós-Ferrer et al., 2016; Jung et al., 2019; Jung and Dorner, 2018). Accordingly. there are two decision alternatives – the left and the right urn. Both urns contain six balls. The balls can be black or white. As depicted in figure 8.3, the composition varies according to the state, namely up and down. The participants have to draw black balls and are informed about the urn composition of each state, the prior probability of each state ($p = 1/2$), and that a state is held constant for two subsequent draws (two-draw decision). Based on the prior probabilities, a state is randomly chosen for the two-draw decision but is not revealed to the participants. This means that the first of the two draws is randomly made. After observing the color of the first draw, participants have to choose an urn for the second draw. By the observation of the first draw's color, the participants can conclude about the likeliness of the current state. Consequently, participants can optimize their decision for the second draw by choosing the urn with the highest expected payoff. In the following, the calculation of the expected payoff is illustrated with an example:

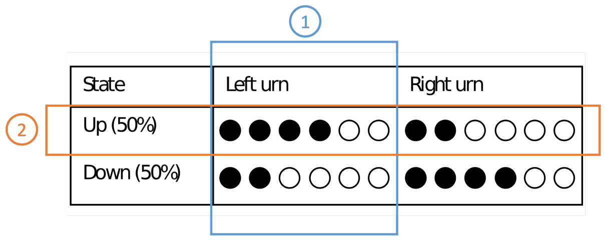



FIGURE 8.4: *Illustrative example of the urn game calculation.*

1. Assuming a black ball is drawn from the left urn. The probability of being in the state *up* is higher than state *down*, since the left urn contains more black balls in the state *up* (see blue box figure 8.4). The probability can also be calculated:

$$P(up) = \frac{\frac{1}{2}x\frac{4}{6}}{\frac{1}{2}x\frac{4}{6} + \frac{1}{2}x\frac{2}{6}} = \frac{2}{3}$$

2. In the next step, staying with the left urn results in a higher probability to draw a further black ball, since the left urn contains more black balls in the state *up* (see orange box figure 8.4). The corresponding probabilities can be calculated using Bayesian updating:

$$P(black \mid left\ urn) = \frac{1}{3}x\frac{2}{6} + \frac{2}{3}x\frac{4}{6} = \frac{5}{9}$$

$$P(black \mid right\ urn) = \frac{1}{3}x\frac{4}{6} + \frac{2}{3}x\frac{2}{6} = \frac{4}{9}$$

The described decision situation is repeated 60 times. The probability to draw a black ball in the first decision equals 50 %. Thus, decision inertia can occur about 30 times.

In the urn game, four possible decision cases arise (see figure 8.3). The divergent and convergent condition summarizes them. The convergent condition (case 1 and case 3) emerges when the first draw is rewarded. Thereby, the choice repetition equals the urn that maximizes the expected payoff. Consequently, decision inertia can not occur, so shifting to the other urn is not optimal. The divergent condition (case 2 and case 4) emerges when the first draw is not rewarded. In this condition, the choice repetition differs from the urn that maximizes the expected payoff. Hence, decision inertia occurs by staying with the urn from the first draw.


FIGURE 8.5: *The situational contexts implemented with Otree (Chen et al., 2016).*

In the appendix C, the design and identification of the study's situational contexts are described. A separate online experiment is implemented with the Otree framework for each situational context (see figure 8.5) (Chen et al., 2016).

## Identification of Nudges

As a reminder, nudges are assumed to vary in their effectiveness depending on the situational context. In order to obtain a high variability across the investigated contexts, nudges are required that are highly likely to lead to a change in decision-making behavior. In the following, the five steps development process of digital nudges (see figure 8.6) is applied (Weinmann et al., 2016).

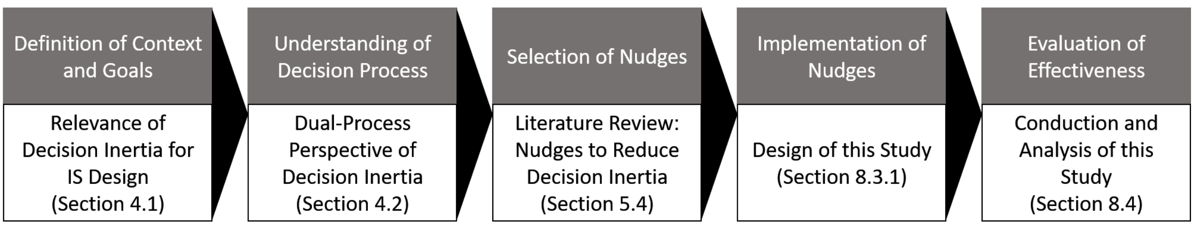



FIGURE 8.6: *The development process of digital nudges to reduce decision inertia (Weinmann et al., 2016)*

The process commences with the definition of context and goals. For this purpose, the relevance of decision inertia in ISs is outlined in section 4.1. The dual-process perspective of decision inertia, which describes the underlying understanding of decision inertia, is described in section 4.2. In the third step, the nudging literature on reducing decision inertia or similar phenomena is reviewed in section 5.4. As a result, defaults, warnings, framings, and social norms are identified as potential interventions. The present study investigates two nudging treatments. Therefore, the two most effective nudges among the identified are used. Hummel and Maedche (2019) conducted a quantitative review of the relative effect sizes of nudges. Accordingly, warnings are the most effective (average relative effect size 107 %), followed by defaults (average relative effect size 87 %). In addition, the theory-based justification for the selection of the two nudges is described in the research model section (8.2). The fourth step is the design of this study (see section 8.3). The process is finalized by the analysis of the study results (see section 8.4)

## Measures

Decision inertia occurs by repeating a previous non-rational choice (Alós-Ferrer et al., 2016; Jung et al., 2019). The dual-choice paradigm is used to measure decision inertia (Charness and Levin, 2005). Accordingly, there is a round-based binary decision, which is repeated. Thus, the dependent variable is a boolean and reflects whether decision inertia occurred (1) or not (0).

The study's independent variables are the categorical variable for the situational contexts (*urn, robo, dating*, and *exam*), categorical variable for the nudges (*baseline, warning*, and *default*), psychologically relevant situational characteristics (DIAMONDS), confidence, and domain expertise (***urn:*** *probabilistic reasoning scale,* ***robo:*** *financial literacy,* ***dating:*** *interpersonal competence*, and ***exam:*** *learning competence*).

The DIAMONDS taxonomy assesses the psychologically relevant situational characteristics with the eight dimensions *duty, intellect, adversity, mating, positivity, negativity, deception,* and *sociality* (Rauthmann et al., 2014). Thereby, the S8-II questionnaire is used (Rauthmann and Sherman, 2015b). It is recommended for the validation of experimental stimuli. There is one item for each dimension, resulting in eight items. The participants rate each item on a 7-point Likert scale ranging from "not at all" to "totally agree".

The true confidence is not measurable since the experimental task is unknown to the participants. Hence, this study assesses the *perceived confidence* with the better-than-average test (Svenson, 1981; Larrick et al., 2007). It is a robust correlate of true confidence (Murphy et al., 2018). The participants have to compare their expected performance with the anticipated performance of the other participants. For this purpose, a slider is used that ranges from the worst (-50) to the best (+50), with the average (0) representing the center of the scale.

Scholars used the perceived confidence measurement after the experimental task to investigate its influence on decision-making (Svenson, 1981; Larrick et al., 2007). However, the causality is questionable since the experimental task affected the outcome of the measurement. In contrast, this study measures before (pre) and after (post) the experimental task. It assumes that the measurement after the experimental task corresponds to the realistic view of participants' abilities. Consequently, analog to the calculation of true confidence, the difference between pre and post task measurements indicates participants' *perceived confidence*.

For the domain expertise of the urn context, the probabilistic reasoning scale is used. It is knowledge based and consists of nine items. Each item is a computational task, capturing statistical literacy or probabilistic reasoning abilities (Primi et al., 2019).

For the domain expertise of the dating context, the interpersonal competence questionnaire measures a person's competence in amicable and romantic relationships (Riemann and Allgöwer, 1993). It consists of 40 items. Each item describes a behavioral situation with other individuals. The participants rate their attitude towards the situation on a 5-point Likert scale ranging from -2 (I would feel so uncomfortable and unable to handle this situation) to +2 (I would feel very comfortable and could handle this situation very well).

In the case of the exam context, the learning competence scale is used. It investigates the learning behavior of students in academic institutions (Villardón-Gallego et al., 2013). The questionnaire consists of 17 items. Each item is a statement about a learning behavior. The participants rate their commitment to the statement on a 5-point Likert scale ranging from -2 (strongly disagree) to +2 (strongly agree).

A financial literacy questionnaire is used to capture the domain expertise of the robo context (Lusardi and Mitchell, 2007). The questionnaire is knowledge-based and consists of three questions concerning compound interest and funds comprehension and two questions for financial retirement planning.

As control variable, the action and state orientation of an individual is measured, since action-oriented individuals are more prone to decision inertia (Jung et al., 2019). The corresponding questionnaire consists of 24 items. Each item describes a state- or action-related situation. The participants rate each item on a 7-point Likert scale ranging from "not at all" to "totally agree" (Kuhl, 1994).

The demographic questionnaire queries gender, age, and field of study.

## Experimental Procedure

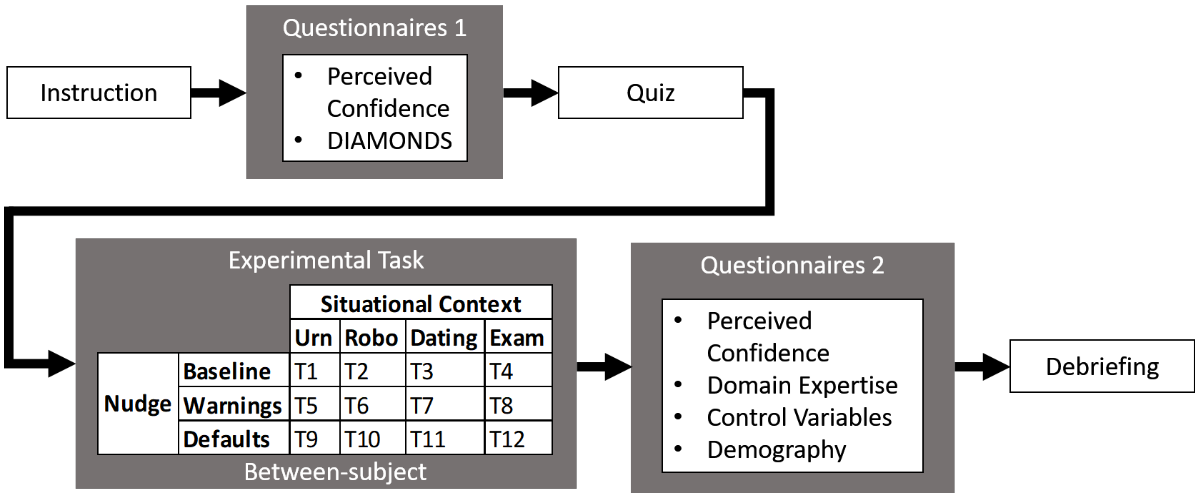



FIGURE 8.7: *Procedure of the study.*

The working hypotheses are addressed with the following experimental procedure (see figure 8.7). In the run-up to the experiment, the participants are randomly allocated to one of twelve treatments. The experiment commences with instructions (Instructions). Participants are informed about the general rules, privacy policy, the tasks' course, and payout. After the general instructions, the instruction differs according to the randomly assigned situational context. The participants are introduced to the randomly assigned situational context and briefed about the rules of the experimental task. Afterward, the participants work on two questionnaires (Questionnaires 1). They rate the perceived situational characteristics (DIAMONDS) and state their perceived confidence about future performance in the experimental task (Perceived Confidence). In the next step (Quiz), participants answer a series of comprehension questions to ensure they understand the experimental task's instructions and rules clearly. The quiz contains nine different multiple-choice questions. If a participant does not correctly answer a question, the question is repeated after giving an explanation for failure. Then, the participants work on the experimental task (Experimental Task). After the experimental task, the participants work on further questionnaires (Questionnaires 2). First, they state their perceived confidence about their past performance in the experimental task (Perceived Confidence). Afterward, domain expertise (Domain expertise) is queried. Then, control variables (Control variables) and demography (Demography) are assessed.

### Participants

The study is conducted with the sample of KD$^2$Lab. It mainly consists of German students from KIT. The participants are acquired via Hroot (Bock et al., 2014). In accordance with the induced value theory (Smith, 1976), participants receive a performance-based payment of 0.10 euro for each successful decision in the experimental task and a flat fee of 1.5 euro for answering the questionnaires. The experimental task consists of 60 rounds with two decisions. Thus, 120 decisions have to be made. The 60 initial decisions are made randomly. Therefore, it is assumed that half of them is successful. The expected payoff for the second decisions is 3.33 euros ($p = 0.56$)[2]. Hence, a total payment of 7.83 (3 + 3.33 + 1.5) euro is expected.

## 8.4 Results

A total of 354 persons participated in the main study. After data cleaning with outliers (processing time $< 10$ minutes, $n = 13$), attention ($n = 12$), and manipulation checks ($n = 4$), 324 complete data sets remained. 207 males and 117 females participated the study. The participants were at the mean 24.4 ($sd = 4.0$) years old, and the experiment lasted on average 33.0 ($sd = 13.9$) minutes. They earned on average 7.67 ($sd = 0.52$) euro.

[2]The calculation of the probability is based on the decision inertia rate of previous studies (Jung et al., 2019; Jung, 2019; Alós-Ferrer et al., 2016)

## Analysis of Treatment Differences

| | Urn | Robo | Dating | Exam |
|---|---|---|---|---|
| Baseline | 26 | 25 | 27 | 28 |
| Warning | 29 | 28 | 27 | 25 |
| Default | 27 | 25 | 30 | 27 |

TABLE 8.2: *Allocation of the participants across the treatments.*

As illustrated in table 8.2, in each treatment, there is a minimum of 25 and a maximum of 30 observations. The treatment differences are examined with the variables *action-orientation, age, gender, field of study,* and *probabilistic reasoning ability*. First, the variables are investigated with respect to their distribution. Then, appropriate tests are used to check the treatment differences.

| **Variable** | **Mean (SD) / Mode** | **Normally Distributed** | **Treatment Difference** |
|---|---|---|---|
| Age | 24.4 (4.0) | No (Shapiro-Wilk normality test: W = 0.91052, P-value < 0.001***) | No (Kruskal-Wallis rank-sum test: chi-squared = 5.9548, P-value = 0.899) |
| Gender | Male = 207 | - | No (Pearson's chi-squared test: X-squared = 15.781, df = 11, p-value = 0.150) |
| Field of Study | Business Engineer = 113 | - | No (Pearson's chi-squared test: X-squared = 13.353, df = 11, p-value = 0.270) |
| Action-orientation | 10.9 (4.6) | No (Shapiro-Wilk Normality test: W = 0.98787, P-value = 0.008**) | No (Kruskal-Wallis rank-sum test: chi-squared = 13.181, P-value = 0.281) |
| Probabilistic Reasoning | 8.3 (1.3) | No (Shapiro-Wilk Normality test: W = 0.6263, P-value < 0.001***) | No (Kruskal-Wallis rank-sum test: chi-squared = 11.12, P-value = 0.433) |

TABLE 8.3: *Analysis of treatment differences.*

The following procedure is used for the analysis. In the case of continuous variables (age, action-orientation, and probabilistic reasoning), The distribution is plotted (see figure 8.8) and tested with a Shapiro-Wilk normality test. Depending on the outcome, parametric (ANOVA) or non-parametric (Kruskal-Wallis rank-sum test) statistical methods are used for the analysis of treatment differences. For nominal variables (gender and field

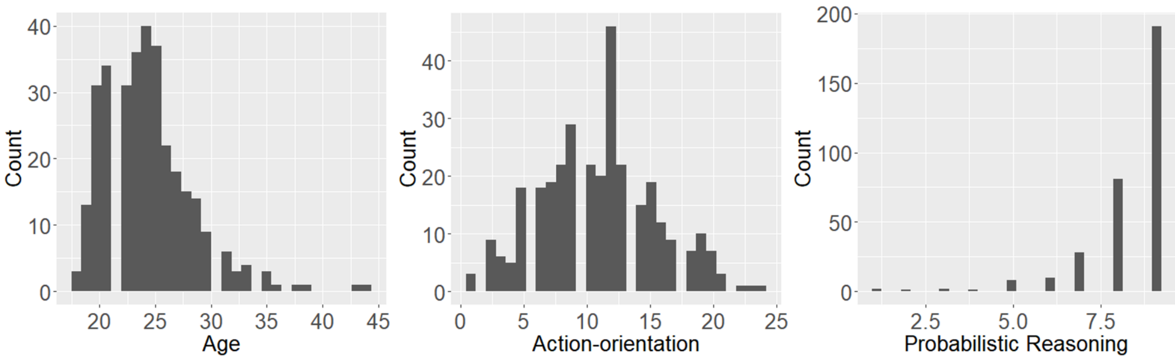



FIGURE 8.8: *Histograms of the variables age, action-orientation, and probabilistic reasoning.*

of study), the modal value is reported and the treatment difference is examined with a Pearson's chi-squared test.

The analysis of the treatment differences are reported in table 8.3. The continuous variables (age, action-orientation, and probabilistic reasoning) are not normally distributed (Shapiro-Wilk normality test, p-values $< 0.1$). Thus, non-parametric Kruskal-Wallis rank-sum tests are performed to investigate their treatment differences. There are no significant treatment differences, since all p-values are greater than 0.1.

As depicted in the histogram of probabilistic reasoning score (see figure 8.8), most participants answered all questionnaire items correctly. Compared with the questionnaire validation study (Primi et al., 2019), this study's participants have a mean of 8.3 (sd $=$ 1.3) correct answers, whereas, in the validation study, the participants reached a mean of 5.9 (sd $=$ 2.2) correct answers. This difference was not expected since, in both studies, students participated. Moreover, the result indicates that the KIT students are highly educated with respect to probabilistic reasoning.

In summary, as expected from random treatment allocation, there is no significant difference across the treatments.

## Analysis of Contextual Differences

This study investigates the effects of different situational contexts (urn, robo, dating, and exam) on decision inertia and nudging. In line with Rauthmann et al. (2015), contexts summarize situations with similar profiles of situational characteristics. Thus, the study's situational contexts have to be significantly different with respect to the perceived situational characteristics. The mean values and standard deviations are depicted in figure 8.9.

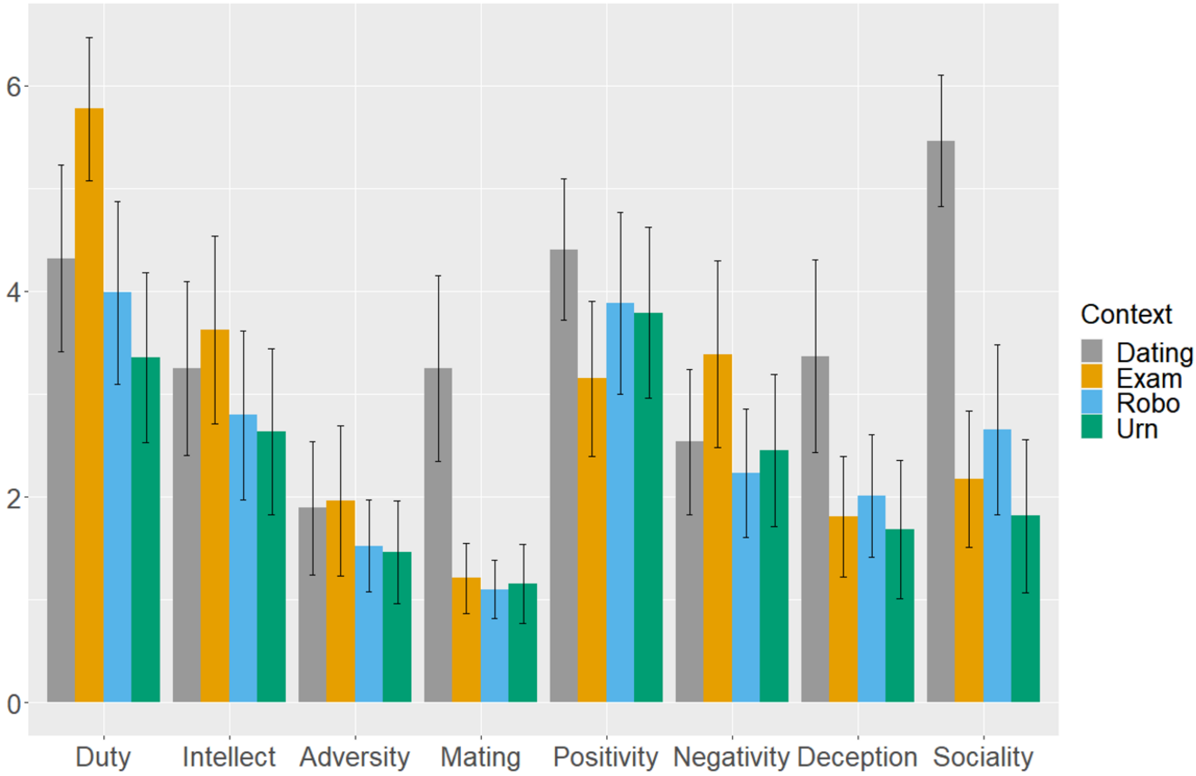



FIGURE 8.9: *Mean values and standard deviations of perceived situational characteristics across the situational contexts.*

First, a MANOVA is performed to check for differences. The result indicates significant differences across situational contexts (n = 325, df = 3, approx F = 19.148, p-value < 0.001). Then, for each pair of situational context, a separate MANOVA is performed (see table 8.4).

Since all pairwise MANOVAs delivered a significant p-value, all situational contexts are differing significantly with respect to their profiles of situational characteristics. Hence, the study successfully induced different situational contexts.

> ***Result A*[3]**: *The situational contexts are different concerning perceived situational characteristics.*

[3] Supplementary results are coded with letters.

| | Robo | Dating | Exam |
|---|---|---|---|
| Urn | Approx F = 2.712, p-value < 0.001 *** | Approx F = 45.853, p-value < 0.001 *** | Approx F = 16.168, p-value < 0.001 *** |
| Robo | | Approx F = 32.268, p-value < 0.001 *** | Approx F = 12.089, p-value < 0.001 *** |
| Dating | | | Approx F = 50.707, p-value < 0.001 *** |

TABLE 8.4: *Pairwise comparison of the situational contexts.*

### Overall Error Rates

In order to analyze decision inertia, the mean error rates of the second decisions are used. There are two conditions. One condition emerges when the first decision leads to a reward and the participant switches the urn (staying with the urn results in greater chance to draw a rewarded ball), namely the convergent condition. The choice repetition and Bayesian updating are aligned. The other condition, namely divergent, emerges when the first decision is not rewarded and the participant stays with the urn (switching results in a greater chance to draw a rewarded ball). In this case, choice repetition and Bayesian updating are not aligned. Hence, decision inertia occurs.

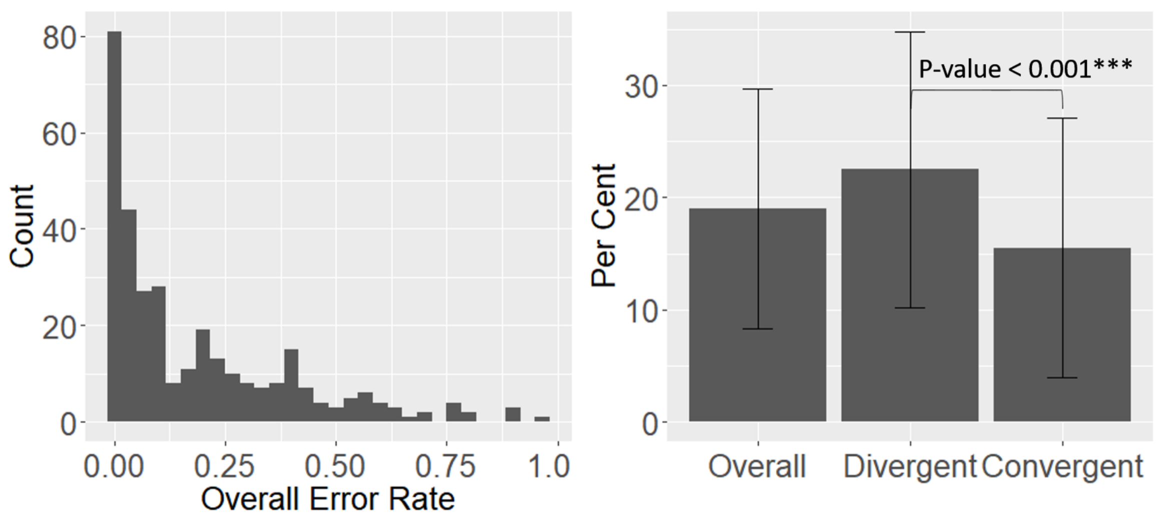



FIGURE 8.10: *Histograms of overall error rate and bar chart of error rates.*

Overall mean error rate of second decisions is 19.1 % (sd = 21.4). The mean error rate in the convergent condition is 15.7 % (sd = 36.5), whereas in divergent condition, the mean error rate amounts to 22.4 % (sd = 41.7). The error rates are not normally distributed (see figure 8.10). Thus, the difference is tested with a non-parametric one-sided Wilcoxon ran-sum test. The result is highly significant (n = 324, v = 27686, p-value < 0.001), indicating that the error rates are higher in the divergent condition. In line with Alós-Ferrer et al. (2016); Jung and Dorner (2018); Jung et al. (2019, 2018), it is assumed that decision inertia is replicated successfully.

## Procedure and Methodology of Analysis

Logistic regressions with participants as random effects and all other variables as fixed effects are chosen since the experimental task was repeated 60 times and learning effects are expected. The dependent variable of the regressions is the binary variable *suboptimal second choice* (1 = true) in each round (*round number*). The interaction of *suboptimal second choice* with the binary variable *divergence* (1 = true) indicates decision inertia (divergent condition). This procedure is in line with Alós-Ferrer et al. (2016); Jung and Dorner (2018); Jung et al. (2019, 2018).

In contrast to existing research on decision inertia (Alós-Ferrer et al., 2016; Jung and Dorner, 2018; Jung et al., 2019, 2018), this study used the average marginal effects to conclude about the influence of investigated variables (Schunck and Nisic, 2020; Brambor et al., 2006). Previous research used the regression coefficients and significances of standard regression tables to investigate hypotheses. However, this approach is solely meaningful in linear additive regressions since average marginal effects are equivalent to the regression coefficients. In the case of logistic regressions and regressions with interactions, the main and interaction effects are conditional, which means that the coefficient and significance of a variable differ for each value of the remaining variables. The standard regression table has to be ignored since it contains the results of a particular condition. In order to conclude about the influence of variables, the average marginal effects have to be used (Schunck and Nisic, 2020; Brambor et al., 2006).

The names of the regression models and results are coded with the same number of the corresponding working hypothesis (see research model 8.1 and experimental design 8.3). Supplementary analyses and results are alphabetically coded.

## Control Variables

Initially, a regression with the control variables *round number, action orientation, gender*, and *business engineer* is performed. The standard regression table can be found in the appendix (see table D.2) because of the complete documentation. However, it is not relevant for the analysis.

| Variable | AME | SE | P-value |
|---|---|---|---|
| Divergence (1 = True) | 0.070 | 0.005 | < 0.001 *** |
| Gender (1 = Male) | -0.065 | 0.026 | 0.012 * |
| Business Engineer (1 = True) | -0.003 | 0.024 | 0.877 |
| Action-orientation | 0.059 | 0.059 | 0.316 |
| Round Number | -0.028 | 0.008 | < 0.001 *** |

TABLE 8.5: *The average marginal effects of the control variables on suboptimal second choice.*

First, the influence on *suboptimal second choice* is investigated. For this purpose, the average marginal effects (AME) are calculated (see table 8.5). *Round number* (AME = -0.028, p-value < 0.001) and *gender* (1 = male, AME = -0.065, p-value = 0.012) decrease *suboptimal second choice* significantly.

***Result B**: With each round, the overall error rate decreases.*

***Result C**: Male participants made less overall errors than female participants.*

In order to analyze the influence on *decision inertia*, the average marginal effects are calculated for the divergence cases (see table 8.6). *Round number* (AME = -0.021, p-value = 0.069) reduces decision inertia significantly. Hence, there is also a learning effect on decision inertia.

***Result D**: With each round, decision inertia decreases.*

Remarkably, in contrast to a previous study (Jung et al., 2019), *action orientation* does not influence *suboptimal second choice* (AME = 0.059, p-value = 0.316) and *decision inertia* (AME = 0.054, p-value = 0.418). However, the results of previous studies are questionable due to the inappropriate analysis method.

***Result E**: Action orientation does not influence decision inertia.*

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.0929 | 0.026 | < 0.001 *** |
| Gender (1 = Male) | 1 | -0.037 | 0.026 | 0.145 |
| Business Engineer (1 = True) | 0 | 0.023 | 0.024 | 0.340 |
| Business Engineer (1 = True) | 1 | -0.030 | 0.024 | 0.211 |
| Action-orientation | 0 | 0.064 | 0.059 | 0.278 |
| Action-orientation | 1 | 0.054 | 0.059 | 0.356 |
| Round Number | 0 | -0.034 | 0.008 | < 0.001 *** |
| Round Number | 1 | -0.021 | 0.008 | 0.006 ** |

TABLE 8.6: *The average marginal effects of the control variables in divergence cases.*

In sum, *round number* and *gender* influence *suboptimal second choice* or *decision inertia*. Hence, these variables are controlled in further analyses.

## Decision inertia across Situational Contexts

*H1* assumes that *decision inertia* varies across *situational contexts*. In order to avoid the interactions with the variable *nudging*, the subsample without nudging (n = 106) is used. The mean values and standard deviations of *decision inertia* across the *situational contexts* are depicted in figure 8.11.

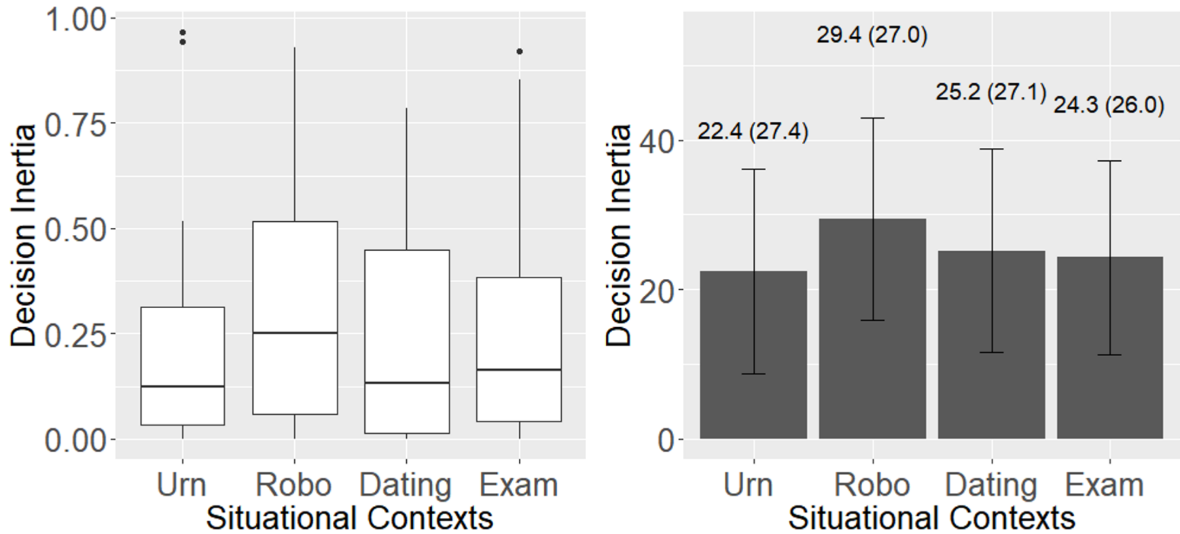



FIGURE 8.11: *Boxplots and bar chart of decision inertia across the situational contexts.*

For the analysis of the differences, three logistic regressions with different reference levels (*urn, robo,* and *dating*) are performed (see appendix D.3). Afterward, for each regression, the average marginal effects are calculated.

