MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„**Context Matters** -
Choice Bracketing as a Complexity and Context-Dependent Decision-Making Phenomenon"

verfasst von / submitted by

Maren Suffel, BSc

gemeinsam mit / in collaboration with

Mounir Derboghossian, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of-

Master of Science (MSc)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 915
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Betriebswirtschaft
Betreut von / Supervisor:	o. Univ.-Prof. Mag. Dr. Rudolf Vetschera
Mitbetreut von / Co-Supervisor:	Aysegül Engin, BSc Bakk. Msc PhD

Table of Contents

Note: Initials in parenthesis name the author of the respective chapter

M.D. - Mounir Derboghossian

M.S. - Maren Suffel

List of Figures

List of Tables

List of Abbreviations

DM Decision Maker

EV Expected Value

MINT Mathematics, Informatics, Natural sciences and Technical studies

EUT Expected Utility Theory

PT Prospect Theory

CEST Cognitive Experiential Self Theory

TC Team Compatibility

PF Position-Fit

CRT Cognitive Reflection Test

REI Rational Experiential Inventory

1 Introduction

Every individual has to do it on a daily basis – make decisions. And each of the approximately 20.000 daily decisions (Pöppel, 2008, p. 19) entails positive or negative consequences for the decision maker. This possibility of a gain or loss makes decision-making an aspect that is not only of high relevance for the personal life of an individual, but also plays a significant role when it comes to their professional life. Especially in an organisational context, where the organisation itself cannot decide on its actions, the decisions undertaken by its many agents may have far-reaching consequences on its long-term survival. Yet, a decision rarely stands on its own but is embedded into an environment of preceding and succeeding decisions, which address the same or a related issue. In this sense, life is a concatenation of decisions where a single choice may have consequences on other decisions, leading to a network of cross-effects and interrelations that together form global patterns. The emergence of such interrelations can severely influence what would constitute an appropriate choice, compared to the isolated assessment of a single decision. Imagine the following situations:

> *Your alarm goes off in the early morning hours so that your sleepy self is faced with the choice of either getting out of bed and starting into the day according to plan or hitting the snooze button for "just 10 more minutes". Although the gain from staying in seems to clearly outweigh the alternative of leaving a cosy bed, the consequences of this choice impact the following decisions that would arise each morning. Reducing the available time for one's morning routine may force you to skip your beloved cup of coffee and leave for work hungrily. Going back to the moment of hitting the snooze button – when considering the whole morning routine instead of the single decision whether to sleep another 10 minutes, would your decision be the same?*

> *You enjoy having your daily bottle of coke after getting home from a long day at work. You know that the sugar in the beverage makes it an unhealthy choice but the damage done by deciding for that single bottle appears to be minuscule, so that it does not outweigh the pleasure gained from its consumption. But repeating that single choice each evening may have serious health consequences in the long-run. The aggregated negative impact of a coke a day may easily outweigh the gain from that single bottle.*

You are planning to invest a share of the savings on your bank account for a longer time period and are offered two options to undertake this intention. You may either invest into a safe portfolio of lower return rate or into a riskier portfolio that could yield better results. You take the safe option because you do not want to lose your savings. However, when considering the long time-horizon of the investment, the risky portfolio resembles a repeating gamble that will lose on some occasions but win on others. From an overall perspective, the development of your assets under this option may seem more volatile, but in the long-term is likely to yield a higher return.

Lastly, you are overseeing several departments of your company. Over the course of the first quarter each of the respective departments awaits your decision of whether to continue their current strategy and take on projects that are in line with what has proven to be successful in the past. Alternatively, each department could launch an innovative project that may either be of high value or great loss to the department. You take the decision for each department whenever it arises and since the responsibility of the department's ruin would be placed on you, you always choose the safe projects for the upcoming year. But when keeping the decisions for the remaining departments in mind whilst deciding on the current one, it might appear considerably more attractive to arrange for the risky innovation in some departments and the safe projects in others, so that the overall risk is still negligible.

What these decision problems have in common is that they are all subject to the concept of choice bracketing (Read et al., 1999). The term refers to the process of grouping several individual decisions into sets. By forming such decision sets, the possible consequences of each choice on the other choices within, but not on those outside of the same set, are considered for the evaluation of each decision. Whether a decision maker is described by broad or narrow bracketing behaviour then depends on the number of such interrelated decisions that are grouped together. If a set consists of one or only a few decisions, an individual is referred to as a narrow bracketer whilst a broad bracketing individual forms sets of larger magnitude (Read et al., 1999). It follows, that only a broad bracketer is able to perceive global patterns and incorporate the respective cross-effects between interrelated choices into his decision-making process.

Taking a more global perspective by bracketing multiple decisions together instead of considering them in isolation generates an enhanced basis of information but also places an increased demand on the information processing capability of a decision maker. Consequently, being able to form decisions based on an extended information pool can lead to superior choices. However, there is a natural limitation to how broad a decision set may be, and this barrier depends on the individual characteristics of the decision maker. The negative consequences of isolated decision-making for an individual or organisation generate the need to identify the drivers of bracketing behaviour. Understanding these determinants is a crucial step towards the active manipulation of bracketing behaviour in such a manner that it results in maximizing the achieved gain. In an organisational context, guiding agents' bracketing behaviour yields an improved decision quality and contributes to profit maximisation, risk diversification, and securing the organisation's existence in the long-run.

The positive effects of applying the appropriate choice bracket are well researched. The most prominent benefits include the ability to recognise aggregate measurements and patterns (Read et. al, 1999; Sabini and Silver, 1982; Haisley et al., 2008; Webb and Shu, 2017), diversifying choices (Simonson, 1990; Read and Loewenstein, 1995; Change et al., 2020), being able to implement trade-offs between choices (Read et al. 1999; Felső and Soetevent, 2014; Abeler and Marklein, 2017), as well as improved self-control (Rachlin, 1995; Koch and Nafziger, 2019). When it comes to what drives the behaviour that is proposed to achieve superior decision-making, no universal framework is given. Nevertheless, several possible determinants are well represented in existing literature. Attention is given to the nature of the decision problem itself, with a focus on risk attitude as well as the framing of decisions (Tversky and Kahneman, 1981; Redelmeier and Tversky, 1992; Kahneman and Lovallo, 1993; Stracke et al., 2015). Moreover, narrow or broad bracketing is attributed to individual characteristics in the form of information processing (Pacini and Epstein, 1999; Koch and Nafziger 2016a) and cognitive limitations (Herrnstein et al., 1993; Stracke et al., 2015). The application of heuristic decision-making (Tversky and Kahneman, 1981; Barbersis et al., 2006) and motivated bracketing to achieve self-control represent further frequently discussed drivers of choice bracketing (Hsiaw, 2018; Koch and Nafziger, 2020). Especially in the context of decision heuristics, the objective complexity of a decision problem as an influencing factor is also noted in research (Stracke et al. 2017).

While literature acknowledges determinants of a decision-specific as well as a person-specific nature, the combination of both aspects is underrepresented. Furthermore, the concept of choice bracketing is most prominently examined in the context of decisions under risk, decisions on personal consumption and expenditure, as well as motivational problems. In contrary, a managerial context is rarely encountered. Thus, the contribution of the underlying thesis lies in the investigation of the following issues. The perceived complexity, that results from applying the psychographic characteristics of a specific decision maker to the properties of the given decision problem, constitutes the first focal point of this study. This perceived complexity is then argued to determine bracketing behaviour together with a direct influence of the individual's psychographics. The additional impact of these characteristics originates from the assumption that they provide a baseline tendency of bracketing, regardless of the specific decision problem.

Consequently, the proposed framework of this thesis connects the characteristics of the decision maker and those of the decision problem with bracketing behaviour and the resulting solution quality. Furthermore, this study examines choice bracketing by incorporating two contexts of different nature. One is given by the well-researched context of lottery choices, whilst the other addresses the underrepresented area of managerial decision-making in the form of personnel selection. The inclusion of a personnel selection context further allows for the investigation of variety seeking behaviour and possible overdiversification in a non-consumer choice setting. Moreover, the examination of two distinct contexts enables this study to not only compare behaviour across different tasks of the same background but also across different contexts. The underlying question of whether context matters is another aspect that is not well represented in existing literature.

Another contribution is based on not relying on the effect of pure decision framing, but instead using information presentation as a means of manipulating bracketing behaviour. This provides the opportunity of exploring the impact of modifying the objective task complexity on bracketing behaviour and the resulting solution quality. By doing so, the opportunity of implementing decision support finds its application in this study.

The thesis is structured as follows. After this first introduction to the thesis and the topic, chapter 2 guides the reader through the theoretical background by elaborating on the development of the concept and the definition of choice bracketing itself. The main effects of choice bracketing as well as the proposed determinants that govern this behaviour are discussed here. Subsequently, chapter 3 presents the research framework along with the main hypotheses that are drawn from it. In chapter 4, the methodology of the thesis is explained as the reader is navigated through the design of the conducted experiment. The following chapter 5 states the results of this experiment by first reporting the descriptive analysis before continuing with the statistical testing of the hypotheses. Chapter 6 provides a discussion of the findings, concludes the thesis, reviews its limitations, and gives a perspective on future research.

2 Theoretical Background

There is an ever-growing body of literature concerning the field of decision-making; driven by the circumstance that it is an essential part of everyday life with the potential to cause gain or loss for the decision maker (DM). Hence, research in this area shifts its focus away from the idea of a rational decision maker towards the behavioural deviations from standard models of rational choice. In consequence, two perspectives on decision-making can be found in literature. The normative perspective takes a traditional approach by having the construct of rationality at its core, whereas a descriptive analysis results from the shift of research focus and is concerned with behaviour as it is observed, instead of prescribed by normative models. It is suggested that the theory of choice bracketing is able to explain a variety of irregularities observed by the normative perspective, such as choosing dominated options, failing to diversify investments, and adopting an unreasonable short-term perspective when evaluating a choice.

The aim of this chapter is to provide an overview of the literature that serves as a basis for the development of the research framework of this thesis. First, since it is the groundwork for every theory of behavioural irregularity, a review of rational decision-making is given. Next, the reader is introduced to a variety of synonymous labels that all address elements of what Read et al. (1999) then labelled choice bracketing. After presenting the general concept of choice bracketing, this chapter continues by providing an overview of the effects of choice bracketing as well as the proposed causes for specific bracketing behaviour.

2.1 Rational Decision-Making

Before one can elaborate on what represents a rational decision-making behaviour, it is necessary to define the underlying decision problem. The model of a typical decision problem consists of the following elements (Tversky and Kahneman, 1981; Simon, 1955): (1) A set of options or alternative behaviours from which the DM can choose. Here it is important to note that a perfectly rational DM is assumed to perceive all possible alternatives instead of a mere subset of such options and hence is able to consider the global model. (2) The possible outcomes and consequences that each option yields as well as a function describing their value for the DM. (3) Since this thesis is concerned with decision-making under risk, the probabilities that link possible outcomes to their respective options. The attractiveness of an option can then be described by its expected value (EV). Kahneman and Tversky (1984) note that for an idealized DM, the experienced value associated with an alternative coincides with the actual decision value. This means that for such an individual, the satisfaction or dissatisfaction experienced from an alternative would correspond to the attractiveness of this alternative based on its anticipated outcome. This is because an idealized DM is assumed to be capable of correctly predicting future experiences.

It is then proposed that rational choices should be governed by the two essential principles of dominance and invariance (von Neumann and Morgenstern, 1944). Dominance requires that option A should be chosen if it is equally good or better than B in all aspects and strictly better in at least one aspect. The second principle demands that the preference order between several alternatives does not change when the alternatives are presented in a different form. This means that two different descriptions of the same decision problem should not lead to a preference reversal for the DM.

Economic theory expects that decisions are made efficiently, so that the expected wealth, or utility when referring to consumption decisions, of a DM is maximized (Herrnstein and Prelec, 1992a; Tversky and Kahneman, 1981; Benjamin et al., 2013). However, the criterions of invariance and dominance are often violated, ultimately leading to non-optimal results and the failure to implement the maximization objective. Literature attributes this failure to the inaccurate assumption that decisions are made from a global perspective that incorporates all the necessary aspects of the current decision problem along with the consideration of expected consequences and future opportunities of other choices (Kahneman, 2003; Kahneman and Lovallo, 1993, Read et al., 1999, Rabin and Weizsäcker, 2009; Gneezy and Potters, 1997). This discrepancy in decision-making when taking a global or local perspective motivates the various labels that describe choice bracketing in literature.

2.2 A Variety of Labels

Before the work of Read et al. (1999) introduced the term choice bracketing, more or less synonymous concepts address the issue of making choices either in a comprehensive and inclusive context or by deciding locally without the consideration of the global consequences. This subchapter aims at providing an overview of the variety of labels that can be found in literature as well as pointing out their contribution to choice bracketing.

Simonson (1990) experiments with purchasing strategies of consumers who have to choose several products of a specific category. In the first strategic setting, consumers have to select multiple products for sequential consumption occasions whereas the second setting requires them to choose one product at each trip to the store. The two strategies of decision-making are then labelled simultaneous and sequential choice (Simonson, 1990). The author argues that a simultaneous choice for sequential consumption causes the DM to select a greater variety; similar to an investor who diversifies his portfolio in order to spread the risks of his investments. The link to choice bracketing can be easily found in the sense that the simultaneous setting bundles several individual decisions at a single point in time. This means that the DM still faces the same number of choices (one for each future consumption occasion) but makes all choices at once and thus takes a global perspective or broad bracket for decisions made in this product category. Accordingly, the sequential choice strategy describes choices made with a focus on the local consequences of each individual consumption decision and therefore matches the idea of narrow bracketing by Read et al. (1999).

Another concept with an obvious connection to choice bracketing are narrow and broad decision frames. Kahneman and Lovallo (1993) argue that a DM tends to view a decision as if it was a unique event. When considering a choice as unique, the DM does not account for future consequences or opportunities and fails to include past decisions in the evaluation process. This narrow decision frame clearly describes local decision-making and can be seen as a synonym to narrow choice bracketing. The authors argue that such a behaviour is an isolation error and has two specific consequences: first, the ignorance of past statistics and focus on current plans causes decisions to be exceedingly optimistic. Second, the ignorance of aggregation effects when considering a single prospect influences the willingness to take risks (Kahneman and Lovallo, 1993). With respect to overly optimistic choices as a result of perceiving a decision problem as unique, Kahneman and Lovallo (1993) propose the approach of distinguishing between an inside and outside view in the field of forecasting. The inside view focuses on the details of the specific case at hand and existing future plans while an outside view does not attempt to predict detailed future development and mainly takes a comparative view of related cases. It is suggested that the adoption of an outside view enables a DM to overcome the consequences of the isolation error. The connection of the two views to choice bracketing might not come at first sight, being that the outside view is argued to not focus on future developments or forecasting its history. Nevertheless, it is the statistical and comparative nature of the outside view that links it to broad bracketing, where the project at hand is one of a greater set that is taken into account. On the other hand, the inside view exclusively focuses on the project at hand, treating it as a unique decision and therefore represents narrow choice bracketing.

Herrnstein and Prelec (1992a) refer to the observation that DMs tend to make their choices on the basis of stable patterns instead of equalizing marginal returns per input unit. Their melioration theory seeks to explain the deviation from maximization in a situation of repeated choice. In this context, the global outcome is an aggregate of multiple smaller decisions, where each choice influences the value of the alternatives in future choices. A possible application would be the decision between two types of cuisines or the allocation of time between two leisure activities (Herrnstein and Prelec, 1992a). The more often each option is selected by the DM, the more its value decreases in future decisions so that the distribution is characterised by diminishing returns for either alternative. Herrnstein and Prelec (1992a) then propose that the typical DM does not calculate the overall value of a specific distribution of choices but instead chooses the option which currently has the higher return. According to the authors, in the long-run, this behaviour leads to a stable distribution of choices in an equilibrium, where the values of the options are either equal or a corner solution (only one option) is reached. In contrary, the optimal distribution between the options would be achieved by considering the option values as weighted average, where the relative frequencies with which each alternative is chosen represent the weights. A subject who meliorates is then defined as failing to consider the cross-effect that the present choice has on the value of the upcoming alternatives and hence chooses the option which currently has the higher return. Thus, the theory is based on the assumption that a DM is able to rank the various options in a direct comparison but is unable to evaluate the resulting distribution in repeated choice settings. Herrnstein and Prelec (1992a, p. 150) call this ignorance of the consequences of a choice for the DM himself a "within-person externality", which results in partial maximization.

The melioration theory nicely illustrates that making a choice today without an eye on the global distribution may negatively impact the value of a choice of tomorrow. By this means, the proposed theory treats the same concept as choice bracketing with a meliorating subject being a narrow bracketer. Heyman (1996) successfully applies the melioration theory to the study of addiction by discussing that a local reference frame causes excessive drug use whereas a global perspective induces more control, ultimately leading to less drug abusive behaviour.

One final approach from the family of choice bracketing is myopic loss aversion (Benartzi and Thaler, 1995; Thaler 1999), which was developed to resolve the phenomenon that people avoid favourable small gambles and received particular recognition for solving the equity premium puzzle. As a short excursion for the reader: the equity premium puzzle refers to the observation that people prefer safe bonds over stocks when building their long-term portfolio despite the fact that stocks would yield a higher return in the long-run. It is found that the behaviour can be explained by the time horizon and frequencies with which investors evaluate their portfolios. Myopic loss aversion then refers to the combination of frequent evaluations and loss aversion provoking a risk averse decision-making behaviour. Therefore, the short-term or local perspective of myopic loss aversion prevents the DM from making more favourable investments as it emphasizes the higher variability of stocks compared to bonds and, hence, corresponds to narrow choice bracketing. In contrast, a global perspective, or broad choice bracket, would point out the superiority of stocks in the long-run and enables the DM to build a more profitable portfolio (Thaler et al., 1997; Benartzi and Thaler, 1995; Thaler, 1999).

The development of a variety of synonymous labels as well as the application in different contexts indicate the importance of the choice bracketing phenomenon in research on decision-making. Taken together, there is a strong consensus that subjects tend to adopt an inside view, treat decision problems as unique instances, meliorate, fail to consider overall utility, take a short-term perspective, and apply a sequential decision strategy. Consequently, narrow bracketing appears to be the prevalent approach for a DM and these observations demand further investigation in search for an explanation.

Literature not only offers a variety of synonymous labels, but one can also find concepts that are closely related to choice bracketing but need to be distinguished as they do not act as synonyms. Before continuing the analysis of the bracketing theory, two related models shall be identified in order to avoid confusing them with choice bracketing.

Outcome editing proposes that a DM either integrates or segregates possible outcomes when evaluating the utility of a decision (Kahneman and Tversky, 1979; Thaler 1985). This is best illustrated by the means of a lottery decision, which are commonly used in the choice bracketing literature (Tversky and Kahneman, 1981). A DM has to make choices concerning two separate lotteries. The first lottery is a choice between a sure gain and a probabilistic prospect with higher EV. The second lottery contains a sure loss option and a probabilistic prospect with a smaller expected loss. This decision problem induces four possible combinations of choices. Now, the DM can either segregate the outcome of each choice combination by considering the sure and probabilistic gain or loss components of the statement separately, as they were presented to him in the decision problem. Alternatively, he can integrate them by combining sure and probabilistic outcomes to one statement of a possible gain or loss respectively. This description appears to be very similar to choice bracketing but Read et al. (1999) argue that the application of the concept depends on the point at which the editing takes place. It is a bracketing effect if a subject does not accomplish the transformation into four choice combinations, but an effect of outcome editing if the subject recognizes the four alternative representations and then fails to integrate the outcomes. Read et al. (1999) further suggest that in the majority of cases, subjects fail at the choice bracketing component; meaning that they evaluate each choice on its own and therefore do not reach the point of possible outcome editing.

A second relative of choice bracketing is the joint versus separate evaluation of alternatives (Hsee et al., 1999). The concept differs from choice bracketing because it concerns the alternatives within a choice and not between several choices. When a DM makes use of separate evaluation, he evaluates each alternative within a choice separately without considering the neighbouring options. Joint evaluation occurs when the DM makes comparisons between the options within a single choice. For the lottery example from above, separate evaluation would take place when the decision for the first lottery is made without explicitly comparing the sure gain to the probabilistic option. In contrary, narrow bracketing occurs when making the decision for the first lottery without considering the options in the second lottery.

2.3 The Concept of Choice Bracketing

Having discussed the problem with rational decision-making that serves as a motivation for a variety of studies, some attention was placed on introducing choice bracketing by first exploring some of its predecessors along with two close relatives. In continuation of this exploration, this chapter intends to give a more profound overview of the bracketing theory as it was proposed by Read et al. (1999).

The underlying principle of choice bracketing is the act of grouping several separate decisions into sets of choices. Decisions within such a set are then said to be bracketed together when the DM considers the consequences and indirect effects of each choice on the other choices in the same set. Possible consequences on choices that are not grouped into the set are ignored by the DM. The terms narrow and broad bracketing refer to the magnitude of a decision set, where broad bracketing occurs when a set contains many individual choices and narrow bracketing takes place when the set includes only one or a few choices (Read et al., 1999). The more choices are grouped into a decision set, the more global is the perspective of the DM as he considers an increasing number of cross-effects between his choices. On the contrary, small decision sets prevent the DM from acknowledging the consequences of his actions, being that his attention shifts from a macro to a micro level. Read et al. (1999, p. 172) draw a fitting comparison, which might as well be one of the oldest applications of bracketing, by describing that narrow bracketing "is like fighting a war one battle at a time with no overall guiding strategy".

As in the preceding chapter, the lottery example offers a good opportunity to demonstrate choice bracketing. Remember, there were two individual lotteries, each comprising two decision alternatives. Lottery 1 could be denoted as the choice between the options {A, B} while the second lottery is described by {C, D}. A broad bracketing subject recognizes the lotteries as a sequence of interrelated events and groups both lotteries into the same decision set. Hence, he would be confronted by a choice between the four alternative choice combinations {AB, AD, BC, BD}. A narrow bracketer, however, fails to group both lotteries together and first decides between A and B, followed by the decision between C and D. When the sequence of choices is important, for example the order of first choosing lottery 1 followed by lottery 2, then temporal bracketing applies. Being that most decisions take place in a sequential order, one can expect that the majority of bracketing behaviour is of a temporal nature (Read et al., 1999).

The responses to the original lottery example from Tversky and Kahneman (1981, p. 454) depict this theoretical behaviour and its negative effect on an individual's welfare. Consider the following decision problem:

Lottery 1: Choose between

A. A sure gain of $240
B. 25% chance to gain $1000, and 75% chance to gain nothing

Lottery 2: Choose between

C. A sure loss of $750
D. 75% chance to lose $1000, and 25% chance to lose nothing

The alternatives A in the first lottery and D in the second decision are chosen by the majority of subjects, with 84% and 87% respectively. This is explained by the circumstance that subjects are loss averse and weight certain and probabilistic outcomes disproportionately (Kahneman and Tversky, 1979; Tversky and Kahneman, 1981). As a consequence, participants in the study behaved risk averse in the first lottery but shifted to a risk seeking behaviour in the second one. Note that, when bracketing the two lotteries together, it should become clear that the choice combination AD is first-order stochastically dominated by the combination BC. This is because the combination BC yields the same probabilities for gains and losses as the combination AD, but with a sure surplus of 10$. Nevertheless, the combination AD is the most frequently chosen pattern (73%), showing that these respondents failed to take a global perspective of the problem. In comparison, the dominant combination BC is only selected by 3%. It is especially interesting that people fail at this task although the instructions to the experiment explicitly stated that both decision problems shall be read before making any choices. This implies that people have a strong tendency to bracket their decisions narrowly that persists even in the presence of a clear invitation to bracket broadly (Rabin and Weizsäcker, 2009). However, when the broad bracketing task is carried out on behalf of the respondents and they are presented with the four choice combinations, none of them choose the dominated pair AD (Tversky and Kahneman, 1981). This result confirms that subjects would have been able to evaluate the prospects but failed to recognize that the lotteries should be treated as a choice set instead of two unique decisions. Other studies were able to replicate the effect of executing the broad bracketing step for their respondents (Thaler et al., 1997, Gneezy and Potters, 1997; Rabin and Weizsäcker, 2009). Ellis and Freeman (2020) criticise the missing information about the bracketing behaviour of individuals who neither select AD nor BC but one of the other two combinations in the design of Tversky and Kahneman (1981) as well as Rabin and Weizsäcker (2009).

Rabin and Weizsäcker (2009) use these results to formulate an argument for narrow bracketing that is based on the Dutch-book theorem (Bruno de Finetti, 1974). According to their bracketing theorem, a non risk-neutral narrow bracketer is tricked into choosing a dominated choice combination in a sequence of related gambles by ignoring potential cross-effects. The essential part is that the attitude of the DM towards risk is not identical over the possible range of outcomes. This is what can be observed in the lottery example, where the risk attitude shifts between lottery 1 and lottery 2. As a result, Rabin and Weizsäcker (2009, p. 1510) conclude that there "exists a pair of independent binary lottery problems where [the DM] chooses a dominated combination".