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.116 | 0.045 | 0.010 * |
| Gender (1 = Male) | 1 | -0.078 | 0.045 | 0.087 + |
| Round Number | 0 | -0.032 | 0.014 | 0.026 * |
| Round Number | 1 | -0.047 | 0.014 | < 0.001 ** |
| Robo (1 = True) | 0 | 0.069 | 0.060 | 0.252 |
| Robo (1 = True) | 1 | 0.074 | 0.060 | 0.220 |
| Dating (1 = True) | 0 | 0.0826 | 0.058 | 0.159 |
| Dating (1 = True) | 1 | 0.001 | 0.058 | 0.977 |
| Exam (1 = True) | 0 | 0.109 | 0.059 | 0.063 + |
| Exam (1 = True) | 1 | 0.012 | 0.059 | 0.832 |

TABLE 8.7: *The average marginal effects of situational contexts in divergence cases with the urn context as the reference level.*

Compared to the *urn* context (see table 8.7), there are no significant differences in the situational context of *robo* (AME = 0.074, p-value = 0.220), *dating* (AME = 0.001, p-value = 0.977), and *exam* (AME = 0.012, p-value = 0.832).

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.116 | 0.045 | 0.010 * |
| Gender (1 = Male) | 1 | -0.078 | 0.045 | 0.087 + |
| Round Number | 0 | -0.032 | 0.014 | 0.026 * |
| Round Number | 1 | -0.047 | 0.014 | 0.001 ** |
| Urn (1 = True) | 0 | -0.069 | 0.060 | 0.252 |
| Urn (1 = True) | 1 | -0.074 | 0.060 | 0.220 |
| Dating (1 = True) | 0 | 0.013 | 0.064 | 0.838 |
| Dating (1 = True) | 1 | -0.072 | 0.064 | 0.261 |
| Exam (1 = True) | 0 | 0.040 | 0.065 | 0.538 |
| Exam (1 = True) | 1 | -0.061 | 0.065 | 0.344 |

TABLE 8.8: *The average marginal effects of situational contexts in divergence cases with the robo context as the reference level.*

The regression with the reference level *robo* context (see table 8.8) indicates no significant differences between the context of *urn* (AME = -0.074, p-value = 0.220), *dating* (AME = -0.072, p-value = 0.261), and *exam* (AME = -0.061, p-value = 0.344).

Finally, the regression with the reference level *dating* context (see table 8.9) shows no

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.116 | 0.045 | 0.010 * |
| Gender (1 = Male) | 1 | -0.078 | 0.045 | 0.087 + |
| Round Number | 0 | -0.032 | 0.014 | 0.026 * |
| Round Number | 1 | -0.047 | 0.014 | 0.001 ** |
| Urn (1 = True) | 0 | -0.082 | 0.058 | 0.158 |
| Urn (1 = True) | 1 | -0.001 | 0.058 | 0.977 |
| Robo (1 = True) | 0 | -0.013 | 0.064 | 0.838 |
| Robo (1 = True) | 1 | 0.072 | 0.064 | 0.261 |
| Exam (1 = True) | 0 | 0.027 | 0.063 | 0.668 |
| Exam (1 = True) | 1 | 0.010 | 0.063 | 0.864 |

TABLE 8.9: *The average marginal effects of situational contexts in divergence cases with the dating context as the reference level.*

significant differences between *urn* (AME = -0.001, p-value = 0.977), *robo* (AME = 0.072, p-value = 0.261), and *exam* (AME = 0.010, p-value = 0.864).

In sum, with three regressions with different reference levels, all situational contexts are compared to each other. Since none of the observed differences is significant, the alternative hypothesis of *H1* is not accepted.

***Result H1**: Decision inertia does not vary across situational contexts.*

## Effectiveness of Nudging

In this section, the general effectiveness of *defaults* (*H5D*) and *warnings* (*H5W*) in reducing *decision inertia* is investigated. The mean values and standard deviations are depicted in figure 8.12. Compared to the *baseline* treatment, in both *nudging* treatments a reduction of *decision inertia* is observed.

In order to examine the significance of the reductions, a logistic regression with baseline (without nudging) as the reference level is performed (see appendix D.4). Afterward, the average marginal effects in the divergence cases (decision inertia: divergence = 1) are calculated (see table 8.10). Accordingly, *defaults* do not reduce *decision inertia* significantly (AME = 0.000, p-value = 0.992).

***Result H5D**: Defaults do not reduce decision inertia.*

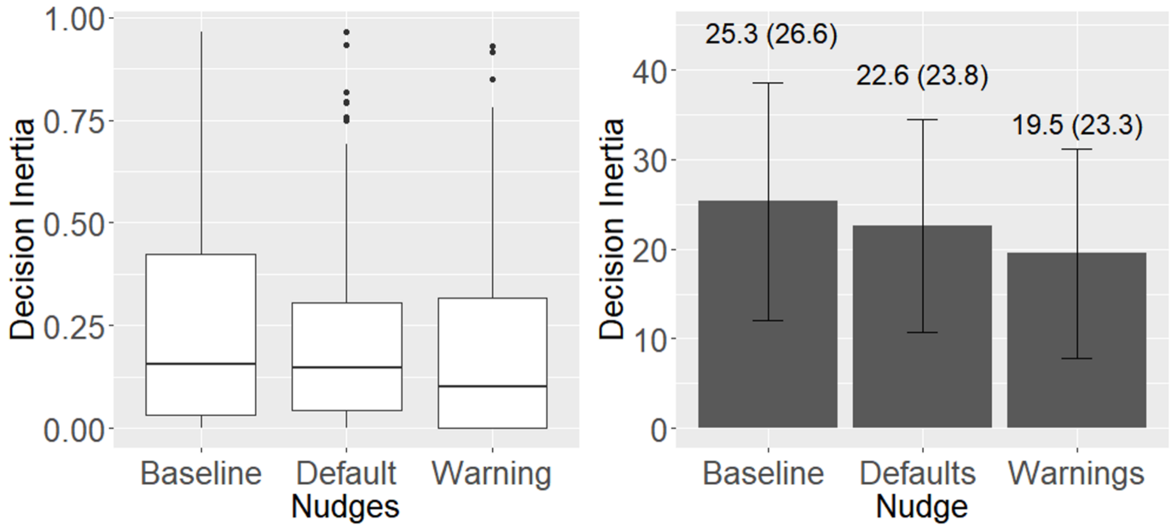



FIGURE 8.12: *Boxplots and bar chart of decision inertia across the nudging treatments.*

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.085 | 0.025 | < 0.001 *** |
| Gender (1 = Male) | 1 | -0.038 | 0.025 | 0.135 |
| Round Number | 0 | -0.035 | 0.008 | < 0.001 *** |
| Round Number | 1 | -0.023 | 0.008 | 0.004 ** |
| Defaults (1 = True) | 0 | 0.007 | 0.029 | 0.787 |
| Defaults (1 = True) | 1 | 0.000 | 0.029 | 0.992 |
| Warnings (1 = True) | 0 | 0.021 | 0.028 | 0.441 |
| Warnings (1 = True) | 1 | -0.048 | 0.028 | 0.089 + |

TABLE 8.10: *The average marginal effects of defaults and warnings in divergence cases with baseline as the reference level.*

In contrast, without differentiating the situational contexts, *warnings* (AME = -0.048, p-value = 0.089) are an effective means in reducing *decision inertia*.

***Result H5W****: Warnings reduce decision inertia.*

## Context-Aware Effectiveness of Nudging

In the following the context-dependent effectiveness of *defaults (H7D)* and *warnings (H7W)* is investigated. First, the descriptive statistics are calculated (see table 8.11) and the corresponding distributions are plotted (see figure 8.13).

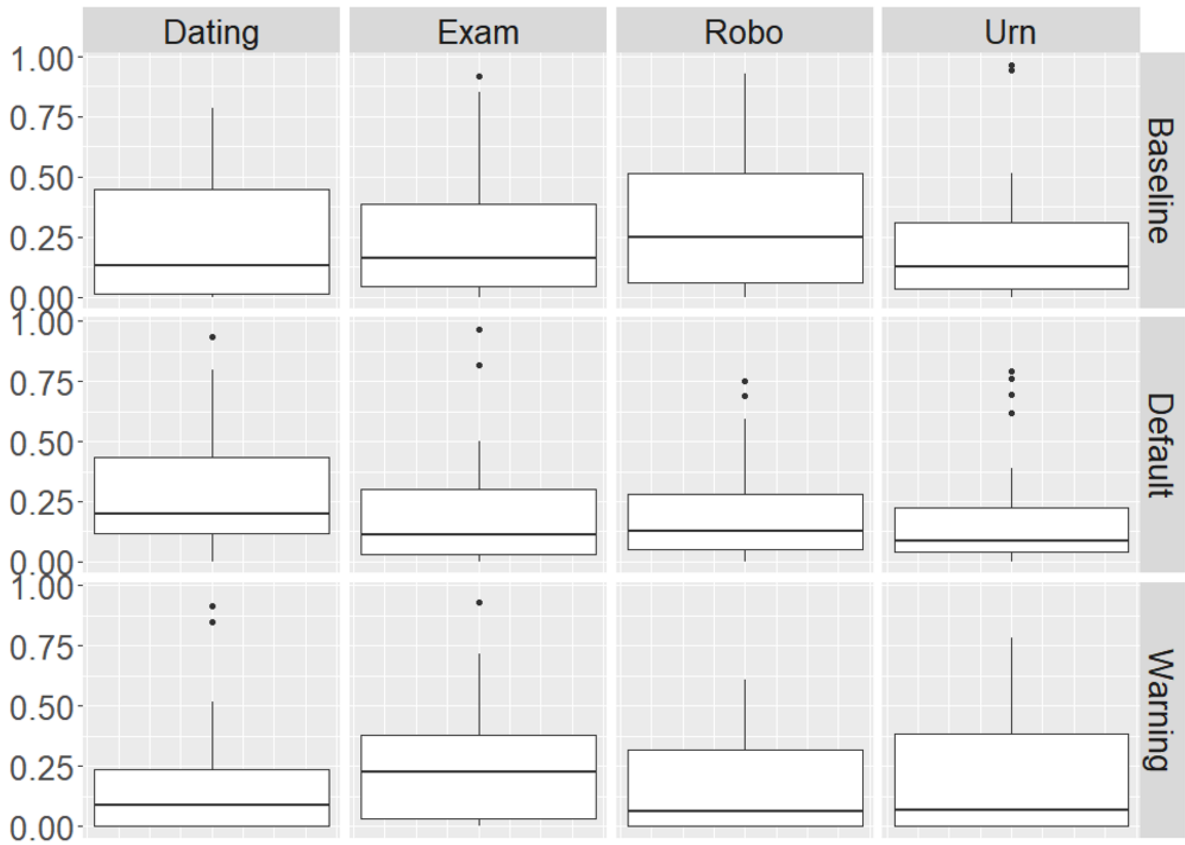



FIGURE 8.13: *Boxplots of decision inertia for each nudging treatment across situational contexts.*

In the *urn* and *robo* context, both *nudges* reduced the mean values of *decision inertia* as expected. Remarkably, in the context of *dating*, *defaults* increased *decision inertia*. Moreover, in the *exam* context, *warnings* have a higher mean value than the *baseline* treatment.

In order to investigate the significance of the differences, for each situational context, a logistic regression with nudges as independent variable and baseline (without nudging) as the reference level is performed (see appendix D.5). With the regression models, the average marginal effects are calculated.

In the *urn* context (see table 8.12), the decrease of *decision inertia* with *defaults* (AME = -0.026, p-value = 0.591) and warnings (AME = -0.037, p-value = 0.419) is not significant.

***Result F***: *Defaults do not significantly reduce decision inertia in the urn context.*

| | Urn | Robo | Dating | Exam |
|---|---|---|---|---|
| Baseline | 22.4 (27.4) | 29.4 (27.0) | 25.2 (27.0) | 24.3 (26.0) |
| Defaults | 19.6 (23.5) | 19.8 (20.3) | 30.1 (25.2) | 21.0 (25.8) |
| Warnings | 18.5 (23.0) | 16.8 20.6) | 17.1 (23.5) | 25.6 (26.0) |

TABLE 8.11: *Mean values and standard deviations of decision inertia in each treatment. Reductions are colored green, whereas increases are colored red.*

| **Variable** | **Divergence** | **AME** | **SE** | **P-value** |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | 0.004 | 0.040 | 0.909 |
| Gender (1 = Male) | 1 | 0.013 | 0.040 | 0.730 |
| Round Number | 0 | -0.009 | 0.014 | 0.502 |
| Round Number | 1 | -0.056 | 0.014 | < 0.001 *** |
| Defaults (1 = True) | 0 | 0.042 | 0.049 | 0.386 |
| Defaults (1 = True) | 1 | -0.026 | 0.049 | 0.591 |
| Warnings (1 = True) | 0 | -0.001 | 0.046 | 0.988 |
| Warnings (1 = True) | 1 | -0.037 | 0.046 | 0.419 |

TABLE 8.12: *The average marginal effects of warnings and defaults in the divergence cases, in the urn context, and baseline as the reference level.*

***Result G**: Warnings do not significantly reduce decision inertia in the urn context.*

| **Variable** | **Divergence** | **AME** | **SE** | **P-value** |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.148 | 0.054 | 0.006 ** |
| Gender (1 = Male) | 1 | -0.141 | 0.054 | 0.009 ** |
| Round Number | 0 | -0.069 | 0.016 | < 0.001 *** |
| Round Number | 1 | -0.025 | 0.016 | 0.137 |
| Defaults (1 = True) | 0 | 0.059 | 0.056 | 0.293 |
| Defaults (1 = True) | 1 | -0.033 | 0.056 | 0.552 |
| Warnings (1 = True) | 0 | -0.012 | 0.050 | 0.811 |
| Warnings (1 = True) | 1 | -0.124 | 0.050 | 0.013 * |

TABLE 8.13: *The average marginal effects of warnings and defaults in divergence cases, in the robo context, and baseline as the reference level.*

In the *robo* context (see table 8.13), *defaults* (AME = -0.033, p-value = 0.552) do not significantly reduced *decision inertia*, whereas *warnings* (AME = -0.124, p-value = 0.013) significantly reduce *decision inertia*.

***Result H**: Defaults do not significantly reduce decision inertia in the robo context.*

***Result I**: Warnings significantly reduce decision inertia in the robo context.*

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.137 | 0.052 | 0.008 ** |
| Gender (1 = Male) | 1 | -0.108 | 0.052 | 0.036 * |
| Round Number | 0 | 0.004 | 0.016 | 0.772 |
| Round Number | 1 | 0.016 | 0.016 | 0.328 |
| Defaults (1 = True) | 0 | 0.031 | 0.059 | 0.598 |
| Defaults (1 = True) | 1 | 0.115 | 0.059 | 0.053 + |
| Warnings (1 = True) | 0 | 0.045 | 0.053 | 0.402 |
| Warnings (1 = True) | 1 | -0.033 | 0.053 | 0.531 |

TABLE 8.14: *The average marginal effects of warnings and defaults in divergence cases, in the dating context, and baseline as the reference level.*

In the *dating* context (see table 8.14), *defaults* (AME = 0.115, p-value = 0.053) significantly backfired (opposite effect than expected). Furthermore, *warnings* (AME = -0.033, p-value = 0.531) do not reduce decision inertia significantly.

***Result J***: *Defaults backfired in the dating context.*

***Result K***: *Warnings do not reduce decision inertia in the dating context.*

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.020 | 0.051 | 0.695 |
| Gender (1 = Male) | 1 | 0.085 | 0.051 | 0.098 + |
| Round Number | 0 | -0.067 | 0.016 | < 0.001 *** |
| Round Number | 1 | -0.025 | 0.016 | 0.120 |
| Defaults (1 = True) | 0 | -0.070 | 0.055 | 0.204 |
| Defaults (1 = True) | 1 | -0.034 | 0.055 | 0.537 |
| Warnings (1 = True) | 0 | 0.086 | 0.065 | 0.182 |
| Warnings (1 = True) | 1 | 0.043 | 0.065 | 0.502 |

TABLE 8.15: *The average marginal effects of warnings and defaults in divergence cases, in the exam context, and baseline as the reference level.*

Finally, in the *exam* context (see table 8.15), neither *defaults* (AME = -0.034, p-value = 0.537) nor *warnings* (AME = 0.043, p-value = 0.502) caused a significant difference in *decision inertia*.

***Result L***: *Defaults do not significantly reduce decision inertia in the exam context.*

***Result M***: *Warnings do not significantly reduce decision inertia in the exam context.*

Remarkably, the effect of *gender* on *decision inertia* varies across situational contexts. In the *urn* context, there is no difference between *males* and *females*, whereas, in the *robo* and *dating* context, females are more prone to *decision inertia* than *males*. In contrast, *males* are more prone decision inertia in the *exam* context.

> ***Result N****: Gender differently influences decision inertia across situational contexts.*

To answer the research question about contextual differences in nudge effectiveness, for each nudge, the situational contexts have to be compared with each other. Hence, three logistic regressions with different reference levels (urn, robo, and dating) are performed for each nudge (see appendix D.6).

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.030 | 0.041 | 0.461 |
| Gender (1 = Male) | 1 | -0.005 | 0.041 | 0.900 |
| Round Number | 0 | -0.033 | 0.013 | 0.015 * |
| Round Number | 1 | -0.004 | 0.013 | 0.757 |
| Robo (1 = True) | 0 | 0.057 | 0.046 | 0.220 |
| Robo (1 = True) | 1 | 0.033 | 0.046 | 0.480 |
| Dating (1 = True) | 0 | 0.065 | 0.050 | 0.199 |
| Dating (1 = True) | 1 | 0.132 | 0.050 | 0.009 ** |
| Exam (1 = True) | 0 | -0.010 | 0.044 | 0.820 |
| Exam (1 = True) | 1 | 0.026 | 0.044 | 0.547 |

TABLE 8.16: *The average marginal effects of situational contexts with the urn context as the reference level and in the subsample of defaults.*

In the case of *defaults* and *urn* context as the reference level (see table 8.16), there is a significant difference with dating context (AME = 0.132, p-value = 0.009) concerning *decision inertia*.

Compared with the *robo* context (see table 8.17), *decision inertia* of the *dating* context (AME = 0.099, p-value = 0.071) is significantly different, whereas no difference is observed in the remaining contexts.

Finally, the average marginal effects of the regression with the *dating* context as reference level (see table 8.18) show significant differences in the *urn* (AME = -0.132, p-value = 0.009), *robo* (AME = -0.099, p-value = 0.071), and *exam* context (AME = -0.105, p-value = 0.046).

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.030 | 0.041 | 0.461 |
| Gender (1 = Male) | 1 | -0.005 | 0.041 | 0.900 |
| Round Number | 0 | -0.033 | 0.013 | 0.015 * |
| Round Number | 1 | -0.004 | 0.013 | 0.757 |
| Urn (1 = True) | 0 | -0.057 | 0.046 | 0.220 |
| Urn (1 = True) | 1 | -0.033 | 0.046 | 0.480 |
| Dating (1 = True) | 0 | 0.008 | 0.055 | 0.885 |
| Dating (1 = True) | 1 | 0.099 | 0.055 | 0.071 + |
| Exam (1 = True) | 0 | -0.067 | 0.048 | 0.159 |
| Exam (1 = True) | 1 | -0.006 | 0.048 | 0.895 |

TABLE 8.17: *The average marginal effects of situational contexts with the robo context as the reference level and in the subsample of defaults.*

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.030 | 0.041 | 0.461 |
| Gender (1 = Male) | 1 | -0.005 | 0.041 | 0.900 |
| Round Number | 0 | -0.033 | 0.013 | 0.015 * |
| Round Number | 1 | -0.004 | 0.013 | 0.757 |
| Urn (1 = True) | 0 | -0.065 | 0.050 | 0.199 |
| Urn (1 = True) | 1 | -0.132 | 0.050 | 0.009 ** |
| Robo (1 = True) | 0 | -0.008 | 0.055 | 0.885 |
| Robo (1 = True) | 1 | -0.099 | 0.055 | 0.071 + |
| Exam (1 = True) | 0 | -0.075 | 0.053 | 0.156 |
| Exam (1 = True) | 1 | -0.105 | 0.053 | 0.046 * |

TABLE 8.18: *The average marginal effects of situational contexts with the dating context as the reference level and in the subsample of defaults.*

In summary, three regressions with different reference levels are performed in the subsample of *defaults* to compare each context with each other. The results show significant differences in *defaults'* effectiveness between *dating* and *urn*, *dating* and *exam*, and *dating* and *robo*.

***Result H7D****: The effectiveness of defaults varies across situational contexts.*

Next, the average marginal effects of *warnings'* effectiveness across situational contexts are investigated. No significant differences are observed with the *urn* context as the reference level (see table 8.19).

The regression with the *robo* context as the reference level (see table 8.20) indicates

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.060 | 0.042 | 0.159 |
| Gender (1 = Male) | 1 | -0.023 | 0.042 | 0.587 |
| Round Number | 0 | -0.037 | 0.013 | 0.005 ** |
| Round Number | 1 | -0.016 | 0.013 | 0.221 |
| Robo (1 = True) | 0 | 0.045 | 0.047 | 0.338 |
| Robo (1 = True) | 1 | -0.027 | 0.047 | 0.567 |
| Dating (1 = True) | 0 | 0.124 | 0.050 | 0.014 * |
| Dating (1 = True) | 1 | -0.001 | 0.050 | 0.988 |
| Exam (1 = True) | 0 | 0.185 | 0.058 | 0.001 ** |
| Exam (1 = True) | 1 | 0.084 | 0.058 | 0.152 |

TABLE 8.19: *The average marginal effects of situational contexts with the urn context as the reference level and in the subsample of warnings.*

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.060 | 0.042 | 0.159 |
| Gender (1 = Male) | 1 | -0.023 | 0.042 | 0.587 |
| Round Number | 0 | -0.037 | 0.013 | 0.005 ** |
| Round Number | 1 | -0.016 | 0.013 | 0.221 |
| Urn (1 = True) | 0 | -0.045 | 0.047 | 0.338 |
| Urn (1 = True) | 1 | 0.027 | 0.047 | 0.567 |
| Dating (1 = True) | 0 | 0.078 | 0.053 | 0.141 |
| Dating (1 = True) | 1 | 0.026 | 0.053 | 0.616 |
| Exam (1 = True) | 0 | 0.140 | 0.061 | 0.022 * |
| Exam (1 = True) | 1 | 0.111 | 0.061 | 0.068 + |

TABLE 8.20: *The average marginal effects of situational contexts with the robo context as the reference level and in the subsample of warnings.*

a significant difference in the *exam* context (AME = 0.111, p-value = 0.068) concerning *decision inertia* (divergence = 1).

Finally, the average marginal effects of the regression with the dating context as the reference level (see table 8.21) show no significant differences concerning *decision inertia*.

In essence, there is a significant difference in warnings' effectiveness between *robo* and *exam*. This case is sufficient to accept the alternative hypothesis of *H7W*.

***Result H7W***: *The effectiveness of warnings varies across situational contexts.*

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.060 | 0.042 | 0.159 |
| Gender (1 = Male) | 1 | -0.023 | 0.042 | 0.587 |
| Round Number | 0 | -0.037 | 0.013 | 0.005 ** |
| Round Number | 1 | -0.016 | 0.013 | 0.221 |
| Urn (1 = True) | 0 | -0.124 | 0.050 | 0.014 * |
| Urn (1 = True) | 1 | 0.001 | 0.050 | 0.989 |
| Robo (1 = True) | 0 | -0.078 | 0.053 | 0.141 |
| Robo (1 = True) | 1 | -0.026 | 0.053 | 0.616 |
| Exam (1 = True) | 0 | 0.061 | 0.063 | 0.330 |
| Exam (1 = True) | 1 | 0.084 | 0.063 | 0.180 |

TABLE 8.21: *The average marginal effects of situational contexts with the dating context as the reference level and in the subsample of warnings.*

## Perceived Confidence

In this section, the analysis of *perceived confidence* is reported. For recall, perceived *confidence* is a self-rated measure of an individual's ability. In the literature, there are the following cases of *confidence*. *Overconfidence* emerges when the judgment of one's ability is greater than the objective reality (Moore and Schatz, 2017). *Underconfidence* is the opposite case, when the judgment of one's ability is smaller than the reality (Moore and Schatz, 2017). Finally, the *realistic view*, when the subjective ability judgement meets the objective reality (Moore and Schatz, 2017).

In cases where true confidence can not be measured, as in this study, *perceived confidence* is used as a robust correlate (Larrick et al., 2007; Svenson, 1981; Murphy et al., 2018). It was measured before (pre) and after (post) the experimental task. It is assumed that the measurement after the experimental task corresponds to the realistic view of participants' abilities. Consequently, the difference between pre and post task measurement indicates participants *perceived confidence*.

First, the descriptive statistics are reported (see figure 8.14) and the distributions are investigated and plotted (see figure 8.15). Shapiro-Wilk normality tests indicate that *pre* (W = 0.9602, p-value $< 0.001$) and *post* task measurement (W = 0.93769, p-value $<$ 0.001) are not normally distributed. Hence, non-parametric tests are used for further analysis. There is a decrease in the mean value after the experimental task (mean$_{pre}$ = 8.3, sd$_{pre}$ = 15.0; mean$_{post}$ = -1.9, sd$_{pre}$ = 14.4). A one-sided Wilcoxon signed rank test

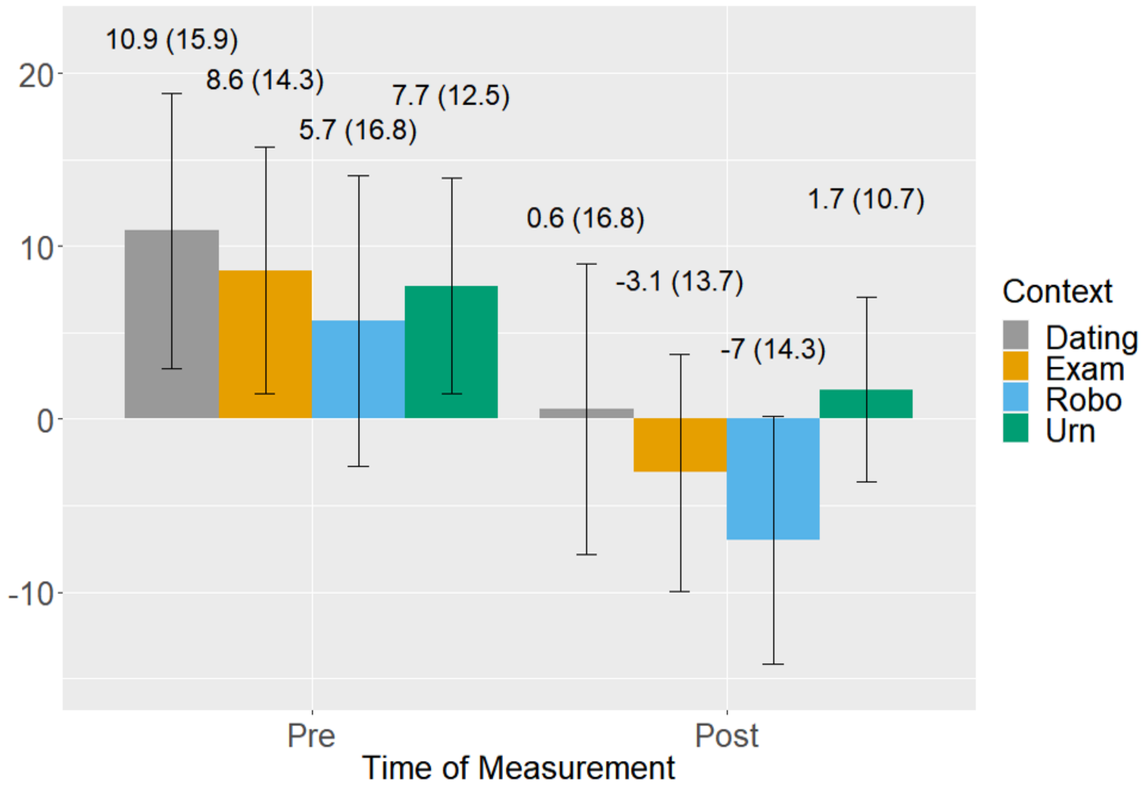



FIGURE 8.14: *Mean values and standard deviations of pre and post task measurement for perceived confidence.*

shows (V = 35037, p-value $< 0.001$) that the decrease is significant. Consequently, most of the participants were overconfident.

***Result O***: *Most of the participants were overconfident.*

Next, *perceived confidence* (difference between pre and post task measurement) is calculated and the distribution is plotted (see figure 8.16). Two hundred twenty-four participants were overconfident (positive difference), 64 were underconfident (negative difference), and 36 had a realistic view (no difference) of their abilities. The variable is not normally distributed (Shapiro-Wilk normality test: W = 0.97573, p-value $< 0.001$). In line with previous research (Sahin and Yilmaz, 2014), *perceived confidence* varies across situational contexts (Kruskal-Wallis rank sum test: chi-squared = 8.2493, df = 3, p-value = 0.041).

***Result P***: *Perceived confidence varies across situational contetxts.*

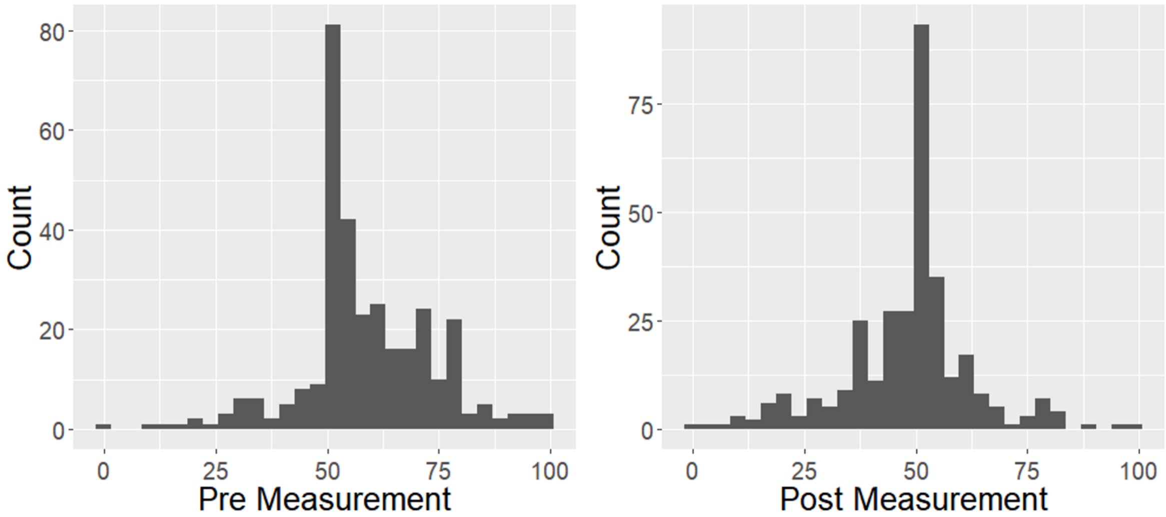



FIGURE 8.15: *Distributions of pre and post task measurement for perceived confidence.*

In order to investigate the hypotheses (*H4O* and *H4U*), two logistic regressions are performed (see appendix D.8). The first regression with *perceived confidence* as the independent variable (linear model). The second regression with *squared perceived confidence* as the independent variable (squared model). If *overconfidence (H4O)* and *underconfidence (H4U)* increase *decision inertia*, the squared model explains more variance than the linear model.

| **Variable** | **Divergence** | **AME** | **SE** | **P-value** |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.089 | 0.025 | < 0.001 *** |
| Gender (1 = Male) | 1 | -0.043 | 0.025 | 0.086 + |
| Round Number | 0 | -0.035 | 0.008 | < 0.001 *** |
| Round Number | 1 | -0.022 | 0.008 | 0.004 ** |
| Perceived Confidence (Linear) | 0 | 0.030 | 0.011 | 0.008 ** |
| Perceived Confidence (Linear) | 1 | 0.035 | 0.011 | 0.002 ** |

TABLE 8.22: *The average marginal effects of the linear model concerning perceived confidence in the cases of divergence.*

First, the average marginal effects of the linear model are reported (see table 8.22). *Perceived confidence (linear)* (AME = 0.035, p-value = 0.002) significantly influences *decision inertia*. *Decision inertia* increases with increasing *perceived confidence (linear)*. The average marginal effects of the squared model (see table 8.23) indicates also a significant and positive influences (AME = 0.032, p-value = 0.004).