It becomes apparent that whether or not a DM applies broad bracketing and benefits from a global perspective depends on three elemental conditions (Kahneman and Lovallo, 1993; Herrnstein et al., 1993). Foremost, the DM needs to know the outcomes of the decision alternatives, either in a certain or probabilistic form. Secondly, an individual needs to be able to group decisions, which may appear unrelated to each other at first glance, into sets. When decision problems are described by many unique characteristics, this task of recognizing problems that should be grouped together grows more challenging as possible commonalities are harder to perceive. Finally, it is necessary to utilize an appropriate evaluation strategy to find the long-term maximizing alternative and avoid falling prey to outcome editing as it was discussed in the previous chapter. Thaler et al. (1997) find that narrow bracketing of decisions and segregated outcome editing are two phenomena that frequently occur together. For a narrowly bracketing individual, only the first requirement, namely the knowledge about the outcomes, applies.

For the reason of preventing the DM to make a welfare maximizing choice, narrow bracketing is commonly regarded as a decision-making error (Read et al., 1999, Rabin and Weizsäcker, 2009; Koch and Nafziger, 2016a). Webb and Shu (2017) show that narrow bracketing causes risk averse behaviour in gambles with a positive EV and risk seeking behaviour on negative EV and pure loss gambles. Although a switch in risk attitude can not be said to be irrational, the combination of risk aversion in the gain domain and risk seeking in the loss domain ultimately leads to the selection of suboptimal alternatives for these gambles. Furthermore, they confirm that broadly bracketing the decisions leads to the value-maximizing choices by reversing the risk attitudes and enabling the DM to choose the gamble in a positive EV lottery and the smaller certain loss in negative EV or pure-loss gambles.

Kahneman and Lovallo (1993) emphasize that broad bracketing is a fundamental component for rational decision-making. Read et al. (1999) also state the main purpose of broad bracketing is welfare maximization, but they note that there are some situations where it is favourable to bracket narrowly. Their argument is that broad bracketing usually entails small compromises or inconveniences in order to achieve long-term welfare. On the contrary, narrow bracketing allows the DM to acknowledge these small inconveniences, which may have a notable effect on welfare in the short-run but can be ignored under a long-term perspective. Which situation requires broad or narrow bracketing should then be determined by comparing the cumulated utility of all these small annoyances with the long-term gain of broad bracketing (Read et al., 1999). Note that, when done by the DM, this overall comparison in order to consciously decide on the more desirable bracketing approach for a specific situation would require broad bracketing itself. This would further imply that only subjects with the ability to bracket broadly are able to distinguish between situations where a change in bracketing behaviour is favourable.

The disagreement in literature whether broad or narrow bracketing is the superior approach to decision-making leads to the classification of narrow bracketing into two phenomena. The first form refers to narrow bracketing as a choice error, as it is observed in the isolated evaluation of lottery choices. On the other hand, the second form perceives narrow bracketing as so-called motivational bracketing in the context of self-control problems. Here, narrow bracketing aims at guiding behaviour to achieve a set goal by providing increased motivation or pressure for overcoming the temptation to deviate from goal-oriented choices. The argument for this is that a step-by-step approach of narrow bracketing may make a large task appear smaller and thus, easier to attain. Telling a dieting person not to eat a piece of cake today seems easier to comply to than the overall goal of not eating cake for the next year. Koch and Nafziger (2019) identify these two forms as distinct phenomena and hence, support the proposition of Read et al. (1999) that some situations may require narrow bracketing.

Despite the initial disagreement about the advantages of either strategy, literature shows a strong consensus regarding the propensity of individuals to employ narrow bracketing (Tversky and Kahneman, 1981; Read et al., 1999; Venkatraman et al., 2006, Kahneman and Lovallo, 1993). Rabin and Weizsäcker (2009) mention that the natural setting of life complicates taking a broad perspective because the majority of decisions in everyday life are presented rather independently, so that they are easily perceived in a segregated manner. Bland (2019) shares the opinion that in some situations it might not be apparent to the DM that grouping choices would be beneficial and that the additional effort of broad bracketing might outweigh the generated benefits. Andreoni et al. (2018) also find that only a small percentage of their respondents benefit from arbitrage by broadly bracketing the decision about their money allocation. The results of Ellis and Freeman (2020) allow for the classification of 75% of their respondents as narrow bracketers while broad bracketing describes 14% of subjects and 4% are assigned to their category of partial-narrow bracketers. The observation of changing fractions of broad bracketers in different decision-making contexts motivates the statement that a classification of subjects into broad and narrow bracketers is not absolute but may change between situations (Ellis and Freeman, 2020). The authors also suggest that, instead of being a mere error, narrow bracketing is a heuristic that is used when a DM is oblivious to the procedure of broad bracketing, its positive effect on welfare, or seeks to facilitate his decision process. Having said that, Read et al. (1999) state that subjects' tendency to bracket narrowly persists even when broad bracketing would be possible for them.

The importance of narrow bracketing in decision-making is not restricted to economic experiments such as the lottery example but proves to be relevant in a variety of contexts. Bland (2019) and Stracke et al. (2015) explore choice bracketing in the context of strategic interactions in the form of games instead of gambles. Felső and Soetevent (2014) find bracketing behaviour in consumers who apply a narrow bracket to cash wealth and gift certificates while the motivated bracketing phenomenon mainly concerns consumption situations that require self-control (Heyman, 1996; Herrnstein and Prelec, 1992b). Hence, the theory of choice bracketing motivates research in various fields and represents an important element in decision-making as it is done by people and not as it is prescribed by models of rational choice. The findings on bracketing behaviour in general stress the importance to investigate the effects of broad or narrow bracketing and the determinants that cause an individual to adopt either strategy.

2.4 Bracketing Effects

Following up on the previous chapter, where it was already established that bracketing behaviour has significant effects on the quality of decision-making, this part of the thesis seeks to elaborate on the different effects observed by research in more detail.

As mentioned before, narrow bracketing of choices often leads to unintended negative consequences on the welfare of a DM. A bracketing effect then exists when a change from narrow to broad bracketing would lead to a different outcome (Read et al., 1999).

In other words, whenever there is a difference in outcomes or behaviour under broad or narrow bracketing, a bracketing effect is observed; and it should be investigated what mechanism causes this alteration.

In general, it is argued that bracketing effects work by changing the weight that a DM places on potential losses. Broad bracketing reduces the weight placed on possible losses for gambles that have a positive EV, while it enhances this weight for gambles with a negative EV. Furthermore, effects of broad bracketing work by enabling the DM to perceive aggregated outcomes and global consequences. Substitutability and complementarity between choices are the main source for the benefits connected to broad bracketing, being that such interactions are not perceivable for a subject who makes choices in isolation (Webb and Shu, 2017; Read and Loewenstein, 1995).

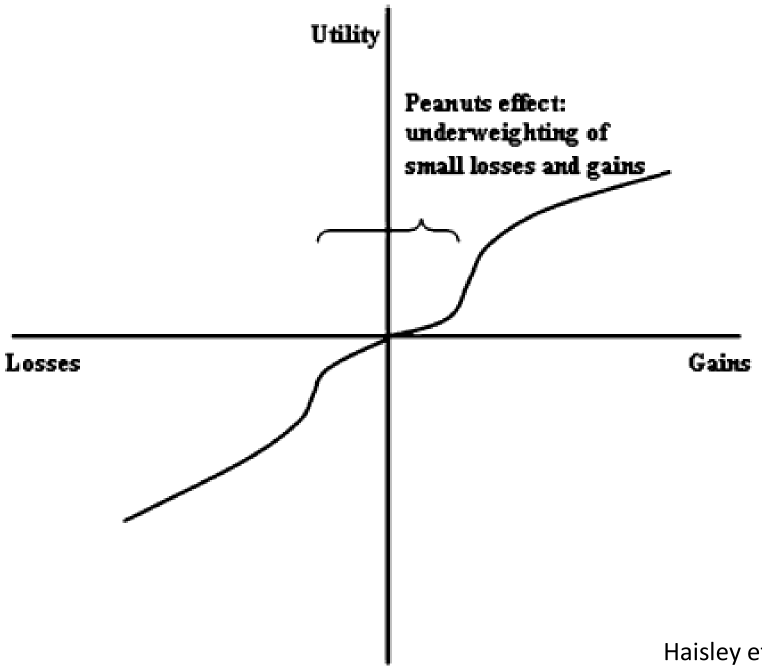
In the following subchapters, five mechanisms for bracketing effects are discussed. The first two, adding-up effects and emergent properties, are considered to be the most prominent effects since they capture the fundamental difference between the two bracketing strategies (Read et al., 1999). Both effects rely on emphasizing global consequences and patterns. The third effect, taste changes, evolves over time and may be seen as a combination of emergent properties and adding-up effects. In decisions that concern resource allocations subjected to capacity constraints of some kind, broad bracketing may produce the effect of trade-offs between alternatives. Lastly, self-control is specific to the application of motivational bracketing. Since the last three effects are of minor importance to this thesis, they shall be mentioned but are not discussed in great detail.

2.4.1 Adding-up-effects

This type of bracketing effect arises whenever a DM perceives the cost of a choice to accumulate at a different rate compared to the benefit of the alternative (Read et al., 1999). Such a situation occurs when the costs or benefits of an option are insignificantly small so that they do not reach an individual's threshold of consideration. But when broad bracketing is applied, the cumulative costs of several repeated choices exceed the threshold for consideration and would cause the DM to reconsider his choices.

The basis of argumentation for the adding-up or aggregation effect is the value function as proposed by Markowitz (1952) and depicted in figure 1. Instead of using absolute levels of wealth, the value function concerns gains and losses. As already noted, gains

Figure 1: Markowitz Utility Function



Haisley et al. (2008, p. 60)

and losses are important measures because choice bracketing operates by changing the weight a DM places on either of them. In addition to being defined over gains and losses, the value function shows three clear kinks; one infliction point in the gain and loss domain and one at the status quo which represents the reference point of a subject. The initial flat segments between the kinks in the gain and loss domain visualise that subjects tend to underweight small amounts of money. Prelec and Loewenstein (1991) refer to this behaviour as the peanuts effect. Aggregating outcomes by the means of broad bracketing then enables individuals to overcome this peanuts effect.

Aggregation is not only an effective tool to mitigate the peanuts effect but is also capable of overcoming loss aversion. For gambles with a positive EV, aggregation enables subjects to recognize that overall losses are rather unlikely and thus encourages taking a riskier gamble as opposed to a safe option of lower EV. The other way around, aggregating negative EV gambles allows subjects to notice the more probable distribution of losses so that gambling is avoided. This important effect of changing the salience of losses was already demonstrated in the lottery example. An example of not taking advantage of the adding-up effect is the popularity of participating in lottery play (Haisley et al., 2008). People severely underweight the aggregated cost of repeatedly buying lottery tickets compared to the benefit of winning the lottery. Bertrand and Morse (2011) examine the adding-up effect in the context of payday loans and find an 11% decline in borrowing from a payday loan agency when respondents consider the aggregated cost instead of focusing on one loan at a time.

In the context of repeated gambling, Webb and Shu (2017) investigate whether the change in risk preference is the result of outcome aggregation or trial aggregation. They refer to outcome aggregation as the process of adding-up the potential outcomes and thereby highlighting the probability distribution of experiencing an overall loss. Instead of explicitly addressing the cumulative outcome distribution, trial aggregation emphasizes the repeated-play nature of the gamble in order to make the multiplicative effect of repeated trials more apparent. Webb and Shu (2017) conclude that it is mainly outcome aggregation that contributes to improved decision-making.

Another example for the peanuts effect and aggregation as an intervention is procrastination. According to Sabini and Silver (1982), procrastination exists due to people underestimating the amount of work that could be done in a seemingly insignificant time period, for example 5 minutes of writing a paper. If an individual acts on such irrationally short time intervals, he will fail to notice the cumulated amount of work that could have been accomplished over several 5-minute time intervals. Adopting a long-term perspective by broadly bracketing the intervals enables the DM to work more efficiently.

2.4.2 Emergent Properties

The bracketing effect of emergent properties refers to characteristics of a decision problem that are not part of the individual alternatives but become apparent when multiple choices are combined through broad bracketing. Being that they operate in the same manner, emergent properties are closely related to the adding-up effect but have different implications. A comparison that facilitates the understanding of emergent properties is that to representativeness where "a representative sample […] has properties that reflect those of its population, yet no single element in the sample can be said to be representative" (Read et al., 1999, p. 177). There are three different forms of emergent properties that will be examined here.

2.4.2.1 Variety Seeking

People generally tend to prefer heterogenous sets of items over homogenous ones and the global perspective of broad bracketing draws their attention to the diversity found in their choices (Read et al., 1999). This can be compared to an investor who prefers to hold a diversified portfolio in order to reduce risks instead of investing all of his resources into a single stock. Simonson (1990) states that variety seeking represents a choice heuristic that is applied whenever an individual is confronted with several simultaneous choices. Read and Loewenstein (1995) agree and suggest that diversification might be the most available heuristic to a DM when asked to make multiple choices. On the other hand, the most straightforward heuristic in a sequential choice setting is to simply choose the alternative that the DM favours the most within this single choice.

Facing several choices simultaneously represents a more challenging decision environment in which three major motives for such a diversification strategy can be identified (Simonson, 1990, Read and Loewenstein, 1995, Mittelman et al., 2014). First of all, individuals face an uncertainty regarding their own preferences for future consumption points. Secondly, the response to possible changes in preferences is to counteract their own uncertainty by selecting a greater variety. The third reason to seek a greater variety can be found in the increased complexity of a simultaneous decision task. In this sense, a DM diversifies in order to reduce the time and effort that he would have to invest into analysing which alternative he favours the most.

Read and Loewenstein (1995) notice that subjects who face multiple choices end up with a more diverse decision outcome than they prefer in retrospect and would wish to change their selection. They label this excess of variety as the diversification bias. Read et al. (1999) also discover that individuals who chose greater variety rate their choices as less satisfying. In this sense, a broad bracketer who is able to perceive a greater selection of alternatives, may be tricked into excessively diversifying his choices by losing focus on the value he experiences from the resulting portfolio. It is then proposed that the favourability of variety seeking under broad bracketing is best described by an inverted-U function (Read et al, 2001). Under such a function, an increase in variety causes the DM to experience increasing satisfaction up to a certain point after which further diversification is defined by a decreasing taste for variety. In consequence, the variety seeking effect of broad bracketing would improve decision-making unless the DM succumbs to the diversification bias and overdiversifies beyond the optimal level of variety.

However, the authors conclude that broad bracketing and its variety seeking property is favourable when there are significant interactions between the various decision alternatives, but that it will lead to a less pleasing outcome when trivial preference interactions are overestimated or imagined. Read et al. (2001) also suggest that the requirement of such genuine interactions together with an increased likelihood to select worse options under narrow bracketing is fulfilled by most investment contexts. This may also indicate that organisational decision-making would find broad bracketing and its variety seeking attribute to be, at least to some extent, beneficial. The global perspective of broad bracketing then allows a DM to maximize his overall experience, which may include some options that would not be the best in an isolated examination (Read et al., 2001).

Several studies confirm that broad bracketing of choices results in a greater variety and that some subjects select more variety than they prefer from a later point of view. Simonson (1990) let his respondents choose three snacks for sequential consumption occasions. One group is forced to bracket broadly by having to choose all three snacks at the same time whereas the narrow bracketing group chose one snack in each of the following three meetings. The results are that 64% of respondents chose three different snacks under broad bracketing. In comparison, merely 9% of the narrow bracketing group chose a different snack for each meeting. When asked to rank the available snacks according to their general preference for them, it is found that broad bracketers selected more of their less favourite snack alternatives as a result of choosing a greater variety.

Simonson and Winer (1992) obtain similar results when their respondents were asked to select several flavours of yoghurt at once for future consumption compared to choosing one flavour at several sequential events. While it is found that subjects select more rare flavours that they usually do not buy, respondents were not asked to rate their liking of the resulting portfolio in retrospect. However, one could suspect that at the point of actual consumption, a similar effect as in the selection of the less liked snack alternatives would be observed. Read and Loewenstein (1995) are also able to replicate the results by dividing trick-or-treaters into a broad and a narrow bracketing group. The broad bracketing children are then asked to select two sweets at a single house while the narrow bracketing children are instructed to select one candy at two consecutive houses. In the simultaneous choice treatment, all participants select two different candies but only 48% do so in the narrow bracketing group. In another experiment, Read and Loewenstein (1995) apply a similar setting to Simonson (1990) where subjects have to choose three snacks either simultaneously or sequentially. They replicate Simonson's (1990) major finding that variety seeking is greater when bracketing the choices broadly. In addition, they find that 84% of respondents in the simultaneous choice setting would prefer to change their selection in favour of less variety. Hence, one may conclude that subjects who broadly bracket consumption choices are prone to the diversification bias.

The variety seeking phenomenon is not constrained to consumer choice settings like those in the above examples but is also applicable to other areas. The behaviour is also found in the selection of retirement funds where a DM is confronted with various options. Benartzi and Thaler (1998) state that subjects in this decision problem employ a 1/n heuristic; meaning that a DM who is offered the choice between n retirement funds, prefers to distribute his capital in equal parts between all options. In such a financial investment context, it can be argued that the variety seeking behaviour is an appropriate approach to diversify risk. Similarly, for the consumer choice context with the snack or yoghurt options, a DM may try to diversify the risk of not knowing which flavour will satisfy his taste in future consumption points. However, while the DM may not be satisfied with a large variety of product flavours, it may be argued that a diversified investment portfolio is not prone to such taste changes. The argument is that all funds yield a monetary outcome, so that each one may be able to satisfy the DM’s taste for a financial gain.

Furthermore, the application of the mechanism to an organizational setting, especially to the context of personnel selection, is of particular interest to this thesis and will be given additional attention. Personnel selection differs from the classical consumer choice discussed above because the DM cannot repeatedly hire the same candidate and people are not consumed as it would be the case with goods. An increasing number of organisations commit themselves to improving the diversity of their workforce. Harrison and Klein (2007) state that a heterogenous workforce provides the organisation with increased creativity, ability for innovation, as well as an integrative problem-solving ability but may also entail an enlarged potential for conflict. As a result, organisations whose teams have access to a diversified pool of information benefit from an enhanced cognitive ability and more efficient decision-making compared to an organisation that makes use of homogenous informational resources (Harrison and Klein, 2007). A diverse workforce represents an emergent property of all hiring decisions within an organisation because diversity is a group-level characteristic; meaning that diversity describes the organisation as a whole and can not be used to describe a single employee (Chang et al., 2020; Harrison and Klein, 2007). Literature offers three main assumptions regarding the construct of variety within an organisation: (1) members of the unit are different with respect to a specific qualitative attribute; (2) the distribution of members over all qualitative characteristics differs across units; (3) diversity in the qualitative attributes of unit members are linked to positive effects on the quality of decision-making and performance, being that diversity represents information (Harrison and Klein, 2007).

Despite these good intentions, it is observed that many organisations continue to be defined by a homogenous workforce. A possible explanation for this is that the top-level executive managers who determine the diversity objectives have a global perspective of the organisation and are therefore able to broadly bracket the diversity found in individual departments. However, these executive managers are usually not the ones who have to implement the determined objective by making hiring choices. This task is faced by employees in lower levels of the hierarchy who lack the natural global perspective of an executive manager. In addition, hiring decisions often occur in a sequential setting, which further impedes the ability of the lower-level manager to broadly bracket the consequences of each hiring decision on the diversity of the organisation (Chang et al., 2020).

In this decision-making context, Chang et al. (2020) explore whether the variety seeking mechanism in personnel selection applies to the gender of candidates. Respondents are separated into a sequential choice treatment where they are asked to choose a single candidate, and a simultaneous choice treatment which comprises multiple hiring decisions. According to what was discussed in the previous paragraph, subjects in the narrow bracketing treatment should fail to evaluate how their current hiring decision affects the diversity of the workforce, being that this is a group-level attribute (Chang et al., 2020). The study finds that the narrow bracketing respondents select women in 7.4% of all choices while the broad bracketing treatment increases the share of women to 18%. Hence, greater variety is chosen when subjects are induced to bracket their decisions broadly. In addition, the authors discover that the effect of isolated choices can be overcome by highlighting the importance of diversity to respondents before asking them to make a decision. This finding is contrary to the standard lottery of Tversky and Kahneman (1981) where the explicit hint to read both lotteries before deciding on the alternatives did not enhance decision behaviour. However, the explanation for these opposing results is likely to be found in the difference of the tasks. It can be assumed that including the information that an organisation values the diversity of its workforce in the experimental instructions is perceived as an explicit order to diversify choices. On the other hand, the instruction to read both lotteries before making a decision does not directly hint at combining the two decisions into a set and, hence, does not represent a pronounced intervention into bracketing behaviour.

The results of Chang et al. (2020) offer a foundation for two main interventions to encourage a broadly bracketed and thereby diverse personnel selection in organisations without making use of explicit directives: (1) multiple hiring decisions could be timed to be made collectively instead of being spread over a time frame of several months; and (2) a single person could be assigned to the supervision of a larger number of hiring decisions to induce a broader frame.

2.4.2.2 Risk Reduction

The study of risk reduction as an effect of broad bracketing is driven by an observation from Samuelson (1963), who offered a lunch colleague a gamble with an equal chance to gain 200 or lose 100. This gamble was rejected but the colleague suggested that he would willingly play 100 repetitions of this gamble. Thaler (1999) replicate this preference pattern by confronting lower-level managers with the decision to undertake a project that entails even odds to generate a loss of 1 million or a gain of 2 million for the division. Out of 25 non-CEO executives, only three responded that they would accept the project. In contrast, the project is eagerly accepted when it is offered to the CEO of the company who is able to bracket the projects of all divisions in a broad manner and perceives a reduced risk compared to the division managers. Likewise, the equity premium puzzle introduced in chapter 2.2 shows that subjects prefer to invest in lower-return bonds compared to the higher-return stock option as a result of evaluating their portfolios too frequently. Such evaluations of narrow time periods emphasise the higher risk of stocks, therefore hiding the benefits that the higher return rate would generate and driving investors towards the safer bonds (Benartzi and Thaler, 1995, Gneezy and Potters, 1997; Thaler et al., 1997).

These examples demonstrate that individuals prefer to avoid narrowly bracketed risky options and by doing so, fail to recognize the risk reduction that can be achieved from the combination of several risky choices. Benartzi and Thaler (1999) find that respondents in a broad bracketing condition are significantly more willing to select the risky options. Similar findings are reported by Gneezy and Potters (1997) as well as Thaler et al. (1997), who show that investors who broadly bracket the time periods between the evaluations of their portfolios displayed reduced risk aversion.

Consequently, it can be stated that combining several risky choices by means of broad bracketing reduces risk averse decision-making and encourages individuals to adapt their choices in the direction of EV maximization. This behaviour stems from the circumstance that the combination of multiple gambles considerably reduces the perceived riskiness of entering the loss domain for the DM. Under broad bracketing, a DM merges the risk of the currently faced decision with his pre-existing and potential future risks before he assesses the attractiveness of the choice. Even a risk averse DM then perceives the diversification of risks over several risky choices as convenient (Barberis et al, 2006; Koch and Nafziger, 2016b). Hence subjects are willing to accept the super-gamble although they would reject its individual components when evaluating them in isolation. Read et al. (1999) argue that this bracketing effect is particularly pronounced when the risks of the individual choices are either uncorrelated or negatively correlated. Webb and Shu (2017) expand the scope of the risk reduction effect of broad bracketing to negative EV and pure-loss gambles. In such a setting, the effect works in an inverse manner compared to positive EV gambles. While preferences are reversed from risk averse to risk seeking in positive EV contexts, they are shifted from risk seeking to risk aversion for negative EV and pure-loss gambles. This is because broadly bracketing such gambles highlights the unlikeliness of gains and increases the weight placed on cumulative losses. Furthermore, as it is suggested by the examples stated above, the same mechanism also applies to the frequency with which a DM evaluates his investments (Thaler, 1999; Benartzi and Thaler, 1999).

In conclusion, it can be claimed that the underlying mechanism of this effect is the adjustment of the situational weight that an individual places on losses. This change in the weighting of losses is proven to be capable of correcting the inefficient behaviour produced by the presence of loss aversion combined with narrow bracketing; the so-called myopic loss aversion (Thaler, 1999; Webb and Shu, 2017; Barberis et al., 2006; Koch and Nafziger, 2019). Due to the fact that missed benefits weight less than averted losses, the utility achieved by the application of broad bracketing surpasses the one under narrow bracketing (Koch and Nafziger, 2019).

2.4.2.3 Scheduling Pleasure and Pain

The scheduling of experiences represents an interesting application of choice bracketing but being that this area is of little relevance to this thesis, it will not be elaborated on in detail.

The fundamental argument is that the value of a specific experience interacts with events that have preceded it as well as those that are yet to take place. In the context of a series of events, an improving sequence represents a gain to the DM whereas a declining sequence is perceived as a loss (Read et al., 1999). Similar to variety being a group-level attribute, the improvement or deterioration of a series of experiences only becomes apparent when the individual events are bracketed together. Loewenstein and Prelec (1993) find that, as long as subjects are not aware of the circumstance that several events represent a sequence of pleasurable or unpleasurable experiences, they prefer to schedule the most preferred events first. A desire for improvement in experiences is then only introduced when subjects perceive multiple events as a sequence. In this context, Read et al. (1999) find that subjects in a narrow bracketing setting, in terms of separately scheduling two categories of events expressed as a sequences, exhibit a strong desire for an improving sequence in each category and therefore have to face an initial time period of purely unpleasant tasks. Contrarily, respondents in a broad bracketing treatment spread the pleasurable and painful experiences of both categories more evenly. Such a distribution of gains and losses can only be achieved if the DM first is aware that events are a sequence and additionally brackets the different types of tasks broadly by acknowledging existing interdependencies.

Combining the findings of Read et al. (1999) and Loewenstein and Prelec (1993) would result in three possible scenarios where a DM either: (1) does not perceive events as a sequence and schedules the best event first, which would fit the definition of narrow bracketing. (2) The DM recognises that events are a sequence but fails to bracket across different categories, leading to an initial period of unpleasant experiences in each category. This could be called a form of partial narrow bracketing as the DM achieves an improving sequence in each category but fails to optimise his overall experience. (3) The DM recognises that events are a sequence and brackets categories broadly so that positive and negative experiences can be scheduled more evenly, which represents broad bracketing in this context.

As already noted, the adding-up effect from subchapter 2.4.1 and emerging properties with its three mechanisms are the essential bracketing effects for the research framework that will be introduced in later parts of this thesis. However, with the aim to provide the reader with a more complete overview of the potential of choice bracketing, the next three subchapters review the bracketing effects that are of great value in other areas of research.

2.4.3 Taste Changes

The concept of a change in taste was already hinted at during the examination of the label melioration as a synonym for choice bracketing. Taste changes describe the effect that interdependencies between alternatives cause the utility of a future choice to change in response to the current selection. This means that the marginal rate of substitution between two options is not constant but changes according to the currently implemented consumption level of either alternative (Read et al., 1999, Herrnstein and Prelec, 1992a).

A narrow bracketing individual ignores these internalities between choices and considers the individual alternatives to be competitors. As a result, he will then choose the option that currently offers a higher return. In the context of repeated choice, this behaviour will provoke a consumption pattern that does not maximize overall utility. In contrast to this, a broad bracketer is able to account for the interdependencies by treating the options as competing aggregates (Herrnstein and Prelec, 1993). For that reason, broadly bracketing multiple consumption decisions allows the DM to accommodate for changes in his taste and implement utility maximizing choice patterns.

Since this bracketing effect depends on the existence of cross-effects between alternatives in a sequential consumption situation, it is specific to temporal bracketing (Read et al., 1999). It is argued that narrowly bracketing such situations is the prevalent behaviour found in everyday life so that individuals are inherently experiencing suboptimal choice patterns instead of benefiting from the taste change effect (Herrnstein, 1990; Herrnstein et al, 1993).

2.4.4 Trade-Off Across Choices

This effect of choice bracketing applies to situations where a DM is confronted with several choices that involve alternatives with multidimensional attributes. When such a decision problem is bracketed broadly, choices can be made so that the favourable components of one alternative compensate for the negative ones of another option (Read et al., 1999). Traditionally, this topic of integrative solutions occurs in negotiations, where multiple subjects are negotiated by two parties. Nevertheless, a DM can also achieve an integrative solution with himself by broadening his perspective to include many decisions and making trade-offs between them.

An illustrative example is the trade-off that cab drivers use between their labour supply and leisure time. Camerer et al. (1997) find that cab drivers quit their work early on days where there are many customers but work extended hours on less successful days. The reason for this is that the drivers act in accordance with an internally set daily income objective and hence fail to bracket several work days broadly. If they would do so, it would become apparent that the daily income threshold does not optimize the trade-off between income and leisure time. Camerer et al. (1997) further shows that the introduction of equal working hours per day with the total hours worked staying constant leads to an increase in income of 10%. In addition, an optimization strategy would require the drivers to trade off working longer on good days for generating 20% more income. Similar behaviour can be observed in the area of resource-allocation, where subjects prefer an allocation that equalizes the gains in utility. In order to achieve equal return, more resources need to be given to the recipient who values it less. Broadly bracketing an allocation decision facilitates the maximization of overall utility by equalizing the marginal rate of return and allocating more of the resources to the individuals who experience greater value from them (Read et al., 1999).

One of the most prominent applications of the trade-off effect is found in the theory of mental accounting. This concept refers to the method of mentally categorizing financial outcomes or budgets into several accounts, where each one is designated to a specific purpose. The motivation for this is the same as for accounting in organisations. Namely, the individual seeks to monitor the income generated by different sources as well as the cost categories where the budget is spent. Tracing the sources of income and outgoing expenses provides an individual with an instrument to control his spending.

According to Tversky and Kahneman (1981), a mental account determines a set of outcomes for joint evaluation along with a reference outcome, which is generally given by the status quo. In this sense, mental accounting specifies the evaluation frame for a choice, together with how the outcome of this choice is perceived with respect to the reference point. Abeler and Marklein (2017) propose that a consumption decision under mental accounting is best described by a two-step process. The first step is the posting phase where all incomes are matched to their respective expenditure categories, but no decision is made about how the allocated budgets will be spent within such an account. The spending phase then refers to the individual spending the budget posted to a specific account on the consumption items that are included by this account. When engaging in mental accounting, bracketing behaviour presents itself in three different forms. Foremost, a DM tends to bracket all expenditures within a specific category broadly but brackets narrowly across the different accounts (Thaler, 1985; Thaler, 1999). Secondly, it is important to note that bracketing behaviour affects how narrow or broad an account is defined in the first place; for example, a DM could broadly bracket all the expenses for leisure activities into one account, or he could bracket them narrowly by setting up separate accounts for entertainment, shopping trips, social gatherings, or favourite hobbies. Finally, following the equity premium puzzle, choice bracketing is also reflected in the frequency with which the DM evaluates his accounts.

How the accounts are bracketed in terms of scope of the categories and evaluation frequency influences the attractiveness of decision alternatives. The underlying mechanism for this effect is that narrow bracketing of mental accounts violates the fundamental economic assumption that money is fungible (Thaler, 1999; Felső and Soetevent, 2014). The fungibility axiom states that units of money should be treated as perfect substitutes for each other. In compliance with this, the budget that is assigned to one account should act as a substitute for the money allocated to another account. The finding that fungibility of money is violated by narrow bracketing causes the budgets assigned to an account to act as a strict constraint on expenditures. When a DM adopts broad bracketing across accounts, he takes his overall budget into consideration and is able to make beneficial trade-offs by transferring money between accounts that would otherwise be strictly tied to a specific account (Abeler and Marklein, 2017; Fels, 2020). The result of such a failure to consider possible trade-offs between accounts is that a contemplated expense in a specific account is labelled as unaffordable although its cost would be moderate in comparison to the overall budget. Heath and Soll (1996) confirm that subjects tend to behave in strict compliance with their implicitly set budget constraints. Thaler (1999) emphasizes that mental accounts are developed as a means to control expenses and reduce the time and cognitive resources involved in the evaluation process and hence, it should be expected that the mechanism is not optimally executed.

Felső and Soetevent (2014) investigate the effect of trade-offs in mental accounting with respect to general income sources and gift certificates. According to the theory, a narrow bracketer strictly distinguishes between the possible uses of gift income and the expenditures that are supposed to be covered by regular sources of income. A broad bracketer, on the other hand, includes all sources of income when evaluating a choice; meaning that such an individual should spend the gift certificate in the same manner as his general income. In contrast to the findings of Heath and Soll (1996), the study reports that the majority of respondents (83%) bracket broadly and treat gift income and other sources of income as substitutes by buying a good that they would have also bought without the gift certificate. Nevertheless, 17% of respondents bracket the gift account and general account narrowly by indicating that they spent the gift certificate on an item that they would not have bought with their regular income (Felső and Soetevent, 2014). A possible explanation for the discrepancy regarding the share of broad bracketers in the two studies might be that a gift income is not strictly assigned to a specific usage but can be posted to an account in a more flexible manner. Moreover, the study of Felső and Soetevent (2014) does not address how strictly subjects comply to the budgets posted to accounts from the other sources of income, being that they focus on whether gift and all other sources of income are used for the same purposes. In contrast, Heath and Soll (1996) concentrate on whether subjects transfer money between narrowly defined expenditure accounts, such as clothes, food, and entertainment. Therefore, the broad bracketing behaviour found by Felső and Soetevent (2014) does not necessarily imply that their respondents would also transfer money between accounts in order to manage a depleted expenditure account in the absence of a gift income. In any case, the study does suggest that the axiom of fungibility holds for gift incomes compared to other sources of income.

2.4.5 Self-Control

The self-control effect of choice bracketing arises whenever a DM actively adopts a specific bracket in order to control his behaviour more efficiently. The adoption of a particular bracketing behaviour is equivalent to the introduction of personal decision-making rules which induce a perceived pain when violated and by this means control behaviour (Read et al., 1999; Abeler and Marklein, 2017). Subsequent to the preceding subchapter, mental accounting is a classic example for improved self-control as a result of choice bracketing. As was just described, narrow bracketing across different accounts as well as a narrow definition or frequent evaluation of accounts intensifies the budget constraints that a DM places on himself. Aside from mental accounting, bracketing choices narrowly as a tool to achieve self-control is a well recognized concept in literature

(Read et al., 1999; Koch and Nafziger, 2016b; Abeler and Marklein, 2017; Koch and Nafziger, 2019).

The underlying justification is that narrow bracketing improves motivation because an overall goal that is reduced to smaller instances seems easier to accomplish. On the other hand, when considering a sequence of demanding choices, broad bracketing reduces the motivation to undertake the choice series because it highlights the cumulated difficulty faced by the DM. This argumentation has interesting implication for a principal-agent setting: when the principal has to confront an agent with a sequence of difficult choices, he would be able to convince the agent to undertake the series with enhanced motivation by treating it in a disconnected manner instead of presenting the tasks simultaneously (Read et al., 1999). Koch and Nafziger (2019) report evidence that mental accounting and setting narrow goals are self-control strategies which are related to each other and differ from the common assumption of narrow bracketing as a pure decision-making error. They additionally state that self-control is not the sole purpose of the two phenomena, but that they also serve to reduce the complexity of a decision problem. However, while this strategy is beneficial when solely considering the enhanced self-control, it might still lead to suboptimal choices with respect to solution quality. An example for this can be easily found in mental accounting where the strict narrow budgets prevent a DM from needless overspending but also impede transfers across accounts in order to meet an essential or beneficial additional expenditure.

Koch and Nafziger (2016b) investigate the superiority of narrow bracketed goals as compared to employing broad bracketing. Participants of the study are either assigned the bracket of a daily work goal in the narrow treatment or a weekly bracket for the broad bracketing group. Note that only the brackets are assigned so that the respective goals are set by the individuals themselves. The results show that the narrow bracketers set higher goals and supply greater effort than the broad bracketing group. The difference is explained by the fear of loss experienced by a narrow bracketer when he might not reach his daily goal. Hence, it is loss aversion that drives an increase in effort supply. In comparison, this fear of loss is less salient for the broad bracketing group since they may compensate for a below expected outcome day by increasing their work on the following days. This knowledge of being able to compensate over the course of the week reduces their motivation (Koch and Nafziger, 2016b).

However, other studies demonstrate the potential of broad bracketing as a self-control mechanism (Rachlin, 1995; Heyman, 1996). The most prominent application for this effect is the study of addiction, where it is argued that broadly bracketing many consumption occasions can be compared to a decision of either being an addict or not. On the contrary, considering each consumption choice in isolation induces more drug abusive behaviour since the attention shifts from the global consequences of many consumptions to the comparatively trivial consequences of a single consumption. Another salient example is dieting. When the diet is bracketed broadly, falling prey to temptation in the current choice equals the expectation to continue to do so and as a consequence, failing the whole diet. Under narrow bracketing, however, the consequence of not controlling oneself at one consumption decision is a negligible amount of extra calories with no effect on expected behaviour in future choices. This mechanism can be generalised in the following way: usually, subjects exhibit a strong preference for the present compared to future alternatives (Andersen et al., 2008). This implies that in a decision problem with existing temporary preferences, an individual can be expected to be oblivious to how he will decide in future choices but might know from his past behaviour that self-control is an issue for him. The next choice faced by the DM can then have substantial implications for his expectations. If the subject acts impulsively and succumbs to his temporal preference, he will expect his failure at achieving self-control for all future choices. In reverse, if he achieves to control his behaviour in this choice, he will expect himself to continue to do so for future choices (Ainslie and Haslam, 1992). Koch and Nafziger (2016b) conclude that narrow bracketing as a self-control strategy is optimal for present-biased subjects who experience little uncertainty about their productivity or future choice while broad bracketing is advantageous for those who face a greater uncertainty.

The dual-self framework presented by Fudenberg and Levine (2006) aims at illustrating the internal conflict for self-control by the means of introducing multiple selves of an individual. According to the model, a DM neither always acts absolutely myopic nor continuously optimises his global return. Instead, each decision problem should be considered to be a contest between several impulsive and myopic selves subject to short-run temptations and one long-run self (Fudenberg and Levine, 2006). The short-run selves as well as the long-run self are described by the same outcome preferences but differ in how they account for future choices. It is then argued that the cost of maintaining self-control is the mediating factor between global maximization and temptation in the short-run.

2.5 Mechanisms Behind and Determinants of Bracketing Behaviour

The previous chapters established the important role of choice bracketing in decision-making, the advantageous effects that are connected to the concept, as well as the potential adverse consequences regarding the welfare of the DM when a bracket is adopted that does not fit the requirements of the decision problem. Consequently, it is crucial to gain an understanding of the characteristics of the decision environment and the DM that drive which bracket is employed by an individual and which one should be adopted.

Some studies note the lack of research on the determinants of choice bracketing (Read et al., 1999; Stracke et al., 2015). Nevertheless, some propositions and empirical work arise in literature but no all-encompassing model has been developed. Common suggestions for sources of bracketing include work on cognitive limitations, the shape of the value function, and the psychological accessibility of the decision consequences (Simon, 1957; Kahneman, 1973; Herrnstein et al., 1993). Herrnstein et. al. (1993) argue that all determinants of bracketing can be perceived to be either of cognitive or of motivational nature.

This chapter has the objective of facilitating the understanding of the influencing factors for how a DM sets his choice by presenting the various scattered propositions found in literature. For the following eight concepts that will be discussed here, two groups of theories can be formed. The first group contains the concepts of risk attitude (chapter 2.5.1) and prospect theory (2.5.2), which describe the underlying mechanisms that lead to different choices under narrow and broad bracketing. This group does not yet address why a specific bracket is adopted by the DM but only focuses on how adopting a specific bracket leads to differences in decision-making. The subsequent six subchapters compose the second group of theories, which discuss the characteristics of the task and subject that may determine which bracketing approach is applied by a specific DM. The most promising concepts in this group are differences in cognitive processing (chapter 2.5.4) as well as limitations in cognitive capacity (chapter 2.5.5). The last subchapter then addresses the motivational source of adopting a specific bracket. However, as this last determinant and its self-control effect is of little importance to this thesis, it will not receive excessive attention.

2.5.1 Attitude towards Risk and Loss

Decision-making under risk refers to choices about gambles that yield multiple outcomes with respective probabilities; meaning that a choice is described by a probability distribution of possible outcomes. In this regard, the term risk describes a situation where the probabilities of outcomes of the various alternatives are known instead of being uncertain. The attitude towards such risk represents the first concept that may be used to address the circumstance that different choice brackets lead to different choices. Literature states that risk preferences of subjects who face decisions in such a context are not homogenous but show individual heterogeneity (Harrison et al., 2007). With respect to this statement, Tversky and Kahneman (1992) note that roughly half of the respondents reject a small-scale gamble where the outcome of a gain and a loss are equally probable, despite the fact that the possible gain would be twice the potential loss. An axiom for risky choices is developed by Kahneman and Lovallo (1993) and states that a DM who would be willing to accept the offer of playing a series of gambles of this kind should also accept the play of a single gamble for the reason that he is likely to be confronted with further similar gambles in the future. However, the aggregation effect that arises from accepting multiple gambles that are not explicitly offered as a series can only be recognized by an individual who employs a broad bracket.

It is generally assumed that the majority of individuals are risk averse in their preferences. Such a DM has a preference for a sure outcome over a gamble that yields an equal EV and chooses a less risky gamble over one with high variance (Kahneman and Lovallo, 1993). An important characteristic of risk preferences is that subjects are proportionately risk averse; meaning that they display nearly as much risk aversion in small-scale gambles as in gambles with higher stakes. Kahneman and Lovallo (1993) identify this as irrational based on two arguments. First, a small-scale gamble is unlikely to pose a threat to the financial survival of the DM so that rejecting it is unreasonable. However, for gambles of larger scale the threat of being ruined represents a conclusive justification for a risk averse attitude. Secondly, since small-scale gambles are likely to occur on a regular basis, a DM should readily exploit the possibility for risk-reduction by means of aggregation.

Nevertheless, it is found that risk aversion on its own is not able to explain the observed behaviour of subjects declining a delayed gamble that would be favourable for them. Although subjects would be expected to combine the delayed gamble with pre-existing risks and achieve an attractive distribution, individuals of a specific risk-attitude make different choices when considering a series of gambles in an aggregated compared to a separated manner (Barberis et al., 2006; Stracke et al., 2015). This indicates that subjects tend to fail at merging several risks and instead evaluate gambles in isolation, so that narrow bracketing is of considerable importance in decision-making under risk.

Stracke et al. (2015) report a positive correlation between risk aversion and narrow bracketing, thereby showing that this risk attitude is related to bracketing behaviour. The emerging proposition is that the degree of risk aversion of a DM is a measure for the tendency to adopt a narrow bracket in decision-making behaviour. Kahneman and Lovallo (1993) state that the risk aversive decision-making of a subject is more pronounced when the DM can be held accountable for his choices. With respect to a managerial setting, where the agent is held accountable for his decisions by a superior, this would imply that the agent naturally displays increased risk aversion. In combination with the finding of Stracke et al. (2015), this intensified risk averse behaviour may encourage the agent to reject even small risks instead of merging them with pre-existing and upcoming decisions. Furthermore, it is observed that individuals who employ narrow bracketing in one situation behave the same in other situations (Stracke et al. 2015). This finding would imply that narrow bracketing is a behavioural characteristic that cannot be easily changed by the DM himself and as a result does not allow him to adapt his bracket in response to the requirements of the decision environment.

Having established that risk aversion alone is unable to solve the rejection of favourable gambles and finding that it is positively related to narrow bracketing, research turns to the concept of myopic loss aversion for an explanation. Loss aversion, in general, describes the circumstance that the loss of an amount x is experienced with a greater weight than gaining the same amount x. As long as the pain that is connected to a potential loss overshadows the potential gain, a DM will decide to avoid this loss. The term myopic refers to the adoption of a short-term evaluation perspective, so that myopic loss aversion is a combination of loss aversion and narrow bracketing (Benartzi and Thaler, 1995; Thaler, 1999). This concept is able to explain the rejection of attractive gambles as well as the selection of safe alternatives as a failure to integrate overall wealth and risk when evaluating a gamble. If a DM would take a long-term perspective and integrate the outcomes of such gambles into overall wealth, losing a single gamble is experienced as tolerable, since it represents a small reduction of overall wealth instead of a pure loss. Consequently, whether a DM evaluates a gamble as attractive or not also depends on the time horizon taken into account. The most prominent example for myopic loss aversion is the equity premium puzzle, which was introduced in chapter 2.2. Thaler et al. (1997) as well as Gneezy and Potters (1997) confirm the prediction that a long-term perspective considerably reduces the tendency of a DM to reject an attractive gamble in the context of investment choices.

A reverse myopic effect is found by Langer and Weber (2005) for negative EV gambles. For such a gamble, myopia induces a risk seeking behaviour instead of increased risk aversion, as it is the case for positive EV gambles. The myopic risk seeking effect then postulates that a long-term perspective reduces the unreasonable risk seeking of individuals. This reverse effect finds its application in the context of lottery tickets, where myopic buyers are more risk seeking since they do not accurately account for the cumulative loss represented by the repeated ticket costs (Haisley et al., 2008).

Venkatraman et al. (2006) declare that it is not the objective riskiness or ambiguity, as it was discussed up to this point, which influences decision-making behaviour. Alternatively, they propose that the perceived riskiness and ambiguity are the determinants of decision behaviour. The rationale behind this claim is that the psychological concepts directly refer to the threat of a loss that is associated with a risky choice, whereas in the normative concept high risk can occur even when the DM does not face a potential loss. Perceived riskiness thus regards an individual's perceived vulnerability to a loss that might be caused by the gamble and considers the magnitude, intensity, and emotional reaction linked to the loss. The difference to the already discussed concept of loss aversion is argued to be that loss aversion addresses the weights placed on losses compared to equivalent gains, while perceived riskiness does not consider gains but only focuses on the perceived exposure to loss (Venkatraman et al., 2016). The argumentation for ambiguity follows the same logic. While the normative concept describes the uncertainty about the probability of an outcome, the psychological construct refers to a situation where the perception of a lack of information causes the DM to perceive the outcome probabilities to be vague (Venkatraman et al., 2016). Kahneman (2003) states that when a decision-making model neglects emotions, it is expected to be unrealistic. Being that both psychological concepts address the pain a DM experiences when he faces a loss, they fulfil this requirement. Furthermore, it is found that both, perceived riskiness and perceived ambiguity, mediate an individual's tendency to accept a gamble (Venkatraman et al., 2016).

Information horizon, the frequency with which gambles are evaluated, as well as the frequency with which decisions occur are reported to be determinants of a subject's risk preferences (Harding and Looney, 2012). Since risk preferences, especially risk aversion, are found to be a major influencing factor for decision-making behaviour, one can conclude that a decision environment which is characterised by a narrow information horizon, high evaluation frequency, or high decision frequency is likely to induce narrow bracketing behaviour. In reality, risky choices are often independent of each other, display themselves in a more separated manner, and are faced by risk averse individuals. Subsequently, one can assume that the natural behaviour in decision-making under risk is narrow bracketing. However, one has to note that the forceful manipulation of behaviour towards broad bracketing has the potential to reverse risk preferences in the direction of value maximization.

2.5.2 Prospect Theory

In order to address the issue of decision-making under risk and capture the roles of risk aversion and loss aversion in decision-making behaviour, various models have been developed.

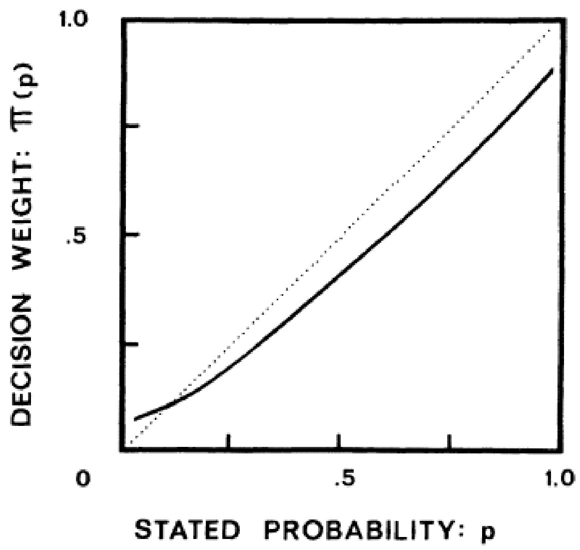
The most dominant model in decision-making under risk is the expected utility theory (EUT), where an individual's moral expectation about states of wealth form the utility of possible decision outcomes (Bernoulli, 1954). A DM thus chooses the prospect with the highest expected utility. It is argued that any individual would seek to behave according to this theory, being that the EUT is referred to as a model of rational choice (Tversky and Kahneman, 1981). However, the behaviour that is generally observed in decision-making under risk is incompatible with EUT. Kahneman and Tversky (1979) seek to include the incompatible observations of behaviour by modifying the EUT and introducing a new model termed prospect theory (PT). This concept is also part of the theories that describe how decision-making under a specific bracket yields different choices and thus, does not yet address why this bracket is adopted by the DM.

Whereas Bernoulli (1954) bases the EUT on states of wealth, Kahneman and Tversky (1979) incorporate the circumstance that subjects tend to consider a decision outcome in terms of a gain or loss compared to a reference point, which is often represented by the status quo. The assumption that changes in wealth instead of states of wealth describe subjective value to the DM is the fundamental feature of PT. The introduction of gain and loss relative to a reference point as determinants of preferences enables the model to describe behaviour as it is exhibited by individuals instead of predicting how a rational DM should behave (Kahneman, 2003).

The rationale for this adaptation is that many decisions require a choice between a deviation from the status quo and maintaining the status quo. Subsequently, the status quo serves as reference point and choosing the deviation can either be evaluated as a gain, when it is favourable, or as a loss in case of a disadvantageous outcome. As already introduced in the preceding subchapter, subjects are loss averse in the sense that a loss weights more than a gain of equal magnitude. Thus, loss aversion represents a natural bias to retain the status quo and reject any deviation as long as the potential loss overshadows possible gains. Research declares that losses outweigh gains by the factor of 2-2.5, so that the pain of losing an amount x is twice the utility from gaining the same amount x. (Tversky and Kahneman, 1981; Kahneman and Tversky, 1984; Kahneman et al., 1990). As a consequence of the greater weight associated with a loss, loss aversion is a contributor to avoiding risks in decision-making in order to maintain the status quo (Kahneman and Lovallo, 1993). Similar to risk aversion in standard theories, loss aversion plays a key role in PT.