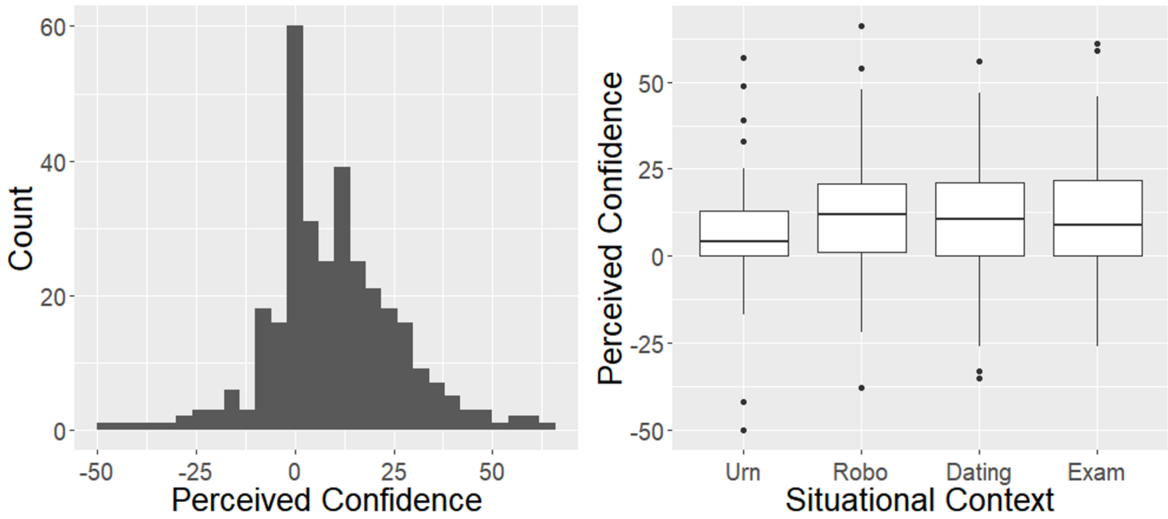



FIGURE 8.16: *Histogram and boxplots of perceived confidence*

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.089 | 0.025 | < 0.001 *** |
| Gender (1 = Male) | 1 | -0.041 | 0.025 | 0.098 + |
| Round Number | 0 | -0.035 | 0.008 | < 0.001 *** |
| Round Number | 1 | -0.022 | 0.008 | 0.004 ** |
| Perceived Confidence (Squared) | 0 | 0.032 | 0.011 | 0.004 ** |
| Perceived Confidence (Squared) | 1 | 0.032 | 0.011 | 0.004 ** |

TABLE 8.23: *The average marginal effects of the squared model concerning perceived confidence in divergence cases.*

For the evaluation of the hypotheses, the models are compared with each other using marginal effect plots and the goodness of fit statistic pseudo-$R^2$. The marginal pseudo-$R^2$ is chosen since it reflects the variance of the fixed effects. The *squared model* (marginal pseudo-$R^2$ = 0.047) explains more variance than the *linear model* (marginal pseudo-$R^2$ = 0.046). Moreover, the marginal effect plots (see figure 8.17) show that *perceived confidence (squared)* has a stronger effect on decision inertia than *perceived confidence (linear)*. Consequently, it is assumed that *under-* and *overconfidence* increase *decision inertia*.

***Result H4O****: Overconfidence increases decision inertia.*

***Result H4U****: Underconfidence increases decision inertia.*

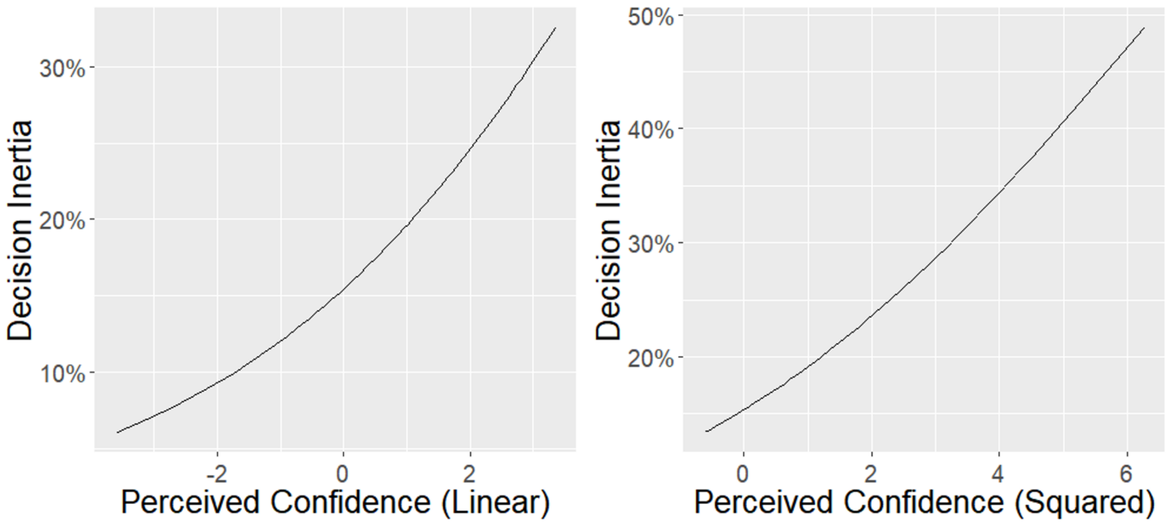



FIGURE 8.17: *Marginal effect plots of linear and squared perceived confidence.*

### Domain Expertise

| | Mean | SD | Scale Min | Scale Max | Shapiro Wilk Normality Test |
|---|---|---|---|---|---|
| **Probabilistic Reasoning** | 8.2 | 1.3 | 0 | 9 | W = 0.6263, P-value 0.001 |
| **Financial Literacy** | 3.9 | 0.8 | 0 | 5 | W = 0.81305, P-value 0.001 |
| **Interpersonal Competence** | 17.8 | 17.9 | -80 | 80 | W = 0.97469, P-value = 0.096 |
| **Learning Competence** | 14.6 | 7.31 | -34 | 34 | W = 0.97483, P-value = 0.115 |

TABLE 8.24: *Descriptive statistics of domain expertise questionnaires*

In this section, the analysis of *domain expertise* is reported. It is hypothesized (*H3*) that higher levels of *domain expertise* decrease *decision inertia*. *Domain expertise* is captured by *probabilistic reasoning* ability (urn), *financial literacy* (robo), *interpersonal competence* (dating), and *learning competence* (exam). The *probabilistic reasoning* and the *financial literacy* questionnaire query the knowledge of the participants, whereas the *interpersonal competence* and *learning competence* scale are designed for subjective self-assessment. All participants answered the *probabilistic reasoning* scale, whereas the remaining questionnaires were solely processed in the corresponding situational context. First, the descriptive statistics are calculated (see table 8.24) and the distributions are plotted (see figure 8.18) and tested.

Remarkably, most participants answered all items of the *probabilistic reasoning* scale correctly. This was not expected, since the questionnaire is evaluated with a student sample

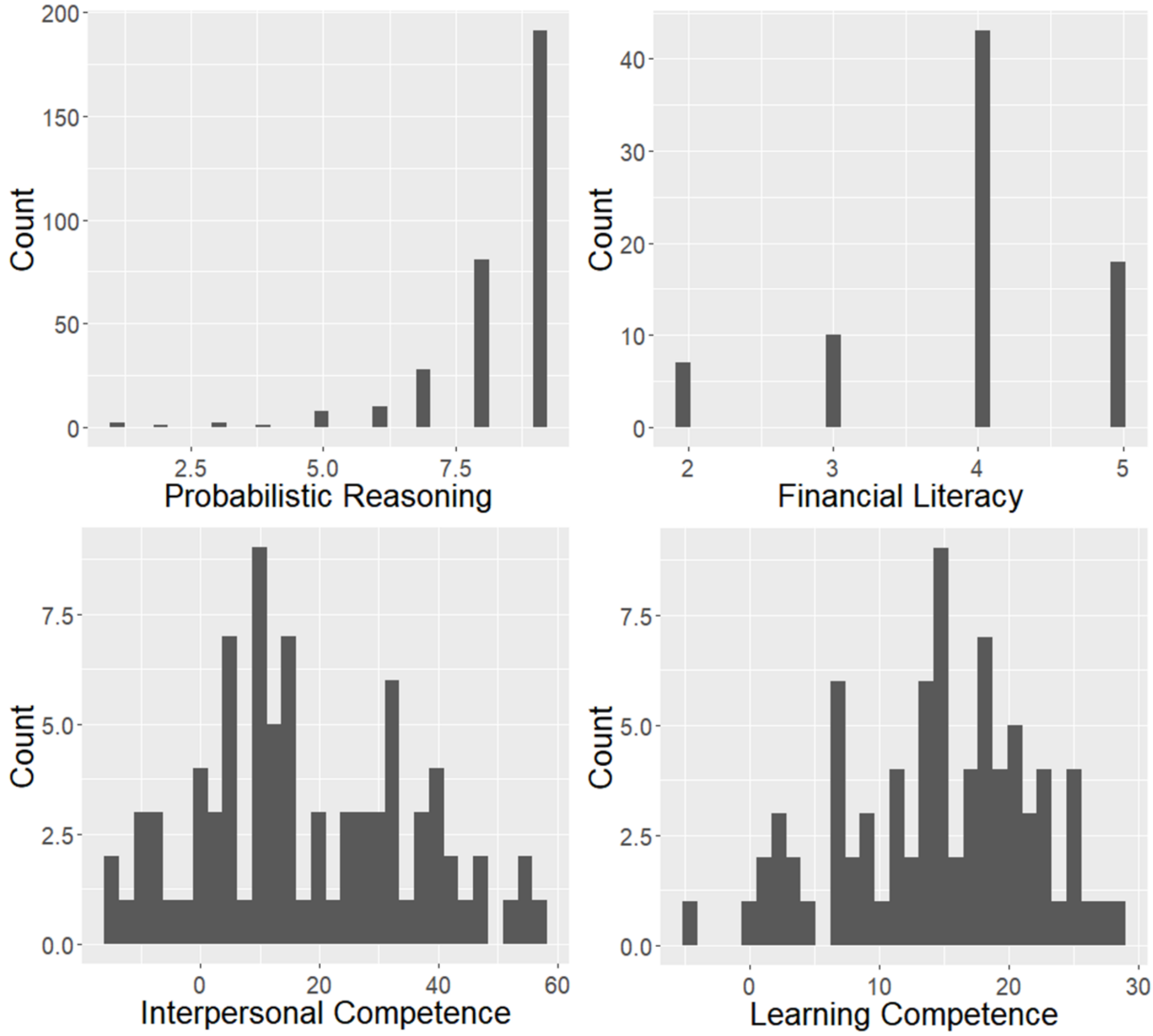



FIGURE 8.18: *Histograms of each domain expertise questionnaire*

and a mean of 5.9 (sd = 2.2) (Primi et al., 2019). Only *learning competence* is normally distributed. For the remaining analysis, the values are normed with a z-transformation so that mean values equal zero and standard deviations equal one.

First, a regression on the *suboptimal second choice* is performed with *probabilistic reasoning* as the predictor and without differentiating of situational contexts (see appendix D.9). Afterward, the average marginal effects in the divergence cases are calculated (see table 8.25). With increasing *probabilistic reasoning* score (AME = -0.046, p-value < 0.001), *decision inertia* decreases significantly.

***Result Q****: Increasing probabilistic reasoning decreases decision inertia.*

| **Variable** | **Divergence** | **AME** | **SE** | **P-value** |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.056 | 0.025 | 0.024 * |
| Gender (1 = Male) | 1 | -0.014 | 0.025 | 0.560 |
| Round Number | 0 | -0.035 | 0.008 | < 0.001 *** |
| Round Number | 1 | -0.022 | 0.008 | 0.004 ** |
| Probabilistic Reasoning | 0 | -0.047 | 0.011 | < 0.001 *** |
| Probabilistic Reasoning | 1 | -0.046 | 0.011 | < 0.001 *** |

TABLE 8.25: *The average marginal effects of probabilistic reasoning in divergence cases and without the differentiation of situational contexts.*

| **Variable** | **Divergence** | **AME** | **SE** | **P-value** |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | 0.035 | 0.040 | 0.372 |
| Gender (1 = Male) | 1 | 0.037 | 0.040 | 0.348 |
| Round Number | 0 | -0.010 | 0.014 | 0.471 |
| Round Number | 1 | -0.056 | 0.014 | < 0.001 *** |
| Probabilistic Reasoning | 0 | -0.039 | 0.018 | 0.037 * |
| Probabilistic Reasoning | 1 | -0.036 | 0.018 | 0.054 + |

TABLE 8.26: *The average marginal effects of probabilistic reasoning in divergence cases and in the urn context.*

Next, a regression for each domain expertise in the corresponding situational context is performed (see appendix D.10). The average marginal effects of the regression models are used to evaluate the hypotheses. In the subsample of *urn* (see table 8.26), *probabilistic reasoning* (AME = -0.036, p-value = 0.054) still negatively and significantly influences *decision inertia*.

> ***Result H3U****: An increasing probabilistic reasoning ability decreases decision inertia in the urn context.*

| **Variable** | **Divergence** | **AME** | **SE** | **P-value** |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.087 | 0.051 | 0.090 + |
| Gender (1 = Male) | 1 | -0.054 | 0.051 | 0.294 |
| Round Number | 0 | -0.067 | 0.016 | < 0.001 *** |
| Round Number | 1 | -0.024 | 0.016 | 0.151 |
| Financial Literacy | 0 | -0.049 | 0.021 | 0.021 * |
| Financial Literacy | 1 | -0.087 | 0.021 | < 0.001 *** |

TABLE 8.27: *The average marginal effects of financial literacy in divergence cases and in the robo context.*

In the *robo* context (see table 8.27), *financial literacy* (AME = -0.087, p-value < 0.001) significantly influences *decision inertia*. With increasing literacy score, *decision inertia* decreases.

> ***Result H3R**: An increasing financial literacy decreases decision inertia in the robo context.*

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.142 | 0.051 | 0.006 ** |
| Gender (1 = Male) | 1 | -0.106 | 0.051 | 0.038 * |
| Round Number | 0 | 0.005 | 0.016 | 0.743 |
| Round Number | 1 | 0.017 | 0.016 | 0.292 |
| Interpersonal Competence | 0 | -0.013 | 0.024 | 0.566 |
| Interpersonal Competence | 1 | -0.025 | 0.024 | 0.282 |

TABLE 8.28: *The average marginal effects of interpersonal competence in divergence cases and in the dating context.*

The average marginal effects of the regression in the *dating* context (see table 8.28) show that *interpersonal competence* (AME = -0.025, p-value = 0.282) does not influence *decision inertia* significantly.

> ***Result H3D**: Interpersonal competence does not influence decision inertia in the dating context.*

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.054 | 0.051 | 0.288 |
| Gender (1 = Male) | 1 | 0.070 | 0.051 | 0.167 |
| Round Number | 0 | -0.065 | 0.016 | < 0.001 *** |
| Round Number | 1 | -0.024 | 0.016 | 0.141 |
| Learning Competence | 0 | -0.033 | 0.024 | 0.172 |
| Learning Competence | 1 | -0.031 | 0.024 | 0.199 |

TABLE 8.29: *The average marginal effects of learning competence in divergence cases and in the exam context.*

Finally, in the *exam* context (see table 8.29), *learning competence* (AME = -0.031, p-value = 0.199) does not influence *decision inertia*.

> ***Result H3E**: Learning competence does not influence decision inertia in the exam context.*

In summary, *probabilistic reasoning* and *financial literacy* significantly influence *decision inertia*. In both cases, *decision inertia* decreases with the increasing literacy.

## Domain Expertise and Nudging

This section reports the interaction effects of *domain expertise* and *nudges*. This study assumes that increasing *domain expertise* leads to higher effectiveness of *defaults (H6D)* and *warnings (H6W)*. *Interpersonal competence* and *learning competence* do not influence *decision inertia* in the corresponding situational context. Hence, they are not further investigated.

In order to investigate the moderation effect, three-way interactions (divergence x domain expertise x nudge) are used as independent variables of the regressions. The R package margins[4], which is used to calculate the average marginal effects, can not handle more than two factors in three-way interactions. Hence, dummy variables are generated for each nudge, and separate regressions are performed.

| Variable | Defaults | Divergence | AME | SE | P-value |
|---|---|---|---|---|---|
| Gender (1 = Male) | 0 | 0 | 0.028 | 0.043 | 0.507 |
| Gender (1 = Male) | 0 | 1 | 0.045 | 0.043 | 0.297 |
| Gender (1 = Male) | 1 | 0 | 0.032 | 0.043 | 0.453 |
| Gender (1 = Male) | 1 | 1 | 0.025 | 0.043 | 0.554 |
| Round Number | 0 | 0 | -0.003 | 0.018 | 0.862 |
| Round Number | 0 | 1 | -0.076 | 0.018 | < 0.001 *** |
| Round Number | 1 | 0 | -0.003 | 0.018 | 0.848 |
| Round Number | 1 | 1 | -0.048 | 0.018 | 0.008 ** |
| Probabilistic Reasoning | 0 | 0 | -0.026 | 0.016 | 0.105 |
| Probabilistic Reasoning | 0 | 1 | -0.007 | 0.016 | 0.644 |
| Probabilistic Reasoning | 1 | 0 | -0.050 | 0.016 | 0.002 ** |
| Probabilistic Reasoning | 1 | 1 | -0.037 | 0.016 | 0.022 * |

TABLE 8.30: *The average marginal effects of probabilistic reasoning and defaults in divergence cases and in the urn context.*

In the *urn* context and in the case of *defaults*, the warning nudge is excluded from the regression analysis (see appendix D.11) resulting in a total of 3300 observations (n =

[4]https://cran.r-project.org/web/packages/margins/vignettes/Introduction.html

55). Afterward, the average marginal effects are calculated (see table 8.30). *Probabilistic reasoning* (AME =-0.037, p-value = 0.022) significantly increases the effectiveness of *defaults*.

***Result H6DU**: In the urn context, probabilistic reasoning increases the effectiveness of defaults to reduce decision inertia.*

| **Variable** | **Warnings** | **Divergence** | **AME** | **SE** | **P-value** |
|---|---|---|---|---|---|
| Gender (1 = Male) | 0 | 0 | 0.030 | 0.040 | 0.457 |
| Gender (1 = Male) | 0 | 1 | 0.061 | 0.040 | 0.129 |
| Gender (1 = Male) | 1 | 0 | 0.026 | 0.040 | 0.510 |
| Gender (1 = Male) | 1 | 1 | 0.047 | 0.040 | 0.241 |
| Round Number | 0 | 0 | -0.014 | 0.019 | 0.443 |
| Round Number | 0 | 1 | -0.026 | 0.019 | 0.161 |
| Round Number | 1 | 0 | -0.013 | 0.019 | 0.497 |
| Round Number | 1 | 1 | -0.022 | 0.019 | 0.238 |
| Probabilistic Reasoning | 0 | 0 | -0.028 | 0.028 | 0.310 |
| Probabilistic Reasoning | 0 | 1 | -0.009 | 0.028 | 0.740 |
| Probabilistic Reasoning | 1 | 0 | -0.033 | 0.028 | 0.238 |
| Probabilistic Reasoning | 1 | 1 | -0.043 | 0.028 | 0.122 |

TABLE 8.31: *The average marginal effects of probabilistic reasoning and warnings in divergence cases and in the urn context.*

In the case of *warnings*, defaults are excluded from the regression in the subsample of the *urn* context resulting in 3180 observations (n = 53) (see appendix D.12). The corresponding average marginal effects table (see table 8.31) shows that *probabilistic reasoning* (AME = -0.043, p-value = 0.122) does not significantly influence the effectiveness of *warnings*.

***Result H6WU**: In the urn context, probabilistic reasoning does not influence the effectiveness of warnings to reduce decision inertia.*

Next, the influence of *financial literacy* on the *defaults'* (observations = 3180, n = 53, see appendix D.13) and *warnings'* effectiveness (observations = 3000, n = 50, see appendix D.14) in the *robo* context is analyzed. With increasing *financial literacy*, the effectiveness of *defaults* (AME = -0.068, p-value = 0.020) increases significantly (see table 8.32).

***Result H6DR**: In the robo context, financial literacy significantly increases the effectiveness of defaults to reduce decision inertia.*

| Variable | Defaults | Divergence | AME | SE | P-value |
|---|---|---|---|---|---|
| Gender (1 = Male) | 0 | 0 | -0.045 | 0.068 | 0.512 |
| Gender (1 = Male) | 0 | 1 | -0.016 | 0.068 | 0.807 |
| Gender (1 = Male) | 1 | 0 | -0.127 | 0.068 | 0.064 + |
| Gender (1 = Male) | 1 | 1 | -0.016 | 0.068 | 0.812 |
| Round Number | 0 | 0 | -0.018 | 0.026 | 0.476 |
| Round Number | 0 | 1 | -0.058 | 0.026 | 0.027 * |
| Round Number | 1 | 0 | -0.084 | 0.026 | 0.001 ** |
| Round Number | 1 | 1 | -0.056 | 0.026 | 0.031 * |
| Financial Literacy | 0 | 0 | -0.016 | 0.029 | 0.570 |
| Financial Literacy | 0 | 1 | -0.110 | 0.029 | < 0.001 *** |
| Financial Literacy | 1 | 0 | -0.003 | 0.029 | 0.917 |
| Financial Literacy | 1 | 1 | -0.068 | 0.029 | 0.020 * |

TABLE 8.32: *The average marginal effects of financial literacy and defaults in divergence cases and in the robo context.*

Finally, the average marginal effect of *warnings* (see table 8.33) indicates that their effectiveness (AME = -0.112, p-value < 0.001) significantly increases with increasing *financial literacy*.

> ***Result H6WR**: In the robo context, financial literacy significantly increases the effectiveness of warnings to reduce decision inertia.*

## Situationally Perceived Negativity

In this section, the analysis of situational *perceived negativity* is reported. It is expected that an increasing *perceived negativity* increases *decision inertia* (*H2*).

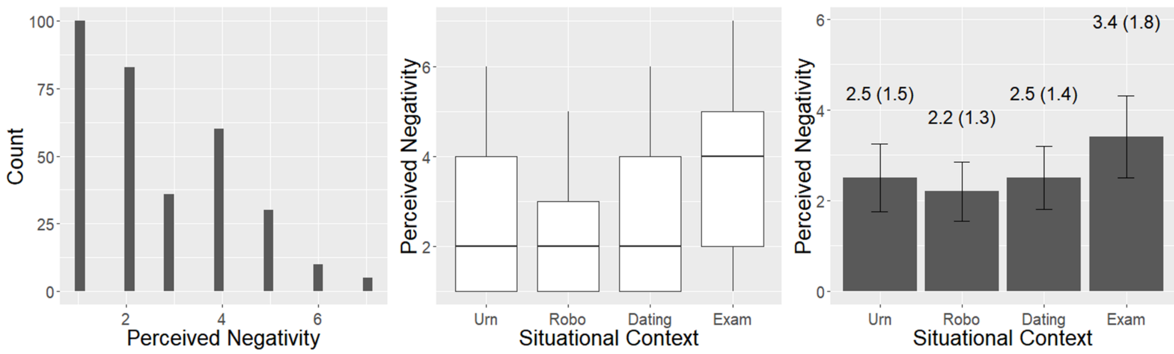



FIGURE 8.19: *Histogram, boxplots, and bar chart of perceived negativity.*

| Variable | Warnings | Divergence | AME | SE | P-value |
|---|---|---|---|---|---|
| Gender (1 = Male) | 0 | 0 | -0.091 | 0.065 | 0.162 |
| Gender (1 = Male) | 0 | 1 | -0.099 | 0.065 | 0.130 |
| Gender (1 = Male) | 1 | 0 | -0.063 | 0.065 | 0.330 |
| Gender (1 = Male) | 1 | 1 | -0.091 | 0.065 | 0.162 |
| Round Number | 0 | 0 | -0.092 | 0.025 | < 0.001 *** |
| Round Number | 0 | 1 | -0.033 | 0.025 | 0.189 |
| Round Number | 1 | 0 | -0.073 | 0.025 | 0.003 ** |
| Round Number | 1 | 1 | -0.030 | 0.025 | 0.221 |
| Financial Literacy | 0 | 0 | -0.093 | 0.030 | 0.002 ** |
| Financial Literacy | 0 | 1 | -0.092 | 0.030 | 0.002 ** |
| Financial Literacy | 1 | 0 | -0.041 | 0.030 | 0.182 |
| Financial Literacy | 1 | 1 | -0.112 | 0.030 | < 0.001 *** |

TABLE 8.33: *The average marginal effects of financial literacy and warnings in divergence cases and in the robo context.*

The *robo* context is rated as least negative, whereas the *exam* context as most negative. This result is plausible since the participants of the study are students. Most of them might have only theoretical knowledge in financial decision-making because they do not have sufficient capital to invest their money. In contrast, they are practically experienced in the case of exams. They may have already experienced negative feelings such as regret. Thus, *perceived negativity* might reflect their experienced emotions.

A Kruskal-Wallis rank sum test shows that there are highly significant differences (chi-square = 19.24, p-value < 0.001), which means the study successfully induced different levels of *perceived negativity*. Moreover, a Pearson's product-moment test indicates no correlation of *perceived negativity* with *perceived positivity* (cor = 0.077, t = 1.360, df = 322, p-value = 0.174). This result is expected. Assuming that negativity and positivity were the two sides of a continuum, the DIAMONDS taxonomy would contain only one characteristic reflecting both.

Next, a random-effects logistic regression on *suboptimal second choice* with *perceived negativity* as the independent variable is performed (see appendix D.15), and the average marginal effects are calculated (see table 8.34). *Perceived negativity* (AME = 0.017, p-value = 0.694) does not significantly influence *decision inertia* (divergence = 1).

***Result H2***: *Perceived negativity does not influence decision inertia.*

| Variable | Divergence | AME | SE | P-value |
|---|---|---|---|---|
| Gender (1 = Male) | 0 | -0.083 | 0.025 | 0.001 ** |
| Gender (1 = Male) | 1 | -0.037 | 0.051 | 0.144 |
| Round Number | 0 | -0.035 | 0.008 | < 0.001 *** |
| Round Number | 1 | -0.022 | 0.016 | 0.004 ** |
| Perceived Negativity | 0 | 0.035 | 0.044 | 0.436 |
| Perceived Negativity | 1 | 0.017 | 0.024 | 0.694 |

TABLE 8.34: *The average marginal effects of perceived negativity in divergence cases.*

## Summary of Results

The following table summarizes the hypotheses results.

| # | Method | Test Result | Insight | Reference |
|---|---|---|---|---|
| H1 | Regression | 0.219 < p-value < 0.998 | Decision inertia does not vary across contexts. | Tables: 8.7, 8.8, 8.9 |
| H2 | Regression | AME = 0.015, P-value = 0.694 | Perceived negativity does not influence decision inertia. | Table: 8.34 |
| H3U | Regression | AME = -0.036, p-value = 0.054 | Decision inertia decreases with increasing probabilistic reasoning. | Table: 8.26 |
| H3R | Regression | AME = -0.087, p-value < 0.001 | Decision inertia decreases with increasing financial literacy. | Table: 8.27 |
| H3D | Regression | AME = -0.025, p-value = 0.282 | Interpersonal competence does not influence decision inertia. | Table: 8.28 |
| H3E | Regression | AME = -0.031, p-value = 0.199 | Learning competence does not influence decision inertia in the exam. | Table: 8.29 |
| H4O | Regression | AME = 0.032, p-value = 0.004 | Overconfidence increases decision inertia. | Table: 8.23 |

| | | | | |
|---|---|---|---|---|
| H4U | Regression | AME = 0.032, p-value = 0.004 | Underconfidence increases decision inertia. | Table: 8.23 |
| H5D | Regression | AME = 0.000, p-value = 0.992 | Defaults do not reduce decision inertia. | Table: 8.10 |
| H5W | Regression | AME = -0.048, p-value = 0.089 | Warnings reduce decision inertia. | Table: 8.10 |
| H6DU | Regression | AME = -0.037, p-value = 0.022 | Probabilistic reasoning increases defaults' effectiveness. | Table: 8.30 |
| H6WU | Regression | AME = -0.043, p-value = 0.122 | Probabilistic reasoning does not influence the warnings' effectiveness. | Table: 8.31 |
| H6DR | Regression | AME = -0.068, p-value = 0.020 | Financial literacy increases defaults' effectiveness. | Table: 8.32 |
| H6WR | Regression | AME = -0.112, p-value $< 0.001$ | Financial literacy increases warnings' effectiveness. | Table: 8.33 |
| H7D | Regression | $0.008 <$ p-value $<$ 0.896 | Effectiveness of defaults varies across contexts. | Tables: 8.16, 8.17, 8.18 |
| H7W | Regression | $0.069 <$ p-value $<$ 0.989 | Effectiveness of warnings varies across contexts. | Tables: 8.19, 8.20, 8.21 |

Finally, the following table summarizes the supplementary results that are relevant for further research.

| # | Method | Test Result | Insight | Reference |
|---|---|---|---|---|
| A | MANOVA | P-values $< 0.001$ | Situational contexts are different. | Table: 8.4 |

| | | | | |
|---|---|---|---|---|
| B | Regression | AME = -0.028, p-value < 0.001 | Overall error decreases with each round. | Table: 8.5 |
| C | Regression | AME = -0.065, p-value < 0.001 | Males have a lower overall error rate than females. | Table: 8.5 |
| D | Regression | AME = -0.021, p-value = 0.006 | Decision inertia decreases with each round. | Table: 8.6 |
| E | Regression | AME = -0.054, p-value = 0.356 | Action orientation does not influence decision inertia | Table: 8.6 |
| F | Regression | AME = -0.026, p-value = 0.591 | Defaults do not reduce decision inertia in the urn. | Table: 8.12 |
| G | Regression | AME = -0.037, p-value = 0.419 | Warnings do not reduce decision inertia in the urn. | Table: 8.12 |
| H | Regression | AME = -0.033, p-value = 0.552 | Defaults do not reduce decision inertia in the robo. | Table: 8.13 |
| I | Regression | AME = -0.124, p-value = 0.013 | Warnings reduce decision inerita in the robo. | Table: 8.13 |
| J | Regression | AME = 0.115, p-value = 0.053 | Defaults backfired in the dating. | Table: 8.14 |
| K | Regression | AME = -0.033, p-value = 0.531 | Warnings do not reduce decision inertia in the dating. | Table: 8.14 |
| L | Regression | AME = -0.034, p-value = 0.537 | Defaults do not reduce decision inertia in the exam. | Table: 8.15 |
| M | Regression | AME = 0.043, p-value = 0.502 | Warnings do not reduce decision inertia in the exam. | Table: 8.15 |

| | | | | |
|---|---|---|---|---|
| N | Regression | 0.008 < p-value < 0.731 | Gender differently influences decision inertia across contexts. | Tables: 8.12, 8.13, 8.14, 8.15 |
| O | One-sided Wilcoxon | P-value < 0.001 | Most of the participants were overconfident. | Section: 8.4 |
| P | Kruskal-Wallis | P-value = 0.041 | Perceived confidence varies across contexts. | Section: 8.4 |
| Q | Regression | AME = -0.046, p-value < 0.001 | Decision inertia decreases with increasing probabilistic reasoning. | Table: 8.25 |

## 8.5 Discussion

Investigating situational characteristics to explain behavioral differences is a quite emerging research stream (Rauthmann, 2012; Rauthmann et al., 2014; Parrigon et al., 2017; Ziegler et al., 2019). This study evaluated the usefulness of situational characteristics in IS design research.

User characteristics and situational cues are not sufficient to classify participants appropriately. First, experimental and real-life situations contain many objective cues, which can not be perceived by everyone (Rauthmann, 2012). Second, the subjective relevance of an objective cue determines the behavior of an individual (Rauthmann, 2012). An illustrative example is the first dinner with another person. If the persons are romantically attracted to each other, the course of the conversation during the dinner will be quite different from amicable relationships. The subjective relevance of objective cues can be captured by psychologically relevant situation situational characteristics (Rauthmann et al., 2014). This study showed that these are capable of situationally classifying experimental stimuli. Moreover, with the help of the classification, situational influences on decision inertia and effectiveness of nudging were investigated.

For recall, different studies with the same research subject and hypotheses delivered mixed results (Jung et al., 2019; Alós-Ferrer et al., 2016). Jung (2019) suspected contextual differences. In contrast to the result of Jung et al. (2019), this study could not replicate the influence of action orientation on decision inertia.