An important difference between a model of loss aversion and PT shall be pointed out. A loss averse model does not pay attention to the distance between outcome and reference point but is mainly interested in whether the DM is above or below the reference point. On the contrary, PT attributes great significance to this distance (Martin, 2017). The effect of this key distinction can be illustrated using the example of the New York cab drivers, who had set themselves daily income targets (Camerer et al., 1997). These income targets represent the reference point for the drivers. Under a loss averse model, the marginal utility is increasing within the loss and gain domain, although at a steeper rate in the loss region. In consequence, the driver's main concern is whether or not he finishes his work day in the loss domain. This means, that not reaching the reference point and finishing the day in the loss region causes the driver to experience pain irrespective of the distance to the reference point. On the other hand, a prospect theoretic driver has a steeper increasing marginal utility the closer he is to the reference point and thus, the driver is more motivated to work as he gets closer to the reference income. Once the reference point is reached, he faces a diminishing marginal utility in the gain domain; meaning that each additional dollar earned causes the driver to be less motivated to continue his work (Martin, 2017).

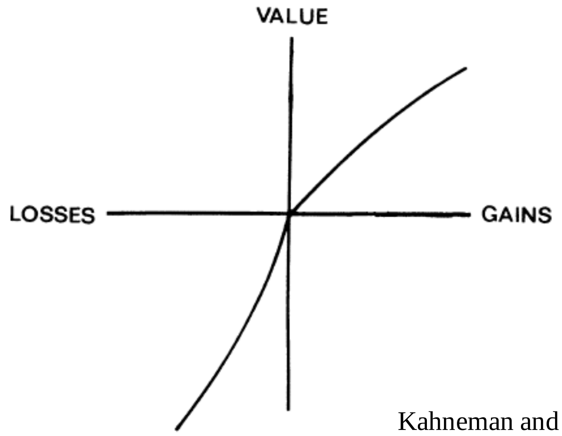
Figure 2: Probability-weighting Function in PT



Kahneman and Tversky (1979, p. 283)

Another crucial adaptation in PT is the introduction of decision weights. In EUT, a prospect is evaluated by summing up the utilities of each outcome alternative multiplied by their respective probability (Bernoulli, 1954). On the contrary, PT accommodates the observation that subjects do not objectively perceive probabilities by including decision weights as multiplicand for the utility of an outcome. Such decision weights are not probabilities since they do not exclusively measure the likelihood of an outcome but incorporate the impact of this outcome on the attractiveness of the prospect. Furthermore, decision weights do not fulfil the probability axiom that the sum of the probabilities of complementary outcomes of a prospect is equal to 1 (Kahneman and Tversky, 1979). The weighting function in figure 2 shows that individuals overweight low probabilities while underweighting higher probabilities (Tversky and Kahneman, 1981). Being that decision weights are a measure for the sensitivity to changes in probabilities, the function implies that a DM grows less sensitive to such probability changes when he moves away from the two certain scenarios of 0% or 100% likelihood. Furthermore, because moderate to high probabilities are underweighted by the DM, he perceives a gamble as less attractive and is more likely to behave in a risk averse manner. The opposite effect is true for risk seeking in disadvantageous gambles (Kahneman and Tversky, 1984).

Figure 3: Prospect-theoretic Value Function



Kahneman and Tversky (1979, p. 279)

The example of the taxi drivers stated above already hints at the shape of the value function that governs decision-making in PT. The value function that arises from the two major adaptations of PT compared to EUT is depicted in figure 3 and can be described by the following attributes (Kahneman and Tversky, 1979):

(1) The value function is of a S-shaped form and defined on gains and losses instead of states of wealth.
(2) It is convex in the loss domain, which represents risk seeking behaviour, while the concavity in the gain domain promotes risk aversion. This describes the diminishing sensitivity as the distance to the reference point increases.
(3) The function for the loss domain is about twice as steep as for gains and the steepest point is found at the reference point.

The attentive reader may notice a similarity in the shape of the prospect theoretic value function to the utility function of Markowitz (1952), which was presented during the discussion of the contribution of the peanuts effect to the adding-up effect (2.4.1) of choice bracketing. Although both functions are defined over gains and losses and incorporate a reference point as their basis for argumentation, the difference between the concepts should be noted. The utility function of Markowitz (1952) has an additional infliction point in the gain domain as well as in the domain of losses. These two kinks portray an effect of wealth on the utility that is associated with a gamble and mark the borders of the peanuts effect. It is proposed that while preferring a smaller sure gain over a gamble, with increasing potential wealth an individual will reach a point of steeper increasing marginal utility, where he is willing to take the gamble before the marginal utility of wealth is of diminishing nature. For the domain of losses, a small certain loss will be preferred at first but as the wealth at stake increases in magnitude, there will come a point where the individual prefers the gamble with its higher expected loss. On the contrary, PT theory does not account for this change in preferences and does not display the initially flatter segments of the function close to the reference point, which represent the peanuts effect. In fact, this is the steepest section for the prospect theoretic value function, representing the conception that the distance to the reference point is of utmost importance. However, Markowitz (1952) indicates that the utility function is a mere proposition while the value function of PT is a widely used concept in literature indicating that it appropriately describes the observed behaviour.

The established theory offers several implications for decision-making behaviour. PT considers the decision process to comprise two phases (Kahneman and Tversky, 1979). The editing phase has the purpose to organise the decision alternatives in such a manner that a simpler presentation facilitates their assessment by the DM. Due to the rephrasing during the editing procedure, invariance of the preference order does not need to hold true. Furthermore, it is also argued that the assessment of strictly positive or negative alternatives is executed by separating the prospect in a riskless and a risky component. The first one describes the certain gain or loss associated with the prospect while the latter component contains the probable additional gain or loss. This decomposition of prospects into a certain and a probable part is termed the isolation effect (Kahneman and Tversky, 1979). Since the riskless component is shared by the alternatives, it is commonly neglected.

The following evaluation phase then aims at determining the most preferred option. PT acknowledges that a dominated alternative might be chosen as a consequence of the isolation effect generating different decomposition forms in the editing phase as well as the decision strategy to select the highest-value option when prospects are considered separately in the evaluation phase. As noted above, the shape of the value function implies risk averse behaviour when evaluating a positive value gamble, whereas the loss domain is characterised by risk seeking choices. Additionally, the weighting function of probabilities contributes to preferring a certain gain over a prospect, considering that the probabilities in a moderate to high range are underweighted by the DM. The same underweighting of probabilities also contributes to choosing a negative gamble compared to a certain loss. This change from risk aversion in gains to risk seeking in losses is named the reflection effect by Kahneman and Tversky (1979).

Following these indications, one can conclude that PT accounts for the behaviour that is typically observed in the choice bracketing literature. Moreover, it strengthens the role of risk aversion and loss aversion as natural determinants for the adoption of a narrow choice bracket. Benartzi and Thaler (1995) agree that loss aversion constitutes the main driver for the observed outcomes.

2.5.3 The Framing of Decisions and Information Presentation

Having introduced the two theories that describe the existence of different decision outcomes under different choice brackets, the following subchapters turn their attentions towards finding an explanation for why a DM employs a specific bracket.

In chapter 2.1, a decision problem was defined to contain different choice alternatives, the outcomes that each alternative may yield, and the probabilities that express the likelihood of a specific outcome. The decision frame then describes how the DM perceives these alternatives, outcomes, and probabilities. Most decisions can be framed in several different ways and the adoption of a frame by the DM depends on the expression and presentation of the decision problem as well as individual characteristics (Tversky and Kahneman, 1981). Engin an Vetschera (2017) address the role of individual characteristics in the form of cognitive style. According to the theory, an individual's cognitive style is a mediator between the external presentation of a decision problem and the mental presentation formed internally by the DM. In this sense, cognitive style influences the perception of a decision problem and thus contributes to the existence of framing effects.

Consequently, different frames may cause an individual to exhibit different preferences for his choices although the objective attributes of the considered decision problem are the same. Such context-dependent reversals of preferences can be caused by changes in the formulation or presentation that would seem to be trivial (Stracke et al., 2015; Tversky and Kahneman, 1981; Kahneman, 2003). Whenever framing affects the decision-making behaviour with respect to a specific choice, the invariance criterion of rational choice is violated. Kahneman and Tversky (1984) demonstrate that the violation of the invariance criterion occurs in the decision-making of naïve individuals as well as sophisticated ones. Furthermore, the presence of preference reversals implies that subjects are commonly oblivious to the existence of several frames and their effects on the perception of the attributes of the decision problem. Moreover, subjects would prefer to comply to the axioms of rational choice but do not know how to address the deviation from rationality (Tversky and Kahneman, 1981).

The fundamental rationale for the existence of framing effects is that individuals have a tendency to passively accept the given formulation, instead of forming a presentation of the decision problem that incorporates all frames. Concerning choice bracketing, this cognitive inertia of individuals causes them to adopt narrow brackets when decisions are encountered in isolation, but they are more inclined to bracket broadly when decisions are presented collectively. The impression that individuals tend to accept any given frame as the basis for their evaluation is commonly shared in literature (Read et al., 1999; Kahneman, 2003; Hilgers and Wibral, 2014; Tversky, 1996). Kahneman (2003) argue that the source of the acceptance principle is not necessarily found in erroneous thinking but is likely to stem from individuals acting intuitively; meaning that an impulsive DM is more likely to experience a framing effect. Hilgers and Wibral (2014) support this by detecting that a change in the framing of a problem has a more intense effect on impulsive individuals.

It is also argued that the prospect theoretic value function with its reference point as central concept contributes to individuals experiencing a framing effect (Kahneman and Tversky, 1984). The reason for this proposition is, that the framing of a decision problem is capable to imply different reference points for the same decision problem. The following example demonstrates the power of decision framing (Tversky and Kahneman, 1981, p. 453). A new disease is expected to cause 600 deaths and there are two alternative programs that can be undertaken to fight the disease. There are two versions of the decision problem that are designed to induce different frames and are given to two separate groups of respondents:

Version 1:

> *"If Program A is adopted, 200 people will be saved.*
> *If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved."*

Version 2:

> *"If Program C is adopted 400 people will die.*
> *If Program D is adopted there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die."*

In the first framing alternative, 72% of respondents chose the sure outcome represented by program A. On the contrary, in version two of the problem, the majority (78%) selected program D (Tversky and Kahneman, 1981). However, the two versions describe the same problem and have equivalent outcomes, but the change in the formulation causes the respondents to reverse their preferences. Moreover, the example nicely demonstrates what was discussed in the previous subchapter; namely, that individuals are risk averse in gains (represented by version one) and risk seeking in losses (version two), even though both versions yield the same result and the perception of gains or losses is entirely due to a positive or negative wording. When translating the framing effect to a managerial context, it is easy to imagine that an organisation would face forgone profits because, for example, possible projects are rejected as a result of an unfavourable presentation.

As already mentioned, decision framing affects the choice bracket that a DM is likely to adopt. Redelmeier and Tversky (1992) discuss that the propensity to bracket decisions narrowly depends on the formulation and presentation of the decision task. Read et al., (1999) agree by stating that the difference between adopting a narrow or a broad bracket is associated with a change in the individual's perspective on the decision problem. This implies that a decision problem can be intentionally formulated in a manner that manipulates the DM to employ the preferred choice bracket. Rabin and Weizsäcker (2009) note that the prevalent frame in everyday life is likely to be separated. Moher and Koehler (2010) share this opinion of people being strongly biased to take a myopic decisions-making perspective. They support their argument by pointing out that the manipulation of bracketing behaviour by means of changing the framing of a choice would not be possible if subjects naturally made use of broad bracketing. The reason for this is that such individuals would account for the different frames instead of being influenced in their decision-making by a framing effect. When combining these statements with the acceptance principle of cognitive inertia, one can suggest that the natural setting of life promotes narrow bracketing and an intervention is needed to achieve more beneficial behaviour.

Hilgers and Wibral (2014) provide an interesting insight on this issue of intentionally exploiting the framing effect on choice bracketing. The study finds that a DM who first encounters a frame that is designed to induce narrow bracketing learns to shift to broad bracketing when he is exposed to an appropriate frame. However, a DM who is first confronted with a frame that facilitates broad bracketing is found to continue to do so even when he faces a narrow bracketing frame. This suggests that subjects may be able to learn to employ broad bracketing and avoid narrow bracketing when it would correspond to a decision-making error. It needs to be noted that this effect was observed for tasks of the same context so that no statement can be made for learning effects across different decision-making contexts. Regarding this suggestion, Fellner and Sutter (2009) find that when subjects were asked to choose between receiving feedback frequently or infrequently, the mere information that the low frequency condition yielded higher returns for participants of previous sessions of the experiment did not prime their decision behaviour for this alternative. Subsequently, individuals seem to need to learn from their own experience (Hilgers and Wibral, 2014).

On the contrary, Kahneman and Tversky (1984) notice that the framing effect does not disappear when subjects answer both framing versions of a problem within the same experimental session. Hence, although the issue of whether respondents learn to adopt the correct bracket is described by inconsistency, the findings are consistently supporting the ability to manipulate choice bracketing by the means of information presentation. In addition to the importance of these results to the research of choice bracketing and decision-making in general, the influence of framing on bracketing behaviour has practical implications for management. Deliberately manipulating the presentation of decision tasks constitutes a powerful instrument for management and decision support systems to guide the behaviour of their subjects into the direction of value maximization by facilitating the adoption of an appropriate bracket.

Three variables emerge as the key mechanisms for the manipulation of bracketing behaviour by changing the manner in which information is presented to the DM (Hardin and Looney, 2012).

(1) Information Horizon

The first variable concerns the information horizon for which the probabilities and possible outcomes of a choice are presented to the DM. In this regard, the presentation of a short-term horizon induces narrow bracketing. On the other hand, providing information for a longer time horizon, especially the provision of probability distributions and aggregated outcomes, facilitates the adoption of a broad bracket (Hardin and Looney, 2012; Webb and Shu, 2017). Webb and Shu (2017) state that the manipulation of the presentation form needs to address the failure of the DM to perceive a single choice as a part of a sequence of choices. Benartzi and Thaler (1995) successfully apply such a manipulation by providing one group of respondents with the distribution of investment returns for the time period of one year while the other group receives a 30-year distribution. The long-term manipulation causes their participants to primarily invest in the attractive risky alternative of stocks. Bertrand and Morse (2011) manipulate their subjects to bracket the consequences of payday loans in a broad manner by providing information about the aggregated fees that have to be paid.

(2) Evaluation Frequency

Allowing a frequent evaluation of outcomes emphasises myopic loss aversion of individuals and causes narrow bracketing behaviour. By binding the evaluation frequency to longer periods of time, broad bracketing is promoted (Hardin and Looney, 2012; Hilgers and Wibral, 2014). Remember that it was such a manipulation that corrected the adverse behaviour in the equity premium puzzle (Benartzi and Thaler, 1995). Hilgers and Wibral (2014) confirm the proposition by restricting the feedback given to investors in order to induce broad bracketing.

(3) Decision Frequency

Frequent decision-making strengthens the tendency of individuals to evaluate each decision in isolation. Hence, reducing the frequency with which a DM is confronted with a choice by presenting several decisions collectively induces broad bracketing behaviour (Gneezy and Potters, 1997; Hardin and Looney, 2012; Moher and Koehler, 2010). Stracke et al. (2015, 2017) demonstrate the application of this variable on immediate versus delayed compensation frames. The latter entails the presentation of an integrated reward over several stages of a tournament whereas the immediate compensation refers to a separate presentation of the stages. The results show that the integrated frame aids broad bracketing while the separate rewards draw the attention to the current stage of the tournament and thus induce narrow bracketing. In the context of organisations, Stracke et al. (2015) suggest that broad bracketing could be achieved by implementing a convex reward structure with low immediate but high future compensation in order to promote foresighted decision-making.

Each of the three variables is capable to induce broad bracketing independently of the other two. However, limiting the frequency of evaluations or decision-making has a stronger effect as mechanism to broaden bracketing behaviour compared to introducing longer information horizons. When comparing the two frequency mechanisms to each other, it is found that they have an equally strong potential to facilitate the adoption of broad bracketing (Harding and Looney, 2012).

One can conclude, that decision framing and information presentation are a major determinant of choice bracketing behaviour. Moreover, the concept can be intentionally exploited to manipulate choice bracketing behaviour in a preferred direction. This instrument will be of further importance in the empirical work of this thesis.

2.5.4 Cognitive Processing Systems

The concept of decision frames as the subjective perception of a decision problem leads the discussion to the issue of cognitive processing as a fundamental element in decision-making and, thus, in determining which bracketing behaviour is exhibited by an individual. In particular, literature provides two suggestions: (1) intuition governs the majority of decision-making (Kahneman, 2003) and (2) intuitive or impulsive decision-making constitutes a source of narrow bracketing (Barberis et al., 2006; Ainslie and Haslam, 1992).

Research offers several models which mainly suggest that cognitive processing is executed by two distinct systems. Although the individual theories make use of different terms, their proposed systems generally describe similar characteristics. The main distinction between the systems found in literature is the confrontation of judgment based on effortful reasoning with behaviour relying on spontaneous intuition. As a result, the systems introduced by different theories can be classified into one of these two categories. A rational DM would then be defined by behaviour that conforms to the first category but with the effortless decision-making cost of the second category (Kahneman, 2003).

Stanovich and West (2000) introduce the concepts of System 1 and System 2. System 1 refers to a process that does not involve focused attention, operates fast and automatically, is slow in learning or modification, and is strongly connected to emotion. On the other hand, System 2 requires effort in its execution and operates slowly, in a rule-governed, but flexible and conscious manner. Hence, System 1 describes an intuitive process while System 2 belongs to the category addressing active reasoning. Another theory was already mentioned during the discussion of self-control in chapter 2.4.5. The dual-self model of Fudenberg and Levine (2006) suggests that behaviour is governed by a long-run and multiple short-run selves of an individual. The long-run self adopts the role of a reasoning system whereas the short-run selves represent a system of impulsive behaviour. The theory that is of particular interest to this thesis is the Cognitive Experiential Self Theory (CEST) developed by Epstein et al. (1996). The interest in this concept results from the fact that it constitutes the basis for the development of a test which finds its application in the empirical part of this thesis.

In accordance with the other theories, CEST assumes that individuals employ two distinct processing systems where one corresponds to the reasoning category and the other refers to intuition. The experiential system represents the latter category, being that it operates in a fast, automatic, impulsive, and generally unconscious manner with minimal effort and negligible demands on cognitive resources. This system learns from experience, is able to react in an adaptive manner, and is closely linked to affect. As a result, it is capable of influencing affect while at the same time being influenced by affect and furthermore guides behaviour towards pleasurable outcomes while avoiding painful ones (Epstein, 2003; Epstein, 2014). Finucane et al., (2000) emphasise the role of affect in information processing and decision-making for the reason of it naturally occurring as a first reaction to a given stimulus. The theory suggests that the capability of guiding behaviour efficiently by requiring little effort and not depleting cognitive resources causes the experiential system to be the default mode in everyday life. Since impulsiveness and intuition were already established as a source for narrow bracketing, the prevalence of experiential processing in most situations supports the claim that narrow bracketing is the more natural behaviour in life. In addition, experiential processing as default setting also strengthens the point of view of Kahneman (2003) that the existence of the acceptance principle in decision framing may not be due to insufficient reasoning but originates from impulsive behaviour.

The second proposed processing system is the so-called rational system. Note that rational in the context of CEST refers to the logical reasoning of an individual in compliance to their principles. This is not necessarily equivalent to the best reasonable behaviour because it may rely on flawed principles or is inappropriate for the given problem. For this reason, one should not make the mistake to assume that rational processing is always the superior approach (Epstein, 2014). The rational processing system thus relies on an individual's comprehension of the principles of logic as well as evidence. It operates in a slow, conscious, and analytical manner based on abstract and general rules, requires effort and places a high demand on the cognitive resources of a DM. Moreover, the rational system is free of affect, oriented towards a long-term perspective, and suited to process substantial degrees of complexity and abstraction. Consequently, it enables the acts of planning, delaying rewards, and understanding causal relations. Note that the consideration of a long-term horizon as well as the ability to delay gratification describe attributes of broad bracketing behaviour. Thus, it may be argued that operating in the rational system promotes the adoption of a broad choice bracket. Furthermore, the rational system is capable of understanding the experiential system (Epstein, 2003). However, the names of the two systems might be deceptive in the sense that one should not assume that the experiential system is unable to reason and the rational system not capable of learning (Epstein, 2014).

According to CEST, the two processing systems operate simultaneously and have a mutual influence on each other. Although both systems jointly determine behaviour, the degree to which either system contributes to the final outcome may vary from minimal to almost complete domination. Research reports differences in gender on this issue, where women are more inclined to engage in experiential processing whereas men use their rational system to reflect on their behaviour (Pacini and Epstein, 1999; Epstein, 2003; Frederick, 2005). Furthermore, it is argued that the degree of relative dominance is not only determined by personal characteristics but is also context-dependent (Epstein et al., 1996; Epstein, 2014). With the objective of measuring the extent to which each system is active, Epstein et al. (1996) developed the Rational Experiential Inventory (REI), which will be applied in the empirical work of the underlying thesis. As a consequence of the fast and unconscious reaction of the experiential system, it is capable to affect the slower workings of the rational system. Since the experiential system learns from experience, it can also act as a provider of information to the rational system. On the other hand, the rational system may use its slower operation in order to correct for inappropriate reactions of the experiential system. This intervention of the rational system is often found when a DM reflects on his initial impulsive idea and in the following chooses a more appropriate behaviour.

Koch and Nafziger (2016a) find that choosing a dominated lottery option is linked to missing cognitive reflection; implying that the rational system failed to correct the experiential system. Considering that the rational system is capable to understand the functioning of its counterpart, it is in a position to educate the experiential system in favour of more appropriate reactions (Pacini and Epstein, 1999; Epstein, 2003). Note that the experiential system does not possess the ability to understand the rational system but may serve as a passive provider of information instead of an active lecturer. Additionally, this influence permits the formation of habits which are then governed by the cognitively less costly experiential system (Smith and DeCoster, 2000). In line with this argumentation, the rational system could also be capable to intervene in the incorrect application of narrow bracketing by controlling the impulsiveness of a DM. Considering that the experiential system is able to form habits, organisations may make use of training the experiential systems of their members to accommodate for broader brackets in the constructed behaviour.

2.5.5 Cognitive Ability and Capacity Limitations

The examination of cognitive processing and its role as an influencing factor for bracketing behaviour point to another important determinant of bracketing, namely the limitation in cognitive capacity. People are generally observed to be constrained in their capacity to perceive and process information as well as in their memory and ability to devote attention to the task at hand (Miller, 1994; Simon, 1957; Baddeley, 1992; Kahneman, 1973). These constraints have a considerable impact on the ability to assess several choices simultaneously; in other words, cognitive capacity limitations may pose a severe threat to broad bracketing in decision-making.

Miller (1994), views a DM as a communication system in the sense that he processes the input information in order to match an appropriate response to the stimuli as an output. When the amount of input increases, for example by including multiple instead of one decision, the DM is at risk of making more errors as he reaches his channel capacity. For the reason of such limitations, a DM might fail to perceive the full extent of a decision task and therefore is unaware of the consequences of a choice. On the other hand, the DM might recognize the remote consequences of his actions but is unable to process their implication on what would constitute the optimal choice. Furthermore, literature argues that the cognitive processing capacity is not a constant but varies with respect to the cognitive effort that was already invested in another task (Pocheptsova et al., 2009). The rationale for this is that each decision task demands cognitive resources for its perception and assessment, and cognitive resources are not infinite. Such a depletion of cognitive resources is then found to induce more impulsive decision-making, which in turn impacts the quality of the present and all succeeding choices unless an individual restocks his cognitive resources by resting (Engin and Vetschera, 2017). This intensified impulsiveness can be explained by applying the Fudenberg and Levine (2006) cognitive processing model: a DM would be more likely to behave impulsively when his long-run self grows exhausted from the high demand in processing, leaving the control to the short-run self which tends to succumb to temptation.

Subsequently, it is suggested that individuals defined by lower cognitive ability are more likely to employ narrow bracketing as means to reduce the complexity of the task and avoid the cognitive cost of broad bracketing (Herrnstein et al., 1993; Stracke et al., 2015). In this sense, a DM trades the potential benefits of broad bracketing against the simplicity associated with narrow bracketing. This trade-off gives rise to the argumentation that narrow bracketing is not necessarily an irrational error, but serves as decision strategy for trivial problems that would not justify the burdensome effort associated with broad bracketing. In this context, the resulting choices can be described as being constrained optimal (Read et al, 1999; Simonson, 1990). Moreover, even in subjects of higher cognitive ability, the repeated application of broad bracketing in past tasks might render them incapable of continuing this behaviour. This may occur when their cognitive resources are depleted to such an extent that the cognitive cost of broad bracketing cannot be met.

A study of Blanco et al. (2010) examines subjects' response in behaviour when hedging opportunities are introduced to the experimental tasks. They find that subjects did not change their behaviour in response to facing a game where a hedging opportunity is present. Only for a coordination game with more obvious hedging opportunities, subjects recognised this benefit and furthermore adapted their behaviour to the anticipation that their opponents would recognise the opportunity as well. This implies that individuals are narrow bracketers as long as the benefit of broad bracketing is not obvious. Bland (2019) draws from this observation that employing narrow bracketing is better described by subjects failing to recognize the potential benefit of broad bracketing as compared to extensive cognitive costs of bracketing broadly. These arguments further support the belief and observations that narrow bracketing is the default strategy of individuals.

Although the reasoning for an effect of cognitive ability on bracketing behaviour appears to be straightforward, research reports mixed results. Koch and Nafziger (2019) investigate narrowly bracketed decisions in the context of lottery tasks, endowments integration, as well as loss aversion but do not find a consistent correlation of the choice errors with a low cognitive ability. However, they find that mental accounting correlates with cognitive ability. Rabin and Weizsäcker (2009) also search for a relation between narrow bracketing in lotteries and mathematics skills. Their results only show weak correlation but find that women are more prone to bracket narrowly. This gender effect is in line with the finding that women are more likely to engage in intuitive processing from the previous chapter. On the contrary, the results of Abeler and Marklein (2017) present a correlation between bracketing narrowly and having low mathematic skills. In compliance to Koch and Nafziger (2019), they state that mental accounting is primarily used by respondents of lower cognitive ability. Burks et al. (2009) also detect a correlation between low cognitive ability and rate of choice errors in the form of narrow bracketing. They find respondents with higher cognitive ability levels to be more patient, thus being less likely to fall prey to temptation. In addition, such individuals exhibit a higher willingness to take the risky option in a positive EV lottery while choosing the certain loss over the risky alternative in a lottery with negative EV; indicating that they aggregate the two gambles.

Similar to these findings, Frederick (2005) states that high cognitive ability in women causes them to be patient while the same attribute induces men to accept more risk. Webb and Shu (2017) conclude that cognitive limitations account for a major part of the observed bracketing effects. Nevertheless, the results of Stracke et al. (2015) suggest that lower cognitive capacity is not a universal indicator of narrow bracketing. Instead, they find an interaction with the complexity of a decision task. In complex decision problems, respondents with higher cognitive ability are capable to learn from repeated rounds by adopting broad bracketing whereas low cognitive ability individuals primarily employ narrow bracketing; seemingly being unable to learn from past rounds. When confronted with simpler decision tasks, no effect for cognitive limitations is found. These findings imply that cognitive limitations determine behaviour as the complexity of the decision environment increases (Stracke et al., 2015).

With respect to learning, Abeler and Marklein (2017) find no evidence that their respondents learn from experience. Herrnstein et al. (1993) also obtain only little evidence of learning while Haisley et al. (2008) find that providing respondents with the opportunity to learn from past rounds intensified the tendency to bracket narrowly. One has to note that Haisley et al. (2008) designed the learning opportunity by providing feedback about the outcome after each round. As it was discussed in the chapter on decision framing (2.5.3), frequent outcome evaluation is one of the main variables that encourages narrow bracketing behaviour. This implies that their adaptation is inappropriate to stimulate learning and the authors observed an effect of information presentation. However, Hilgers and Wibral (2014) note that when the task effectively guides towards learning, subjects of higher cognitive skill show a superior ability for learning to bracket gambles broadly. Therefore, an important implication for management and decision support systems is that manipulations designed to guide bracketing behaviour should provide information that facilitates processing and neutralises cognitive constraints without simultaneously generating a framing effect.

2.5.6 Complexity

The preceding chapter has already hinted at the complexity of the decision environment representing an important determinant of bracketing behaviour, which may also be linked to the topic of cognitive ability of the DM, in particular, the constraints placed on it when facing increasingly complex problems.

It is suggested that the complexity associated with the perception and computational evaluation of a series of decisions interferes with an individual's ability to interpret the integrative decision environment. The prevention of a correct interpretation of the decision problem causes preferences to shift and is likely to yield an outcome that is inconsistent with rational choice (Redelmeier and Tversky, 1992). Burks et al. (2009) describe this complexity as noisy component in the perception of the decision value; implying that complexity moderates the perception of the objective utility that various outcome states of a problem may yield. This noisy component increases along with complexity and, as a consequence, distorts the interpretation of the DM to a greater extent and induces a shift towards behaviour based on a less precise perception. With respect to the degree of complexity, it is found that the number of different alternatives or decisions is the major influencing factor (Payne, 1976).

According to theory, when a DM is confronted with a decision problem that would involve processing a complex environment, he will choose to decompose the problem into serval components which are then evaluated in isolation. Although this behaviour is likely to yield suboptimal outcomes, the natural attractiveness of it is assumed to be of considerable importance in many areas of decision-making (Stracke et al., 2017). In this sense, an individual resorts to narrow bracketing as a decision strategy whenever a decision task is defined by an excessively complex decision environment. A DM may then bracket complex decisions narrowly for three reasons: either (1) the potential benefits gained from adopting broad bracketing are not sufficiently high to compensate for the effort required to assess the decision problem in its entirety; or (2) he would prefer to bracket broadly but is not capable of implementing this approach because the given degree of complexity exceeds his cognitive ability. Alternatively (3), he may fail to recognize that the decision problem has indirect consequences which should be integrated to an aggregate problem and therefore adopts a narrow bracket (Herrnstein et al., 1993; Read et al., 1999; Stracke et al., 2015).

Stracke et al. (2017) investigate the impact of complexity on bracketing behaviour in the context of a two-stage pairwise elimination contest where the winner of stage 1 earns the right to participate in stage 2. The experiment consists of two treatments, which in turn incorporate two presentation forms each. One form is designed to induce narrow bracketing while the other facilitates broad bracketing by presenting either separate or integrated rewards over the two stages. This base treatment is of a higher complexity since broad bracketing requires subjects to consider their opponents' investments in both stages as well as their own investment in the second stage in order to decide upon their investment in the first stage. The control treatment then reduces the complexity of the task by eliminating the strategic interaction between stages through the replacement of the second stage investment with a lottery choice. Their results confirm that subjects employ narrow bracketing as a strategy to reduce complexity as well as the moderating role of cognitive ability. While it is found that subjects tend to bracket broadly in both presentation modes of the less complex control treatment, an interaction of complexity and cognitive ability is found for the baseline treatment. When confronted with the higher degree of complexity, respondents with greater cognitive ability continue to bracket broadly whereas a low cognitive ability induces narrow bracketing behaviour for the complex environment. Note that in the simplified treatment, cognitive ability does not seem to affect bracketing behaviour (Stracke et al., 2017).

2.5.7 Adopting Heuristic Decision-Making

The empirical finding of the preceding subchapter that respondents employ narrow bracketing in a complex setting while doing so to a lesser extent in a more simple environment fortifies the assumption that narrow bracketing may serve as a heuristic. Determining the preferred option in a decision problem regularly demands the investment of time and effort and may cause the DM to experience discomfort and conflicts in his processing of the task. On this basis, research on decision-making offers a variety of heuristics which aim at adjusting the difficulty of processing the decision problem to the available cognitive resources while retaining reasonable results.

Nevertheless, heuristic decision-making also has the potential to cause considerable errors in choices by excluding essential aspects of the problem from consideration. In line with the argument that cognitive resources are depleted by continuous and effortful decision-making, it is suggested that the application of heuristics depends on the cognitive ability of the DM as well as the characteristics of the task (Payne, 1976; Simonson, 1990). In addition, the framing versions of the decision problem may induce the application of different heuristics in the same individual (Tversky and Kahneman, 1973). Regarding the impact of framing in choice bracketing, Read and Loewenstein (1995) propose that when confronting a DM with a simultaneous presentation frame of choice, he prefers to apply a diversification heuristic. On the contrary, the choice heuristic that is argued to be most natural in a separate frame is the selection of the most preferred option within each problem in the narrow frame.

When taking the theory on cognitive processing into account, CEST considers decision-making heuristics to be inherent to the experiential systems (Epstein et al., 1996). Note that the experiential system as processing mode as well as narrow bracketing were already introduced to be the default modes in decision-making. Connecting these findings with the statement that heuristic responses are directly linked to the experiential system, supports the argument that narrow bracketing is a decision-making heuristic as well. The rationale is that judgment, in particular of probabilities, based on intuition does neither involve an extensive examination of everyday events in connection to probabilities nor the use of single probabilities to calculate and assess compound ones (Tversky and Kahneman, 1983). Accordingly, it is found that subjects high in experiential processing are more inclined to use heuristics in their decision-making and perceive their resulting behaviour as reasonable. Although the primary source of heuristic response is the experiential system, the analysis of Epstein et al. (1996) indicated that it is also influenced by the rational system. This reflects the theory that both systems jointly determine the behaviour of a DM.

As a consequence of the spontaneous nature of intuitive processing, the induced thoughts and behaviours are strongly dependent on the situation-specific attributes that can easily be accessed by the individual; meaning that they come to mind conveniently and effortlessly. This relationship is captured by the availability heuristic which describes the tendency of individuals to evaluate the frequency or probability of an event based on the lack of difficulty in recalling relevant examples. The principle of ease of retrieval can also be transferred to the context of information presentation, so that an individual tends to act in response to the most accessible frame (Tversky and Kahneman, 1981; Barbersis et al., 2006). Hence, the availability heuristic is a likely contributor to the acceptance principle in decision framing as well as narrow bracketing behaviour; being that narrow bracketing involves the neglect of indirect consequences and thereby acts on the information that comes to mind easily.

It is important to note that the presented interaction between lower cognitive ability and higher complexity of the decision environment does not necessarily imply that heuristic decision-making is exclusively found in naïve individuals. Research reports that the application of heuristics is also undertaken by professionals and statistically trained subjects. Although such a sophisticated DM may apply more advanced heuristics and therefore avoid elemental mistakes, they still succumb to choice errors caused by heuristic responses when they judge intuitively (Tversky and Kahneman, 1974).

2.5.8 Motivated Choice Bracketing

The final determinant of choice bracketing discussed here concerns the deliberate adoption of a specific bracket with the objective to achieve a self-control effect. Setting personal goals creates expectations about future outcomes, which then serve as a reference point to the individual. Subsequently, achieving an outcome that lies above the goal is perceived as satisfaction whereas failing to meet the reference point is experienced as a loss in form of dissatisfaction that increases with the distance to the goal. In this sense, a DM can manipulate his behaviour by entering a self-enforcing contract with his future selves in the form of an objective. Self-set goals are argued to be one of the most adaptable and readily employable instruments to achieve commitment to a preferred behaviour but may also entail adverse consequences (Koch and Nafziger, 2020; Ainslie and Haslam, 1992; Hsiaw, 2018).

In this context, it is suggested that bracketing goals narrowly, i.e. setting incremental instead of aggregate goals, enhances the perception of how easily goals can be accomplished. Hence, such incremental goals increase the immediate motivation of a DM as he seeks to avoid the pending perceived loss of not meeting a goal in the near future (Read et al., 1999; Hsiaw, 2018). On the contrary, having an overall goal provides less motivation since the DM is confronted with the temptation of effort substitution; meaning that the broader time frame associated with an overall goal offers the possibility to compensate for a lack of effort at the beginning of an evaluation period by increased effort supply at a later point in time. Koch and Nafziger (2020) observe this tendency to procrastinate in respondents who are subjected to a weekly goal instead of daily goals.

Hence, the effect of effort substitution induces a skewed effort supply and thus represents suboptimal behaviour. With respect to the effort distribution, the suboptimality of this behaviour arises from the circumstance that the costs to supply effort are considered to be convex, so that the DM would prefer an equal distribution of effort over time instead of extensively high effort before the goal evaluation is due (Koch and Nafziger, 2020). Furthermore, the adopted bracket has strong implications on how an individual responds to feedback from previous decisions. When adopting interim goals, the DM perceives the payoffs to be independent from each other so that his future decisions are not influenced by past outcomes. A broad bracketer, however, views each outcome as a contribution to his progress towards the overall goal and therefore may adapt his behaviour for subsequent decisions based on present outcomes (Hsiaw, 2018).

While the imminent threat of experiencing a loss motivates the DM to stick to narrowly bracketed goals, broad bracketing may also serve the purpose of self-control in another context. Especially when the DM faces an uncertainty about the degree of effort that he will be able to provide, incremental goals cause an increased perception of potential dissatisfaction from being unable to accomplish one of the frequently evaluated goals. Hence, in the context of uncertainty about one's own ability to meet narrow goals, a broad bracketing approach may be favourable as it allows the subject to benefit from a risk aggregation effect (Hsiaw, 2018; Koch and Nafziger, 2020). Despite providing an insurance against the threat of an immediate loss, the overall goal still holds the temptation of effort substitution (Koch and Nafziger, 2016b). However, broad bracketing is found to be the superior strategy with regard to self-control in consumption behaviour, in particular, when addressing addictive behaviour (Rachlin, 1995; Heyman 1996). The underlying rationale is that the long-term interest of the DM conveys such importance to him so that he would not abandon it in favour of a short-term temptation. The existence of the aggregate goal then depends on the DM perceiving himself capable of fulfilling the long-term interest at each decision point. If he abandons this interest at one point, he expects that he will continue to fail in restraining his behaviour for future decision points and the long-term interest cedes to prevail (Heyman, 1996; Ainslie and Haslam, 1992).

The two preceding paragraphs lead to the conclusion that optimal motivated bracketing is represented by a trade-off between motivation and risk aggregation (Koch and Nafziger, 2016b; Hsiaw, 2018). When multiple decisions arise simultaneously, the favourability of broadly bracketed goals increases alongside the uncertainty regarding the potential effort supply or the occurrence of unforeseen adverse events. The reason for this is, that the risk reducing effect associated with broad goals becomes increasingly advantageous. In relation to decisions that occur sequentially, the optimal bracketing strategy depends on which segment of the evaluation period is characterised by uncertainty. Setting an aggregate goal is favourable when the uncertainty concerns the early section of the time frame, since the risk aggregation property enables the DM to react to unforeseen developments without experiencing the disutility from not achieving incremental goals. However, when the uncertainty is found in the later part of the time period, the DM is more prone to supply little effort in the earlier stage; implying that the motivational power is lost and the fulfilment of the overall goal is at risk (Hsiaw, 2018).

Koch and Nafziger (2020) report a finding that offers valuable implications for organisations. For the reason that organisations are often confronted with a skewed effort supply, with high effort levels immediately before an evaluation point, the impact of interim requirements is investigated. It is found that an externally set work requirement in addition to self-induced narrow goals lead to a lower effort supply compared to the sole use of incremental goals. This implies that self-set incremental goals provide enough motivational power while the externally set minimum requirement has demotivating properties. On the other hand, for individuals with an aggregate goal, the introduction of the frequent work requirements causes a significantly higher effort level, being that effort substitution is prevented. Hence, an organisation may benefit from re-framing e.g. a monthly performance target adopted by its employees into multiple intermediate goals in favour of generating a higher motivational power. However, in response to the set work requirement, a considerable increase in the drop-out rate was noticed during the experimental period. This finding might translate to a higher turnover rate in an organisation.

3 Research Framework and Hypotheses

Subsequent to analysing the literature that is the foundation for choice bracketing, with its proposed determinants as well as effects that stem from applying a specific bracket to a decision problem, this chapter aims at introducing the reader to the research framework of this thesis and its corresponding main hypotheses.

Three main findings emerge from the literature review and represent the groundwork for the development of the research framework. First, chapter 2.5.4 on cognitive processing states that the application of the experiential system is the default mode for a DM. The subsequent chapter addressing cognitive ability (2.5.5) finds that the trade-off between benefiting from adopting the appropriate bracket and the cognitive demand of doing so is likely to cause narrow bracketing to be the default mode in decision-making. Chapter 2.5.7 then presented narrow bracketing as a heuristic in order to facilitate the assessment of choices that are defined by a higher degree of complexity, which is discussed in chapter 2.5.6. Moreover, when excluding findings on the theory of motivational bracketing, the common consensus found in several chapters states that employing narrow bracketing yields suboptimal outcomes. This proposition of task complexity as major determinant and the resulting heuristic response of individuals leads to the first research question.

> *RQ1: In which manner does varying the degree of task complexity influence a decision maker's choice bracketing behaviour?*

In order to introduce several different complexity dimensions, chapter 2.5.3 on decision framing effects provides the basis for the idea of manipulating how the DM perceives the decision problem. However, in contrast to pure decision framing which would not change the underlying task complexity across different frames, the aim here is to manipulate the objective complexity by the means of information presentation and the initial state of the task. The influence of task complexity on bracketing behaviour has not yet received extensive attention in empirical studies. Stracke et. al (2015) find support for the assumption that high task complexity induces more individuals to bracket narrowly. The authors are also able to link this adoption of narrow bracketing to a lower cognitive ability measured by self-reported math grades. However, the complexity manipulation was undertaken by eliminating a strategic interaction between two stages of a gamble, so that one treatment plays a strategic two-stage game whereas the other treatment plays a lottery. Hence, the context of the task type between the treatments was changed and subjects of the different treatments are not confronted with comparable tasks. On the contrary, this thesis seeks to establish multiple complexity dimensions of the decision problem while maintaining the basic type of the task that subjects need to perform. We suggest that manipulating the way and amount of information that is presented as well as the initial state of the task without an actual alteration in the underlying type of the problem might be preferable with respect to maintaining comparability between the treatments. In the context of Stracke et al. (2015), this would mean the introduction of two games or two lotteries of different complexity but not the mixture of both task types.

This is where the second research question finds its starting point. The role of cognitive capability, e.g. in the form of math grades, in guiding the final behaviour is easy to argue for as it is such skill that opposes the complexity of a task. While a majority of studies uses such self-reported math grades as a proxy for cognitive ability, one can argue, that, especially in the context of decision-making under risk, the concept of numeracy constitutes an appropriate measure. This is because decision-making under risk is inherently linked to the application of probabilities or proportions that need to be accurately assessed by the DM. The concept of statistical numeracy describes an individual's ability to comprehend probabilistic or statistical calculations. From this ability to understand probabilistic operations subsequently originates an individual's competence to interpret information about risk and adapt their choices accordingly (Lipkus et al., 2001; Cokely et al., 2012). While math grades certainly incorporate numeracy to some degree, it may be favourable to utilise a more accurate measure of this skill. Cokely et al. (2012) present the Berlin Numeracy Test in order to measure this skill and highlight its appropriateness for testing educated or highly educated respondents. Considering that most studies use university students as participants, this characteristic of the Berlin Numeracy Test supports the argument to use it as a measure for cognitive ability instead of eliciting math grades. Thus, the psychographic of numeracy represents the chosen measure for cognitive ability in the framework of this study.

However, while acknowledging the importance of cognitive ability as an intermediating factor, we propose that it is not solely the objective task complexity that influences bracketing behaviour. Instead, we suggest that emphasis should be placed on the perceived complexity that results from the objective complexity, cognitive ability, and the personal processing of the task by an individual. The assumption is that cognitive ability limitations determine the basic threshold of when a task is objectively too complex for a specific individual; and that this threshold is then subjected to the individual's cognitive processing tendencies. In this sense, two subjects with equal cognitive ability may perceive the same decision problem to be of a different complexity level, due to operating in different cognitive processing systems. Therefore, we argue that alongside cognitive ability, the other selected psychographics that measure cognitive processing are an intermediating force between objective complexity and perceived complexity and deserve to be investigated in research.

Based on the findings of chapter 2.5.4, the influence of primarily applying the experiential system should receive particular attention. Our motivation for this is, that the implications of the experiential system on intuitiveness and impulsiveness may interfere with an otherwise sophisticated cognitive ability, causing an individual to perceive a higher complexity. On the other hand, primarily relying on the rational system may be of supporting nature with respect to the utilization of the given cognitive capacity and as a result provokes an upwards shift of the complexity threshold.

> *RQ2: How and to what extent do a decision maker's psychographics in the form of numeracy, rationality, experientiality, intuitiveness, emotionality, imagination, and impulsiveness act as intermediators on perceived task complexity and the resulting bracketing behaviour?*

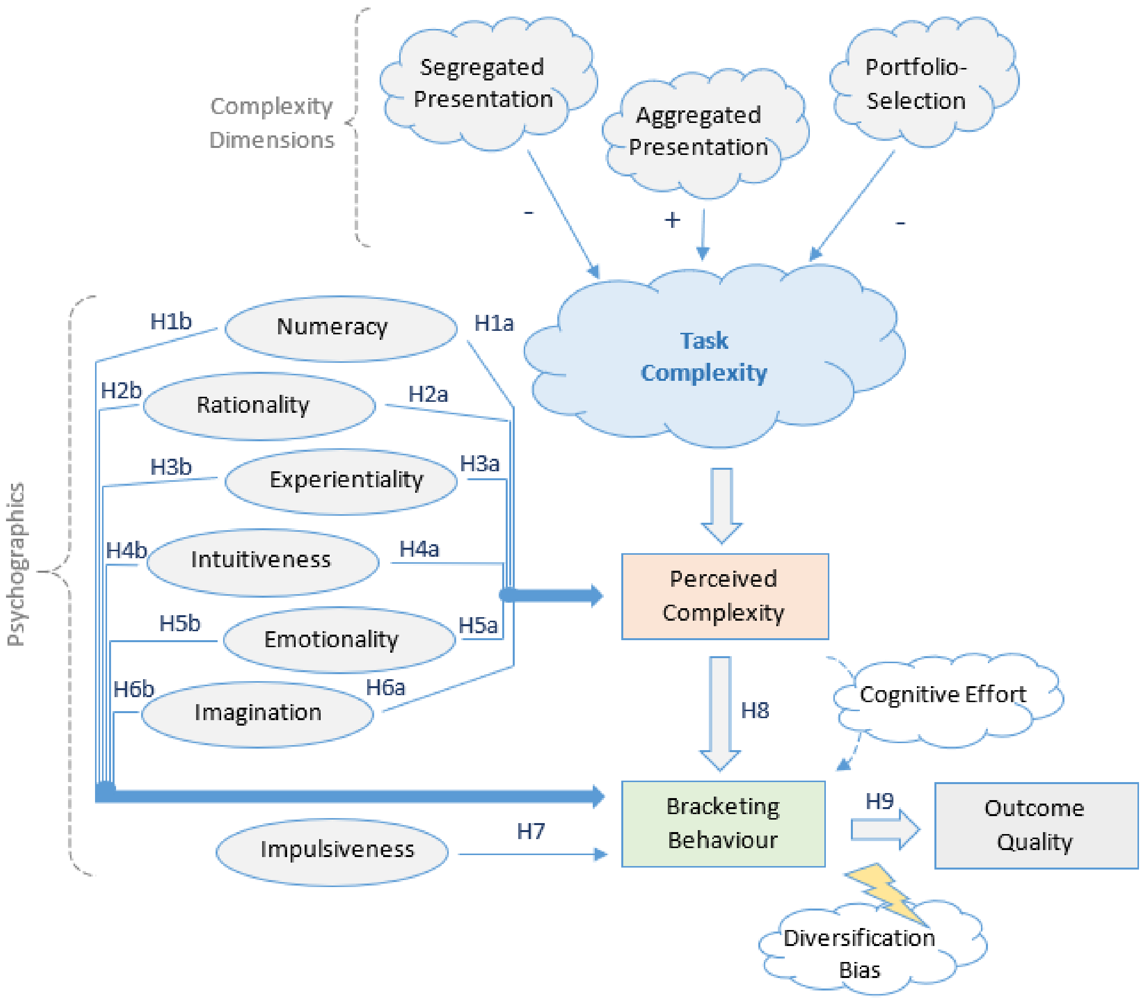
A third research question is specific to the context in which the main task of the experiment is set and addresses the concept of variety seeking under broad bracketing, which is discussed in chapter 2.4.2 as part of the emergent properties effect. According to the findings in literature, broad bracketers are subject to the diversification bias (Read and Loewenstein, 1995); meaning that they are prone to selecting a degree of variety that exceeds their optimal allocation and causes them to experience dissatisfaction. The key decision problem that respondents of our designed experiment face is a personnel selection task. This setting is chosen since, to our knowledge, up to the point in time where the experimental design was undertaken, no other study investigated choice bracketing in a personnel selection task. One has to note that during the implementation phase of the experiment, Chang et al. (2020) published their findings on variety seeking behaviour in personnel selection.

However, Chang et al. (2020) investigate the variety seeking effect on gender diversity whereas this thesis explores diversification with respect to the skill level on several attributes that characterise each candidate. Furthermore, although the pure variety seeking effect for gender can be well argued for, it may be disputable if the adverse consequence of seeking an excessively high degree of variety is applicable to gender diversity. Our argumentation for this is that the maximum degree of diversity that gender as a traditionally dichotomous variable may achieve is an equal distribution between male and female employees in a team. As soon as such an equal distribution is reached, the addition of another employee of either gender shifts the distribution back towards the dominance of one gender and thus, less variety. In the case of exchanging one team member for a new employee, the distribution either remains equal or also shifts in the direction of dominance of one gender. Hence, the maximum degree of variety in gender is 50%, which may not be labelled as excessively diversified, so that negative effects of the diversification bias are unlikely. Nevertheless, one has to note that a more modern view on the gender variable leads away from it being a variable with only two states. This means that a higher degree of variety may be observed when a diverse gender is included in the measurements.