The differences in the results are understandable in retrospect since previous studies did not analyze their results properly (Alós-Ferrer et al., 2016; Jung, 2019; Jung et al., 2019). The Scholars used the results of regression tables to evaluate the hypotheses. This is not sufficient because in regressions with interactions and logistic regressions, estimators and the significance of the estimators are conditional. Hence, the estimator and significance of a variable differ for each value of the remaining variables. Regression tables only reflect a particular condition. Consequently, the average marginal effects of the desired conditions have to be calculated and used for hypothesis evaluation (Brambor et al., 2006). This study makes a valuable contribution by providing a suitable analysis method. Currently, the findings of previous studies are questionable. Therefore, it is recommended to re-analyze the results properly and revise the findings.

There are indices in the existing literature for the situational dependency of decision inertia. This study investigated four different situational contexts of decision inertia. However, no significant differences could be observed. Nevertheless, the question of contextual differences has not yet been answered definitively.

The initial assumption that perceived negativity is a driver of decision inertia has not proven true. However, the investigated situational contexts were designed to be different concerning perceived negativity. Moreover, in the identification study, differences between the designed contexts were used to identify the most contrasted among them (see appendix C). Hence, there could be other situational contexts that have greater differences in their characteristics profiles, resulting in differences in decision inertia. This study provides a procedure to design and identify different situational contexts of an experimental task. In addition, it provides an experimental design to investigate contextual differences.

This study investigated the influence of perceived confidence on decision inertia since the measurement of true confidence is not practicable in the case of decision inertia. In line with the literature on confidence, perceived confidence is context-dependent and changes after context-specific experience (Larrick et al., 2007). Most participants lowered their confidence ratings after the experimental task, which could be attributed to the novelty of the task for them.

Scholars used the perceived confidence measurement after the experimental task in order to investigate its influence on decision-making (Larrick et al., 2007; Svenson, 1981; Murphy et al., 2018). However, the causality is questionable since the experimental task influences the measurement. In contrast, this study used the difference between the confidence measurements before and after the experimental task to identify over- and underconfidence. Thereby, the measurement after the task is assumed to be the realistic view of participants' abilities. Hence, a positive difference between prior and post task measurement indicates overconfidence, whereas a negative difference indicates underconfidence. The results indicate that both increase decision inertia.

Another examined context-specific user characteristic is domain expertise. It was assumed that increasing expertise decreases decision inertia. In the baseline context (urn) and the robo-advisor context (robo), the corresponding domain expertise, namely probabilistic reasoning ability and financial literacy, negatively influence decision inertia as expected. However, in the remaining contexts (exam: choosing learning strategies for exams, dating: choosing topics for dating conversation), the corresponding expertise (learning and interpersonal competence) does not influence decision inertia. It should be mentioned that the questionnaire of probabilistic reasoning and financial literacy is knowledge-based. In contrast, the remaining domain expertise questionnaires are self-rated measurements. Hence, the knowledge-based questionnaire might reflect the overall intelligence of the participants. The literature about financial literacy and probabilistic reasoning indeed shows that intelligence-related factors, such as education, are correlated to them. However, both questionnaires are sufficiently validated and capable of explaining unique variance.

Individual preferences might cause the differences concerning domain expertise. In the urn and robo context, preferences for choice alternatives may be less important than in the exam and dating context. It is undisputed that individuals have preferences regarding topics of conversation and learning strategies. In contrast, in the urn and robo context, participants had to choose between different urns or portfolios. It is assumed that in these contexts, preferences play less of a role, but rather individuals try to maximize their profits or payoffs.

This study also investigated the overall effectiveness of nudges (defaults and warnings) in reducing decision inertia. In both cases, a reduction is observed. However, only warnings reduce decision inertia significantly without differentiating the situational contexts. In order to provide a deeper understanding, the effectiveness of the nudges was investigated in the situational contexts.

In line with Jung (2019), warnings significantly reduced decision inertia in the robo context, whereas in the remaining contexts, the differences are not significant. In none of the investigated contexts do defaults lead to a significant reduction in decision inertia. In the context dating, defaults even significantly backfired. See et al. (2013) have discovered that individuals' attitude determines the effectiveness of nudges. In their study, nudges, which have mismatched the attitude of the participants, are more likely to backfire. Hence, participants' prior preferences in the dating context might cause the opposite effect of defaults.

Moreover, the study also showed that the effectiveness of defaults and warnings varies across situational contexts. Consequently, it is recommended to assess situational characteristics in future nudging studies to determine which nudges are effective in which situational circumstances. Thereby, the appropriateness of other comprehensive situational characteristics taxonomies could be examined, such as Situation 5 or Caption (Ziegler et al., 2019; Parrigon et al., 2017).

Furthermore, the interaction of domain expertise and nudging was investigated, as it was hypothesized that the effectiveness of nudges increases with increasing expertise. Defaults are more effective with increasing probabilistic reasoning, whereas warnings are more effective with increasing financial literacy and probabilistic reasoning. Hence, for financial ISs warnings are recommended, presupposed the users are financially literate. Otherwise, users financial knowledge could be improved.

Overall, the study delivered a valuable contribution to situation-aware IS design. Further studies are needed to investigate other situational contexts and contextual drivers of decision inertia and nudging. Moreover, the individual preferences in situational contexts have to be examined in order to understand their relation to decision inertia and nudges.

Last but not least, the external validity of the study is restricted. The participants were students from the KIT and mostly from the field of business engineering and computer science. A replication of the study with a broader sample could deliver different insights because the populace does not have a comparable understanding of statistics, and moreover, other situational contexts could be more relevant for them.

# Part V

# Finale

# Chapter 9

# Discussion and Conclusion

> "It always seems impossible until it is done."
>
> NELSON MANDELA, 2001

DIGITIZATION is advancing steadily. Companies create new business models or replace older ones with the incorporation of data by producing, analyzing, visualizing, or making decisions based on it. The resulting ISs support various aspects of our lives, from leisure activities, such as traveling and dating, to vital life decisions, such as retirement planning and health insurance. In this process, the role of ISs is growing steadily and permeating more and more all levels of our society. The existing approach to creating a unified UI is reaching its limits.

This thesis stresses out the importance of personalized ISs by investigating two major aspects of decision-making. Firstly, the comprehension of decision-relevant information by investigating adaptive information visualization. Secondly, rational and internally consistent information integration by investigating situation-aware decision inertia and situation-aware adaptive interventions, in the form of nudges. This chapter summarizes the main contributions, discusses its implications, and outlines directions for future research.

## 9.1 Contributions

Chapter 2 reviews and organizes the relevant literature as basis for the experimental investigation in chapter 6, which contributes to the first research question:

**Research Question 1:** *How do users comprehend visualizations, and what factors influence their comprehension?*

Chapter 6 reports on an exploratory eye-tracking study. It was conducted to understand how visualizations are comprehended and which user characteristics might be relevant for personalized information visualization. Participants eye movements were recorded during a digitalized presentation containing different visualization. Thereby, it has been shown that eye-tracking assesses perceptual and cognitive processes with a relatively small sample. The study reveals line charts and tables as common representations in the financial domain. However, none of the investigated visualizations can be preferred across the board since some participants had issues perceiving and comprehending certain visualizations, whereas others adequately comprehended them.

> **Contribution 1.1:** *IS design needs to personalize presented information according to the needs of the users.*

Additionally, chapter 6 shows a possible relation between users' profession and decision-making competency with particular visualizations. This insight is underlined by the review of adaptive information systems in chapter 2.4, in which experience and knowledge are identified as possibilities for visualization recommendation.

> **Contribution 1.2:** *User characteristics are needed that reflect users' prior experience or competencies with particular viualizations.*

The insights of chapters 2 and 6 were incorporated into the design of two experimental studies presented in chapter 7. Both studies were conducted to answer the second research question empirically:

**Research Question 2:** *How do user characteristics influence the comprehension of information with different visualizations?*

The first study evaluated the experimental design and investigated the main assumptions generated by the insights in chapters 6 and 2. Participants had to choose between different investment options that are either represented in a line chart or a table. In addition, the financial and graphical literacy were assessed, and the influences on the visualizations were investigated. The second study replicated the first study and investigated adaptive visualizations more thoroughly by adding the boxplot visualization and additional user characteristics. Both studies show that none of the investigated visualizations are generally preferable. Moreover, the studies reveal the positive influence of financial, graphical, and statistical literacy on financial decision-making.

**Contribution 2.1:** *Financial, graphical, and statistical literacy reflect users' financial decision-making competency.*

The studies in chapter 7 confirm the expectation that line charts and tables are common visualizations in financial decision-making. Furthermore, they show that financial literacy positively influences decision-making with line charts and tables. Most importantly, the effect of financial literacy is stable in both studies despite differences in samples and incentives.

**Contribution 2.2:** *Financial literacy improves decision-making competency with line charts and tables.*

In addition to the effectiveness of decision-making, both studies also examined efficiency by means of cognitive load. Thereby, the subjective cognitive load was measured after the decision-making. As a result, a negative relationship has been shown between statistical literacy and cognitive load.

**Contribution 2.3:** *Statistical literacy lowers subjective cognitive load in financial decision-making.*

Both studies examined the influences of visualizations on cognitive load. While no difference could be observed between the line chart and the table, boxplots cause less subjective cognitive load than tables.

**Contribution 2.4:** *Boxplots cause less subjective cognitive load than tables.*

There are differences in the results between the studies. The first study shows a negative relationship between decision quality and cognitive load. Moreover, it shows a positive influence of graphical literacy on decision-making with line charts. The second study does not replicate the results. It is assumed that the differences in the samples and incentives lead to inconsistencies. In addition, situational characteristics are assumed to measure sampling and incentive differences in participants' situational experiences. This translates into the third research question.

The foundations of situational research and decision inertia are described in the chapters 3 and 4. Moreover, chapter 4 gives an overview of the relevant literature on situational insights and dependencies of decision inertia. Both chapters contribute to the clarification of the third research question:

**Research Question 3:** *How do situational characteristics influence decision inertia?*

The study in chapter 8 hypothesizes that decision inertia varies across situational contexts. In order to investigate the assumption, an established experimental task, namely the urn game, was systematically transferred to different situational contexts in a preliminary study, and a survey was conducted to identify the most different among them (see appendix C).

> **Contribution 3.1:** *Decision situations can be transferred to different contexts using situational characteristics.*

In chapter 8, besides the urn game as the baseline context, three further contexts (robo-advisor, partnership dating, and writing exams) were examined concerning differences in decision inertia. No significant differences across the investigated contexts could be observed. In addition, the initial assumption that situationally perceived negativity increases decision-making inertia is not true.

> **Contribution 3.2:** *Perceived negativity does not influence decision inertia.*

The study reveals gender-specific differences across the contexts. In the robo-advisor and partnership dating context, males are less prone to decision inertia than females. Females are less prone than males in the writing exams context.

> **Contribution 3.3:** *Gender differently influences decision inertia across contexts.*

The study determines perceived confidence as a driver of decision inertia. Perceived confidence varies across situational contexts and has a twofold influence on decision inertia. Overconfidence, the greater judgment of one's abilities than the objective reality, increases decision inertia. Underconfidence, which is the opposite case, also increases decision inertia.

**Contribution 3.4:** *Under- and overconfidence increase decision inertia.*

Finally, the study demonstrates the negative influence of knowledge-based domain expertise. Probabilistic reasoning decreases decision inertia in the urn game context. Additionally, financial literacy decreases decision inertia in the robo-advisor context. Hence, knowledge-based expertise improves users' decision-making competency by reducing decision inertia in the corresponding situational context.

**Contribution 3.5:** *Probabilistic reasoning decreases decision inertia in the context of the urn game.*

**Contribution 3.6:** *Financial literacy decreases decision inertia in the context of robo-advisor.*

Besides investigating situational dependencies of decision inertia, the study in chapter 8 examined the situational dependencies of nudging. For this purpose, the foundations of nudging and the relevant literature on nudging to reduce decision inertia or inertia-related problems are provided in chapter 5. The situational contexts, reported in appendix C, were used to answer the fourth research question:

**Research Question 4:** *How do situational characteristics influence the effectiveness of nudging?*

Appropriate nudges can support rational and internally consistent information integration. The study presented in chapter 8 investigated the effectiveness of warnings and defaults in reducing decision inertia. In contrast to defaults, warnings are an effective means of reducing decision inertia without the differentiation of situational contexts.

**Contribution 4.1:** *Warnings reduce decision inertia.*

The study also examined the situational effectiveness of defaults and warnings. Warnings reduce decision inertia in the robo-advisor context, while defaults do not significantly reduce decision inertia in any of the situational contexts. In the dating context, defaults even cause a significant adverse effect.

**Contribution 4.2:** *Warnings reduce decision inertia in the robo-advisor context.*

**Contribution 4.3:** *Defaults increase decision inertia in the partnership dating context.*

Besides the effectiveness within the situational contexts, the study examined defaults' and warnings' effectiveness across contexts. In both cases, the are significant differences.

**Contribution 4.4:** *The effectiveness of defaults varies across situational contexts.*

**Contribution 4.5:** *The effectiveness of warnings varies across situational contexts.*

The study in chapter 8 relates domain expertise to nudging as a driver of their effectiveness. Probabilistic reasoning increases the effectiveness of warnings, whereas financial literacy positively influences both nudges' effectiveness.

**Contribution 4.6:** *Probabilistic reasoning increases the effectiveness of warnings.*

**Contribution 4.7:** *Financial literacy increases the effectiveness of defaults and warnings.*

Prior studies produced mixed results concerning drivers of decision inertia (Alós-Ferrer et al., 2016; Jung et al., 2019; Jung, 2019). Alós-Ferrer et al. (2016) explored a positive effect of preference for consistency, while Jung et al. (2019) did not find any association despite the same experimental design and comparable sample. This study suggests that the discrepancies are caused by inadequate analysis. Hence, the regression-based insights of prior studies are questionable and must be revised with appropriate analysis (Alós-Ferrer et al., 2016; Jung et al., 2019; Jung, 2019). This study provides an appropriate regression-based analysis method by investigating average marginal effects (Schunck and Nisic, 2020; Brambor et al., 2006).

**Contribution 4.8:** *Provision of an appropriate procedure for the regression analysis of decision inertia.*

## 9.2 Limitations

This thesis contributes to user-adaptive information visualization, situation-aware decision inertia, and situational dependencies of nudging. Due to the experimental methodology, there are limitations and implications for IS design, which are outlined below.

### Adaptive Information Visualization

Chapters 2, 6, and 7 clearly show that visualizations need to be personalized according to the users' needs. Hence, ISs should implement adaptive personalization since users are unaware of their needs, and learning of adaptation possibilities increases users' complexity (Milligan, 2019; Mackay, 1991; Langley, 1999).

Financial ISs can evaluate users' decision-making competency by assessing users' statistical literacy (Cokely et al., 2012), financial literacy (Lusardi and Mitchell, 2007), or graphical literacy (Okan et al., 2012). The processing of the corresponding questionnaires does not require much time. Moreover, the disclosed information is not sensitive in terms of privacy. Especially the statistical literacy questionnaire is promising since it contains only four questions and allows conclusions about users' decision-making competency and cognitive load. If users of financial ISs do not have the required literacy, their knowledge could be built up with micro-videos before the decision-making (Frydenberg and Andone, 2016). Alternatively, the IS could support their users with nudging during the decision-making. Chapter 8 identified the displaying of warning messages as an efficient intervention to alter suboptimal behavior in the financial context.

In the case of visualization recommendations based on user characteristics, the studies in chapter 7 produced mixed results concerning graphical literacy. The first study shows with a broad sample of the population that graphical literacy increases users' decision-making competency with line charts. This result could not be replicated In the second study with a specific student sample. Hence, the question remains whether graphical literacy is an appropriate recommendation characteristic for non-student participants. However, this insight underlines the robustness of financial literacy results. Both studies show that users' decision-making competency with line charts and tables increases with increasing financial literacy. Hence, financial literacy can be used by a broad spectrum of society to recommend line charts or tables in ISs.

When cognitive limitations play a role in decision-making, ISs can measure users' statistical literacy to assess their cognitive limitations and to recommend line charts (chapter 6) or boxplots (chapter 7) since these are less cognitively demanding.

## Decision Inertia

Chapter 8 did not observed contextual differences in decision inertia. However, the question has not yet been answered finally. The preliminary study in appendix C used relative distances of perceived situational characteristics to identify situational contexts that are as different as possible. It is unclear how much contexts have to be different so that differences in decision inertia can occur. Consequently, there could be other contexts with larger differences so that differences in decision inertia can be observed.

The investigated situational contexts could also be examined with a broader or different sample to show differences in decision inertia. After all, the investigation of domain expertise has shown that decision inertia decreases as financial and statistical literacy increases. The study's participants were KIT students and showed relatively high competence with low variance. Hence, a broader sample with more variations concerning domain expertise could lead to significant differences in decision inertia across situational contexts.

The study revealed significant gender-specific differences in decision inertia across situational contexts. General support of all users is out of the question since interventions in already optimal decisions can lead to adverse effects (Czajkowski et al., 2019). This means that ISs, especially in essential financial decisions, should support their users in a gender-specific way. Thereby, financial literacy could be assessed in front of decision-making and, if necessary, built up with learning material (Frydenberg and Andone, 2016). Another possibility is gradual support depending on the literacy level. Besides nudging, especially warning messages, for the short-term reduction of decision inertia, boosting is a promising intervention mechanism to foster competencies in the long term (Hertwig and Grüne-Yanoff, 2017). Hoffrage et al. (2000) showed that boosting by representing statistics in terms of natural frequencies improves Bayesian updating. Hence, the gender-specific representation of the decision situations could be used in ISs to increase decision-making competency sustainably.

Furthermore, both overconfidence and underconfidence have been shown to increase decision inertia. As shown in chapter 8, confidence is a subjective and context-dependent experience in decision-making. It increases for correct decisions and decreases for error decisions (Insabato et al., 2010). In the case of overconfident decision-makers, ISs could give immediate feedback concerning the correctness of made decisions. Also, in this case, warning messages are a proven way in financial ISs. They give information about the current decision and encourage redeliberating it with adjusted confidence.

For underconfident users, ISs should gradually guide them to the decision situation. The difficulty level of the decisions could be gradually increased using practical exercises. With increasing difficulty, the user's confidence can be built up until the level of difficulty needed for the upcoming decision is reached.

### Nudging

Chapter 8 shows that warning messages are an effective means of reducing decision inertia without differentiating situational contexts. Moreover, it shows that their effectiveness depends on the situational context. Hence, when designing IS, and it is unknown which nudges are effective in the decision context, warning messages should be used. The differentiate analysis of the contexts shows that the warnings only lead to a significant reduction of decision inertia in the robo-advisor context, which is in line with Jung (2019). It is of utmost importance to avoid decision inertia in financial decisions, and therefore, the design of ISs that support users to avoid the trap of repeating non-profitable investment decisions. Financial decision support with choice architecture interventions, such as warnings, is a promising and relatively low-cost approach to alter non-favorable behavior without restricting users' freedom of choice.

In contrast to Jung (2019), this research does not support the reduction of decision inertia with default nudges in financial decision-making. Defaults do not reduce decision inertia in any of the investigated situational contexts. On the contrary, they have strong adverse effect in the dating contexts. See et al. (2013) also reported on the adverse effects of nudges and related them to user preferences. They suggest negative effects can occur when the nudge recommendation strongly opposes the user's preferences. It is undeniable that individuals have preferences regarding conversation topics during dates. Sharing similar preferences leads to sympathy and thus increases the probability of a partnership. For instance, couples were found to correlate their political orientation strongly (Klofstad et al., 2013). Consequently, in certain decision situations, individual preferences may override rational cognitive processes to the extent that nudges do not have an effect or have the opposite effect.

Moreover, the results show that knowledge-based domain expertise, namely financial literacy and probabilistic reasoning, improves the effectiveness of defaults and warnings. The underlying assumption is that literate individuals are more likely to trust nudge recommendations (Nekmat, 2020). Building on this, improving the literacy of users is a promising way to ensure the effectiveness of nudges. Alternatively, ISs could implement adaptive interventions based on literacy. For literate users, nudges could be used to reduce suboptimal choice repetition. For novice users, boosting is an alternative approach to avoid undesired decisions and simultaneously improve users' competencies in the given decision situation (Hertwig and Grüne-Yanoff, 2017).

This work contributes to nudging research by providing empirical evidence for the situational effectiveness of defaults and warnings. The insights emphasize the need for further research on the situational effectiveness of nudging. It implies that situational characteristics should be used as a matter of principle in nudging research. Overall, it represents a cornerstone of situational nudging research.

## 9.3 Outlook

This thesis provides a deeper understanding of improving users' decision-making competency by investigating user-adaptive information visualization, situation-aware IS design for decision inertia, and situational effectiveness of nudges in reducing decision inertia. The ideas for future research are presented in the following.

## Future Direction of Adaptive Information Visualization

The studies of chapter 7 are carried out in the context of financial decision-making and show that users' decision-making competency is influenced by financial, statistical, and graphical literacy. However, the generalizability of the findings is limited because they were generated in experimental settings to avoid and control confounding factors. Hence, further research could evaluate the findings in field experiments to increase their external validity.

This thesis generated mixed results concerning the influence of graphical literacy. In the first study of chapter 7, which was conducted with a broad sample of the population, graphical literacy increases users' decision-making competency with line charts. In the second study with a specific student sample, the influence of graphical literacy on decision-making with line charts could not be replicated. The participants of the second study are students from the KIT. The graphical literacy questionnaire revealed that they are highly competent in decision-making with visualizations. Hence, the question arises whether graphical literacy is purposeful for less literate or novice users. Consequently, a further study with less literate participants is needed to clarify the question.

Moreover, expertise is identified as a relevant driver of financial decision-making competency. Building on this insight, increasing users' competency seems to be a promising future pathway. Education and training in front of decision-making is additional effort and increases users' complexity. Hence, user support is needed that enhance adaptive expertise acquisition during decision-making. In an educational settings, collaborative learning models boost students' expertise (Bishop-Clark et al., 2006; Nurhayati et al., 2018). Anthropomorphic chatbots could replace the learning partner in a digital setting. The chatbot could support decision-making with adaptive recommendations depending on the users' competency. Additionally, decision-relevant information could be conveyed adaptively by the chatbot in order to improve users' competency. Morana et al. (2020) already show that anthropomorphic design increases chatbots' social presence and the users' likeliness to follow chatbots' recommendations. Hence, an anthropomorphic chatbot could be integrated in the experimental design, presented in chapter 7. For the first step, participants' competence could be queried in front of the experimental task so that the chatbot adaptively conveys the information needed to fill the competence gap.

## Future Directions of Decision inertia

As already mentioned in section 9.2, there could be other situational contexts with larger differences in situational characteristics so that differences in decision inertia comes into play. The systematic design and identification process, presented in appendix C, can be used to create and identify further situational contexts. Thereby, certain situational characteristics could be focused. Promising characteristics are perceived duty and deception. In certain decision situations, such as the selection of conversation topics in dating scenarios, individuals have strong preferences concerning choice alternative. These preferences could override rational decision-making processes resulting in greater decision inertia. Perceived duty can capture whether someone is intrinsically motivated in a particular situation or perceives the situation as a job that has to be done. Deciding against one's preferences could also be perceived as a kind of cheating, so the characteristic deception could reflect this perception.

This thesis revealed that probabilistic reasoning and financial literacy decrease decision inertia. The participants of the study were students from the KIT. They performed well in both expertise questionnaires. The question arises whether the investigated situational contexts would result in differences in decision inertia with more diverse participants concerning the expertise. Consequently, the experiment could be replicated with another sample.

| State (Probability) | Left Urn | Right Urn | State (Frequency) | Left Urn | Right Urn |
|---|---|---|---|---|---|
| Up (50 %) | ●●●●○○ | ●●○○○○ | Up (1 of 2) | 4 of 6 | 2 of 6 |
| Down (50 %) | ●●○○○○ | ●●●●○○ | Down (1 of 2) | 2 of 6 | 4 of 6 |

FIGURE 9.1: *Natural frequency representation of the urn game.*

This work revealed gender-specific differences in decision inertia across situational contexts. Therefore, the most common situational contexts have to be examined concerning gender differences. Subsequently, possibilities for gender-specific IS support could be explored. Support during decision-making could be achieved by research on gender-specific nudging. Another promising approach is gender-specific representation. Bayesian inferences especially of females can be boosted by natural frequencies (Hoffrage et al., 2000). It is suggested that cognitive statistical reasoning processes evolved to work with natural frequencies. As depicted in figure 9.1, the probabilities of the states and the urn composition can be represented with frequencies. This alternative representation can be coupled with the experimental design of chapter 8 to examine its effect on gender and across situational contexts.

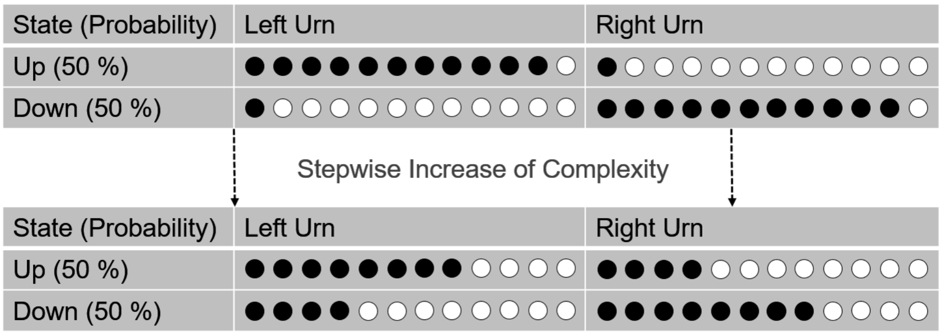



FIGURE 9.2: *Increasing complexity of the urn game.*

Finally, this work identified under- and overconfidence as a driver of decision inertia. This insight implies that users with a realistic view of their abilities are least prone to decision inertia. Confidence is situation-dependent and increases or decreases with experience (Moore and Schatz, 2017). Decision situations with positive outcomes increase it, whereas negative outcomes decrease it. For overconfident individuals, warnings can be used to adjust their confidence during decision-making. For underconfident individuals, increasing their confidence with incremental complexity increase seems promising. For this purpose, the experimental task (urn game) can be modified. In order to gain different levels of complexity, more balls can be added to each urn. During the experiment, beginning with the lowest complexity (see figure 9.2 - top urn composition), the complexity can be stepwise increased by changing the color of one ball in each urn until the original complexity is reached (see figure 9.2 - bottom urn composition).

## Future Directions of Nudging

This thesis identified warnings as an effective means of reducing decision inertia, especially in the context of robo-advisory. However, the generalizability of the insights is restricted since the results are generated in experimental settings. Hence, future field studies are needed to ensure external validity.

Default nudges increased decision inertia in the context of partnership dating. It is assumed that in certain decision situations, where users have preferences regarding choice alternatives, strongly opposing nudge recommendations have no effect and may even have a negative effect in the worst case See et al. (2013). Thus, further research is needed which relates individual preferences to effectiveness of nudges. For this purpose, the experimental design presented in chapter 8 could be used and enriched with questionnaires for preference identification. The preference for the choice alternative should be evaluated before the experimental task, as processing the task could change the prior preferences.

Another finding of this thesis that needs to be investigated in future research is the situational effectiveness of defaults and warnings. This finding indicates that further fundamental research on situational nudging is necessary. The present results only contribute to the situational dependency of defaults and warnings in reducing decision inertia. Hence, future studies are needed to examine the situational dependencies of other nudges, such as simplifications, disclosures, and social references. Particularly, the investigation of social references seems promising since Lehner et al. (2016) already observed large differences in their effectiveness across domains.

In addition, the situational effectiveness of nudges can be investigated with other cognitive biases. Fleischmann et al. (2014) reviewed the extant literature on cognitive biases in ISs and identified framing, anchoring, and negativity bias as the most relevant. For the situation-aware research of nudge effectiveness, the quantitative literature review of Hummel and Maedche (2019) provides valuable information on the most important contexts in ISs and the most used nudges in the respective contexts. For the investigation of further nudges and the investigation of other cognitive biases, the experimental design of chapter 8 can be adapted.

Finally, this thesis revealed the positive influence of financial literacy on the effectiveness of warnings and defaults. It is interesting to know whether financial literacy influences other nudges' effectiveness. Social references and reminders could be studied next, as they are widely used in financial decision-making (Hummel and Maedche, 2019).

# Part VI

# Appendices

# Appendix A

# Slides and Eye-Tracking Metrics

This appendix contains the presentation slides and the corresponding eye-tracking metrics of the experiment in chapter 6. As the content of the presentation is subject to confidentiality agreements, the slides has been redesigned for the purpose of publication. The type of elements and their arrangement were retained. The texts and the contents of the graphics have been modified.

## A.1 Presentation Slides

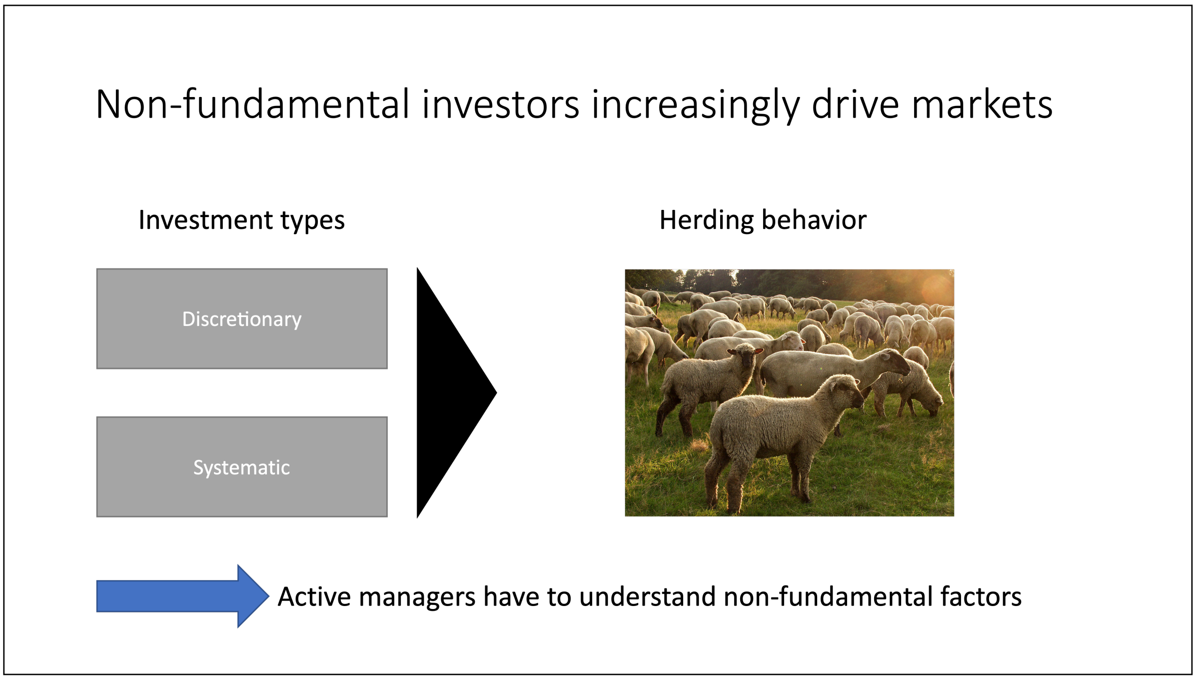



FIGURE A.1: *Slide two of the presentation*


FIGURE A.2: *Slide three of the presentation*

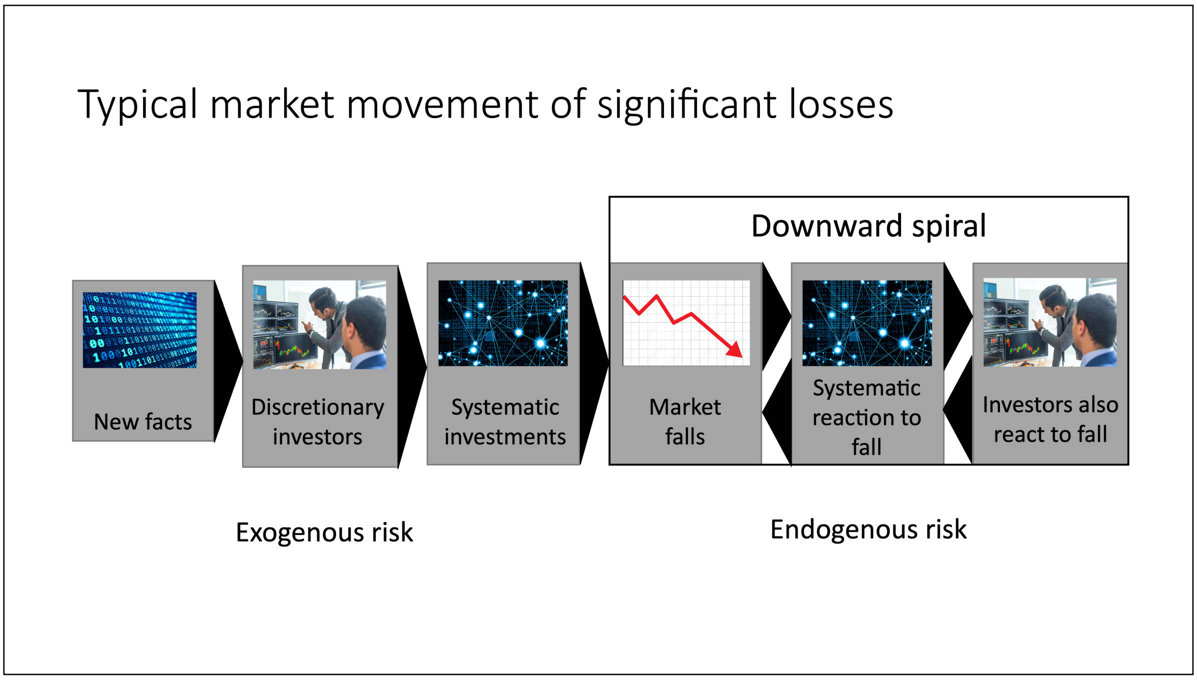



FIGURE A.3: *Slide four of the presentation*

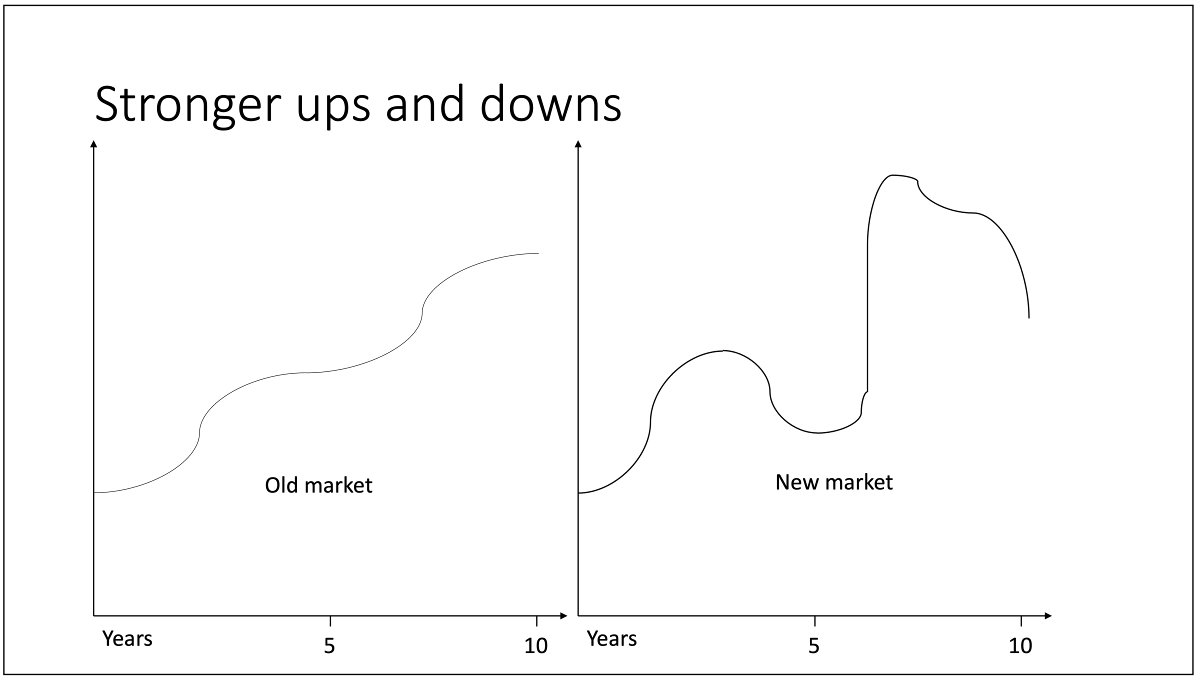



FIGURE A.4: *Slide five of the presentation*

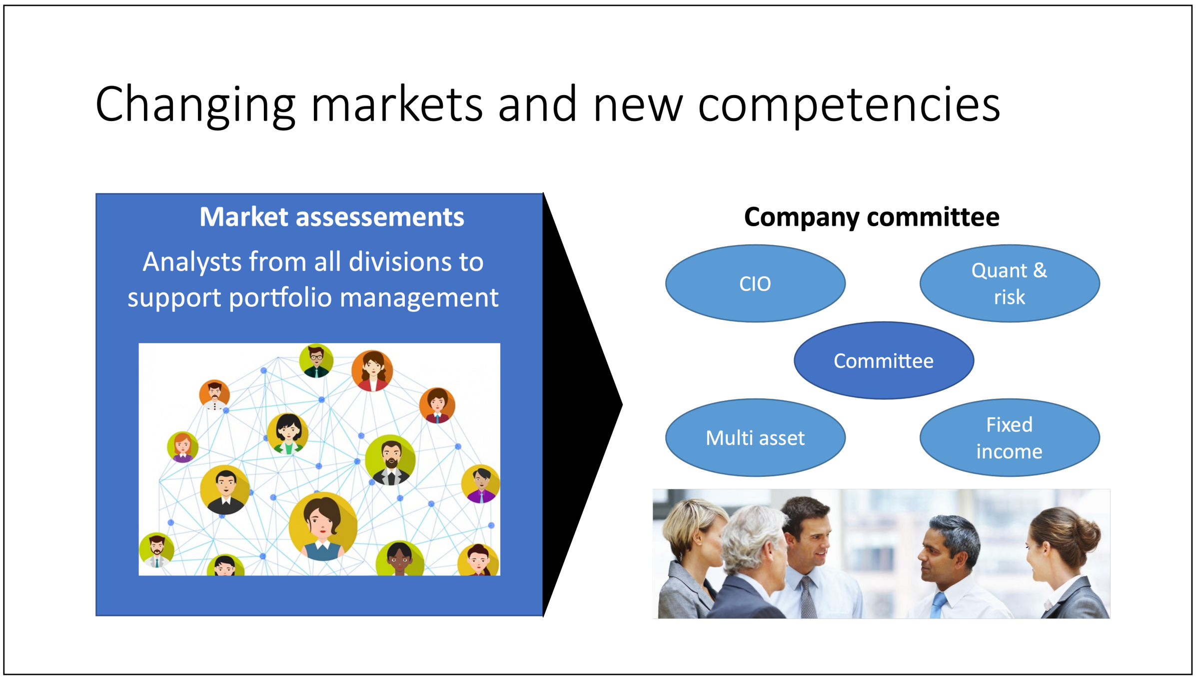



FIGURE A.5: *Slide six of the presentation*

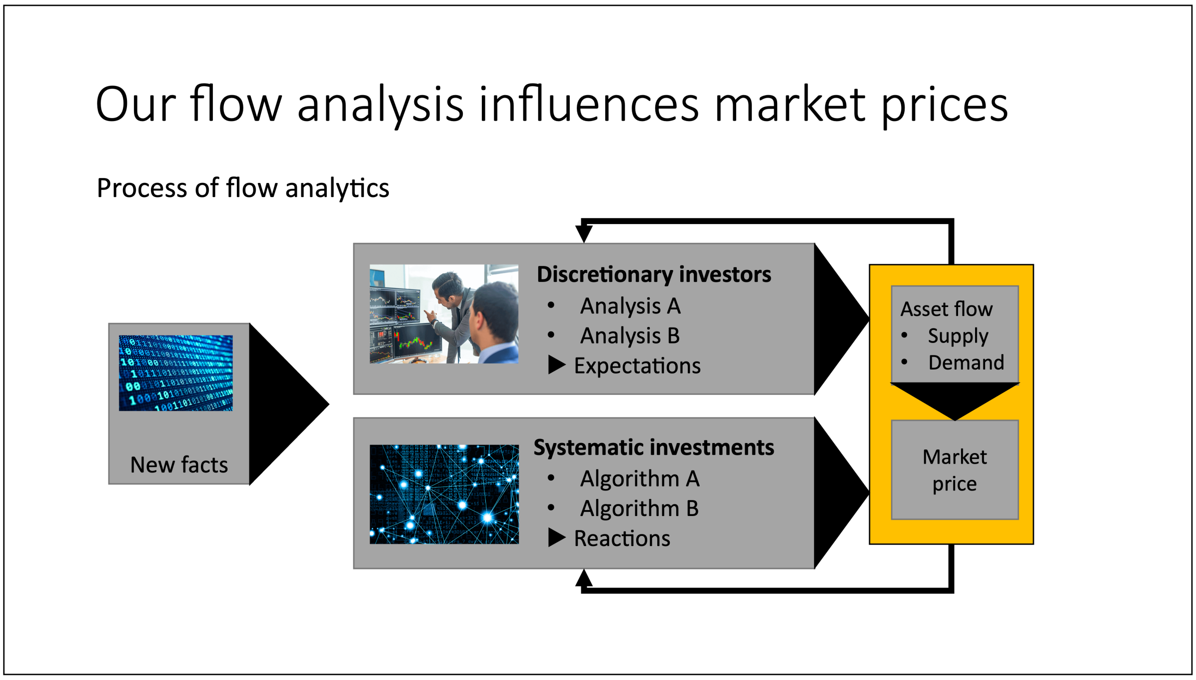



FIGURE A.6: *Slide seven of the presentation*

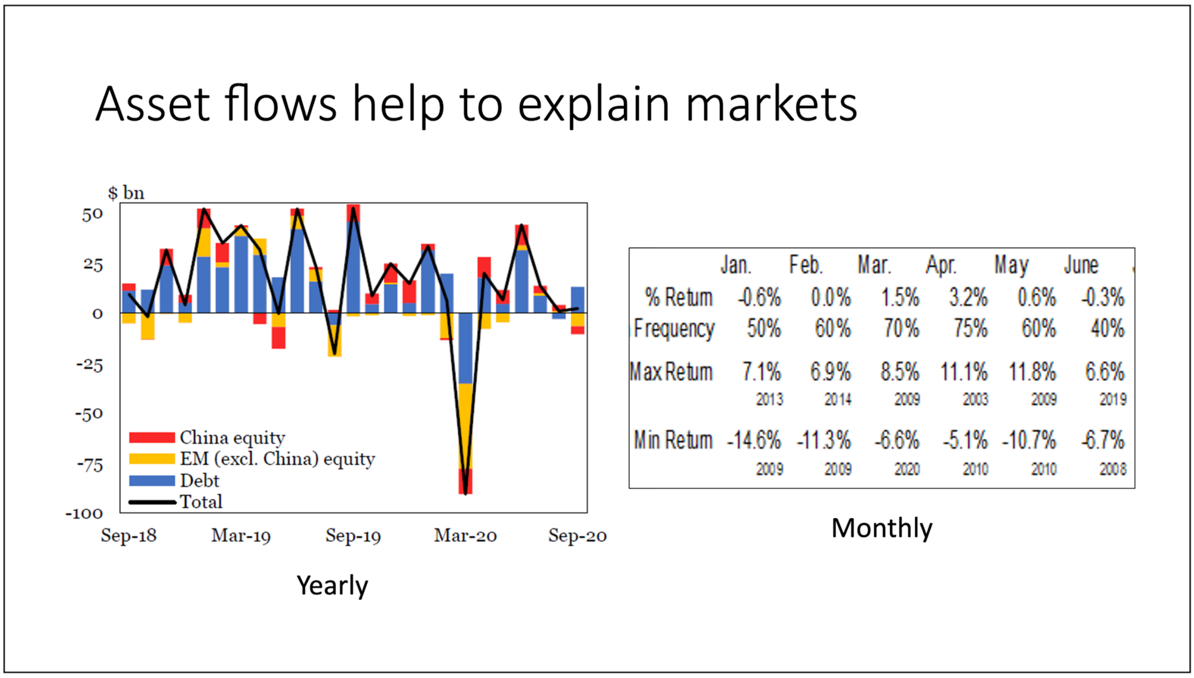



FIGURE A.7: *Slide eight of the presentation*


FIGURE A.8: *Slide nine of the presentation*

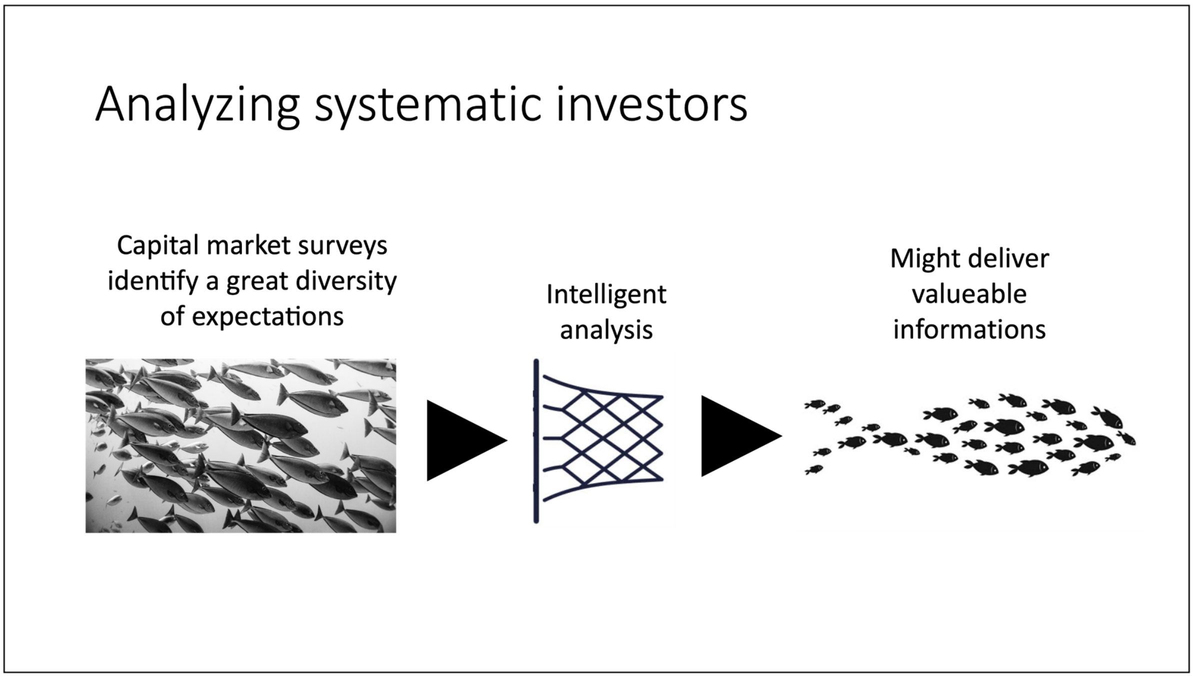



FIGURE A.9: *Slide ten of the presentation*

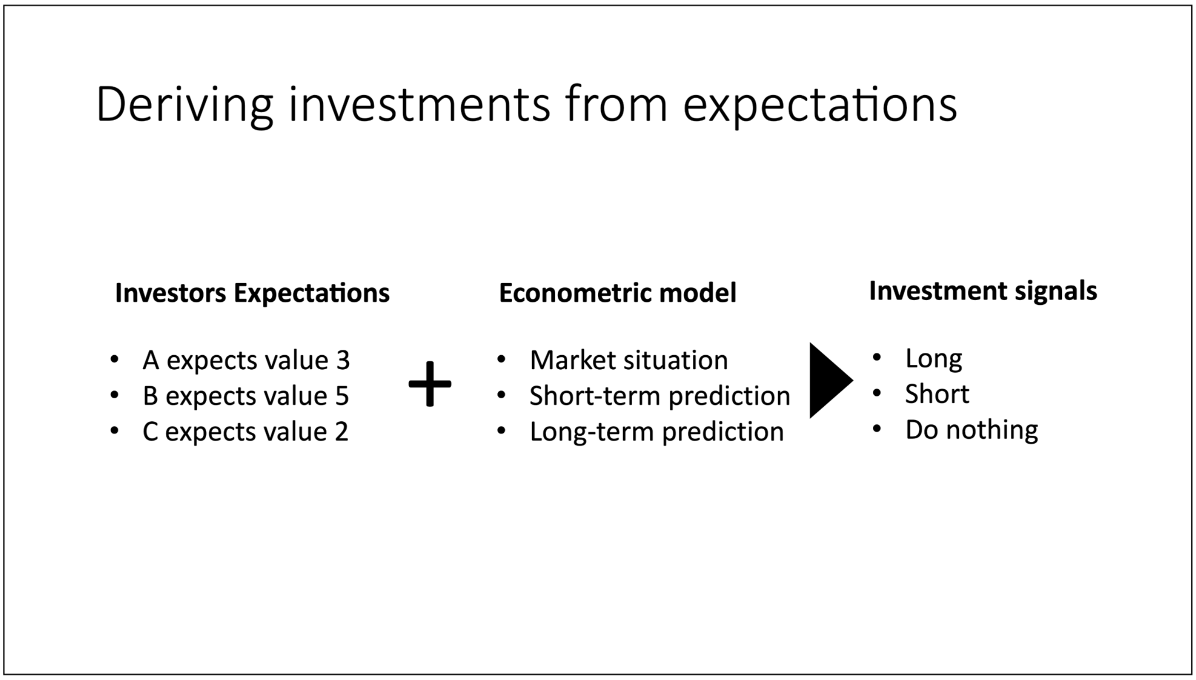



FIGURE A.10: *Slide eleven of the presentation*

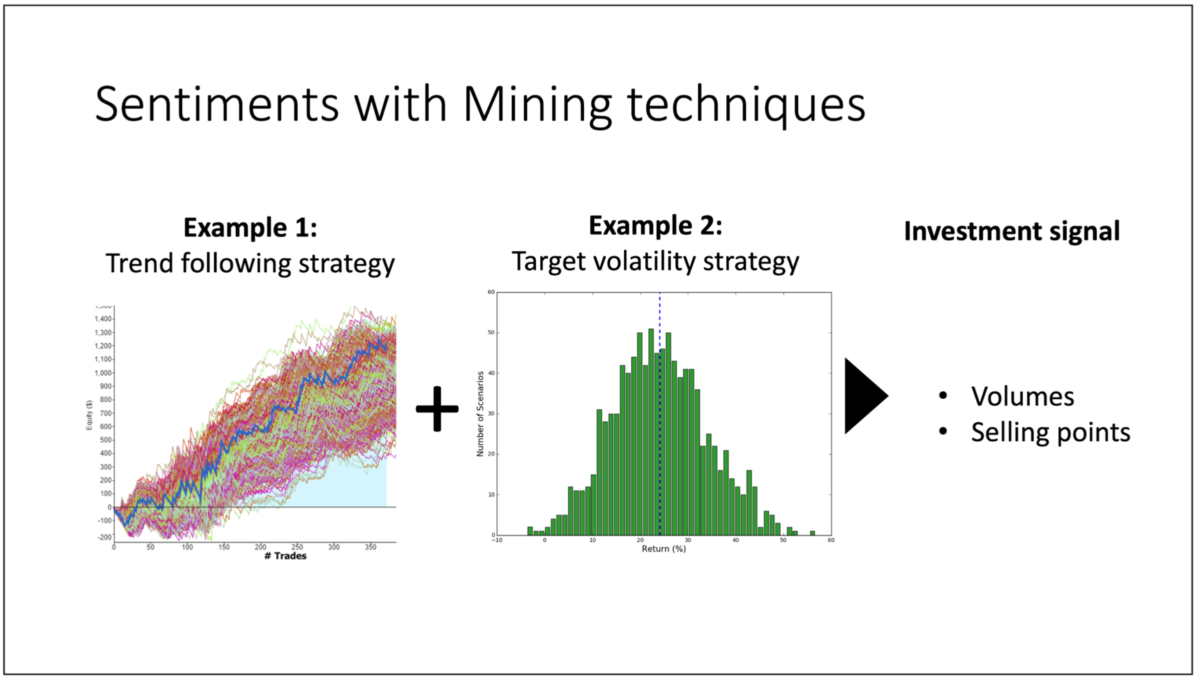



FIGURE A.11: *Slide twelve of the presentation*


FIGURE A.12: *Slide thirteen of the presentation*

## A.2 Eye-Tracking Metrics

The following table provides the detailed eye-tracking metrics of the presentation's visual elements used to investigate perceptual and cognitive processes.

TABLE A.1: *The eye-tracking metrics of each visual element.*

| | **Part 1** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Slide 2** | | | | **Slide 3** | | | | **Slide 4** | | | |
| **Area of Interest** | **Fixated** | **Relative duration** | **Relative frequency** | **Sequence** | **Fixated** | **Relative duration** | **Relative frequency** | **Sequence** | **Fixated** | **Relative duration** | **Relative frequency** | **Sequence** |
| Element 1 | 6 | 0.04 | 0.09 | 4 | 6 | 0.025 | 0.045 | 4 | 6 | 0.06 | 0.07 | 1 |
| Element 2 | 6 | 0.05 | 0.08 | 2 | 6 | 0.14 | 0.19 | 1 | 6 | 0.065 | 0.055 | 2 |
| Element 3 | 6 | 0.27 | 0.36 | 1 | 6 | 0.22 | 0.22 | 3 | 6 | 0.1 | 0.09 | 3 |
| Element 4 | 6 | 0.34 | 0.3 | 3 | 6 | 0.26 | 0.23 | 5 | 6 | 0.115 | 0.095 | 3 |
| Element 5 | 6 | 0.28 | 0.14 | 5 | 6 | 0.16 | 0.145 | 6 | 5 | 0.05 | 0.05 | 7 |
| Element 6 | 5 | 0.02 | 0.03 | 6 | 6 | 0.11 | 0.11 | 7 | 6 | 0.14 | 0.16 | 6 |
| Element 7 | - | - | - | - | - | - | - | - | 6 | 0.21 | 0.2 | 8 |
| Element 8 | - | - | - | - | - | - | - | - | 6 | 0.17 | 0.15 | 9 |
| Element 9 | - | - | - | - | - | - | - | - | 6 | 0.08 | 0.1 | 5 |
| Element 10 | - | - | - | - | - | - | - | - | 4 | 0.01 | 0.03 | 10 |
| Element X | - | - | - | - | 6 | 0.085 | 0.06 | 2 | - | - | - | - |

| | **Part 2** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Slide 5** | | | | **Slide 6** | | | | **Slide 7** | | | |
| **Area of Interest** | **Fixated** | **Relative duration** | **Relative frequency** | **Sequence** | **Fixated** | **Relative duration** | **Relative frequency** | **Sequence** | **Fixated** | **Relative duration** | **Relative frequency** | **Sequence** |
| Element 1 | 5 | 0.01 | 0.06 | 2 | 6 | 0.01 | 0.17 | 3 | 6 | 0.06 | 0.12 | 1 |
| Element 2 | 6 | 0.25 | 0.31 | 1 | 6 | 0.41 | 0.37 | 2 | 6 | 0.36 | 0.38 | 2 |
| Element 3 | 6 | 0.56 | 0.44 | 3 | 6 | 0.54 | 0.4 | 1 | 6 | 0.42 | 0.23 | 5 |
| Element 4 | 6 | 0.18 | 0.19 | 4 | 5 | 0.04 | 0.06 | 4 | 6 | 0.07 | 0.09 | 6 |
| Element 5 | - | - | - | - | - | - | - | - | 4 | 0.02 | 0.03 | 7 |
| Element X | - | - | - | - | - | - | - | - | 5 | 0.03 | 0.06 | 4 |
| Element Y | - | - | - | - | - | - | - | - | 6 | 0.04 | 0.09 | 3 |

| | **Part 3** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Slide 8** | | | | **Slide 9** | | | | **Slide 10** | | | |
| **Area of Interest** | **Fixated** | **Relative duration** | **Relative frequency** | **Sequence** | **Fixated** | **Relative duration** | **Relative frequency** | **Sequence** | **Fixated** | **Relative duration** | **Relative frequency** | **Sequence** |
| Element 1 | 6 | 0.09 | 0.14 | 1 | 5 | 0.08 | 0.27 | 2 | 5 | 0.07 | 0.27 | 1 |
| Element 2 | 6 | 0.28 | 0.23 | 3 | 6 | 0.39 | 0.21 | 1 | 6 | 0.29 | 0.22 | 2 |
| Element 3 | 6 | 0.53 | 0.26 | 2 | 6 | 0.34 | 0.31 | 3 | 6 | 0.38 | 0.3 | 4 |
| Element 4 | 6 | 0.02 | 0.31 | 4 | 6 | 0.19 | 0.21 | 4 | 6 | 0.26 | 0.21 | 3 |
| Element 5 | 6 | 0.08 | 0.06 | 5 | - | - | - | - | - | - | - | - |

| | **Part 4** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Slide 11** | | | | **Slide 12** | | | | **Slide 13** | | | |
| **Area of Interest** | **Fixated** | **Relative duration** | **Relative frequency** | **Sequence** | **Fixated** | **Relative duration** | **Relative frequency** | **Sequence** | **Fixated** | **Relative duration** | **Relative frequency** | **Sequence** |
| Element 1 | 6 | 0.02 | 0.09 | 2 | 6 | 0.02 | 0.13 | 1 | 5 | 0.01 | 0.01 | 3 |
| Element 2 | 6 | 0.38 | 0.405 | 1 | 6 | 0.46 | 0.41 | 2 | 6 | 0.37 | 0.41 | 1 |
| Element 3 | 6 | 0.43 | 0.305 | 3 | 6 | 0.41 | 0.32 | 3 | 6 | 0.14 | 0.18 | 4 |
| Element 4 | 6 | 0.17 | 0.2 | 4 | 6 | 0.11 | 0.14 | 4 | 6 | 0.39 | 0.24 | 5 |
| Element X | - | - | - | - | - | - | - | - | 5 | 0.09 | 0.16 | 2 |

# Appendix B

# Data Catalogues and Regression Tables of Chapter 7

This appendix contains data catalogues and regression tables of chapter 7. In addition, the compact disc attached to the last page of this thesis contains both the data and the analyses using the R programming language.

## B.1 Data Catalogue of Study 1

The following table matches the variables of the thesis with the variables of the R-script. Moreover, it contains a short description and the source of variables.

| Name CSV / R-script | Name Thesis | Meaning | Source | Note / Comment |
|---|---|---|---|---|
| id..Antwort.ID | - | Unique identifier of participant | - | - |
| Treatment | Representation | Treatment variable for representation | - | Treatments: Tabular line chart |

| | | | | |
|---|---|---|---|---|
| GK | Graphical literacy | Participant's score of questionnaire | Galesic and Garcia-Retamero (2012) | Number of correct answers |
| FW | Financial literacy | Participant's score of questionnaire | Lusardi et al. (2007) | Number of correct answers |
| Weiblich | Gender | Dummy variable of gender | - | - |
| Age | Age | Age of participant | - | - |
| Time | - | Processing time of experiment | - | Seconds |
| DE3..Sind. Sie.Student. | Student | Dummy variable of student | - | - |
| Quiz | Quiz result | Questions to ensure comprehension of experimental task | - | Number of correct answers |
| CL | Cognitive Load | Subjective cognitive load (NASA-TLX) | Hoonakker et al. (2011) | Average of six subscales |
| high_FW | Financially literate | Dummy variable of financially literate participants | Iacobucci et al. (2015) | Participant's score above median |
| low_FW | Financiall non-literate | Dummy variable of financially non-literate | Iacobucci et al. (2015) | Participant's score below median |
| high_GK | Grahically literate | Dummy variable of graphically literate | Klapper et al. (2012) | Participant's score above median |

| | | | | |
|---|---|---|---|---|
| low_GK | Graphically non-literate | Dummy variable of graphically non-literate | Klapper et al. (2012) | Participant's score below median |
| v | Line chart | Dummy variable of line chart | - | - |
| t | Tabular representation | Dummy variable of tabular representation | - | - |
| Correct | Rational investment decisions | Participant's score of experimental task | Rudolph et al. (2009) | Number of correct investment decisions |

## B.2 Data Catalogue of Study 2

The following table matches the variables of the thesis with the variables of the R-script. Moreover, it contains a short description and the source of variables.

| **Name CSV / R-script** | **Name Thesis** | **Meaning** | **Source** | **Note / Comment** |
|---|---|---|---|---|
| ProbandID | Participant_ID | Unique identifier of participant | - | - |
| Ausbildung | Education | Highest education of participants | - | - |
| Geschlecht | Gender | Gender of participants | | |
| Altersgruppe | Age | Age group of pariticipants | - | - |
| Quiz1 - Quiz4 | Quiz questions | Comprehension question for experimental task | - | - |

| | | | | |
|---|---|---|---|---|
| quiz1r - quiz4r | Quiz result | Correctly answered question | - | - |
| F1 - F6 | 6 Investment scenarios line chart | Answer of participants | - | - |
| TT1 - TT6 | 6 Investment scenarios tabular | Answers of participants | - | - |
| B1 - B6 | 6 Investment scenarios boxplot | Answers of participants | - | - |
| NASATLX [SQ001] - NASATLX [SQ006] | Cognitive Load Items | Subjective cognitive load (NASA-TLX) | Hoonakker et al. (2011) | Pariticpant's Likert scale answer |
| Balkendiagramm [SQ001] | Graphical literacy item 1 | - | Galesic and Garcia-Retamero (2012) | Participant's answer |
| Balkendiagramm [SQ002] | Graphical literacy item 2 | - | Galesic and Garcia-Retamero (2012) | Participant's answer |
| Kuchendiagramm [SQ001] | Graphical literacy item 3 | - | Galesic and Garcia-Retamero (2012) | Participant's answer |
| Kuchendiagramm [SQ002] | Graphical literacy item 4 | - | Galesic and Garcia-Retamero (2012) | Participant's answer |

| | | | | |
|---|---|---|---|---|
| GeradeAbso [SQ001] | Graphical literacy item 5 | - | Galesic and Garcia-Retamero (2012) | Participant's answer |
| GeradeAbso [SQ002] | Graphical literacy item 6 | - | Galesic and Garcia-Retamero (2012) | Participant's answer |
| GeradeAbso2 | Graphical literacy item 7 | - | Galesic and Garcia-Retamero (2012) | Participant's answer |
| Bankenvgl | Graphical literacy item 8 | - | Galesic and Garcia-Retamero (2012) | Participant's answer |
| BankvglLine | Graphical literacy item 9 | - | Galesic and Garcia-Retamero (2012) | Participant's answer |
| Kreditvgl | Graphical literacy item 10 | - | Galesic and Garcia-Retamero (2012) | Participant's answer |
| Studienabbruch | Graphical literacy item 11 | - | Galesic and Garcia-Retamero (2012) | Participant's answer |
| FraMaZahl [SQ001] | Graphical literacy item 12 | - | Galesic and Garcia-Retamero (2012) | Participant's answer |

| | | | | |
|---|---|---|---|---|
| FraMaZahl [SQ002] | Graphical literacy item 13 | - | Galesic and Garcia-Retamero (2012) | Participant's answer |
| FK1 - FK8 | 8 financial literacy items | - | Lusardi et al. (2007) | Participant's answers |
| MiniIPPIP [SQ001] - MiniIPPIP [SQ021] | 20 Big-Five items | - | Donnellan et al. (2006) | Pariticpant's Likert scale answer |
| BN1 - BN4 | 4 statistical literacy items | - | Cokely et al. (2012) | Participant's answers |
| MS[SQ001] - MS[SQ013] | 12 maximizing items | - | Schwartz et al. (2002) | Pariticpant's Likert scale |
| Interviewtime | Processing time | - | - | Seconds |
| Treatment | Treatment variable of representation | - | - | Graph, Table or Box |
| F1C - F6C | 6 rational investment decisions line charts | Correcntess of each scenario in line chart treatment | - | - |
| T1C - T6C | 6 rational investment decisions tabular | Correcntess of each scenario in tabular treatment | - | - |

| | | | | |
|---|---|---|---|---|
| B1C - B6C | 6 rational investment decisions boxplots | Correcntess of each scenario in boxplot treatment | - | - |
| Round1 - Round6 | Round number | Boolean variable for each round | - | - |
| correct_Sum | Rational investment decision | Aggregation of rational investment decisions | | Number of rational investment decisions |
| GL1 - GL13 | Graphical literacy | Boolean variable for the correctness of each graphical literacy item | Galesic and Garcia-Retamero (2012) | - |
| GL_Sum | Graphical literacy | Aggregation of graphical literacy items | Galesic and Garcia-Retamero (2012) | Number of correct answers |
| FL1 - FL8 | Financial literacy | Aggregation of financial literacy items | Lusardi et al. (2007) | Number of correct answers |
| Extraversion1 - Extraversion4 | 4 Extraversion items of Big Five | - | Donnellan et al. (2006) | Pariticpant's Likert scale answer |
| Agreeableness1 - Agreeableness4 | 4 Agreeableness items of Big Five | - | Donnellan et al. (2006) | Pariticpant's Likert scale answer |
| Neuroticism1 - Neuroticism4 | 4 Neuroticism items of Big Five | - | Donnellan et al. (2006) | Pariticpant's Likert scale answer |

| | | | | |
|---|---|---|---|---|
| Conscientiousness1 - Conscieniousness4 | 4 Conscientiousness items of Big Five | - | Donnellan et al. (2006) | Pariticpant's Likert scale answer |
| Openness1 - Openness4 | 4 Openness items of Big Five | - | Donnellan et al. (2006) | Pariticpant's Likert scale answer |
| Extraversion | Extraversion | - | Donnellan et al. (2006) | Mean value of items |
| Agreeableness | Agreeableness | - | Donnellan et al. (2006) | Mean value of items |
| Neuroticism | Neuroticism | - | Donnellan et al. (2006) | Mean value of items |
| Conscientiousness | Conscientiousness | - | Donnellan et al. (2006) | Mean value of items |
| Openness | Openness | - | Donnellan et al. (2006) | Mean value of items |
| BN1C - BN4C | Statistical literacy | Correctntess of each item | Cokely et al. (2012) | Boolean for correct |
| BN_correct | Statistical literacy | Aggregation of statistical literacy | Cokely et al. (2012) | Number of correct answers |
| MS1 - MS13 | Maximizing | Likerts scale points of each item | Schwartz et al. (2002) | Participants answer |
| MS | Maximizing | Aggregated value | Schwartz et al. (2002) | Summation of Likerts scale points |
| Box | Boxplot treatment | Dummy variable | - | - |
| Graph | Line chart treatment | Dummy variable | - | - |

| | | | | |
|---|---|---|---|---|
| Table | Tabular treatment | Dummy variable | - | - |