Excess variety in such non-dichotomous demographics or individual payoff-relevant characteristics may pose a source of conflict and hence, negatively impact the team spirit or performance. In line with this, in the design of our experiment, an excessive focus on variety regarding the skill composition of a team may lead participants to fail to maximise the utility generated by the selected team. Thus, the experimental design allows for the existence of the diversification bias. In this sense, our work is of complementary character and should not be seen as a replication study.

> *RQ3: How does bracketing behaviour influence the implemented degree of variety as well as achieved utility with respect to the distribution of skill level across multiple attributes of applicants chosen to form a team in a personnel selection process?*

Figure 4: Research Framework



Based on these research questions and their underlying argumentation, the research framework illustrated in figure 4 is constructed with the aim of providing a comprehensive overview of the proposed interrelations. The framework encompasses three complexity stages generated by the manipulation of the initial state and presentation of the underlying personnel selection task. Each of these stages then creates a different starting point with respect to the task complexity that a respondent of the experiment is confronted with. This objective task complexity represents the context-dependent component for the complexity that is then perceived by an individual.

Following this objective task complexity, the complexity that an individual respondent perceives acts as an intermediating factor between the task complexity induced by a specific dimension and the resulting bracketing behaviour. This perceived complexity is also dependent on the stated psychographic attributes of an individual, which in combination yield either a high or low perception of the objective complexity of the decision environment. These psychometric characteristics represent the subject-dependent component of the perceived complexity. For the reason that psychometrics are individual characteristics, these influencing factors remain constant across different decision-making contexts. A pronounced perceived complexity is then argued to induce narrow bracketing as a result of the higher required cognitive effort for accurately processing the task; which in turn exceeds the available cognitive resources. On the contrary, a low perceived complexity facilitates an optimal use of an individual's cognitive capacity and enables the employment of a broad bracket.

In addition, psychometrics as individual characteristics do not only impact the complexity that a DM experiences in a specific decision problem but may also form a baseline tendency for either applying broad or narrow brackets to a given problem. In this sense, following the statement that narrow bracketing is the default mode for decision-making, we suggest that psychographics contribute to the determination of this default mode for a specific DM. Thus, the psychographics of the framework are depicted as influencing factor for the perceived complexity as well as the bracketing behaviour itself. Note that the attribute of impulsiveness is displayed to be directly linked only to behaviour instead of influencing the perception of complexity as well. The rationale for this is that impulsiveness is a feature that strongly affects the fundamental decision-making attitude of an individual and does so independently from the complexity of the decision problem. In this sense, an impulsive subject is suggested to be prone to adopting a narrow bracket in high as well as low complexity settings.

Finally, decision-making under a specific choice bracket influences which alternatives are chosen and thus, the utility that an individual gains from the decision. Regarding the quality of the solution, it should be noted that our experimental setting does not address the issue of self-control problems and therefore does not incorporate the concept of motivational bracketing. Hence, narrow bracketing may not be viewed as an optimal strategy but as a decision error or heuristic in order to facilitate decision-making in complex task environments. In this sense, broad bracketing is the favourable strategy as long as individuals do not fall prey to the diversification bias by allowing the tendency of selecting an excessive degree of variety to overshadow the aim of value maximization for the choice.

The proposed framework gives rise to the formulation of the following main hypotheses. These are complemented by task-specific hypotheses, which will be introduced in the subsequent chapter on the design of the experiment. The first two hypotheses address an individual's cognitive ability measured by the numeracy variable as well as the tendency to process information by the means of the rational system. As suggested above, cognitive ability determines the threshold of when a task is not processible and thus, has an important role in the degree of perceived complexity. For the reasons that the rational system operates in a conscious and analytical manner and is furthermore characterised by a long-term perspective and an ability to process complexity and causal interactions, we hypothesise that it has a decreasing effect on perceived complexity. With respect to bracketing behaviour, a high cognitive ability and rational processing is hypothesised to form a tendency for broad bracketing as default mode.

H1a: If a decision maker scores highly in numeracy, he will perceive a lower complexity compared to an individual with low numeracy measurements.

H1b: If a decision maker scores highly in numeracy, he is more likely to apply a broad bracket to a decision-problem compared to an individual with low numeracy measurements.

H2a: If a decision maker achieves a high score on rationality, he will perceive a lower complexity compared to an individual with a low rationality score.

H2b: If a decision maker achieves a high score on rationality, he is more likely to apply a broad bracket to a decision-problem compared to an individual with a low rationality score.

The following five hypotheses address the psychographics associated with the inclination to primarily operate in the experiential system. The argumentation for the hypotheses bases on the automatic, impulsive, and affective properties of this system as well as its aim of placing minimal demand on cognitive resources while guiding behaviour towards pleasurable outcomes. Consequently, we hypothesise that the employment of the experiential system as predominant system does not support an efficient processing of complexity so that its perceived magnitude is increasing. In accordance with not having a complexity reducing effect, primarily processing information by means of the experiential system is hypothesised to facilitate the decision process by setting narrow bracketing as default mode.

H3a: If a decision maker scores highly on experientiality, he will perceive a higher complexity compared to an individual characterised by a low experientiality score.

H3b: If a decision maker scores highly on experientiality, he is more likely to apply a narrow bracket to a decision-problem compared to an individual characterised by a low experientiality score.

H4a: If a decision maker scores highly on intuitiveness, he will perceive a higher complexity compared to an individual who scores low on intuitiveness.

H4b: If a decision maker scores highly on intuitiveness, he is more likely to apply a narrow bracket to a decision-problem compared to an individual who scores low on intuitiveness.

H5a: If a decision maker is characterised by a high score on emotionality, he will perceive a higher complexity compared to an individual with a low emotionality score.

H5b: If a decision maker is characterised by a high score on emotionality, he is more likely to apply a narrow bracket to a decision-problem compared to an individual with a low emotionality score.

H6a: If a decision maker scores highly on imagination, he will perceive a higher complexity compared to an individual who is described by low imagination score.

H6b: If a decision maker scores highly on imagination, he is more likely to apply a narrow bracket to a decision-problem compared to an individual who is described by low imagination score.

H7: If an individual is an impulsive decision maker, he is more likely to adopt narrow bracketing in comparison to a decision maker who is low on impulsiveness.

The final two main hypotheses consider the impact of the aggregate perceived complexity, which arises from the individual psychographics as subject-dependent and the objective complexity as situation-dependent component, on the adopted bracketing behaviour as well as the quality of the decision outcome. Here it is hypothesised that a higher degree of perceived complexity encourages the adoption of narrow bracketing as a heuristic in order to facilitate the decision process. In addition, the individual traits and processing tendencies of a DM may favour the adoption of heuristic decision-making. For the reason that heuristic decision-making is commonly regarded as suboptimal, we hypothesise that broad bracketing achieves outcomes that are closer to an optimal allocation. This accounts for the fact that, although broad bracketing commonly leads to superior outcomes, one should not assume that a broad bracketer is an entirely rational value maximiser, but that he is still subjected to cognitive constraints. Therefore, a broad bracketer in the context of this thesis is an individual that behaves close to value maximisation but may not necessarily find the optimal allocation.

H8: A high perceived complexity influences behaviour in favour of narrow bracketing compared to a low perceived complexity.

H9: Decisions made under broad bracketing yield a higher utility compared to narrowly bracketed decisions.

4 Methodology

Following the presentation of the research framework and the derived main hypotheses, the purpose of this chapter lies in navigating the reader through the design of an experiment before finishing with explaining its implementation as well as providing a first insight into the descriptive statistics regarding the respondents in the sample.

4.1 Experimental Design

In pursue of investigating the introduced research questions and supporting the developed framework, the approach of designing and executing an experiment is chosen. This subchapter presents the main components and individual tasks of the experiment as well as their sequencing in the implementation phase. Furthermore, the main hypotheses will be complemented by task-specific hypotheses.

The experiment consists of two main parts, namely the lottery part and the personnel part, which each comprise several tasks. The first part is set in the context of lottery choices like they are commonly found in the choice bracketing literature and have been used to illustrate behaviour in the theoretical chapters of this thesis on multiple occasions. However, the motivation to include lottery tasks in the experiment does not solely stem from the circumstance that they are regularly used in this research area but is based upon the intention that this part serves two purposes. For the reason that existing studies report a clear bracketing effect, the lotteries can be used to measure the basic bracketing attitudes of the respondents and subsequently serve as a check whether choice bracketing is applicable to personnel selection. The personnel selection part is developed with the aim of examining a decision-making context that has received little attention by choice bracketing research. The selection of personnel represents an important decision-making problem for any organisation since the wellbeing of the organisation depends on the decision quality of its agents. Furthermore, by incorporating two different contexts into the experiment and observing choice bracketing in both parts, the lottery tasks provide a measurement for investigating the consistency of bracketing behaviour across different areas. The full instructions and tasks can be found in Appendix A.

4.1.1 Part 1: Lotteries

This part of the experiment contains four independent lottery tasks. After each task, subjects are requested to state the difficulty that they would assign to the respective task in percent, ranging from 0% for simple to 100% for difficult.

The first task is a single lottery choice, where subjects have to decide between a sure profit and a gamble of higher EV. This task aims at eliciting a baseline risk preference of respondents in the absence of choice bracketing. The absence of bracketing results from the fact that choice bracketing requires decision-making with respect to multiple decision problems. In order to maintain the storyline of the experiment, where respondents were assigned the role of head of department, the wording of the single lottery task refers to potential projects that may be undertaken by the department.

Choose between:

A) A sure profit of 19.000
B) A 20 percent chance to gain 100.000, and 80 percent chance to gain nothing

Subsequent to the single lottery, respondents are confronted with the first double lottery. Both decisions that are part of such a double lottery are presented simultaneously and subjects are instructed to read both decisions before making their choice for either one. This decision problem introduces bracketing behaviour and is based on the original lottery example of Tversky and Kahneman (1981, p. 454), but multiplies the values by the factor of 100 in order to generate prospects that are more appropriate in a managerial context. Furthermore, the prospects describe the potential profit or loss generated by the respective candidate to a vacancy in the department of the DM instead of a contextless gamble. Concerning the negative EV lottery, respondents receive the instruction that they need to fill the position and thus, are required to make a decision. In accordance with prospect theory and the findings of existing studies, subjects are expected to select A) for the first decision and D) as their second choice. This results from individuals being risk averse in the domain of gains (Decision 1) while behaving in a risk seeking manner in the domain of losses (Decision 2). When bracketing both decisions as a set, the broad bracketing individual would be able to benefit from the risk reducing effect and recognizes that the choice combination AD is dominated by the combination of BC.

Please read both decisions carefully before making your choices for each one.

Decision (1): Choose between:

A) A candidate with a sure profit of 24.000 generated
B) A candidate with a 25 percent chance to generate a profit of 100.000, and 75 percent chance to generate nothing

Decision (2): Choose between:

C) A candidate with a sure loss of 75.000 generated
D) A candidate with a 75 percent chance to cause a loss of 100.000, and 25 percent chance to lose nothing

The third task describes a double lottery similar to the one above but with slightly adapted prospects. The main purpose of this lottery is to serve as a consistency check for the bracketing behaviour of respondents. Furthermore, it might be possible to observe fluctuations in the perceived complexity of subjects for this task compared to the first double lottery, being that subjects now face such a lottery for the second time instead of the first occasion. In consequence of repeating the task, more individuals might be able to bracket broadly in the consistency lottery. On the other hand, a subject bracketing broadly in the first double lottery but failing to do so in the consistency lottery might hint at the first result being a lucky guess instead of an actual evaluation of the task as a broad decision set. Alternatively, such an observation may indicate that a DM is cognitively depleted by the first double lottery and thus does not dispose of sufficient cognitive resources to continue broad bracketing for a second task.

Please read both decisions carefully before making your choices for each one.

Decision (1): Choose between:

A) A candidate with a sure profit of 19.000 generated
B) A candidate with a 20 percent chance to generate a profit of 100.000, and 80 percent chance to generate nothing

Decision (2): Choose between:

C) A candidate with a sure loss of 80.000 generated
D) A candidate with an 80 percent chance to cause a loss of 100.000, and 20 percent chance to lose nothing

The end of the lottery part is represented by a second single lottery. In contrary to the first task, this lottery does not aim at measuring risk preferences but has the objective of forcing broad bracketing of the first double lottery by explicitly stating the possible choice combinations. In this sense, the formulation of the decision problem executes broad bracketing on behalf of respondents so that their task is reduced to the assessment of the overall outcomes. However, the combinations AC and BD are not included in the presentation. This is done because the purpose of the task is to directly investigate the ability of respondents to select the dominant combination (BC) instead of the choice combination that literature finds to be the prevalent choice (AD). This aggregate presentation drastically reduces the complexity of the original lottery and hence, should enable subjects to recognize the dominant option. In particular, this task manipulation facilitates the observation that both combinations offer equal probabilities but option B (describing the combination BC) includes a sure surplus of 1.000 compared to option A. The observation of this hypothesised change in choice preference would support the following assumption: subjects choose a dominated combination in the double lottery as a result of them being unable to first integrate both decisions into a set and then evaluate the total outcomes. They would not choose an inferior combination as a result of being unable to calculate outcomes in general. This means that subjects fail at the bracketing task and not the following outcome evaluation task of the decision-making process.

Choose between:

A) A candidate with a 25 percent chance of generating a profit of 24.000 and a 75 percent chance of generating losses of 76.000
B) A candidate with a 25 percent chance of generating a profit of 25.000 and a 75 percent chance of losing 75.000

The lottery part of the experiment allows for the formulation of three task-specific hypothesis.

H10: For the double lottery, a lower fraction brackets broadly by selecting the dominant choice combination of B and C, compared to the share of narrow bracketing subjects, which follow prospect theoretic predictions and select the dominated combination of A and D.

H11: Behaviour of decision makers is consistent for the double lottery and consistency lottery, so that if a decision maker brackets broadly in the double lottery, he will maintain the adopted bracket in the consistency lottery.

H12: When explicitly combining the prospects of the double lottery, more decision makers bracket broadly and choose option B, which represents the dominant BC combination.

4.1.2 Part 2: Personnel Selection

The second part constitutes the focal point of the experiment, being that these tasks are specifically designed for this experiment and, to our knowledge, cannot be found in other studies addressing choice bracketing. The personnel selection segment comprises three tasks and, as it was implemented for the lotteries, requests participants to enter their perceived complexity after each of them.

The basic setting for this part is that the respondents, as heads of department, need to fill in three vacancies and may choose from nine different applicants to fulfil their task. All of the three vacancies have a similar profile, so that each candidate may be considered for any of the positions. Subjects are not given any specific job description or information on skill requirements. In turn, the candidates are described by varying levels on four skill items as well as one team compatibility (TC) item. The skill items define the position-fit (PF) of a candidate where each item may range from 0% to 100%, representing an extremely low and exceptionally high ability respectively. The same ranking scheme applies to the TC factor, which states the ability of a specific candidate to function properly in a team environment. In order to calculate the overall-fit and thereby the utility of a specific applicant, the sum of the skill scores needs to be multiplied with the TC attribute. Consequently, and except for the case of a 100% TC score, this multiplication will decrease the utility of an applicant. This reduction is based on the assumption that even a well qualified candidate will not live up to his full potential if he is not able to properly integrate himself into a team setting. Hence, the TC factor accounts for the loss of potential productivity in candidates who would represent a poor fit for the team. The utility of a potential candidate j, who is characterised by different levels on the skill items i, is then expressed by means of the following equation:

$$u_j = TC_j * \sum_{k=1}^{4} i_k$$

With respect to the skill distribution across the nine candidates, participants in the experiment encounter candidates that are defined by average scores on all skill items as well as candidates who are highly skilled in some items but score poorly on others. These two types of applicants are then called generalists and specialists. The risky component of the decision problem is found in the vagueness of the job description. In this sense, it is assumed that the organisation does not yet know which particular skills will be of greater relevance in upcoming projects, so that the participants as heads of department encounter the threat of building a team that is highly qualified in the wrong skills. Concerning the bracketing behaviour of subjects, one may argue that a narrow bracketer will choose to bracket each vacancy as an individual choice set and fails to optimise the skill distribution amongst the resulting team. In this sense, employing narrow bracketing should cause the DM to select generalists, which as a team yield a suboptimal outcome. On the contrary, a broad bracketing individual is expected to recognize the potential of specialists and builds the team in such a manner that its members compensate for the specific weaknesses of each other, thereby generating a higher utility as a team. Hence, this setting allows broad bracketers to exhibit variety seeking with respect to the skill distribution amongst the members of a team.

Having said that, the experimental design also accounts for the diversification bias, which states that subjects who are confronted with an aggregate presentation as broad bracketing manipulation tend to select more variety than they would find to be optimal in retrospect. In this sense, the value maximising team composition in this experiment is not achieved by the selection of three specialists, but consists of the most specialised applicant combined with two rather generalised candidates. This setting separates broadly bracketing respondents into two groups: (1) those who lose focus of their task of maximising value by exclusively concentrating on variety and thus, falling prey to the diversification bias, and (2) those who maintain an overall evaluation even when they are presented with an invitation to overdiversify their portfolio. Hence, this enables us to explore the proposition that broad bracketers are inclined to excessively focus on variety instead of value maximisation and, as a result diversify their choices to a higher degree than would be beneficial.

The first task presented to respondents is a control question (table1-3) in order to assess whether the participants correctly understood the instructions concerning this part of the experiment, in particular the explanation of how to evaluate a given candidate. Note that during the subsequent two tasks, respondents are enabled to repeat the instructions at any time. The control task then requests participants to compare two candidates in a simplified setting by deciding whether the second candidate is better, equally good, or worse than the first one. This procedure is repeated for three pairs of candidates with varying difficulty. However, since the candidates are already presented with their PF scores instead of four distinct skill items and simple values are chosen, none of the three pairwise comparisons may be considered as complex.

Table 1: Candidate Table Control Question 1

	Candidate 1	Candidate 2
Sum Skill Scores	70	80
Team Compatibility	65	65

Candidate 1 is … Candidate 2:

- Better than
- Worse than
- Equally good as

Sub-question two and three of the control question are described by the following two tables:

Table 2: Candidate Table Control Question 2

	Candidate 1	Candidate 2
Sum Skill Scores	100	50
Team Compatibility	50	100

Table 3: Candidate Table Control Question 3

	Candidate 1	Candidate 2
Sum Skill Scores	150	100
Team Compatibility	50	80

The remaining two tasks of the personnel selection part are embedded in two treatments, which are played in a between-subjects design. Participants are randomly allocated to one of the treatments. Each treatment then comprises a specific presentation form and initial state of a team building task followed by a team selection task. The latter task is the same for both treatments and thus, is of a within-subjects design.

Both presentation forms display the same nine candidates to the participant. However, at this point we have to make the following remark. Contrary to the lab-experiment approach that was initially intended, the pandemic of the Corona virus required the transition to an online experiment. For the reason that the control of the lab setting regarding the communication of participants among each other is lost, an additional group for each treatment was introduced. Respondents are randomly assigned to either group A or group B of one of the treatments. This action provides a countermeasure by reducing the chances that two participants of the same treatment and group communicate with each other. In order to retain the basic comparability between the two treatments, group A and group B of each treatment are essentially the same but present a different order of slightly adjusted candidates.

(1) <u>Treatment 1: Segregated Presentation</u>

The segregated presentation mode provides the DM with a pre-allocation of the nine candidates to the three vacancies. As a result, subjects need to make three choices between three candidates each, instead of considering all available candidates for each open position. Thus, the pre-allocation and presentation as sequential choices reduces the complexity of the task along with the cognitive demand that it places on an individual. This simplification of the task should then enable more participants to achieve an enhanced solution quality. However, the segregated presentation of the pre-allocated applicants also induces a narrow perspective on decision-making with respect to the three vacancies. The underlying effect of this configuration would then tempt participants to bracket each position narrowly without considering the possible effects on the skill composition and emerging team utility. In consequence, the potential benefits of variety seeking are less recognisable to participants.

Position 1:

Table 4: Candidate Table Position 1, Group A

Skill	Person 1	Person 2	Person 3
Analytics	40	100	95
IT	70	55	70
Languages	85	20	75
Negotiation	55	45	60
Team Compatibility	20	45	35

Please choose your candidate for position 1 out of the candidates 1 – 3: ___

Position 2:

Table 5: Candidate Table Position 2, Group A

Skill	Person 4	Person 5	Person 6
Analytics	65	45	20
IT	10	55	90
Languages	90	50	40
Negotiation	15	40	50
Team Compatibility	45	60	40

Please choose your candidate for position 2 out of the candidates 4 – 6: ___

Position 3:

Table 6: Candidate Table Position 3, Group A

Skill	Person 7	Person 8	Person 9
Analytics	70	10	85
IT	30	80	50
Languages	65	30	10
Negotiation	55	40	55
Team Compatibility	25	75	35

Please choose your candidate for position 3 out of the candidates 7 – 8: ___

Note that, due to having to rearrange the candidates in order to create group B of both treatments, the score of the skill item negotiation of candidate 7 needed to be adapted from 55 to 75. This results in a change of the candidate-utility from 55 to 60. The adjustment is necessary in order to maintain the difficulty encountered in group A for the evaluation of the new pre-allocation in group B. Otherwise, the pre-allocation for position 1 of group B (candidates 3, 4, and 7) would portray a clear dominance of candidate 3, being that this candidate is superior to the others in three out of four skills. When adjusting candidate 7, this candidate is superior in one skill so that the dominance of candidate 3 is less obvious. However, the change in utility does not induce a change of the optimal choice. Subsequently, the rearrangement of the pre-allocation is performed in such a manner that the same three candidates as in group A yield the optimal team constellation. With respect to the presentation, the candidates in group B are re-named to reflect an order from number one to nine. The resulting candidate tables for both treatments of group B can be found in Appendix A.

This segregated presentation of the portfolio building task is succeeded by the portfolio selection task, which is the same for both treatments. The task establishes a third complexity dimension by confronting respondents with three potential team constellations from which the DM is required to select one team. In addition, the achieved skill scores of each applicant are already given as a sum. Therefore, this task significantly reduces the complexity of the decision environment compared to both configurations of the preceding task by two means. First, the presentation as portfolio selection instead of a portfolio building task decreases the number of necessary choice comparisons to three, in contrast to 27 and 84 in the segregate and aggregate modes. Furthermore, the presentation of the sum for the skill scores minimises the computational capacity required to efficiently assess the proposed teams.

One may argue the task does not allow for bracketing behaviour to exist, as it is reduced to a single decision instead of multiple ones. However, while this assumption is true, the purpose of the task is comparable to the combined lottery task in the first part of the experiment. It serves to test whether respondents are capable of recognizing the dominant solution when the complexity of the task is sufficiently reduced. Additionally, the presentation of pre-designed teams introduces the role of decision support systems with the respondent as final instance to the experiment. Hence, the underlying task also provides information on the influence of decision support on the quality of the solution compared to two settings of different complexity where the DM needs to construct his portfolio without major support.

Table 7: Candidate Table Portfolio Selection, Group A

Team 1	Person A	Person B	Person C
Sum Skill Scores	190	200	180
Team Compatibility	50	60	55

Team 2	Person D	Person E	Person F
Sum Skill Scores	240	165	225
Team Compatibility	30	80	40

Team 3	Person G	Person H	Person I
Sum Skill Scores	220	200	280
Team Compatibility	40	45	35

Please choose your preferred team to hire: ___

For the creation of group B, the values were adjusted in such a manner that the team utilities of both groups are almost identical. As a result, the best team in group A generates a utility of 314 while the best team in group B has a utility of 312. The second-best teams are defined by a utility of 294 (A) and 291 (B), whereas the worst teams offer a utility of 276 (A) and 273 (B). Hence, the difficulty of comparing the three teams remains equal in both groups.

(2) Treatment 2: Aggregated Presentation

The second configuration of the team building task displays all nine candidates at once and without the support of a pre-allocation to vacancies. Consequently, a DM needs to consider all candidates simultaneously. This design yields 84 possible team constellations, as compared to 27 possible teams in the segregated mode of treatment 1. This task entails a considerably higher degree of complexity so that the required cognitive effort for accurately assessing the decision problem may exceed the cognitive capacity of a DM and causes him to resort to heuristic decision-making. However, the aggregate presentation encourages the consideration of the three vacancies as a team instead of unique decisions as well as placing an emphasis on the possibility to seek variety in skills by selecting specialists. Remember that such an invitation to seek variety also emphasises a possible tendency of broad bracketers to behave in compliance with the diversification bias and thus, losing focus of the resulting team utility by overdiversifying their solution.

Table 8: Candidate Table Aggregated, Group A

Skill	Person 1	Person 2	Person 3
Analytics	40	100	95
IT	70	55	70
Languages	85	20	75
Negotiation	55	45	60
Team Compatibility	20	45	35
	Person 4	**Person 5**	**Person 6**
Analytics	65	45	20
IT	10	55	90
Languages	90	50	40
Negotiation	15	40	50
Team Compatibility	45	60	40
	Person 7	**Person 8**	**Person 9**
Analytics	70	10	85
IT	30	80	50
Languages	65	30	10
Negotiation	55	40	55
Team Compatibility	25	75	35

Please choose your first candidate: ___

Please choose your second candidate: ___

Please choose your third candidate: ___

Again, the portfolio building task is followed by the portfolio selection task as it was presented above.