## B.3 Regression Tables of Study 2

This section provides the regression tables of the second study.

| | H1 | | | H2 | | |
|---|---|---|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** |
| Intercept | -1.068 | 0.656 | 0.103 | -0.437 | 0.808 | 0.588 |
| Round Number | -0.151 | 0.045 | 0.001 *** | -0.151 | 0.045 | 0.001 *** |
| Financial Literacy | 0.280 | 0.089 | 0.001 ** | - | - | - |
| Graphical Literacy | - | - | - | 0.130 | 0.075 | 0.084 + |
| **Deviance** | **1019.7** | | | **1026.4** | | |

Number of Observations: 774; Random-effect: Participant-id; Number of Participants: 129; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE B.3: *Logistic regression on rational investment decision for the analysis of financial literacy (H1) and graphical literacy (H2)*

| | H3 | | | H4 | | |
|---|---|---|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** |
| Intercept | -0.517 | 0.898 | 0.564 | -0.712 | 0.785 | 0.364 |
| Round Number | -0.151 | 0.045 | 0.001 *** | -0.151 | 0.045 | 0.001 *** |
| Graphical Literacy | 0.138 | 0.084 | 0.102 | - | - | - |
| Financial Literacy | - | - | - | 0.229 | 0.107 | 0.033 * |
| Line Chart | 0.407 | 2.004 | 0.838 | -1.114 | 1.369 | 0.415 |
| Graphical Literacy x Line Chart | -0.038 | 0.190 | 0.837 | - | - | - |
| Financial Literacy x Line Chart | - | - | - | 0.158 | 0.191 | 0.405 |
| **Deviance** | **1026.3** | | | **1019** | | |

Number of Observations: 774; Random-effect: Participant-id; Number of Participants: 129; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE B.4: *Logistic regressions on rational investment decision for the interaction analysis of line chart with graphical literacy (H3) and financial literacy (H4).*

| | H5 | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | -1.211 | 0.811 | 0.135 |
| Round Number | -0.151 | 0.045 | 0.001 *** |
| Financial Literacy | 0.318 | 0.112 | 0.004 ** |
| Tabular | 0.334 | 1.299 | 0.796 |
| Financial Literacy x Tabular | -0.096 | 0.180 | 0.593 |
| **Deviance** | **1015.5** | | |

Number of Observations: 774; Random-effect: Participant-id; Number of Participants: 129; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE B.5: *Logistic regression on rational investment decision for the interaction analysis of tabular with financial literacy (H5).*

| | H8 | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | 0.929 | 0.185 | $<$ 0.001 *** |
| Round Number | -0.151 | 0.045 | $<$ 0.001 *** |
| Maximizing | 0.006 | 0.090 | 0.947 |
| **Deviance** | **1029.3** | | |

Number of Observations: 774; Random-effect: Participant-id; Number of Participants: 129; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE B.6: *Logistic regression on rational investment decision for the analysis of maximizing (H8).*

| | H10 | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | 0.811 | 0.471 | 0.085 + |
| Round Number | -0.151 | 0.045 | < 0.001 *** |
| Conscientiousness | 0.030 | 0.110 | 0.785 |
| **Deviance** | | **1029.3** | |

Number of Observations: 774; Random-effect: Participant-id; Number of Participants: 129; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE B.7: *Logistic regression on rational investment decision for the analysis of conscientiousness (H10).*

| | H12 | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | 0.411 | 0.226 | 0.069 + |
| Round Number | -0.151 | 0.045 | < 0.001 *** |
| Statistical Literacy | 0.229 | 0.063 | < 0.001 *** |
| **Deviance** | | **1029.3** | |

Number of Observations: 774; Random-effect: Participant-id; Number of Participants: 129; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE B.8: *Logistic regression on rational investment decision for the analysis of statistical literacy (H12).*

| | H15 | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | -0.846 | 0.924 | 0.359 |
| Round Number | -0.151 | 0.045 | 0.001 *** |
| Graphical Literacy | 0.160 | 0.087 | 0.066 + |
| Boxplot | 1.927 | 1.802 | 0.284 |
| Graphical Literacy x Boxplot | -0.147 | 0.170 | 0.385 |
| **Deviance** | | **1022.1** | |

Number of Observations: 774; Random-effect: Participant-id; Number of Participants: 129; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE B.9: *Logistic regression on rational investment decision for the interaction analysis of boxplot with graphical literacy (H15).*

# Appendix C

# Design and Identification of Situational Contexts

In front of the study in chapter 8, a preliminary study was conducted in order to construct and identify different situational contexts of the urn game.

The urn game is based on the dual choice paradigm (see section 8.3), and is established to investigate decision inertia (Alós-Ferrer et al., 2016; Jung et al., 2019; Jung and Dorner, 2018). The goal of the main study is to investigate decision inertia across situational contexts. Jung et al. (2018) transferred the urn game to the robo-advisor context. Their approach is used to develop further situational contexts. The instructions are changed as little as possible and as much as necessary to obtain comparability.

Three different situational contexts of the urn game are needed for the main study. The robo-advisor context is adapted from Jung et al. (2018) to obtain comparability. Moreover, the urn game is included as the baseline context. This study aims to identify two highly contrasted contexts with respect to the situational characteristics (DIAMONDS) (Rauthmann et al., 2014). The DIAMONDS taxonomy measures the situational characteristics with the dimensions *duty, intellect, adversity, mating, positivity, negativity, deception,* and *sociality*. According to Rauthmann and Sherman (2015a), situations are perceived as different when their profile of situational characteristics differs significantly. Moreover, contexts summarize situations with similar profiles of characteristics (Rauthmann et al., 2014).

## C.1 Procedure of Design

As orientation for the identification of possible contexts, the context taxonomy of Van Heck (1989) is used. It is the most useful taxonomy to classify everyday life situations (Ten Berge and De Raad, 1999) and contains ten contexts (*interpersonal conflict, trading, joint working and information exchange, serving, intimacy and interpersonal relations, excesses, recreation, sport, traveling,* and *rituals*).

| Urnenspiel | Robo-Advisor | Dating | Klausur |
|---|---|---|---|
| Stellen Sie sich vor, Sie nehmen an einem Urnenspiel teil. | Stellen Sie sich vor, Sie investieren Ihr Geld mit einem Robo-Advisor. | Stellen Sie sich vor, Sie nutzen eine Blind Dating App. | Stellen Sie sich vor, Sie absolvieren ein Studium. |
| Das Urnenspiel hat 60 Runden. | Es gibt 60 Investitionsrunden. | Es gibt 60 Chatrunden. | Es gibt 60 Lehrveranstaltungen. |
| In jeder Runde gibt es zwei Ziehungen (mit Zurücklegen). | In jeder Investitionsrunde gibt es zwei Investitionen. | In jeder Chatrunde gibt es zwei Chats. | Jede Lehrveranstaltung wird anhand von zwei Klausuren geprüft. |
| Bei jeder Ziehung wählen Sie zwischen zwei Urnen. | Bei jeder Investition wählen Sie zwischen zwei Portfolios. | Bei jedem Chat wählen Sie zwischen zwei Themenmixen. | Bei jeder Klausur wählen Sie zwischen zwei Lernstrategien. |
| Die linke und die rechte Urne enthalten schwarze und weiße Kugeln in einem unterschiedlichen Verhältnis. | Die zwei Portfolios, das Anleihenportfolio und das Aktienportfolio, enthalten Anleihen und Aktien in einem unterschiedlichen Verhältnis. | Die zwei Themenmixe, der Arbeitsmix und der Modemix, enthalten Arbeit und Mode in einem unterschiedlichen Verhältnis. | Die zwei Lernstrategien, Theoriestrategie und Transferstrategie, enthalten Theorie und Transfer in einem unterschiedlichen Verhältnis. |
| Wenn Sie eine schwarze Kugel ziehen, erhöht sich Ihre Auszahlung um 10 Cent. | Wenn Sie erfolgreich investieren, erhöht sich Ihre Auszahlung um 10 Cent. | Wenn Ihrer/m Chatpartner:in der Chat gefällt, erhöht sich Ihre Auszahlung um 10 Cent. | Wenn Sie eine Klausur bestehen, erhöht sich Ihre Auszahlung um 10 Cent. |
| Welche Urne eine höhere Wahrscheinlichkeit dafür birgt, eine schwarze Kugel zu ziehen, hängt vom Umweltzustand ab. | Welches Portfolio eine höhere Wahrscheinlichkeit dafür birgt, erfolgreich zu investieren, hängt vom Marktzustand ab. | Welcher Themenmix eine höhere Wahrscheinlichkeit dafür birgt, Ihrer/m Chatpartner:in zu gefallen, hängt vom Beruf der/des Chatpartner:in ab. | Welche Lernstrategie eine höhere Wahrscheinlichkeit dafür birgt, eine Klausur zu bestehen, hängt vom Aufgabentyp ab. |
| Es gibt zwei mögliche Umweltzustände, Up und Down. | Es gibt zwei mögliche Marktzustände, Bärenmarkt und Bullenmarkt. | Es gibt zwei mögliche Berufe der Chatpartner:innen, Influencer:in und Business Manager:in. | Es gibt zwei mögliche Aufgabentypen, Theorieaufgaben und Transferaufgaben. |
| Sie treten zufällig und gleich häufig auf. | Sie treten zufällig und gleich häufig auf. | Sie treten zufällig und gleich häufig auf. | Sie treten zufällig und gleich häufig auf. |
| Der Umweltzustand wird vor jeder Runde neu festgelegt und ändert sich innerhalb einer Runde nicht. | Der Marktzustand wird vor jeder Investitionsrunde neu festgelegt und ändert sich innerhalb einer Investitionsrunde nicht. | Die/der Chatpartner:in wird vor jeder Chatrunde neu festgelegt und ändert sich innerhalb einer Chatrunde nicht. | Der Aufgabentyp wird für jede Lehrveranstaltung neu festgelegt und ändert sich innerhalb einer Lehrveranstaltung nicht. |
| Das heißt, Sie können innerhalb einer Runde von der ersten Ziehung auf die zweite Ziehung Rückschlüsse auf den Umweltzustand ziehen, jedoch nie von Runde zu Runde. | Das heißt, Sie können innerhalb einer Investitionsrunde von der ersten Investition auf die zweite Investition Rückschlüsse auf den Marktzustand ziehen, jedoch nie von Investitionsrunde zu Investitionsrunde . | Das heißt, Sie können innerhalb einer Chatrunde vom ersten Chat auf den zweiten Chat Rückschlüsse auf den Beruf der/des Chatpartner:in ziehen, jedoch nie von Chatrunde zu Chatrunde . | Das heißt, Sie können innerhalb einer Lehrveranstaltung von der ersten Klausur auf die zweite Klausur Rückschlüsse auf den Aufgabentyp ziehen, jedoch nie von Lehrveranstaltzung zu Lehrveranstaltung. |
| Im Umweltzustand Up sind mehr schwarze Kugeln in der linken als in der rechtern Urne. Im Umweltzustand Down sind mehr schwarze Kugeln in der rechten Urne. | In einem Bullenmarkt führen Aktienportfolios häufiger zu einer erfolgreichen Investition als Anleihenportfolios. In einem Bärenmarkt führen Anleihenportfolios häufiger zu einer erfolgreichen Investition. | Influencer:innen gefällt der Modemix häufiger als der Arbeitsmix. Business Manager:innen gefällt der Arbeitsmix häufiger. | Bei Theorieaufgaben führt die Theoriestrategie häufiger zum Bestehen der Klausur als die Transferstrategie. Bei Transferaufgaben führt die Transferstrategie häufiger zum Bestehen der Klausur. |
| Die folgende Tabelle fasst die Entscheidungssituation für Sie zusammen. | Die Erfolgschancen sind in der folgenden Tabelle als schwarze Kugeln dargestellt. | Die Erfolgschancen sind in der folgenden Tabelle als schwarze Kugeln dargestellt. | Die Erfolgschancen sind in der folgenden Tabelle als schwarze Kugeln dargestellt. |
| Wenn Sie sich für die die linke Urne entscheiden und ... | Wenn Sie sich für das Anleihenportfolio entscheiden und ... | Wenn Sie sich für den Arbeitsmix entscheiden und ... | Wenn Sie sich für die Theoriestrategie entscheiden und ... |
| der Umweltzustand Up ist, werden Sie in 4 von 6 Fällen eine schwarze Kugel ziehen | der Marktzustand Bärenmarkt ist, werden Sie in 4 von 6 Fällen erfolgreich investieren. | die/der Chatpartner:in Business Manager:in ist, wird ihr/ihm der Themenmix in 4 von 6 Fällen gefallen. | die Klausur überwiegend aus Theorieaufgaben besteht, werden Sie in 4 von 6 Fällen die Klausur bestehen |
| der Umweltzustand Down ist, werden Sie in 2 von 6 Fällen eine schwarze Kugel ziehen. | der Marktzustand Bullenmarkt ist, werden Sie in 2 von 6 Fällen erfolgreich investieren. | die/der Chatpartner:in Influencer:in ist, wird ihr/ihm der Themenmix in 2 von 6 Fällen gefallen. | die Klausur überwiegend aus Transferaufgaben besteht, werden Sie in 2 von 6 Fällen die Klausur bestehen. |
| Wenn Sie sich für die rechte Urne entscheiden und ... | Wenn Sie sich für das Aktienportfolio entscheiden und ... | Wenn Sie sich für den Modemix entscheiden und ... | Wenn Sie sich für die Transferstrategie entscheiden und ... |
| der Umweltzustand Up ist, werden Sie in 2 von 6 Fällen eine schwarze Kugel ziehen. | der Marktzustand Bärenmarkt ist, werden Sie in 2 von 6 Fällen erfolgreich investieren. | die/der Chatpartner:in Business Manager:in ist, wird ihr/ihm der Themenmix in 2 von 6 Fällen gefallen. | die Klausur überwiegend aus Theorieaufgaben besteht, werden Sie in 2 von 6 Fällen die Klausur bestehen. |
| der Umweltzustand Down ist, werden Sie in 4 von 6 Fällen eine schwarze Kugel ziehen. | der Marktzustand Bullenmarkt ist, werden Sie in 4 von 6 Fällen erfolgreich investieren. | die/der Chatpartner:in Influencer:in ist, wird ihr/ihm der Themenmix in 4 von 6 Fällen gefallen. | die Klausur überwiegend aus Transferaufgaben besteht, werden Sie in 4 von 6 Fällen die Klausur bestehen. |
| Gewinnmöglichkeit: Ihre Auszahlung im Experiment hängt von der Anzahl der gezogenen schwarzen Kugeln ab. Je häufiger Sie eine schwarze Kugel ziehen, desto mehr Geld erhalten Sie. Für eine schwarze Kugel erhalten Sie 10 Cent. | Gewinnmöglichkeit: Ihre Auszahlung im Experiment hängt von der Anzahl der erfolgreichen Investitionen ab. Je häufiger Sie erfolgreich investieren, desto mehr Geld erhalten Sie. Für eine erfolgreiche Investitionen erhalten Sie 10 Cent. | Gewinnmöglichkeit: Ihre Auszahlung im Experiment hängt von der Anzahl der Chats ab, die Ihren Chatpartner:innen gefallen. Je häufiger ein Chat gefällt, desto mehr Geld erhalten Sie. Für jeden Chat, das gefällt, erhalten Sie 10 Cent. | Gewinnmöglichkeit: Ihre Auszahlung im Experiment hängt von der Anzahl der bestandenen Klausuren ab. Je häufiger Sie eine Klausur bestehen, desto mehr Geld erhalten Sie. Für eine bestandene Klausur erhalten Sie 10 Cent. |

FIGURE C.1: *Identification and transferring of key terms.*

In a series of focus group sessions with four experts from the field of experimental economics, the following decision situations were created: *Selection of travel (traveling),* *conversation topics during dates (dating), posting social media (social media), learning strategies for exams (exam),* and *selection of internet contracts (contracts)*. The key terms of the urn game's instruction were identified and equivalents in the corresponding contexts were selected (see figure C.1). Solely, the key terms were replaced to maintain equal induction of the decision situations. In the example of the robo-advisor context, there are two different portfolios, which are different with respect to the ratio of bonds and stocks. These are equivalent to the left and right urn of the urn game. Furthermore, there are two different states of the market, namely bear and bull market, which are equivalent to urn game's states of the world. In the bull market, the portfolio with a larger share of stocks performs better than the other portfolio with smaller share of stocks, whereas the portfolio with a larger share of bonds performs better in the bear market.

## C.2 Procedure of Identification

The situational contexts were investigated in an online study with a within-subject design. The participants processed the situational contexts in random order to control learning and fatigue effects (see figure C.2). The contexts are operationalized with the DIAMONDS taxonomy. For the assessment of DIAMONDS, the S8-II questionnaire is used (Rauthmann and Sherman, 2015b). After each context, participants rated the eight perceived situational characteristics on a 7-point Likert scale ranging from "not at all" to "totally agree".

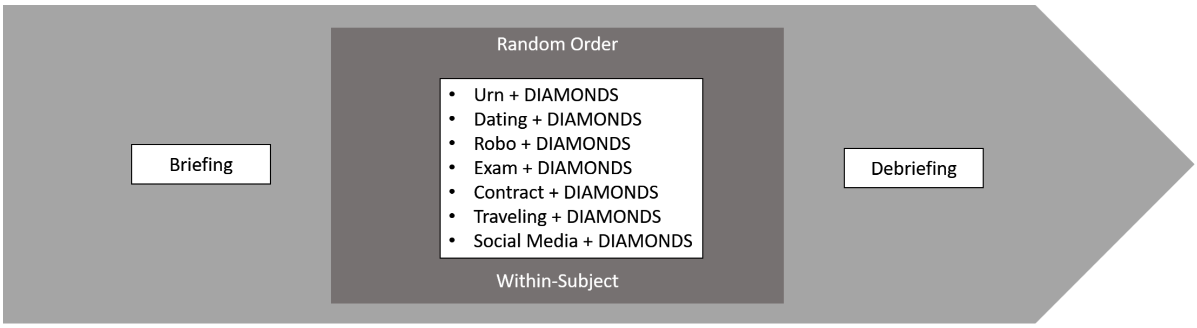



FIGURE C.2: *Procedure of the study*

## C.3 Results

The study was conducted with the sample of KD$^2$Lab. The participants were acquired via Hroot (Bock et al., 2014). In total, 36 persons participated the online survey, 21 male and 15 female participants. On average, the participants were 23.4 (sd = 3.8) years old. Most of them were students in the field of economics and engineering.

Table C.1 provides the mean values and the standard deviations of the situational characteristics for each situational context.

| | **Duty** | **Intellect** | **Adversity** | **Mating** | **Positivity** | **Negativity** | **Deception** | **Sociality** |
|---|---|---|---|---|---|---|---|---|
| **Urn** | 3.0 (1.7) | 3.1 (2.0) | 2.3 (1.6) | 1.7 (1.6) | 4.2 (1.8) | 3.6 (1.9) | 2.5 (1.8) | 2.7 (1.9) |
| **Robo** | 5.4 (1.6) | 3.7 (1.7) | 3.2 (1.7) | 1.5 (1.2) | 4.9 (1.4) | 4.3 (1.5) | 3.0 (1.9) | 3.5 (1.8) |
| **Dating** | 4.8 (2.1) | 4.3 (1.6) | 2.5 (1.7) | 4.3 (2.0) | 5.0 (1.5 | 3.2 (1.4) | 3.3 (1.8) | 6.1 (1.2) |
| **Exam** | 6.6 (1.7) | 3.8 (2.2) | 2.9 (2.0) | 1.3 (1.2) | 3.4 (1.7) | 4.5 (1.6) | 2.4 (1.5) | 2.8 (1.8) |
| **Contract** | 5.4 (1.7) | 2.3 (1.8) | 2.8 (1.7) | 1.6 (1.4) | 3.9 (2.0) | 4.4 (1.6) | 3.4 (2.0) | 4.7 (1.8) |
| **Traveling** | 2.5 (1.7) | 3.4 (1.9) | 2.5 (1.9) | 3.5 (2.3) | 5.9 (1.4) | 3.8 (1.9) | 2.5 (1.9) | 4.9 (2.0) |
| **Social Media** | 5.6 (1.7) | 4.7 (1.9) | 2.9 (1.5) | 2.7 (1.7) | 4.9 (1.5) | 3.0 (1.5) | 3.1 (1.8) | 6.2 (1.3) |

TABLE C.1: *The mean values and standard deviations of situational characteristics across situational contexts.*

For the remaining analysis, the values of situational characteristics are normed with a z-transformation (mean = 0, sd = 1). Otherwise, dimensions with higher variances would be higher weighted in multidimensional analysis.

In order to evaluate the distinctiveness of the DIAMONDS taxonomy, a decision tree is trained with 75 % of the data and the model is tested with the remaining 25 %. This is repeated 1000 times with different samples and the mean values of the models are calculated. The models have a mean accuracy of 98.12 %. Each dimension of the DIAMONDS taxonomy is used for the classification, and there is no big drop in the feature importance. The mean feature importance in descending order: Duty = 0.17, sociality = 0.165, positivity = 0.132, intellect = 0.128, negativity = 0.111, mating = 0.107, deception = 0.097, and adversity = 0.09). This result underlines the validity of the DIAMONDS taxonomy.

Next, the identification of the situational contexts for the main study is described. The decision criteria for the selection are:

- The situational contexts should differ as much as possible from the robo-advisor context and urn game.

- The situational contexts have to be different with respect to all the situational characteristics.
- As perceived negativity is the subject of a working hypothesis (see section 8.1), the situational contexts have to be different with respect to negativity.

First, the analysis for all situational characteristics is performed. In order to investigate the differences, the pairwise Euclidean distances are calculated (see figure C.3).

| | Traveling | Dating | Social media | Exam | Contracts | Urn | Robo |
|---|---|---|---|---|---|---|---|
| Urn | 0.1503 | 0.2552 | 0.1756 | 0.2069 | 0.1438 | 0 | 0.0772 |
| Robo | 0.1074 | 0.3024 | 0.1727 | 0.2136 | 0.1266 | 0.0772 | 0 |

| | Traveling | Dating | Social Media | Exam | Contracts |
|---|---|---|---|---|---|
| Dating | 0.1709 | 0 | 0.167 | 0.1941 | 0.1716 |
| Exam | 0.1612 | 0.1941 | 0.1045 | 0 | 0.2011 |

FIGURE C.3: *Pairwise Euclidean distances of all situational characteristics*

Figure C.3 shows that the dating and exam context are the most different from urn game and robo-advisor context. Compared with the dating context, exam and contracts context are most different, whereas dating and contracts contexts are most different from the exam context.

| | Traveling | Dating | Social media | Exam | Contracts | Urn | Robo |
|---|---|---|---|---|---|---|---|
| Urn | 0.1424 | 0.3522 | 0.0691 | 0.2603 | 0.1416 | 0 | 0.0344 |
| Robo | 0.143 | 0.372 | 0.069 | 0.2495 | 0.1286 | 0.0344 | 0 |

| | Traveling | Dating | Social Media | Exam | Contracts |
|---|---|---|---|---|---|
| Dating | 0.1243 | 0 | 0.1625 | 0.206 | 0.1367 |
| Exam | 0.1495 | 0.206 | 0.1683 | 0 | 0.1104 |

FIGURE C.4: *Pairwise Euclidean distances of negativity*

Next, the Euclidean distances of perceived negativity are analyzed. Figure C.4 illustrates that the dating and exam context are most different from the urn game and robo-advisor. Compared with the dating context, social media and exam context are most different. Dating and social media context are most different from the exam context.

## C.4 Discussion

In both distance analyses, exam and dating context are identified as most different from the urn game and robo-advisor context. Dating and exam context are most different from each other in three of four cases. Thus, the situational contexts for the main study are robo-advisor, urn game, dating, and exam context.

There are no comparative values in the literature, as the study is the first one, which uses situational characteristics to identify different situational contexts of a decision situation. Thus, a limitation of the study is that the relative similarities are used to identify different situational contexts of the urn game. It could not be ensured that the differences are sufficient to analyze decision inertia in different situational contexts.

The generalizability of the results is restricted. The KD$^2$Lab sample consists mainly of students of the KIT. The values of the situational characteristics could be different with another sample.

There are certainly other situational contexts in which decision inertia could be investigated. In the design phase, it was ensured that the created contexts correspond to the everyday life situations of the KD$^2$Lab sample's participants. The participants should be able to put themselves into the context as best they can to maximize the induction effect.

# Appendix D

# Data Catalogue and Regression Tables of Chapter 8

This appendix contains the data catalogue and regression tables of chapter 8. In addition, the compact disc attached to the last page of this thesis contains both the data and the analyses using the R programming language.

## D.1 Data Catalogue

The following table matches the variables of the thesis with the variables of the R-script. Moreover, it contains a short description and the source of variables.

| Name CSV / R-script | Name Thesis | Meaning | Source | Note / Comment |
|---|---|---|---|---|
| df_relevant_ cleaned | - | Dataset in wide format | - | - |
| df_long_ cleaned | - | Dataset in long format | - | - |
| DI | - | Boolean variable of decision inertia | | 1 = Decision inertia |

| | | | | |
|---|---|---|---|---|
| di_rate | - | Average decision inertia rate of participant | | Average of divergence = 1 decision inertia = 1 |
| participant .code | - | Unique identifier of participant | - | - |
| round | Round number | Indicator experimental task's round | | 60 Rounds |
| first_choice | First decison of urn game | Values: left and right | - | - |
| second choice | Second decison of urn game | Values: left and right | - | - |
| nudged_ choice | - | Values: left and right | | Decision after warnings |
| second_choice _result | - | indicator for reward (black ball) | - | 1 = reward |
| optimal_second- _choice | - | Optimal second decision | | Based on Bayesian updating |
| divergence | divergent case | First choice not rewarded | - | - |
| ID | Random Effect | ParticipantID for regressions | - | - |
| participant .payoff | Payoff | Rewarded decision of urn game | | Without flat fee (1.5 euro) |
| session.config .treatment | Treatments | 12 treatments: nudge (3) x situational contexts (4) | | - |

| | | | | |
|---|---|---|---|---|
| false_quiz_total | Quiz result | Number of incorrect answers | - | - |
| kontext | situational contexts | context treatmets | - | Values: urn, robo, dating, exam |
| nudge | nudges | nudge treatments | - | Values: baseline, warning, default |
| duty | Duty | Situational characteristics | Rauthmann et al. (2014) | Likert scale value of item |
| intellect | Intellect | Situational characteristics | Rauthmann et al. (2014) | Likert scale value of item |
| adversity | Adversity | Situational characteristics | Rauthmann et al. (2014) | Likert scale value of item |
| mating | Mating | Situational characteristics | Rauthmann et al. (2014) | Likert scale value of item |
| positivity | Positivity | Situational characteristics | Rauthmann et al. (2014) | Likert scale value of item |
| negativity | Negativity | Situational characteristics | Rauthmann et al. (2014) | Likert scale value of item |
| deception | Deception | Situational characteristics | Rauthmann et al. (2014) | Likert scale value of item |
| sociality | Sociality | Situational characteristics | Rauthmann et al. (2014) | Likert scale value of item |
| di_count | - | Number of decision inertia cases | - | Count |
| divergence_count | - | Number of divergence cases | - | - |
| di_rate | decision inertia rate | - | - | di_rate = di_count / divergence_count |

| | | | | |
|---|---|---|---|---|
| fl_total | Financial literacy | Particiant's score in questionnaire. 5 items | Lusardi et al. (2007) | Particiant's score in questionnaire |
| prs_total | Probabilistic reasoning | Particiant's score in questionnaire 9 items | Primi et al. (2019) | Particiant's score in questionnaire |
| ho_total | Action orientation | Particiant's score in questionnaire. 24 items | Kuhl (1994) | Summation of Likert Scale ratings |
| icq_total | Interpersonal competence | Particiant's score in questionnaire. 40 items | Riemann and Allgöwer (1993) | Summation of Likert Scale ratings |
| lcs_total | Learning competence | Particiant's score in questionnaire. 40 items | Villardón-Gallego (2013) | Summation of Likert Scale ratings |
| pre_oc | Perceived prior confidence | Participant's confidence in front of task | Svenson (1981) | Slider value |
| post_oc | Perceived post confidence | Participant's confidence after the task | Svenson (1981) | Slider value |
| attention _failed | - | Failed attention tests | - | Number of failed tests |
| manipulation _failed | - | failed manipulation test | - | Boolean 1 = True |
| time_prs | - | Processing time of Probbailistic reasoning | - | Seconds |
| urn | Urn game | Situational context | - | Boolean 1 = True |
| robo | Robo-advisor | Situational context | - | Boolean 1 = True |

| | | | | |
|---|---|---|---|---|
| dating | Dating game | Situational context | - | Boolean 1 = True |
| exam | Exam game | Situational context | - | Boolean 1 = True |
| baseline | Without nudging treatment | Nudging treatment | - | Boolean 1 = True |
| default_nudge | Defaults | Nudging treatment | - | Boolean 1 = True |
| warning | Warnings | Nudging treatment | - | Boolean 1 = True |
| normed_prs | Probabilistic reasoning | Standardized Value | - | Z-standardization |
| normed_fl | Financial literacy | Standardized Value | - | Z-standardization |
| normed_flc | Learning competence | Standardized Value | - | Z-standardization |
| normed_ic | Interpersonal competence | Standardized Value | - | Z-standardization |
| dom_expert | Domain Expertise | Financial literacy, probabilistic reasoning, interpersonal competence, and learning competence as one variable | - | normed values |
| mn_round | Round number | Normed value | - | Min-max |
| mn_pre_oc | Perceived prior confidence | Normed value | - | Min-max |

| | | | | |
|---|---|---|---|---|
| mn_post_oc | Perceived post confidence | Normed value | - | Min-max |
| mn_negativity | Negativity | Normed value | - | Min-max |
| mn_positivity | Positivity | Normed value | - | Min-max |
| mn_duty | Duty | Normed value | - | Min-max |
| mn_intellect | Intellect | Normed value | - | Min-max |
| mn_adversity | Adversity | Normed value | - | Min-max |
| mn_mating | Mating | Normed value | - | Min-max |
| mn_deception | Deception | Normed value | - | Min-max |
| mn_sociality | Sociality | Normed value | - | Min-max |
| mn_ho_total | Action orientation | Normed value | - | Min-max |
| correct_second_choice | - | boolean variable of correct second choice | - | Boolean 1 = True |
| difference_pre_post_oc | Over- / Underconfidence | Positive values = oc, negative values = uc | - | pre_oc - post_oc |
| difference_pre_post_oc_squared | Over- / Underconfidence | variable of squared regression analysis | - | $(pre\_oc - post\_oc)^2$ |
| zn_difference_pre_post_oc | Over- / Underconfidence | Normed value | - | Z-standardization |
| zn_difference_pre_post_oc_squared | Over- / Underconfidence | Normed value | - | Z-standardization |
| male | Male | Indicator of males | - | Boolean 1 = True |
| wiing | Business engineer | Indicator of business engineers | - | Boolean 1 = True |

## D.2 Regression Tables

This section provides the regression tables of the analysis section.

| | $\mathbf{M}_{control}$ | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | -2.211 | 0.288 | 0.001 *** |
| Divergence (1 = True) | 0.469 | 0.134 | 0.001 *** |
| Round Number | -0.357 | 0.111 | 0.001 ** |
| Action-orientation | 0.675 | 0.562 | 0.229 |
| Gender (1 = Male) | -0.941 | 0.232 | 0.001 *** |
| Business Engineer (1 = True) | ’0.244 | 0.233 | 0.296 |
| Divergence x Round Number | 0.179 | 0.148 | 0.228 |
| Divergence x Action-orientation | -0.229 | 0.232 | 0.324 |
| Divergence x Gender | 0.636 | 0.092 | 0.001 *** |
| Divergence x Business Engineer | -0.499 | 0.093 | 0.001 *** |
| **Deviance** | | **14213.8** | |

Number of Observations: 19440; Random-effect: Participant-id; Number of Participants: 324; Significance Codes: ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’+’ 0.1

TABLE D.2: *Logistic regression on suboptimal second choice with control variables.*

TABLE D.3: *Logistic regressions for the analysis of decision inertia across situational contexts.*

| | $\mathbf{H1}_{urn}$ | | | $\mathbf{H1}_{robo}$ | | | $\mathbf{H1}_{dating}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** |
| Intercept | -2.622 | 0.497 | 0.001 *** | -1.795 | 0.463 | 0.001 *** | -1.668 | -1.668 | 0.001 *** |
| Divergence (1 = True) | 1.181 | 0.245 | 0.001 *** | 0.908 | 0.210 | 0.001 *** | 0.240 | 0.196 | 0.221 |
| Round Number | -0.337 | 0.193 | 0.081 + | -0.337 | 0.193 | 0.081 + | -0.337 | 0.193 | 0.081 + |
| Gender (1 = Male) | -1.191 | 0.400 | 0.002 ** | -1.191 | 0.400 | 0.002 ** | -1.191 | 0.400 | 0.002 ** |
| Urn ( 1 = True) | - | - | - | -0.826 | 0.574 | 0.150 | -0.953 | 0.570 | 0.094 + |
| Robo (1 = True) | 0.826 | 0.575 | 0.150 | - | - | - | -0.127 | 0.558 | 0.819 |
| Dating (1 = True) | 0.953 | 0.57 | 0.094 + | 0.127 | 0.558 | 0.819 | - | - | - |
| Exam (1 = True) | 1.199 | 0.558 | 0.031 * | 0.373 | 0.550 | 0.498 | 0.245 | 0.544 | 0.651 |
| Divergence x Round Number | -0.022 | 0.255 | 0.929 | -0.022 | 0.255 | 0.929 | -0.022 | 0.255 | 0.929 |
| Divergence x Gender | 0.598 | 0.160 | 0.001 *** | 0.598 | 0.160 | 0.001 *** | 0.598 | 0.16 | 0.001 *** |
| Divergence x Urn | - | - | - | 0.273 | 0.239 | 0.254 | 0.940 | 0.239 | 0.001 *** |
| Divergence x Robo | -0.273 | 0.239 | 0.254 | - | - | - | 0.667 | 0.211 | 0.001 ** |
| Divergence x Dating | -0.940 | 0.239 | 0.001 *** | -0.667 | 0.211 | 0.001 ** | - | - | - |
| Divergence x Exam | -1.101 | 0.233 | 0.001 *** | -0.828 | 0.211 | 0.001 *** | -0.160 | 0.207 | 0.438 |
| **Deviance** | **4806** | | | **4806** | | | **4806** | | |

Number of Observations: 6360; Random-effect: Participant-id; Number of Participants: 106; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

| | H5D/H5W | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | -1.979 | 0.237 | 0.001 *** |
| Divergence (1 = True) | 0.473 | 0.113 | 0.001 *** |
| Round Number | -0.367 | 0.11 | 0.001 *** |
| Gender (1 = Male) | -0.863 | 0.229 | 0.001 *** |
| Defaults (1 = True) | 0.084 | 0.272 | 0.756 |
| Warnings (1 = True) | 0.228 | 0.27 | 0.399 |
| Divergence x Round Number | 0.179 | 0.148 | 0.228 |
| Divergence x Gender | 0.553 | 0.090 | 0.001 *** |
| Divergence x Defaults | -0.082 | 0.110 | 0.457 |
| Divergence x Warnings | -0.629 | 0.107 | 0.001 *** |
| **Deviance** | | **14204.5** | |

Number of Observations: 19440; Random-effect: Participant-id; Number of Participants: 324; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE D.4: *Logistic regression for the analysis of warnings' and defaults' effectiveness in reducing decision inertia.*

TABLE D.5: *Logistic regressions for the analysis of nudge' effectiveness in each situational context.*

| | $\mathbf{M}_{NudgeUrn}$ | | | $\mathbf{M}_{NudgeRobo}$ | | | $\mathbf{M}_{NudgeDating}$ | | | $\mathbf{M}_{NudgeExam}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** |
| Intercept | -3.554 | 0.528 | 0.001 *** | -1.431 | 0.432 | 0.001 *** | -1.776 | 0.422 | 0.001 *** | -1.776 | 0.422 | 0.001 *** |
| Divergence (1 = True) | 1.718 | 0.289 | 0.001 *** | 0.782 | 0.226 | 0.001 *** | 0.239 | 0.206 | 0.247 | 0.239 | 0.206 | 0.247 |
| Round Number | -0.146 | 0.262 | 0.575 | 0.702 | 0.226 | 0.001 ** | 0.043 | 0.204 | 0.829 | 0.043 | 0.204 | 0.829 |
| Gender (1 = Male) | 0.069 | 0.468 | 0.882 | -1.407 | 0.436 | 0.001 ** | -1.214 | 0.421 | 0.003 ** | -1.214 | 0.421 | 0.003 ** |
| Defaults (1 = True) | 0.596 | 0.549 | 0.278 | 0.570 | 0.492 | 0.246 | 0.301 | 0.516 | 0.559 | 0.301 | 0.516 | 0.559 |
| Warnings (1 = True) | -0.01 | 0.563 | 0.984 | -0.135 | 0.503 | 0.787 | 0.424 | 0.506 | 0.401 | 0.424 | 0.506 | 0.401 |
| Divergence x Round Number | -0.341 | 0.332 | 0.304 | 0.501 | 0.300 | 0.094 + | 0.081 | 0.279 | 0.769 | 0.081 | 0.279 | 0.769 |
| Divergence x Gender | 0.051 | 0.213 | 0.809 | 0.303 | 0.185 | 0.100 | 0.377 | 0.173 | 0.029 * | 0.377 | 0.173 | 0.029 * |
| Divergence x Defaults | -0.822 | 0.244 | 0.001 *** | -0.820 | 0.217 | 0.001 *** | 0.551 | 0.209 | 0.008 ** | 0.551 | 0.209 | 0.008 ** |
| Divergence x Warnings | -0.316 | 0.253 | 0.212 | -0.905 | 0.224 | 0.001 *** | -0.725 | 0.209 | 0.001 *** | -0.725 | 0.209 | 0.001 *** |
| **Deviance** | **3091.6** | | | **3496.8** | | | **3948.2** | | | **3948.2** | | |
| **Number of Observations** | **4920** | | | **4680** | | | **5040** | | | **5040** | | |
| **Number of Participants** | **82** | | | **78** | | | **84** | | | **84** | | |

Random-effect: Participant-id; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE D.6: *Logistic regressions for the analysis of defaults' effectiveness across situational context.*

| | **H7D$_{urn}$** | | | **H7D$_{robo}$** | | | **H7D$_{dating}$** | | |
|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** |
| Intercept | -2.564 | 0.425 | 0.001 *** | -1.888 | 0.478 | 0.001 *** | -1.808 | 0.437 | 0.001 *** |
| Divergence (1 = True) | 0.414 | 0.227 | 0.068 + | 0.039 | 0.233 | 0.867 | 0.712 | 0.221 | 0.001 ** |
| Round Number | -0.380 | 0.203 | 0.062 + | -0.380 | 0.203 | 0.062 + | -0.380 | 0.203 | 0.062 + |
| Gender (1 = Male) | -0.342 | 0.398 | 0.389 | -0.342 | 0.398 | 0.389 | -0.342 | 0.398 | 0.389 |
| Urn ( 1 = True) | - | - | - | -0.675 | 0.494 | 0.171 | -0.755 | 0.489 | 0.122 |
| Robo (1 = True) | 0.676 | 0.494 | 0.171 | - | - | - | -0.079 | 0.489 | 0.87 |
| Dating (1 = True) | 0.755 | 0.489 | 0.122 | 0.079 | 0.489 | 0.870 | - | - | - |
| Exam (1 = True) | -0.145 | 0.524 | 0.781 | -0.821 | 0.510 | 0.107 | -0.901 | 0.519 | 0.082 + |
| Divergence x Round Number | 0.345 | 0.264 | 0.190 | 0.345 | 0.264 | 0.19 | 0.345 | 0.264 | 0.190 |
| Divergence x Gender | 0.300 | 0.173 | 0.083 + | 0.300 | 0.173 | 0.083 + | 0.300 | 0.173 | 0.083 + |
| Divergence x Urn | - | - | - | 0.374 | 0.219 | 0.088 + | -0.298 | 0.215 | 0.165 |
| Divergence x Robo | -0.374 | 0.219 | 0.088 + | - | - | - | -0.673 | 0.206 | 0.001 ** |
| Divergence x Dating | 0.298 | 0.215 | 0.165 | 0.673 | 0.206 | 0.001 ** | - | - | - |
| Divergence x Exam | 0.392 | 0.255 | 0.125 | 0.767 | 0.242 | 0.001 ** | 0.093 | 0.242 | 0.698 |
| **Deviance** | **4670.4** | | | **4670.4** | | | **4670.4** | | |

Number of Observations: 6540; Random-effect: Participant-id; Number of Participants: 109; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE D.7: *Logistic regressions for the analysis of warnings' effectiveness across situational context.*

| | $H7W_{urn}$ | | | $H7W_{robo}$ | | | $H7W_{dating}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** |
| Intercept | -3.006 | 0.492 | 0.001 *** | -2.359 | 0.494 | 0.001 *** | -1.567 | 0.455 | 0.001 *** |
| Divergence (1 = True) | 0.875 | 0.245 | 0.001 *** | -0.040 | 0.223 | 0.854 | -0.569 | 0.212 | 0.007 ** |
| Round Number | -0.386 | 0.185 | 0.037 * | -0.386 | 0.185 | 0.037 * | -0.386 | 0.185 | 0.037 * |
| Gender (1 = Male) | -0.595 | 0.400 | 0.136 | -0.595 | 0.400 | 0.136 | -0.595 | 0.400 | 0.136 |
| Urn ( 1 = True) | - | - | - | -0.646 | 0.575 | 0.260 | -1.439 | 0.542 | 0.008 ** |
| Robo (1 = True) | 0.646 | 0.575 | 0.260 | - | - | - | -0.792 | 0.548 | 0.148 |
| Dating (1 = True) | 1.439 | 0.542 | 0.008 ** | 0.792 | 0.548 | 0.148 | - | - | - |
| Exam (1 = True) | 1.938 | 0.556 | 0.001 *** | 1.291 | 0.561 | 0.021 * | 0.499 | 0.526 | 0.343 |
| Divergence x Round Number | 0.237 | 0.260 | 0.361 | 0.237 | 0.260 | 0.361 | 0.237 | 0.260 | 0.361 |
| Divergence x Gender | 0.390 | 0.157 | 0.013 * | 0.390 | 0.157 | 0.013 * | 0.390 | 0.157 | 0.013 * |
| Divergence x Urn | - | - | - | 0.916 | 0.243 | 0.001 *** | 1.445 | 0.232 | 0.001 *** |
| Divergence x Robo | -0.916 | 0.244 | 0.001 *** | - | - | - | 0.528 | 0.220 | 0.016 * |
| Divergence x Dating | -1.445 | 0.232 | 0.001 *** | -0.528 | 0.220 | 0.016 * | - | - | - |
| Divergence x Exam | -1.238 | 0.226 | 0.001 *** | -0.322 | 0.214 | 0.133 | 0.206 | 0.202 | 0.306 |
| **Deviance** | **4617.8** | | | **4617.8** | | | **4617.8** | | |

Number of Observations: 6540; Random-effect: Participant-id; Number of Participants: 109; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE D.8: *Logistic regressions for the analysis of perceived confidence.*

| | $\mathbf{H4}_{linear}$ | | | $\mathbf{H4}_{squared}$ | | |
|---|---|---|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** |
| Intercept | -1.846 | 0.185 | 0.001 *** | -1.85 | 0.185 | 0.001 *** |
| Divergence (1 = True) | 0.239 | 0.098 | 0.015 * | 0.237 | 0.098 | 0.016 * |
| Round Number | -0.366 | 0.110 | 0.001 *** | -0.366 | 0.110 | 0.001 *** |
| Gender (1 = Male) | -0.905 | 0.225 | 0.001 *** | -0.902 | 0.224 | 0.001 *** |
| Perceived Confidence | 0.32 | 0.108 | 0.003 ** | - | - | - |
| Squared Perceived Confidence | - | - | - | 0.342 | 0.106 | 0.001 ** |
| Divergence x Round Number | 0.180 | 0.148 | 0.224 | 0.181 | 0.148 | 0.222 |
| Divergence x Gender | 0.556 | 0.089 | 0.001 *** | 0.567 | 0.089 | 0.001 *** |
| Divergence x Perceived Confidence | -0.030 | 0.043 | 0.487 | - | - | - |
| Divergence x Squared Perceived Confidence | - | - | - | -0.077 | 0.039 | 0.051 + |
| **Deviance** | **14235.4** | | | **14232.2** | | |

Number of Observations: 19440; Random-effect: Participant-id; Number of Participants: 324; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

| | $\mathbf{M}_{ProbabilisticReasoning}$ | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | -2.065 | 0.187 | 0.001 *** |
| Divergence (1 = True) | 0.321 | 0.102 | 0.001 ** |
| Round Number | -0.37 | 0.111 | 0.001 *** |
| Gender (1 = Male) | -0.576 | 0.228 | 0.011 * |
| Probabilistic Reasoning | -0.498 | 0.106 | 0.001 *** |
| Divergence x Round Number | 0.184 | 0.148 | 0.215 |
| Divergence x Gender | 0.458 | 0.093 | 0.001 *** |
| Divergence x Probabilistic Reasoning | 0.125 | 0.038 | 0.001 ** |
| **Deviance** | | **14217.3** | |

Number of Observations: 19440; Random-effect: Participant-id; Number of Participants: 324; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE D.9: *Logistic regression for the analysis of probabilistic reasoning without the differentiation of situational contexts.*

TABLE D.10: *Logistic regressions for the analysis of domain expertise in each situational contexts.*

| | H3U | | | H3R | | | H3D | | | H3E | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** | **Beta** | **SE** | **P-value** |
| Intercept | -3.704 | 0.413 | 0.001 *** | -1.628 | 0.368 | 0.001 *** | -1.513 | 0.330 | 0.001 *** | -1.656 | 0.395 | 0.001 *** |
| Divergence (1 = True) | 1.576 | 0.261 | 0.001 *** | 0.115 | 0.214 | 0.590 | 0.171 | 0.180 | 0.341 | -0.493 | 0.203 | 0.015 * |
| Round Number | -0.156 | 0.261 | 0.549 | -0.674 | 0.224 | 0.002 ** | 0.049 | 0.203 | 0.808 | -0.660 | 0.215 | 0.002 ** |
| Gender (1 = Male) | 0.553 | 0.489 | 0.258 | -0.844 | 0.443 | 0.056 + | -1.257 | 0.415 | 0.002 ** | -0.527 | 0.481 | 0.272 |
| Domain Expertise | -0.577 | 0.201 | 0.004 ** | -0.494 | 0.206 | 0.016 * | -0.125 | 0.210 | 0.549 | -0.339 | 0.232 | 0.143 |
| Divergence x Round Number | -0.339 | 0.331 | 0.306 | 0.479 | 0.298 | 0.108 | 0.085 | 0.278 | 0.759 | 0.458 | 0.297 | 0.124 |
| Divergence x Gender | -0.223 | 0.244 | 0.360 | 0.417 | 0.206 | 0.042 * | 0.448 | 0.165 | 0.006 ** | 1.142 | 0.181 | 0.001 *** |
| Divergence x Domain Expertise | 0.263 | 0.080 | 0.001 ** | -0.208 | 0.094 | 0.027 * | -0.074 | 0.091 | 0.412 | 0.001 *** | 0.088 | 0.404 |
| **Deviance** | | **3089** | | | **3506.8** | | | **3987.1** | | | **3544.9** | |
| **Number of Observations** | | **4920** | | | **4680** | | | **5040** | | | **5040** | |
| **Number of Participants** | | **82** | | | **78** | | | **84** | | | **84** | |

Random-effect: Participant-id; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

| | H6DU | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | -3.921 | 0.604 | 0.001 *** |
| Divergence (1 = True) | 2.013 | 0.352 | 0.001 *** |
| Round Number | -0.045 | 0.317 | 0.885 |
| Gender (1 = Male) | 0.438 | 0.613 | 0.474 |
| Probabilistic Reasoning | -0.386 | 0.303 | 0.202 |
| Defaults (1 = True) | 0.512 | 0.551 | 0.352 |
| Divergence x Round Number | -0.63 | 0.402 | 0.117 |
| Divergence x Gender | -0.066 | 0.301 | 0.825 |
| Divergence x Probabilistic Reasoning | 0.318 | 0.113 | 0.004 ** |
| Divergence x Defaults | -0.826 | 0.255 | 0.001 ** |
| Probabilistic Reasoning x Defaults | -0.273 | 0.409 | 0.503 |
| Divergence x Probabilistic Reasoning x Defaults | -0.189 | 0.149 | 0.203 |
| **Deviance** | | **2099.1** | |

Number of Observations: 3300; Random-effect: Participant-id; Number of Participants: 55; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE D.11: *Logistic regressions in the subsample of the treatment urn to analyze the influence of probabilistic reasoning on the effectiveness of defaults.*

| | H6WU | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | -3.802 | 0.594 | 0.001 *** |
| Divergence (1 = True) | 1.578 | 0.355 | 0.001 *** |
| Round Number | -0.196 | 0.341 | 0.564 |
| Gender (1 = Male) | 0.425 | 0.599 | 0.478 |
| Probabilistic Reasoning | -0.382 | 0.295 | 0.195 |
| Warnings (1 = True) | 0.150 | 0.559 | 0.788 |
| Divergence x Round Number | -0.053 | 0.422 | 0.898 |
| Divergence x Gender | 0.142 | 0.302 | 0.636 |
| Divergence x Probabilistic Reasoning | 0.295 | 0.113 | 0.008 ** |
| Divergence x Warnings | -0.386 | 0.257 | 0.134 |
| Probabilistic Reasoning x Warnings | -0.119 | 0.732 | 0.870 |
| Divergence x Probabilistic Reasoning x Warnings | -0.274 | 0.357 | 0.443 |
| **Deviance** | | **1974.8** | |

Number of Observations: 3180; Random-effect: Participant-id; Number of Participants: 53; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE D.12: *Logistic regressions in the subsample of the treatment urn to analyze the influence of probabilistic reasoning on the effectiveness of warnings.*

| | H6DR | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | -1.772 | 0.460 | 0.001 *** |
| Divergence (1 = True) | 0.497 | 0.287 | 0.083 + |
| Round Number | -0.891 | 0.273 | 0.001 ** |
| Gender (1 = Male) | -1.194 | 0.545 | 0.028 * |
| Financial Literacy | -0.796 | 0.268 | 0.002 ** |
| Defaults (1 = True) | 0.866 | 0.452 | 0.055 + |
| Divergence x Round Number | 0.508 | 0.358 | 0.156 |
| Divergence x Gender | 1.084 | 0.287 | 0.001 *** |
| Divergence x Financial Literacy | 0.069 | 0.139 | 0.620 |
| Divergence x Defaults | -1.161 | 0.249 | 0.001 *** |
| Financial Literacy x Defaults | 0.764 | 0.439 | 0.081 + |
| Divergence x Financial Literacy x Defaults | -0.498 | 0.217 | 0.022 * |
| **Deviance** | | **2479.8** | |

Number of Observations: 3180; Random-effect: Participant-id; Number of Participants: 53; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE D.13: *Logistic regressions in the subsample of the treatment robo to analyze the influence of financial literacy on the effectiveness of defaults.*

| | H6WR | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | -2.027 | 0.512 | 0.001 *** |
| Divergence (1 = True) | 0.975 | 0.290 | 0.001 *** |
| Round Number | -0.865 | 0.289 | 0.002 ** |
| Gender (1 = Male) | -0.853 | 0.558 | 0.126 |
| Financial Literacy | -0.877 | 0.315 | 0.005 ** |
| Warnings (1 = True) | 0.303 | 0.533 | 0.569 |
| Divergence x Round Number | 0.621 | 0.378 | 0.100 |
| Divergence x Gender | 0.112 | 0.244 | 0.645 |
| Divergence x Financial Literacy | 0.194 | 0.137 | 0.157 |
| Divergence x Warnings | -1.083 | 0.249 | 0.001 *** |
| Financial Literacy x Warnings | 0.397 | 0.512 | 0.437 |
| Divergence x Financial Literacy x Warnings | -0.603 | 0.216 | 0.005 ** |
| **Deviance** | | **2210.2** | |

Number of Observations: 3000; Random-effect: Participant-id; Number of Participants: 50; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE D.14: *Logistic regression in the subsample of the treatment robo to analyze the influence of financial literacy on the effectiveness of warnings.*

| | H2 | | |
|---|---|---|---|
| **Variable** | **Beta** | **SE** | **P-value** |
| Intercept | -1.983 | 0.227 | 0.001 *** |
| Divergence (1 = True) | 0.306 | 0.112 | 0.006 ** |
| Round Number | -0.367 | 0.110 | 0.001 *** |
| Gender (1 = Male) | -0.842 | 0.227 | 0.001 *** |
| Perceived Negativity | 0.365 | 0.423 | 0.387 |
| Divergence x Round Number | 0.181 | 0.148 | 0.221 |
| Divergence x Gender | 0.542 | 0.089 | 0.001 *** |
| Divergence x Perceived Negativity | -0.222 | 0.172 | 0.199 |
| **Deviance** | | **14242** | |

Number of o Observations: 19440; Random-effect: Participant-id; Number of Participants: 324; Significance Codes: '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1

TABLE D.15: *Logistic regression to analyze the influence of perceived negativity on decision inertia.*

# References


Achtziger, A. and C. Alós-Ferrer (2014). Fast or rational? a response-times study of bayesian updating. *Management Science 60*(4), 923–938.

Ahn, J.-w. and P. Brusilovsky (2013). Adaptive visualization for exploratory information retrieval. *Information Processing & Management 49*(5), 1139–1164.

Alessie, R., A. Lusardi, and M. Van Rooij (2007). *Financial literacy and stock market participation*. National Bureau of Economic Research.

Alison, L., N. Power, C. van den Heuvel, M. Humann, M. Palasinksi, and J. Crego (2015). Decision inertia: Deciding between least worst outcomes in emergency responses to disasters. *Journal of Occupational and Organizational Psychology 88*(2), 295–321.

Alós-Ferrer, C., S. Hügelschäfer, and J. Li (2016). Inertia and decision making. *Frontiers in psychology 7*, 169.

Alós-Ferrer, C., S. Hügelschäfer, and J. Li (2017). Framing effects and the reinforcement heuristic. *Economics Letters 156*, 32–35.

Álvarez, M. A., C. Dafonte, M. Manteiga, D. Garabato, and R. Santoveña (2022). Guasom: an adaptive visualization tool for unsupervised clustering in spectrophotometric astronomical surveys. *Neural Computing and Applications 34*(3), 1993–2006.

Alves, T., J. Natálio, J. Henriques-Calado, and S. Gama (2020). Incorporating personality in user interface design: A review. *Personality and Individual Differences 155*, 109709.

Anderson, E. W., K. C. Potter, L. E. Matzen, J. F. Shepherd, G. A. Preston, and C. T. Silva (2011). A user study of visualization effectiveness using eeg and cognitive load. In *Computer graphics forum*, Volume 30, pp. 791–800. Wiley Online Library.

Anscombe, F. J. (1973). Graphs in statistical analysis. *The american statistician 27*(1), 17–21.

Awais, M., M. F. Laber, N. Rasheed, and A. Khursheed (2016). Impact of financial literacy and investment experience on risk tolerance and investment decisions: Empirical evidence from pakistan. *International Journal of Economics and Financial Issues 6*(1).

Bai, X., D. White, and D. Sundaram (2013). Context adaptive visualization for effective business intelligence. In *2013 15th IEEE International Conference on Communication Technology*, pp. 786–790. IEEE.

Balzer, W. K., M. E. Doherty, et al. (1989). Effects of cognitive feedback on performance. *Psychological bulletin 106*(3), 410.

Bansal, G., D. Gefen, et al. (2010). The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision support systems 49*(2), 138–150.

Barber, B. M. and T. Odean (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The journal of Finance 55*(2), 773–806.

Barrick, M. R., G. L. Stewart, and M. Piotrowski (2002). Personality and job performance: test of the mediating effects of motivation among sales representatives. *Journal of applied psychology 87*(1), 43.

Baskin, E., M. Gorlin, Z. Chance, N. Novemsky, R. Dhar, K. Huskey, and M. Hatzis (2016). Proximity of snacks to beverages increases food consumption in the workplace: A field study. *Appetite 103*, 244–248.

Baumann, N. and J. Kuhl (2005). Positive affect and flexibility: Overcoming the precedence of global over local processing of visual information. *Motivation and Emotion 29*(2), 123–134.

Bechara, A., H. Damasio, D. Tranel, and A. R. Damasio (1997). Deciding advantageously before knowing the advantageous strategy. *Science 275*(5304), 1293–1295.

Belanche, D., L. V. Casaló, and C. Flavián (2019). Artificial intelligence in fintech: understanding robo-advisors adoption among customers. *Industrial Management & Data Systems*.

Benjamini, Y. (1988). Opening the box of a boxplot. *The American Statistician 42*(4), 257–262.

Benyon, D. and D. Murray (1993). Applying user modeling to human-computer interaction design. *Artificial Intelligence Review 7*(3), 199–225.

Bernheim, B. D., D. M. Garrett, and D. M. Maki (2001). Education and saving:: The long-term effects of high school financial curriculum mandates. *Journal of public Economics 80*(3), 435–465.

Beshears, J. and H. Kosowsky (2020). Nudging: Progress to date and future directions. *Organizational behavior and human decision processes 161*, 3–19.

Bhandari, G., K. Hassanein, and R. Deaves (2008). Debiasing investors with decision support systems: An experimental investigation. *Decision Support Systems 46*(1), 399–410.

Bhattacharyya, O., K. Mossman, L. Gustafsson, E. C. Schneider, et al. (2019). Using human-centered design to build a digital health advisor for patients with complex needs: persona and prototype development. *Journal of medical Internet research 21*(5), e10318.

Bishop-Clark, C., J. Courte, and E. V. Howard (2006). Programming in pairs with alice to improve confidence, enjoyment, and achievement. *Journal of educational computing research 34*(2), 213–228.

Bock, O., I. Baetge, and A. Nicklisch (2014). hroot: Hamburg registration and organization online tool. *European Economic Review 71*, 117–120.

Bolton, G. E., E. Dimant, and U. Schmidt (2019). When a nudge backfires: Using observation with social and economic incentives to promote pro-social behavior. Technical report, CeDEx Discussion Paper Series.

Brambor, T., W. R. Clark, and M. Golder (2006). Understanding interaction models: Improving empirical analyses. *Political analysis 14*(1), 63–82.

Brehm, J. W. and A. R. Cohen (1962). Explorations in cognitive dissonance.

Brody, N. (1965). The effect of commitment to correct and incorrect decisions on confidence in a sequential decision-task. *The American journal of psychology 78*(2), 251–256.

Brown, N. A., R. Neel, and R. A. Sherman (2015). Measuring the evolutionarily important goals of situations: Situational affordances for adaptive problems. *Evolutionary Psychology 13*(3), 1474704915593662.

Brusilovsky, P., J.-w. Ahn, T. Dumitriu, and M. Yudelson (2006). Adaptive knowledge-based visualization for accessing educational examples. In *Tenth International Conference on Information Visualisation (IV'06)*, pp. 142–150. IEEE.

Brusilovsky, P. and T. D. Loboda (2006). Wadein ii: A case for adaptive explanatory visualization. In *Proceedings of the 11th annual SIGCSE conference on Innovation and technology in computer science education*, pp. 48–52.

Burlacu, S., S. R. Patarlageanu, A. Diaconu, and G. Ciobanu (2021). E-government in the era of globalization and the health crisis caused by the covid-19 pandemic, between standards and innovation. In *SHS Web of Conferences*, Volume 92. EDP Sciences.

Camerer, C. F. and R. M. Hogarth (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of risk and uncertainty 19*(1), 7–42.

Carenini, G., C. Conati, E. Hoque, B. Steichen, D. Toker, and J. Enns (2014). Highlighting interventions and user differences: Informing adaptive information visualization support. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1835–1844.

Châlons, C. and N. Dufft (2016). Die rolle der it als enabler für digitalisierung. In *Was treibt die Digitalisierung?*, pp. 27–37. Springer.

Chan, N. H. (2004). *Time series: applications to finance*. John Wiley & Sons.

Charness, G. and D. Levin (2005). When optimal choices feel wrong: A laboratory study of bayesian updating, complexity, and affect. *American Economic Review 95*(4), 1300–1309.

Chen, D. L., M. Schonger, and C. Wickens (2016). otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance 9*, 88–97.

Cialdini, R. B., M. R. Trost, and J. T. Newsom (1995). Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications. *Journal of personality and social psychology 69*(2), 318.

Cleveland, W. S. and R. McGill (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American statistical association 79*(387), 531–554.

Cokely, E. T., M. Galesic, E. Schulz, S. Ghazal, and R. Garcia-Retamero (2012). Measuring risk literacy: the berlin numeracy test. *Judgment and Decision making*.

Comba, J. L. (2020). Data visualization for the understanding of covid-19. *Computing in Science & Engineering 22*(6), 81–86.

Conati, C., G. Carenini, D. Toker, and S. Lallé (2015). Towards user-adaptive information visualization. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*.

Czajkowski, M., K. Zagórska, and N. Hanley (2019). Social norm nudging and preferences for household recycling. *Resource and Energy Economics 58*, 101110.

Dain, S. J. (2004). Clinical colour vision tests. *Clinical and Experimental Optometry 87*(4-5), 276–293.

Dane, E., K. W. Rockmann, and M. G. Pratt (2012). When should i trust my gut? linking domain expertise to intuitive decision-making effectiveness. *Organizational behavior and human decision processes 119*(2), 187–194.

Deck, C. and S. Jahedi (2015). The effect of cognitive load on economic decision making: A survey and new experiments. *European Economic Review 78*, 97–119.

DeLone, W. H. and E. R. McLean (1992). Information systems success: The quest for the dependent variable. *Information systems research 3*(1), 60–95.

Dennett, D. C. and D. Roy (2015). Our transparent future. *Scientific American 312*(3), 64–70.

Donnellan, M. B., F. L. Oswald, B. M. Baird, and R. E. Lucas (2006). The mini-ipip scales: tiny-yet-effective measures of the big five factors of personality. *Psychological assessment 18*(2), 192.