Taken together, the segregated presentation of treatment 1 facilitates superior decision-making for all respondents who are assigned to this decision mode, whereas the aggregated presentation in treatment 2 is hypothesised to only be processible by individuals of higher cognitive ability. The two configuration modes of the portfolio building task then give rise to several task-specific hypotheses stated below.

H8a: Due to the reduced complexity in the segregated presentation mode of treatment 1, more individuals find the optimizing team constellation by bracketing broadly compared to the aggregate presentation of treatment 2.

H8b: Due to reduced complexity, the average achieved team utility is higher in the segregated presentation form of treatment 1 than it is in the aggregated one implemented in treatment 2.

H13: On average, the perceived complexity in treatment 1 is lower in comparison to the aggregate configuration of treatment 2.

H13a: In the more complex presentation mode of treatment 2, the average numeracy score of broad bracketers is higher compared to the numeracy score of broad bracketers in the less complex treatment 1.

H13b: In the more complex presentation mode of treatment 2, the average rationality score of broad bracketers is higher compared to the rationality score of broad bracketers in the less complex treatment 1.

H14: On average, the variety in skill distribution of built teams is greater in the aggregated presentation mode of treatment 2 compared to the segregated form in treatment 1.

With respect to the portfolio selection task, two additional hypotheses are formulated.

H15: Due to lower complexity, a greater fraction of individuals achieves an optimal solution in the portfolio selection task compared to the aggregate presentation form of treatment 2.

H16: Due to minimal complexity, a greater fraction of individuals achieves an optimal solution in the portfolio selection task compared to the segregate configuration of treatment 1.

4.1.3 Complementing Tests and Task Sequencing

The two main parts of the experiment are complemented by the integration of three commonly used instruments, which aim at measuring the psychographic attributes of the respondents. The flow chart depicted in figure 5 gives the sequencing of all main tasks as well as the complementary testing procedures.

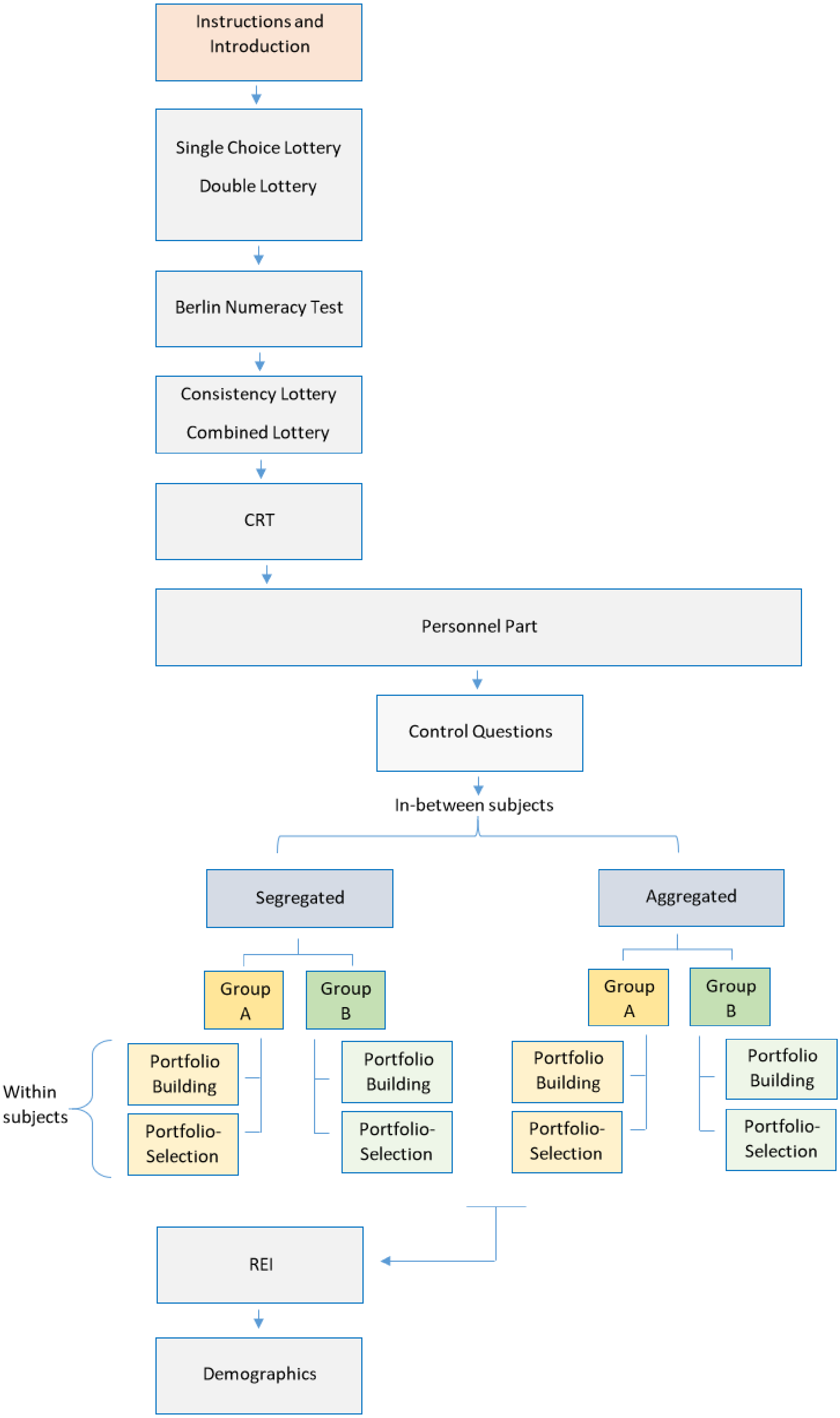
For the reason that the experiment needed to be transferred from a lab-setting to an online experiment, the instruction page is adapted to explicitly state a set of rules for the experiment. Participants are informed that they are not allowed to use a calculator, paper and pen, practise group work, any form of communication or other influence of a third party, as well as the use of any search engine.

Having started the experiment, the first task set belongs to the lottery part of the design and includes the single choice lottery, which elicits the basic risk attitude of an individual, followed by the first double lottery task. However, the lottery part is interrupted by the implementation of the first psychometric test. The Berlin Numeracy Test consists of four items and measures an individual’s numeracy attribute, which is introduced in chapter 3 and refers to the ability to properly interpret information about risk as well as adapting one’s behaviour accordingly (Cokely et al., 2012). Furthermore, the test is specifically developed for assessing the numeracy of educated respondents and thus, represents a suitable questionnaire for a sample of university students. Beyond evaluating the numeracy of an individual, the placement of the Berlin Numeracy Test in the task sequence serves the purpose of passing some time between the double lottery and the consistency lottery.

Subsequent to completing the questionnaire, the lottery part advances with its remaining two tasks before leading to the Cognitive Reflection Test (CRT) introduced by Frederick (2005). This test consists of three items which seem simple at first but require respondents to suppress their first intuitive answer in order to be able to reply correctly. In this manner, the CRT measures the tendency of a DM to be patient and reflect on his first impulses for behaviour. Frederick (2005) finds that individuals who score highly on the CRT are more likely to accept a gamble in the domain of gains and avoid risk seeking in the domain of losses, whereas low CRT individuals behave according to prospect theoretic predictions. Therefore, the CRT provides the measurement for the impulsiveness psychographic of an individual.

After the completion of the CRT items, respondents advance to the personnel task as focal part of the experiment. Participants first reach a separate introduction page where the basic setting and candidate evaluation procedure is explained before they proceed to the control questions. The flow chart then displays the 2x2 factorial design of the treatments with their two presentation configurations, groups A and B as between-subjects architecture, as well as the two consecutive task types in a within-subjects arrangement. The tasks of the personnel selection part are followed by the last psychometric questionnaire of the experiment. The Rational Experiential Inventory (REI) is a 42-item questionnaire which evaluates an individual’s scores on the attributes of rationality, experientiality, intuitiveness, emotionality, and imagination (Epstein et al., 1996). Thereby, the REI assesses to which extent a DM operates in the experiential or rational system of cognitive processing. The experiment concludes by asking participants to provide information on some demographic variables.

Figure 5: Flow Chart Experimental Design



4.2 Implementation and Data

Having outlined the tasks that are designed for the experiment, the architecture of the treatments, as well as their final sequencing, this subchapter concerns the implementation and execution with the sampled respondents.

The experiment is programmed in the form of an oTree project (Chen et al., 2016), using the programming language python and is deployed over a Heroku server. The experiment was conducted with Bachelor students from a university in Austria who were granted performance-contingent bonus course credits for their participation. Since the experiment could not be conducted in the planned lab setting, subjects received access via a link provided on the learning platforms of the relevant courses and their individual identifier acting as a password so that no third-party individual could start the experiment. By means of the learning platform, students were also provided with an instructions sheet (Appendix B) stating the rules of the experiment as well as an explanation of how to correctly enter the experiment using their identifier. Access to the experiment was open for one week, starting from 08.06.2020 and ending on 15.06.2020 at midnight.

During this period of time, 387 data entries were collected. From this total number, nine entries were excluded during the process of cleaning the data frame. Two of those entries corresponded to testing accounts while seven entries were eliminated for the reason that the respondents did not reach the final page of the experiment. None of the participants needed to be excluded for entering nonsensical answers in the free text entry fields of the demographic questionnaire and no participant was eliminated as a consequence of merely clicking through the experiment in order to earn the extra class credits. This was examined by checking the time stamps as well as answer patterns of the respondents. Furthermore, no data entry was excluded for failing the control questions.

During the first inspection of the data, it became obvious that a remarkable fraction of respondents (52.91%) did not pass all three control questions, which is a circumstance that calls for further investigation. When examining the decomposed data, we find that 80.69% of respondents pass at least two out of three control questions and it becomes visible that control question three represented the major obstacle, since it is answered incorrectly by 179 (47.35%) respondents. In comparison, the first and second control questions are missed by 26 (6.88%) and 70 (18.52%) subjects respectively. The high failure rate in control question three may result from the fact that it requires the highest effort, being that the answer is not as obvious compared to question 1 and question 2.

With respect to the time that participants spent on the task, it is noticeable that 75% of participants needed 2.5 minutes or less to complete all three questions with the median being 1.7 minutes. Hence, one can argue that subjects tended to not spend enough time on the task and in consequence failed the non-intuitive third question.

In addition, the online setting of the experiment does not allow for the thorough explanation of the personnel task and the answering of possible questions, as it could have been done in a lab-setting. Only two individuals (0.53%) failed all three control questions, which would usually lead to their elimination from the data frame. However, it is decided to take a closer look at the two data entries for the following three general circumstances: (1) a majority of respondents is found to struggle with the control questions; (2) respondents display a general tendency to not spend an extensive amount of time on the task and (3) are not able to ask questions in case of not understanding the proper procedure in the personnel task. On closer inspection, we find that both individuals perform well during the main tasks of the personnel part, indicating that they understand the evaluation process stated in the instructions for the task. Furthermore, both individuals are between the first and third quartile of the time spent on the task (1.3 and 1.7 minutes) and therefore take little time to complete the tasks. In consequence of these general tendencies, it is decided to exclude none of the participants for their behaviour in the control questions. Out of the remaining 378 valid entries, treatment 1 is played by 93 respondents in group A and 95 in group B, while treatment 2 has 95 remaining participants in each group.

The resulting sample possesses the following characteristics. All graphical presentations are created using the R package ggplot2 (Wickham, 2016). Figure 6 displays the gender composition of the sample, where 60.05% of respondents are female, 39.68% are male, and 1 subject identifies as diverse. This distribution in the sample reflects the one found amongst all enrolled students of the underlying university, where 62.49% of students in the winter semester of 2019 were female. All groups of the treatments (table 9), except group B of treatment 2, display almost equal gender distributions when compared to each other. In compliance with this observation, the deviation across the four groups is found to be not significant (p-value = 0.7861) by means of Pearson's Chi-square test.

Figure 6: Gender Distribution

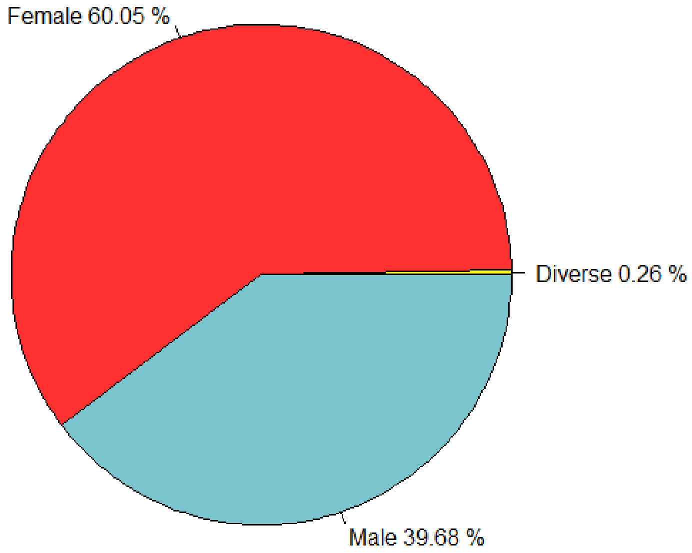


Table 9: Gender Distribution Treatments

Gender in %	Segregate A	Segregate B	Aggregate A	Aggregate B
Male	37.63 %	37.89 %	38.95 %	44.21 %
Female	61.29 %	62.1%	61.05 %	55.79 %
Diverse	1.08%	0 %	0 %	0 %

The variable age is illustrated by a histogram in figure 7 and the corresponding boxplot in figure 8. The age distribution is defined by a mean age of 22.99 years with a minimum of 18 and a maximum of 43 years. However, 75% of respondents are at most 24 years old and the distributions across all treatment groups do not reveal considerable deviations from each other. The confirmation is received by first conducting a Shapiro-Wilk test to check for a normal distribution (p-value < 2.2e-16) in addition to checking the distributions with histograms. This is followed by an F-test in order to test for significant differences (p-value = 0.642) in connection with a pairwise comparison using t-tests revealing no significances.

Figure 7: Histogram Age

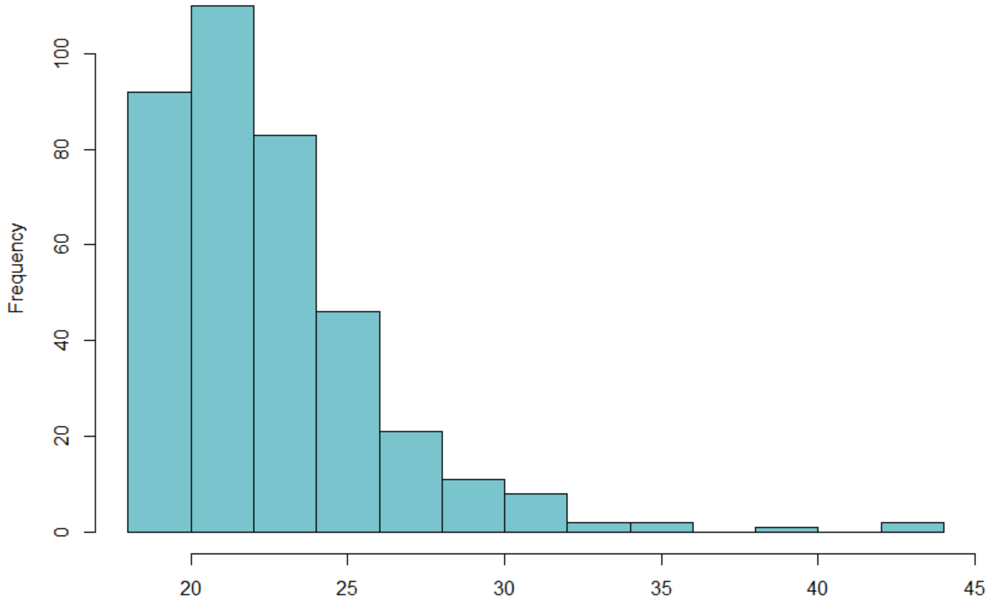
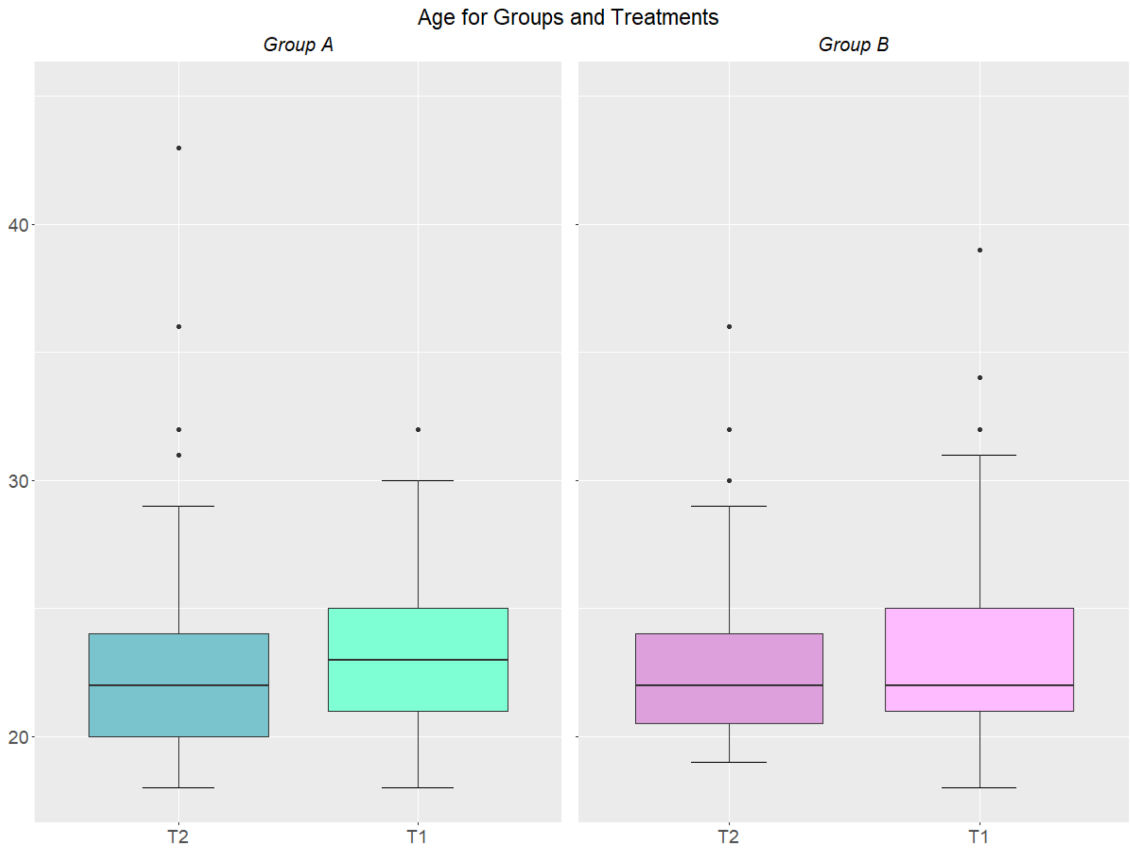


Figure 8: Grouped Boxplots Age



With respect to the main study field of participants, figure 9 presents that the majority of 70.37% are students of economic studies, which contains the study fields of (international) business administration and economics. Communication studies, comprising all studies of languages, as well as students of MINT studies (Mathematics, Informatics, Natural sciences and Technical) represent the succeeding two distinguishable fractions with 11.9% and 7.94% respectively. The remaining study fields are grouped under the category others and compose 9.79% of the sample. Concerning the distribution across treatments (table 10), minor deviations can be found. In particular, group B of treatment 1 contains noticeable less students of MINT studies (2.11%) but a higher percentage of the category for other study fields (14.74%). On the other hand, group A of the same treatment displays a larger share of MINT students (12.90%) as well as students in the communication fields (16.13%) than the other groups. The application of Pearson's Chi-squared test reports a significance above the 5% degree (p-value = 0.08603), which results from the differences described above.

Figure 9: Distribution Study Field

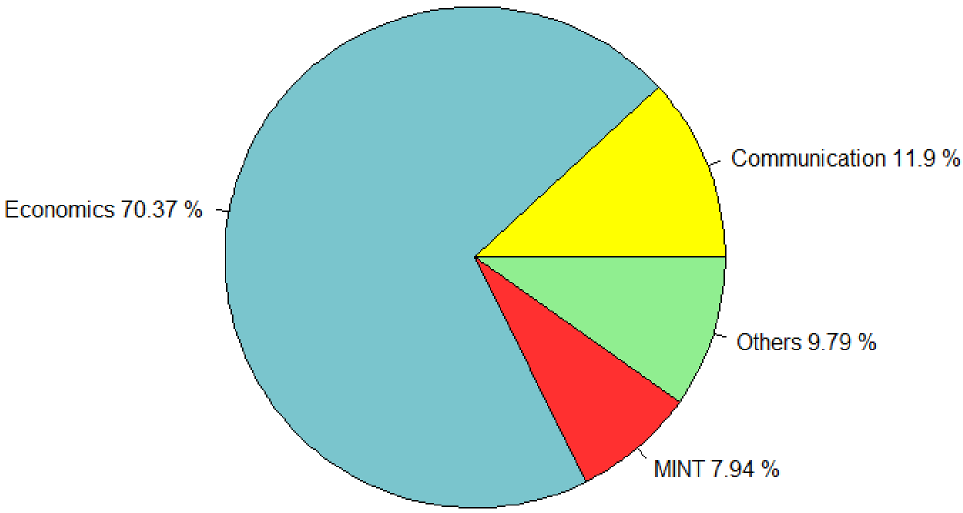


Table 10: Distribution Study Field Treatments

Studyfield in %	Segregate A	Segregate B	Aggregate A	Aggregate B
Economics	65.59%	71.58%	75.79%	68.42%
Communication	16.13%	11.58%	9.47%	10.53%
MINT	12.9%	2.11%	7.37%	9.47%
Others	5.38%	14.74%	7.37%	11.58%

Regarding the mother language of participants, depicted in figure 10, German represents the largest share of respondents (50.26%). The remaining languages are categorised into either Slavic languages (29.89%) or others (19.84%), for the reason that no other single language distinctly emerged from the data. No suspicious deviations in the distribution across groups can be detected, which is confirmed by Pearson's Chi-squared test (p-value

= 0.9431). The corresponding countries of origin of participants are sorted into six main categories and presented in figure 11. The largest share of the sample is Austrian (42.59%), followed by other EU-countries (18.78%), and European non-EU countries (16.67%). Furthermore, Germany (14.29%) and Turkey (3.97%) emerge as distinct countries while the remaining share of the sample is represented by non-European respondents (3.70%). As for the mother language, no significant deviations across treatments are revealed by Pearson's Chi-squared test (p-value = 0.8739).