Dudley, N. M., K. A. Orvis, J. E. Lebiecki, and J. M. Cortina (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology 91*(1), 40.

Emens, E. F. (2007). Changing name changing: Framing rules and the future of marital names. *U. Chi. L. Rev. 74*, 761.

Fein, M. L. (2015). Robo-advisors: A closer look. *Available at SSRN 2658701*.

Findlater, L. and J. McGrenere (2004). A comparison of static, adaptive, and adaptable menus. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 89–96.

Finucane, M. L., P. Slovic, J. H. Hibbard, E. Peters, C. Mertz, and D. G. MacGregor (2002). Aging and decision-making competence: An analysis of comprehension and consistency skills in older versus younger adults considering health-plan options. *Journal of Behavioral Decision Making 15*(2), 141–164.

Fischer, G. (2001). User modeling in human–computer interaction. *User modeling and user-adapted interaction 11*(1), 65–86.

Fischer, M., S. Lotz, et al. (2014). Is soft paternalism ethically legitimate?-the relevance of psychological processes for the assessment of nudge-based policies. *German Review of Social Policy*.

Fleischmann, M., M. Amirpur, A. Benlian, and T. Hess (2014). Cognitive biases in information systems research: A scientometric analysis.

Forwood, S. E., A. L. Ahern, T. M. Marteau, and S. A. Jebb (2015). Offering within-category food swaps to reduce energy density of food purchases: a study using an experimental online supermarket. *International Journal of Behavioral Nutrition and Physical Activity 12*(1), 1–10.

Friedman, S., D. Friedman, and S. Sunder (1994). *Experimental methods: A primer for economists*. Cambridge university press.

Frydenberg, M. and D. Andone (2016). Creating micro-videos to demonstrate technology learning and digital literacy. *Interactive Technology and Smart Education*.

Funder, D. C. (2009). Persons, behaviors and situations: An agenda for personality psychology in the postwar era. *Journal of Research in Personality 43*(2), 120–126.

Galesic, M. and R. Garcia-Retamero (2011). Graph literacy: A cross-cultural comparison. *Medical decision making 31*(3), 444–457.

Geller, E. S. and G. F. Pitz (1968). Confidence and decision speed in the revision of opinion. *Organizational Behavior and Human Performance 3*(2), 190–201.

Geller, E. S. and G. F. Pitz (1970). Effects of prediction, probability, and run length on choice reaction speed. *Journal of Experimental Psychology 84*(2), 361.

Gerpott, F., D. Balliet, and R. de Vries (2017). Thinking about interdependence: The situational interdependence scale. *Manuscript submitted*.

Geukes, K., S. Nestler, R. Hutteman, A. C. Küfner, and M. D. Back (2017). Trait personality and state variability: Predicting individual differences in within-and cross-context fluctuations in affect, self-evaluations, and behavior in everyday life. *Journal of Research in Personality 69*, 124–138.

Ghazal, S., E. T. Cokely, and R. Garcia-Retamero (2014). Predicting biases in very highly educated samples: numeracy and metacognition. *Judgment and decision making*.

Golemati, M., C. Halatsis, C. Vassilakis, A. Katifori, and G. Lepouras (2006). A context-based adaptive visualization environment. In *Tenth International Conference on Information Visualisation (IV'06)*, pp. 62–67. IEEE.

Gowthami, K. and M. P. Kumar (2017). Study on business intelligence tools for enterprise dashboard development. *International Research Journal of Engineering and Technology (IRJET) 4*(04), 2987–2992.

Grabitz, H.-J. (1971). Zur beziehung von inertia effekt und sequentieller position widersprechender ereignisse bei der revision subjektiver wahrscheinlichkeiten. *Psychologische Forschung 35*(1), 35–45.

Grabitz, H.-J. and G. Grabitz-Gniech (1972). The inertia effect as a function of the likelihood ratio of given information. *Zeitschrift für Experimentelle und Angewandte Psychologie*.

Halevy, N., T. A. Kreps, and C. K. De Dreu (2019). Psychological situations illuminate the meaning of human behavior: Recent advances and application to social influence processes. *Social and Personality Psychology Compass 13*(3), e12437.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review 103*(7), 2643–82.

Hanelt, A., R. Bohnsack, D. Marz, and C. Antunes Marante (2021). A systematic review of the literature on digital transformation: insights and implications for strategy and organizational change. *Journal of Management Studies 58*(5), 1159–1197.

Hart, S. G. (2006). Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Volume 50, pp. 904–908. Sage publications Sage CA: Los Angeles, CA.

Hastie, R. and R. M. Dawes (2009). *Rational choice in an uncertain world: The psychology of judgment and decision making*. Sage Publications.

Heinström, J., E. Sormunen, and S. Kaunisto-Laine (2014). Spanning information behaviour across the stages of a learning task: Where do personality and approach to studying matter? *Journal of Documentation*.

Hertwig, R. and T. Grüne-Yanoff (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science 12*(6), 973–986.

Hindal, H., N. Reid, and M. Badgaish (2009). Working memory, performance and learner characteristics. *Research in Science & Technological Education 27*(2), 187–204.

Hoffrage, U., S. Lindsey, R. Hertwig, and G. Gigerenzer (2000). Communicating statistical information.

Hollenbeck, J. R., D. R. Ilgen, D. J. Sego, J. Hedlund, D. A. Major, and J. Phillips (1995). Multilevel theory of team decision making: Decision performance in teams incorporating distributed expertise. *Journal of applied psychology 80*(2), 292.

Hoonakker, P., P. Carayon, A. P. Gurses, R. Brown, A. Khunlertkit, K. McGuire, and J. M. Walker (2011). Measuring workload of icu nurses with a questionnaire survey: the nasa task load index (tlx). *IIE transactions on healthcare systems engineering 1*(2), 131–143.

Horstmann, K., J. Ziegler, and M. Ziegler (2018). Assessment of situational perceptions: Measurement issues and a joint taxonomization of persons/situations.

Horstmann, K. T., J. F. Rauthmann, and R. A. Sherman (2018). Measurement of situational influences. *The SAGE handbook of personality and individual differences*, 465–484.

Huber, J., D. Jung, E. Schaule, and C. Weinhardt (2019). Goal framing in smart charging-increasing bev users'charging flexibility with digital nudges.

Huh, Y. E., J. Vosgerau, and C. K. Morewedge (2014). Social defaults: Observed choices become choice defaults. *Journal of Consumer Research 41*(3), 746–760.

Hummel, D. and A. Maedche (2019). How effective is nudging? a quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics 80*, 47–58.

Hussain, J., A. U. Hassan, H. S. M. Bilal, R. Ali, M. Afzal, S. Hussain, J. Bang, O. Banos, and S. Lee (2018). Model-based adaptive user interface based on context and user experience evaluation. *Journal on Multimodal User Interfaces 12*(1), 1–16.

Iacobucci, D., S. S. Posavac, F. R. Kardes, M. J. Schneider, and D. L. Popovich (2015). The median split: Robust, refined, and revived. *Journal of Consumer Psychology 25*(4), 690–704.

Iansiti, M. and K. R. Lakhani (2014). Digital ubiquity:: How connections, sensors, and data are revolutionizing business. *Harvard business review 92*(11), 19.

Ingendahl, M., D. Hummel, A. Maedche, and T. Vogel (2021). Who can be nudged? examining nudging effectiveness in the context of need for cognition and need for uniqueness. *Journal of Consumer Behaviour 20*(2), 324–336.

Inibhunu, C. and S. Langevin (2016). Adaptive visualization of complex networks with focalpoint: A context aware level of details recommender system. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Volume 60, pp. 233–237. Sage Publications Sage CA: Los Angeles, CA.

Insabato, A., M. Pannunzi, E. T. Rolls, and G. Deco (2010). Confidence-related decision making. *Journal of neurophysiology 104*(1), 539–547.

Issa, N., M. Schuller, S. Santacaterina, M. Shapiro, E. Wang, R. E. Mayer, and D. A. DaRosa (2011). Applying multimedia design principles enhances learning in medical education. *Medical education 45*(8), 818–826.

Jang, I. and H. S. Yoo (2009). Personal information classification for privacy negotiation. In *2009 Fourth International Conference on Computer Sciences and Convergence Information Technology*, pp. 1117–1122. IEEE.

Jermias, J. (2001). Cognitive dissonance and resistance to change: the influence of commitment confirmation and feedback on judgment usefulness of accounting systems. *Accounting, Organizations and Society 26*(2), 141–160.

Johnson, E. J. and D. Goldstein (2003). Do defaults save lives?

Johnson, E. J., S. B. Shu, B. G. Dellaert, C. Fox, D. G. Goldstein, G. Häubl, R. P. Larrick, J. W. Payne, E. Peters, D. Schkade, et al. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters 23*(2), 487–504.

Jung, D. (2019). *Robo-Advisory and Decision Inertia - Experimental Studies of Human Behaviour in Economic Decision-Making*. Ph. D. thesis, Karlsruher Institut für Technologie (KIT).

Jung, D. and V. Dorner (2018). Decision inertia and arousal: Using neurois to analyze bio-physiological correlates of decision inertia in a dual-choice paradigm. In *Information systems and neuroscience*, pp. 159–166. Springer.

Jung, D., V. Dorner, C. Weinhardt, and H. Pusmaz (2018). Designing a robo-advisor for risk-averse, low-budget consumers. *Electronic Markets 28*(3), 367–380.

Jung, D., E. Erdfelder, A. Bröder, and V. Dorner (2019). Differentiating motivational and cognitive explanations for decision inertia. *Journal of Economic Psychology 72*, 30–44.

Jung, D., E. Erdfelder, and F. Glaser (2018). Nudged to win: Designing robo-advisory to overcome decision inertia.

Jung, D. and C. Weinhardt (2018). Robo-advisors and financial decision inertia: How choice architecture helps to reduce inertia in financial planning tools.

Just, M. A. and P. A. Carpenter (1980). A theory of reading: from eye fixations to comprehension. *Psychological review 87*(4), 329.

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *American psychologist 58*(9), 697.

Kahneman, D. and G. Klein (2009). Conditions for intuitive expertise: a failure to disagree. *American psychologist 64*(6), 515.

Kahneman, D., S. P. Slovic, P. Slovic, and A. Tversky (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge university press.

Kahneman, D. and A. Tversky (1979). On the interpretation of intuitive probability: A reply to jonathan cohen.

Kahneman, D. and A. Tversky (2013). Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*, pp. 99–127. World Scientific.

Kenny, D. A., C. D. Mohr, and M. J. Levesque (2001). A social relations variance partitioning of dyadic behavior. *Psychological bulletin 127*(1), 128.

Kim, S., L. J. Lombardino, W. Cowles, and L. J. Altmann (2014). Investigating graph comprehension in students with dyslexia: An eye tracking study. *Research in developmental disabilities 35*(7), 1609–1622.

Kirk, A. (2012). *Data Visualization: a successful design process*. Packt publishing LTD.

Klapper, L. F., A. Lusardi, and G. A. Panos (2012). Financial literacy and the financial crisis. Technical report, National Bureau of Economic Research.

Klofstad, C. A., R. McDermott, and P. K. Hatemi (2013). The dating preferences of liberals and conservatives. *Political Behavior 35*(3), 519–538.

Kostov, V. and S. Fukuda (2001). Development of man-machine interfaces based on user preferences. In *Proceedings of the 2001 IEEE International Conference on Control Applications (CCA'01)(Cat. No. 01CH37204)*, pp. 1124–1128. IEEE.

Kozielecki, J. (1966). The mechanism of self-confirmation of hypothesis in a probabilistic situation. In *Symposium*, Volume 25, pp. 86–95.

Kretzschmar, K., H. Tyroll, G. Pavarini, A. Manzini, I. Singh, and N. Y. P. A. Group (2019). Can your phone be your therapist? young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. *Biomedical informatics insights 11*, 1178222619829083.

Kronthaler, F. (2021). Normalverteilung und weitere testverteilungen. In *Statistik angewandt mit dem R Commander*, pp. 141–156. Springer.

Kuhl, J. (1994). Action versus state orientation: Psychometric properties of the action control scale (acs-90). *Volition and personality: Action versus state orientation 47*(56).

Langley, P. (1999). User modeling in adaptive interface. In *UM99 User Modeling*, pp. 357–370. Springer.

Langner, R., T. Horak, and R. Dachselt (2015). Information visualizations with mobile devices: Three promising aspects. In *Workshop on Data Exploration for Interactive Surfaces DEXIS 2015*, pp. 20.

Larrick, R. P., K. A. Burson, and J. B. Soll (2007). Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes 102*(1), 76–94.

Lee, S., B. C. Kwon, J. Yang, B. C. Lee, and S.-H. Kim (2019). The correlation between users' cognitive characteristics and visualization literacy. *Applied Sciences 9*(3), 488.

Lehner, M., O. Mont, and E. Heiskanen (2016). Nudging–a promising tool for sustainable consumption behaviour? *Journal of Cleaner Production 134*, 166–177.

LePine, J. A., J. A. Colquitt, and A. Erez (2000). Adaptability to changing task contexts: Effects of general cognitive ability, conscientiousness, and openness to experience. *Personnel psychology 53*(3), 563–593.

Lewin, K. (1936). A dynamic theory of personality: Selected papers. *The Journal of Nervous and Mental Disease 84*(5), 612–613.

Lewin, K. (2013). *Principles of topological psychology*. Read Books Ltd.

Li, A., M. Feng, Y. Li, and Z. Liu (2016). Application of outlier mining in insider identification based on boxplot method. *Procedia Computer Science 91*, 245–251.

Li, H., Q. Yang, Y. Cao, J. Yao, and P. Luo (2020). Cracking tabular presentation diversity for automatic cross-checking over numerical facts. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2599–2607.

Li, J., M. Liu, and X. Liu (2016). Why do employees resist knowledge management systems? an empirical study from the status quo bias and inertia perspectives. *Computers in Human Behavior 65*, 189–200.

Lichtenstein, S. and P. Slovic (2006). *The construction of preference*. Cambridge University Press.

Loeb, K. L., C. Radnitz, K. Keller, M. B. Schwartz, S. Marcus, R. N. Pierson, M. Shannon, and D. DeLaurentis (2017). The application of defaults to optimize parents' health-based choices for children. *Appetite 113*, 368–375.

Lupton, D. (2016). Digital health technologies and digital data: new ways of monitoring, measuring and commodifying human bodies. In *Research handbook on digital transformations*. Edward Elgar Publishing.

Lusardi, A. and O. S. Mitchell (2007). Financial literacy and retirement planning: New evidence from the rand american life panel. *Michigan Retirement Research Center Research Paper No. WP 157*.

Lusardi, A., O. S. Mitchell, and V. Curto (2010). Financial literacy among the young. *Journal of consumer affairs 44*(2), 358–380.

Lux, M. (1997). Visualization of financial information. In *Proceedings of the 1997 workshop on New paradigms in information visualization and manipulation*, pp. 58–61.

Mackay, W. E. (1991). Triggers and barriers to customizing software. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 153–160.

Mahdzan, N. S., A. A. Mohd-Any, and M.-K. Chan (2017). The influence of financial literacy, risk aversion and expectations on retirement planning and portfolio allocation in malaysia. *Gadjah Mada International Journal of Business 19*(3), 267–288.

Majooni, A., M. Masood, and A. Akhavan (2015). Scientific visualizations based on integrated model of text and picture comprehension via eye-tracking. *Procedia-Social and Behavioral Sciences 176*, 52–59.

Majooni, A., M. Masood, and A. Akhavan (2018). An eye-tracking study on the effect of infographic structures on viewer's comprehension and cognitive load. *Information Visualization 17*(3), 257–266.

Malloy, T. R., R. S. Wigton, J. Meeske, and T. G. Tape (1992). The influence of treatment descriptions on advance medical directive decisions. *Journal of the American Geriatrics Society 40*(12), 1255–1260.

Massaro, D., F. Savazzi, C. Di Dio, D. Freedberg, V. Gallese, G. Gilli, and A. Marchetti (2012). When art moves the eyes: a behavioral and eye-tracking study. *PloS one 7*(5), e37285.

Maybury, M. and W. Wahlster (1998). *Readings in intelligent user interfaces*. Morgan Kaufmann.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. *The Cambridge handbook of multimedia learning 41*, 31–48.

Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning 33*(5), 403–423.

Mertens, S., M. Herberz, U. J. Hahnel, and T. Brosch (2022). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences 119*(1).

Miller, C. A., H. Funk, R. Goldman, J. Meisner, and P. Wu (2005). Implications of adaptive vs. adaptable uis on decision making: Why "automated adaptiveness" is not always the right answer. In *Proceedings of the 1st international conference on augmented cognition*, pp. 22–27.

Milligan, J. N. (2019). *Learning Tableau 2019: Tools for Business Intelligence, data prep, and visual analytics*. Packt Publishing Ltd.

Miraz, M. H., M. Ali, and P. S. Excell (2021). Adaptive user interfaces and universal usability through plasticity of user interface design. *Computer Science Review 40*, 100363.

Moore, D. A. and D. Schatz (2017). The three faces of overconfidence. *Social and Personality Psychology Compass 11*(8), e12331.

Morana, S., U. Gnewuch, D. Jung, and C. Granig (2020). The effect of anthropomorphism on investment decision-making with robo-advisor chatbots. In *ECIS*.

Moreno, R. and R. E. Mayer (2000). A learner-centered approach to multimedia explanations: Deriving instructional design principles from cognitive theory. *Interactive multimedia electronic journal of computer-enhanced learning 2*(2), 12–20.

Murphy, S. C., F. K. Barlow, and W. von Hippel (2018). A longitudinal test of three theories of overconfidence. *Social Psychological and Personality Science 9*(3), 353–363.

Nekmat, E. (2020). Nudge effect of fact-check alerts: source influence and media skepticism on sharing of news misinformation in social media. *Social Media+ Society 6*(1), 2056305119897322.

Nielsen, J. and T. K. Landauer (1993). A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pp. 206–213.

Nivala, A.-M. and T. L. Sarjakoski (2007). User aspects of adaptive visualization for mobile maps. *Cartography and Geographic Information Science 34*(4), 275–284.

Novak, T. P., D. L. Hoffman, and A. Duhachek (2003). The influence of experiential and goal-directed activities on online flow experiences. *Journal of Consumer Psychology 13*(1-2), 3–16.

Nurhayati, N., R. Rosmaiyadi, and B. Buyung (2018). Efforts to improve student's self confidence using collaborative learning model. *JPMI (Jurnal Pendidikan Matematika Indonesia) 2*(2), 57–62.

Okan, Y., R. Garcia-Retamero, E. T. Cokely, and A. Maldonado (2012). Individual differences in graph literacy: Overcoming denominator neglect in risk comprehension. *Journal of Behavioral Decision Making 25*(4), 390–401.

Oppermann, R. (1994). Adaptively supported adaptability. *International Journal of Human-Computer Studies 40*(3), 455–472.

Parker, A. M., W. B. De Bruin, and B. Fischhoff (2007). Maximizers versus satisficers: Decision-making styles, competence, and outcomes. *Judgment and Decision making 2*(6), 342.

Parrigon, S., S. E. Woo, L. Tay, and T. Wang (2017). Caption-ing the situation: A lexically-derived taxonomy of psychological situation characteristics. *Journal of personality and social psychology 112*(4), 642.

Payne, E. H. M., A. J. Dahl, and J. Peltier (2021). Digital servitization value co-creation framework for ai services: a research agenda for digital transformation in financial service ecosystems. *Journal of Research in Interactive Marketing*.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & cognition 17*(4), 398–422.

Pitz, G. F. (1969). An inertia effect (resistance to change) in the revision of opinion. *Canadian Journal of Psychology/Revue canadienne de psychologie 23*(1), 24.

Polites, G. L. and E. Karahanna (2012). Shackled to the status quo: The inhibiting effects of incumbent system habit, switching costs, and inertia on new system acceptance. *MIS quarterly*, 21–42.

Poole, A. and L. J. Ball (2006). Eye tracking in hci and usability research. In *Encyclopedia of human computer interaction*, pp. 211–219. IGI global.

Power, N. and L. Alison (2017). Redundant deliberation about negative consequences: Decision inertia in emergency responders. *Psychology, Public Policy, and Law 23*(2), 243.

Primi, C., M. Donati, S. Massino, E. Borace, E. Franchi, and K. Morsanyi (2019). Measuring probabilistic reasoning: the development of a brief version of the probabilistic reasoning scale (prs-b). In *Eleventh Congress of the European Society for Research in Mathematics Education (CERME11)*, Number 31. Freudenthal Group; Freudenthal Institute; ERME.

Qin, X., Y. Luo, N. Tang, and G. Li (2018). Deepeye: An automatic big data visualization framework. *Big data mining and analytics 1*(1), 75–82.

Raoelison, M., V. A. Thompson, and W. De Neys (2020). The smart intuitor: Cognitive capacity predicts intuitive rather than deliberate thinking. *Cognition 204*, 104381.

Rauthmann, J. F. (2012). You say the party is dull, i say it is lively: A componential approach to how situations are perceived to disentangle perceiver, situation, and perceiver× situation variance. *Social Psychological and Personality Science 3*(5), 519–528.

Rauthmann, J. F. (2016). Motivational factors in the perception of psychological situation characteristics. *Social and Personality Psychology Compass 10*(2), 92–108.

Rauthmann, J. F., D. Gallardo-Pujol, E. M. Guillaume, E. Todd, C. S. Nave, R. A. Sherman, M. Ziegler, A. B. Jones, and D. C. Funder (2014). The situational eight diamonds: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology 107*(4), 677.

Rauthmann, J. F. and R. A. Sherman (2015a). Measuring the situational eight diamonds characteristics of situations. *European Journal of Psychological Assessment*.

Rauthmann, J. F. and R. A. Sherman (2015b). Ultra-brief measures for the situational eight diamonds domains. *European Journal of Psychological Assessment*.

Rauthmann, J. F. and R. A. Sherman (2016). Situation change: Stability and change of situation variables between and within persons. *Frontiers in Psychology 6*, 1938.

Rauthmann, J. F. and R. A. Sherman (2018). The description of situations: Towards replicable domains of psychological situation characteristics.

Rauthmann, J. F., R. A. Sherman, and D. C. Funder (2015). Principles of situation research: Towards a better understanding of psychological situations. *European Journal of Personality 29*(3), 363–381.

Reichenbacher, T., S. Shekhar, and X. Hui (2008). Mobile usage and adaptive visualization. *Springer reference*, 677–682.

Riemann, R. and A. Allgöwer (1993). Eine deutschsprachige fassung des „interpersonal competence questionnaire "(icq). *Zeitschrift für differentielle und diagnostische Psychologie 14*(3), 153–163.

Rose, J. M., F. D. Roberts, and A. M. Rose (2004). Affective responses to financial data and multimedia: The effects of information load and cognitive load. *International Journal of Accounting Information Systems 5*(1), 5–24.

Rothrock, L., R. Koubek, F. Fuchs, M. Haas, and G. Salvendy (2002). Review and reappraisal of adaptive interfaces: toward biologically inspired paradigms. *Theoretical issues in ergonomics science 3*(1), 47–84.

Rudolph, S., A. Savikhin, and D. S. Ebert (2009). Finvis: Applied visual analytics for personal financial planning. In *2009 IEEE symposium on visual analytics science and technology*, pp. 195–202. IEEE.

Saal, M., S. Starnes, and T. Rehermann (2017). Digital financial services.

Sahin and Yilmaz (2014). Overconfidence in finance wıth different domains: An interdisciplinary experimental approach. *Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi 15*(1), 253–268.

Salo, M. and H. Haapio (2017). Robo-advisors and investors: Enhancing human-robot interaction through information design. In *Trends and Communities of Legal Informatics. Proceedings of the 20th International Legal Informatics Symposium IRIS*, pp. 441–448.

Saucier, G., T. Bel-Bahar, and C. Fernandez (2007). What modifies the expression of personality tendencies? defining basic domains of situation variables. *Journal of personality 75*(3), 479–504.

Sautua, S. I. (2017). Does uncertainty cause inertia in decision making? an experimental study of the role of regret aversion and indecisiveness. *Journal of Economic Behavior & Organization 136*, 1–14.

Schneider-Hufschmidt, M., U. Malinowski, and T. Kuhme (1993). *Adaptive user interfaces: Principles and practice*. Elsevier Science Inc.

Schunck, R. and N. Nisic (2020). Vorsicht bei vorschlägen zur interpretation von regressionen mit interaktionsvariablen. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie 72*(1), 109–119.

Schwartz, B., A. Ward, J. Monterosso, S. Lyubomirsky, K. White, and D. R. Lehman (2002). Maximizing versus satisficing: happiness is a matter of choice. *Journal of personality and social psychology 83*(5), 1178.

Seddon, P., M.-Y. Kiew, et al. (1996). A partial test and development of delone and mclean's model of is success. *Australasian Journal of Information Systems 4*(1).

See, Y. H. M., G. Valenti, A. Y. Ho, and M. S. Tan (2013). When message tailoring backfires: The role of initial attitudes in affect–cognition matching. *European Journal of Social Psychology 43*(6), 570–584.

Seifert, A. (2020). The digital exclusion of older adults during the covid-19 pandemic. *Journal of Gerontological Social Work 63*(6-7), 674–676.

Selker, T. and W. Burleson (2000). Context-aware design and interaction in computer systems. *IBM systems Journal 39*(3.4), 880–891.

Shah, P. (2002). Graph comprehension: The role of format, content and individual differences. In *Diagrammatic representation and reasoning*, pp. 173–185. Springer.

Sharafi, Z., T. Shaffer, B. Sharif, and Y.-G. Guéhéneuc (2015). Eye-tracking metrics in software engineering. In *2015 Asia-Pacific Software Engineering Conference (APSEC)*, pp. 96–103. IEEE.

Sherman, R. A., C. S. Nave, and D. C. Funder (2010). Situational similarity and personality predict behavioral consistency. *Journal of personality and social psychology 99*(2), 330.

Sherman, R. A., J. F. Rauthmann, N. A. Brown, D. G. Serfass, and A. B. Jones (2015). The independent effects of personality and situations on real-time expressions of behavior and emotion. *Journal of personality and social psychology 109*(5), 872.

Shi, L., N. Cao, S. Liu, W. Qian, L. Tan, G. Wang, J. Sun, and C.-Y. Lin (2009). Himap: Adaptive visualization of large-scale online social networks. In *2009 IEEE Pacific Visualization Symposium*, pp. 41–48. IEEE.

Shipley, N. L., M. J. Jackson, and S. Segrest (2010). The effects of emotional intelligence, age, work experience, and academic performance.

Smith, V. L. (1976). Experimental economics: Induced value theory. *The American Economic Review 66*(2), 274–279.

Spence, S. (1995). Descartes' error: Emotion, reason and the human brain. *BMJ 310*(6988), 1213.

Stebbins, R. A. (2001). *Exploratory research in the social sciences*, Volume 48. Sage.

Steichen, B., G. Carenini, and C. Conati (2013). User-adaptive information visualization: using eye gaze data to infer visualization tasks and user cognitive abilities. In *Proceedings of the 2013 international conference on Intelligent user interfaces*, pp. 317–328.

Story, M. F. (1998). Maximizing usability: the principles of universal design. *Assistive technology 10*(1), 4–12.

Strack, F. and R. Deutsch (2004). Reflective and impulsive determinants of social behavior. *Personality and social psychology review 8*(3), 220–247.

Streit, M. and N. Gehlenborg (2014). Bar charts and box plots: creating a simple yet effective plot requires an understanding of data and tasks. *Nature methods 11*(2), 117–118.

Stryja, C., G. Satzger, and V. Dorner (2017). A decision support system design to overcome resistance towards sustainable innovations.

Stutzer, A., L. Goette, and M. Zehnder (2011). Active decisions and prosocial behaviour: a field experiment on blood donation. *The Economic Journal 121*(556), F476–F493.

Sunstein, C. R. (2018). Nudging: a very short guide. In *The Handbook of Privacy Studies*, pp. 173–180. Amsterdam University Press.

Sunstein, C. R. and R. H. Thaler (2008). Privatizing marriage. *The Monist 91*(3/4), 377–387.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta psychologica 47*(2), 143–148.

Ten Berge, M. A. and B. De Raad (1999). Taxonomies of situations from a trait psychological perspective. a review. *European Journal of Personality 13*(5), 337–360.

Thaler, R. H. and C. R. Sunstein (2003). Libertarian paternalism. *American economic review 93*(2), 175–179.

Thaler, R. H. and C. R. Sunstein (2009). Nudge: Improving decisions about health, wealth, and happiness.

Toker, D. and C. Conati (2014). Eye tracking to understand user differences in visualization processing with highlighting interventions. In *International Conference on User Modeling, Adaptation, and Personalization*, pp. 219–230. Springer.

Toker, D., C. Conati, G. Carenini, and M. Haraty (2012). Towards adaptive information visualization: on the influence of user characteristics. In *International conference on user modeling, adaptation, and personalization*, pp. 274–285. Springer.

Tseng, K. (2006). Behavioral finance, bounded rationality, neuro-finance, and traditional finance. *Investment Management and Financial Innovations* (3, Iss. 4), 7–18.

Van Heck, G. L. (1989). Situation concepts: Definitions and classification.

Van Rooij, M., A. Lusardi, and R. Alessie (2011). Financial literacy and stock market participation. *Journal of Financial economics 101*(2), 449–472.

Velez, M. C., D. Silver, and M. Tremaine (2005). Understanding visualization through spatial ability differences. In *VIS 05. IEEE Visualization, 2005.*, pp. 511–518. IEEE.

Villardón-Gallego, L., C. Yániz, C. Achurra, I. Iraurgi, and M. del Carmen Aguilar (2013). Learning competence in university: Development and structural validation of a scale to measure//la competencia para aprender en la universidad: Desarrollo y validación de un instrumento de medida. *Revista de Psicodidáctica 18*(2).

Von Neumann, J. and O. Morgenstern (2007). *Theory of games and economic behavior*. Princeton university press.

Wahlster, W. and A. Kobsa (1989). User models in dialog systems. *User models in dialog systems*, 4–34.

Wang, P. and J. Wang (2021). Research on the application of data visualization in the ui interface design of health apps. In *2021 International Wireless Communications and Mobile Computing (IWCMC)*, pp. 2013–2019. IEEE.

Weinmann, M., C. Schneider, and J. Vom Brocke (2016). Digital nudging. *Business & Information Systems Engineering 58*(6), 433–436.

Wright, H. H., R. A. Downey, M. Gravier, T. Love, and L. P. Shapiro (2007). Processing distinct linguistic information types in working memory in aphasia. *Aphasiology 21*(6-8), 802–813.

Yarbus, A. L. (2013). *Eye movements and vision*. Springer.

Yates, J. F. (1990). *Judgment and decision making.* Prentice-Hall, Inc.

Yelizarov, A. and D. Gamayunov (2014). Adaptive visualization interface that manages user's cognitive load based on interaction characteristics. In *Proceedings of the 7th International Symposium on Visual Information Communication and Interaction*, pp. 1–8.

Zhang, B. and H. Xu (2016). Privacy nudges for mobile applications: effects on the creepiness emotion and privacy attitudes. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*, pp. 1676–1690.

Zhang, S., J. F. Cornwell, and E. T. Higgins (2014). Repeating the past: Prevention focus motivates repetition, even for unethical decisions. *Psychological science 25*(1), 179–187.

Zhang, X., S. Liu, X. Chen, L. Wang, B. Gao, and Q. Zhu (2018). Health information privacy concerns, antecedents, and information disclosure intention in online health communities. *Information & Management 55*(4), 482–493.

Zhang, Y., S. Hou, H. LIU, S. WANG, and X. ZHANG (2020). Research on application of data visualization in finance. *DEStech Transactions on Engineering and Technology Research,(acaai)*.

Zhu, F., W. Lei, Y. Huang, C. Wang, S. Zhang, J. Lv, F. Feng, and T.-S. Chua (2021). Tat-qa: A question answering benchmark on a hybrid of tabular and textual content in finance. *arXiv preprint arXiv:2105.07624*.

Ziegler, M., K. T. Horstmann, and J. Ziegler (2019). Personality in situations: Going beyond the ocean and introducing the situation five. *Psychological Assessment 31*(4), 567.

Ziemkiewicz, C., R. J. Crouser, A. R. Yauilla, S. L. Su, W. Ribarsky, and R. Chang (2011). How locus of control influences compatibility with visualization style. In *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 81–90. IEEE.