Figure 10: Distribution Mother Language

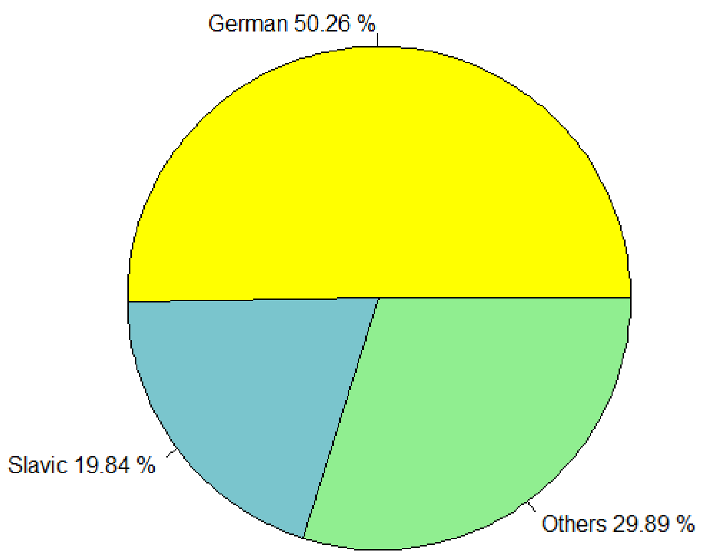
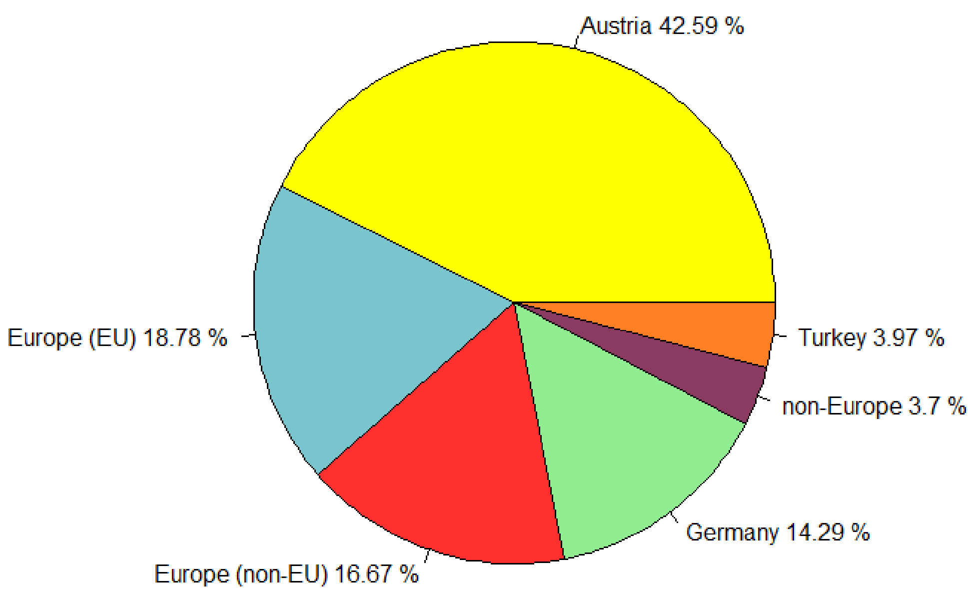
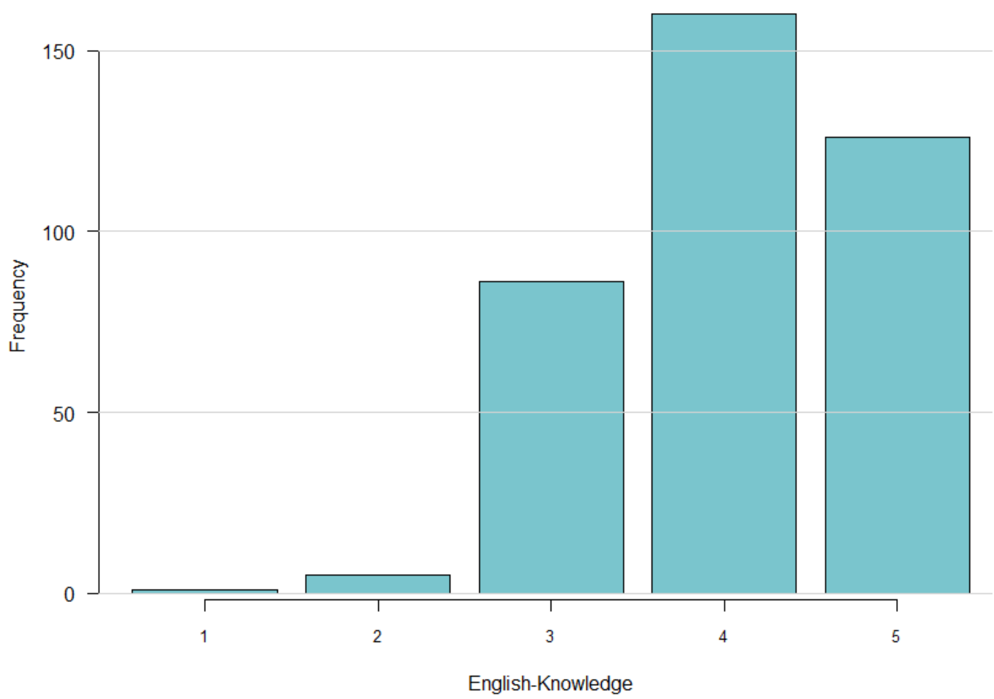


Figure 11: Distribution Country



In addition to these commonly retrieved demographics, participants are requested to estimate their knowledge of the English language on a five-point Likert scale (figure 12). This is implemented for the reason that the courses of the bachelor's program in which the experiment was conducted are primarily taught in German, whereas the experiment is in English language. However, the instructions and tasks are phrased as simple as possible, so that no student of a bachelor's programme should be expected to experience any difficulties with the formulation. Furthermore, all potential participants were informed that the study takes place in English in advance. Only 25% of respondents select an English level below a score of four and on closer inspection, only 5 individuals selected the score of two and one subject estimated his languages skills to correspond to the score 1. The application of Pearson's Chi-squared test yields no significant differences across groups (p-value = 0.2233).

Figure 12: Distribution English Knowledge



5 Results

After examining the research framework as well as the corresponding design and implementation procedure of the experiment, this chapter presents the gathered data and the consecutive analysis.

The first three subchapters will outline the descriptive results for variables that address the experiment in its entirety as well as those that are specific to the two main parts of the experiment. Subchapter 5.4 then gives a combined perspective of the bracketing behaviour that was observed in the two main parts. The consecutive subchapter continues with the statistical testing of the introduced hypotheses and presents the resulting models.

5.1 Descriptive Results: Holistic Overview

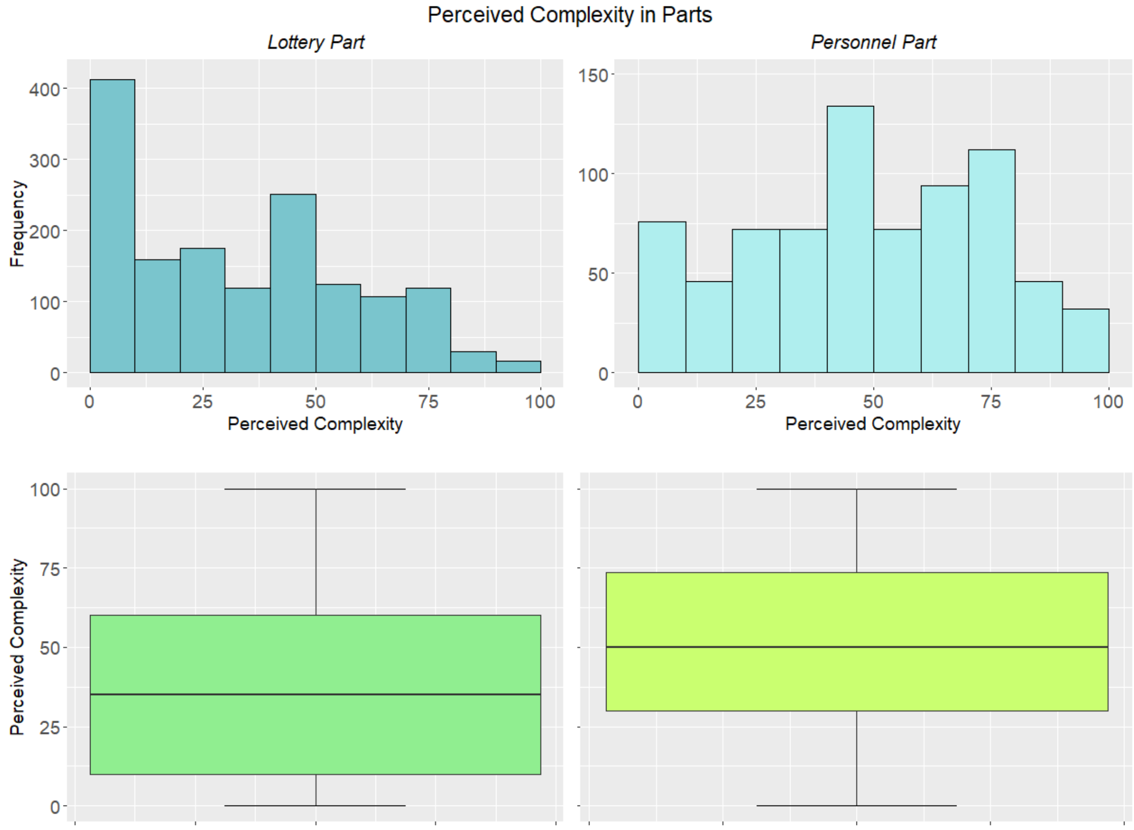
Before turning the attention to the results of the two main parts of the experiment, this subchapter introduces the reader to holistic results, which should be presented together instead of being assigned to a specific part of the experiment. This includes the overall perceived complexity that was observed in the experiment as well as data on the psychometric measurements of the participants who compose the sample. All tables containing the descriptive statistics for this subchapter can be found in Appendix C.

(1) <u>Perceived Complexity</u>

As already mentioned whilst reviewing the experimental design, respondents were asked to report the difficulty for each task of the lottery and personnel parts of the experiment. The resulting variable ranges from 0% to 100% and represents the perceived complexity, which, according to the research framework, is a moderator between the objective task complexity and the resulting bracketing behaviour.

Figure 13 provides a comparison of the overall perceived complexity of the four lottery tasks in the lottery part and the perceived complexity in the personnel part with its two tasks, two treatments, and the two groups A and B. When comparing the two histograms, it becomes visible that the lottery part is of lower perceived complexity since most observations are located in the lower range of the complexity scale. In contrary, the perceived complexity of the personnel part takes a shape that resembles a normal distribution with its peak in the middle range of the perceived complexity scale. The corresponding boxplots further offer a clear impression of the differences in perceived

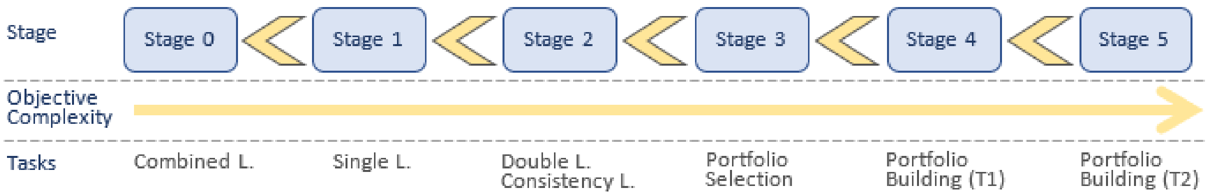
Figure 13: Perceived Complexity in Main Parts



complexity encountered in the two parts of the experiment. The median in the lottery part lies at a perceived complexity of 35%, whereas the personnel part has a median of 50%. In fact, when inspecting the quartiles, it becomes apparent that the first quartile of the personnel part (30%) is almost at the same level as the median of the lottery part. Hence, the lower 25% of respondents in the personnel part perceive almost the same degree of complexity as half of them did in the lottery part. The standard deviations in both parts are almost equal at 26.75 for the lottery part and 25.95 for the personnel part. Furthermore, both parts cover the full range from 0% to 100% perceived complexity.

The comparison of the two parts allows for a first impression but in order to get a more detailed view on the perceived complexity, the data is grouped according to five stages of objective task complexity. These objective complexity stages are argued to be the basis for the perceived complexity and express an objective order of difficulty for the tasks in the experiment. This order is a result of assessing each task with respect to its length, content, requirements on processing, as well as decision support by means of information presentation. The resulting tasks to stages assignment are displayed by figure 14. The stages zero to two refer to the lottery part, whereas the stages three to five address the tasks of the personnel part. The combined lottery task is rated as the least complex for the

Figure 14: Objective Complexity Stages

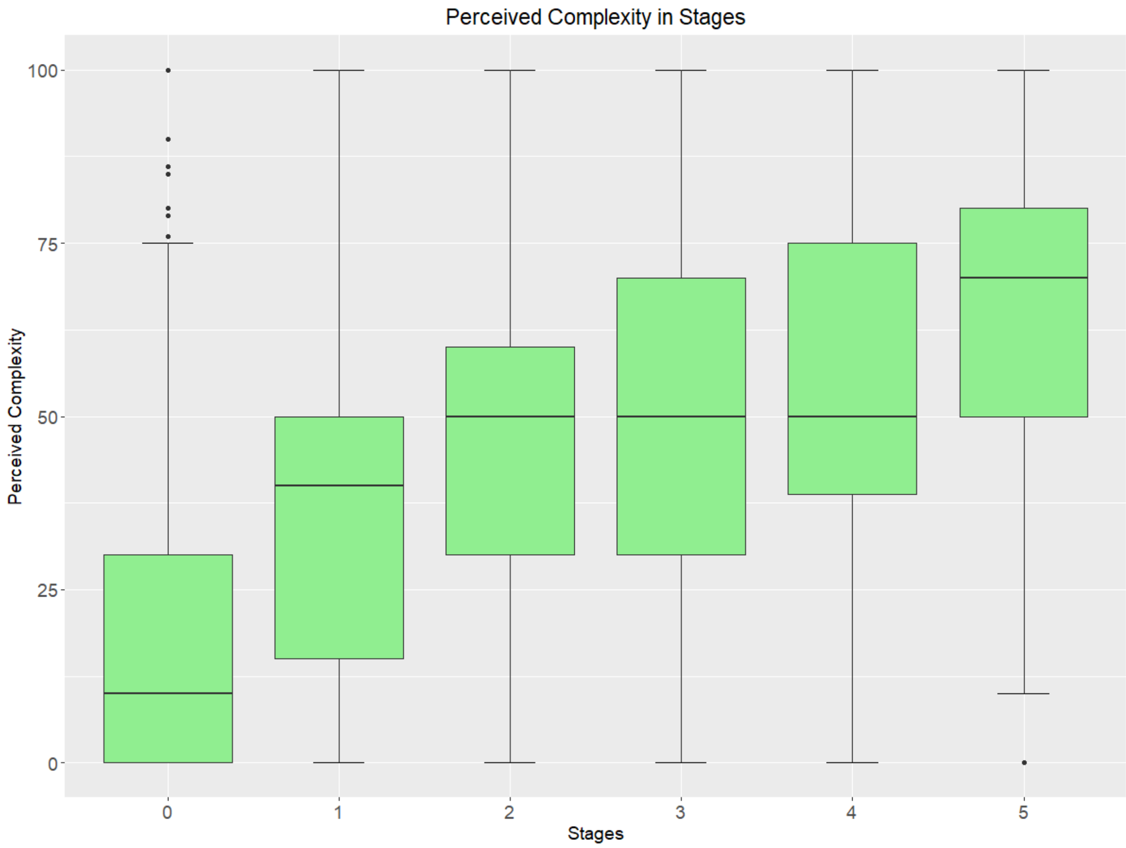


reason that it is a lottery decision where the two decision alternatives are stated in such a manner that strongly supports the direct comparison of the options. The single lottery task in stage 1 also asks for a decision between two alternatives but this decision is between a sure gain and a gamble. As a result, it is not a simple comparison of two similar options but has a more pronounced focus on the risk attitude of respondents. Hence, this task is evaluated to be more complex than the combined lottery. Stage 2 incorporates the double lottery task as well as the consistency lottery, since there are only minor differences in the prospects. Both tasks are double lotteries with two decisions that each comprise two alternatives. Hence, the decision process requires the integration of the alternatives of both decisions instead of evaluating each in an isolated manner. Remember that these are also the first tasks where bracketing behaviour is possible.

The following stage 3 then refers to the personnel part of the experiment and consists of the portfolio selection task. The evaluation of the three teams is considered to be more demanding than the decision between stated prospects as they are encountered in the first part of the experiment. However, the presentation as already formed teams facilitates the evaluation so that the task is less complex than both treatments of the portfolio building task. Treatment 1 of the portfolio building task then represents stage 4 of the objective complexity for the reason that the evaluation of the separate candidates requires more processing than the evaluation of teams but the pre-allocation of three candidates to each vacancy constitutes a supportive measure. This leaves the second treatment of the portfolio building task for the highest complexity stage. Here the simultaneous evaluation of the nine candidates without decision support places the highest demand on the DM. In this sense, the pattern that emerges from the perceived complexity, where the personnel part is experienced as more complex than the lottery part, also reflects the order of objective complexities.

After reviewing the development of the complexity stages, figure 15 illustrates the perceived complexity grouped with respect to the stages variable. The general trend of an increasing perceived complexity is certainly noticeable. The most pronounced jump in

Figure 15: Perceived Complexity grouped by Objective Complexity Stages



perceived complexity occurs between stage 0 and stage 1, where the mean value increases from 18.48 to 37.27. The means of the stages 2, 3, and 4 follow closer together (45.45 | 47.82 | 52.53) whilst the increase from stage 4 to stage 5 (62.51) is again of a more distinctive character. The proximity of stage 2, 3, and 4 is also visible when inspecting their median values, which are 50 for all three stages. However, the increasing development of the perceived complexity becomes noticeable again when considering the quartiles of the three stages. Although the Q1 for stage 2 (30) and stage 3 (30) is identical, stage 4 shows an increase to 38.75. The third quartile is increasing across all three stages, with the values of 60, 70, and 75 respectively. Regarding the standard deviations, one may note that all stages report similar values that range between 23.96 in stage 2 to 25.71 in stage 4. In Chapter 5.5.1 the differences in perceived complexity between the stages will be tested for statistical significance.

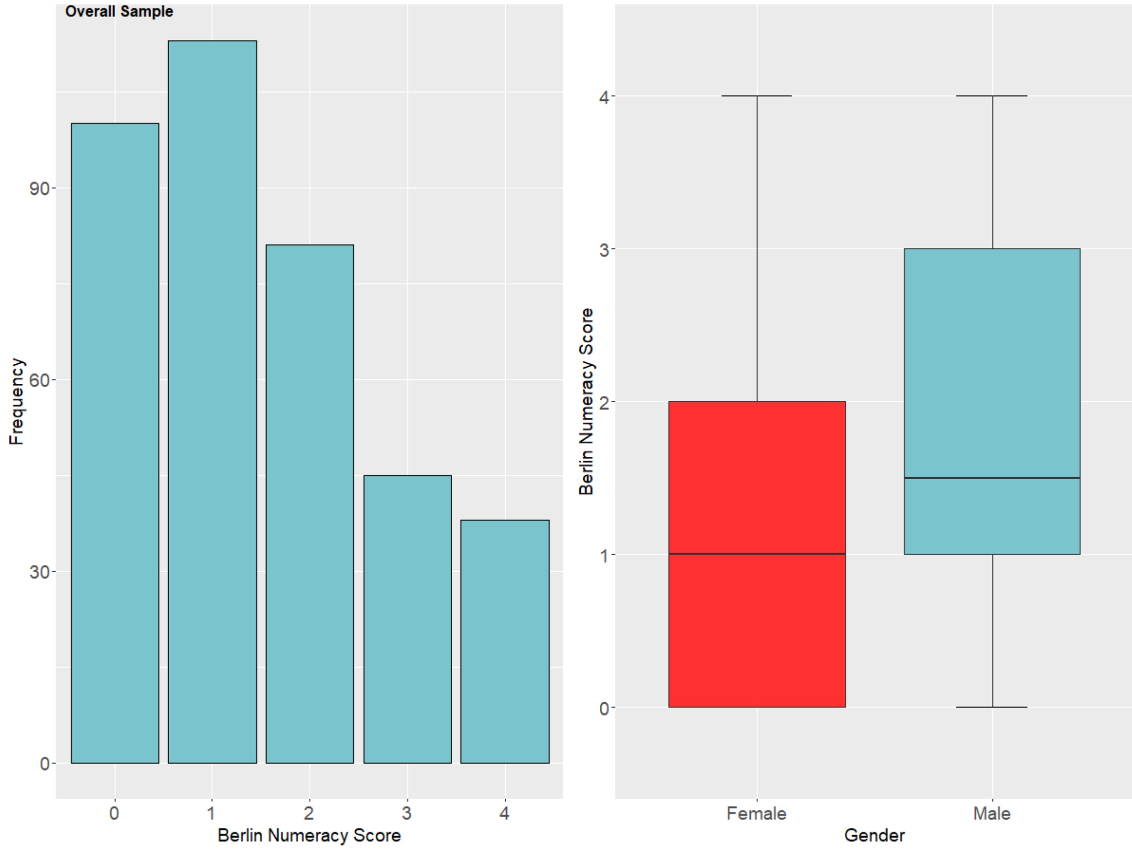
All five stages contain data entries that cover the full perceived complexity scale from 0% to 100%. One has to note that stage 0 and stage 5 span the full range as a result of distinct outliers. The outliers in stage 0 are to be found upwards from a perceived complexity of 75% whereas the outlier in stage 5 is located at a perceived complexity of 0%. The circumstance that all stages report the 0% and 100% marks may be attributed to a suboptimal formulation of the request that asks respondents to enter the degree of difficulty after each task. This potential problem as well as a proposal of an improved design are discussed in more detail in the limitations of this study (chapter 6).

(2) Psychometrics

Alongside the objective complexity as a basis, the research framework (chapter 3) also emphasises the psychographic characteristics of the respondents as influencing factors for the perceived complexity as well as a baseline bracketing tendency. The Berlin Numeracy Test, Cognitive Reflection Test, and the Rational-Experiential Inventory were then selected for the purpose of measuring these psychometrics. The subsequent paragraphs are dedicated to presenting the results of these surveys by illustrating the distribution in the form of either bar charts or histograms for the overall sample. These charts are accompanied by boxplots that are grouped by gender. The grouping is done in order to visualise possible gender differences, which literature generally suggests to be present in psychometric characteristics as well as bracketing behaviour (Rabin and Epstein, 1999; Epstein, 2003; Frederick, 2005; Rabin and Weizsäcker, 2009).

The Berlin Numeracy Test measures an individual’s capability to assess and act on information regarding risks and probabilities. The variable Berlin Numeracy Score is created to report the number of correct answers for the four items of the test. The bar chart (figure 16) for the overall sample reveals that a majority of 213 (56.50%) of respondents gives at most one correct answer. In comparison, only 45 (11.94%) give three correct answers and 38 (10.08%) answer all 4 items correctly. This distribution indicates that the capability to interpret risks is not extensively developed amongst the sampled participants. The corresponding boxplot shows a difference in gender, where 75% of female respondents achieve a score of at most two correct questions while 50% of men achieve a score which lies at two or higher. This observation might result in men achieving better decisions or perceive a task as less complex. Especially the lottery part of the experiment might be affected by this psychometric, considering that these tasks directly require the assessment of risky prospects.

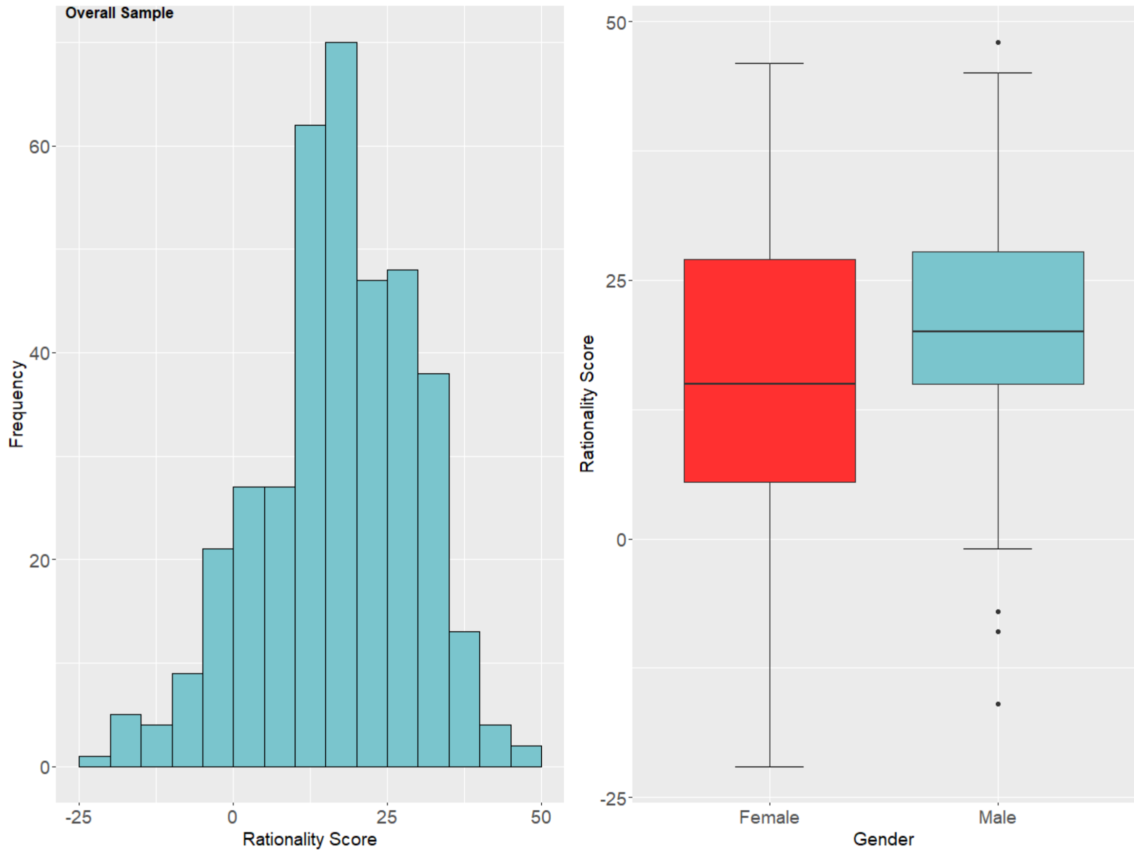
Figure 16: Berlin Numeracy Score - Bar Chart and Boxplot



The following five graph-pairs describe the data gathered by the REI questionnaire, which depicts an individual's tendency to operate in the rational or experiential system. At this point, it shall be noted again that the two systems are not mutually exclusive but operate together. However, the extent to which either system is applied may vary.

The rationality score (figure 17) shows a distribution that is close to a normal distribution but slightly skewed to the left (-0.39). The mean score of respondents is 17.43 while the median lies at 18. Hence, the majority of the sample score positively on rationality, with only a few achieving a high score and a longer tail towards lower and negative values. Again, the boxplot presents a notable gender difference. Women cover an extraordinary wide range of scores whereas the scores of male participants are, apart from three outliers, focused in the positive value range. The first quartile of men lies at a score of 15, which is the same level as the median of female participants. This means that 75% of men score higher in rationality than 50% of women. In contrary, the third quartile of both groups are almost identical at 27 for women and 27.75 for men. Furthermore, the maximum values achieved by female participants are with a score of 46 almost equal to the maximum male

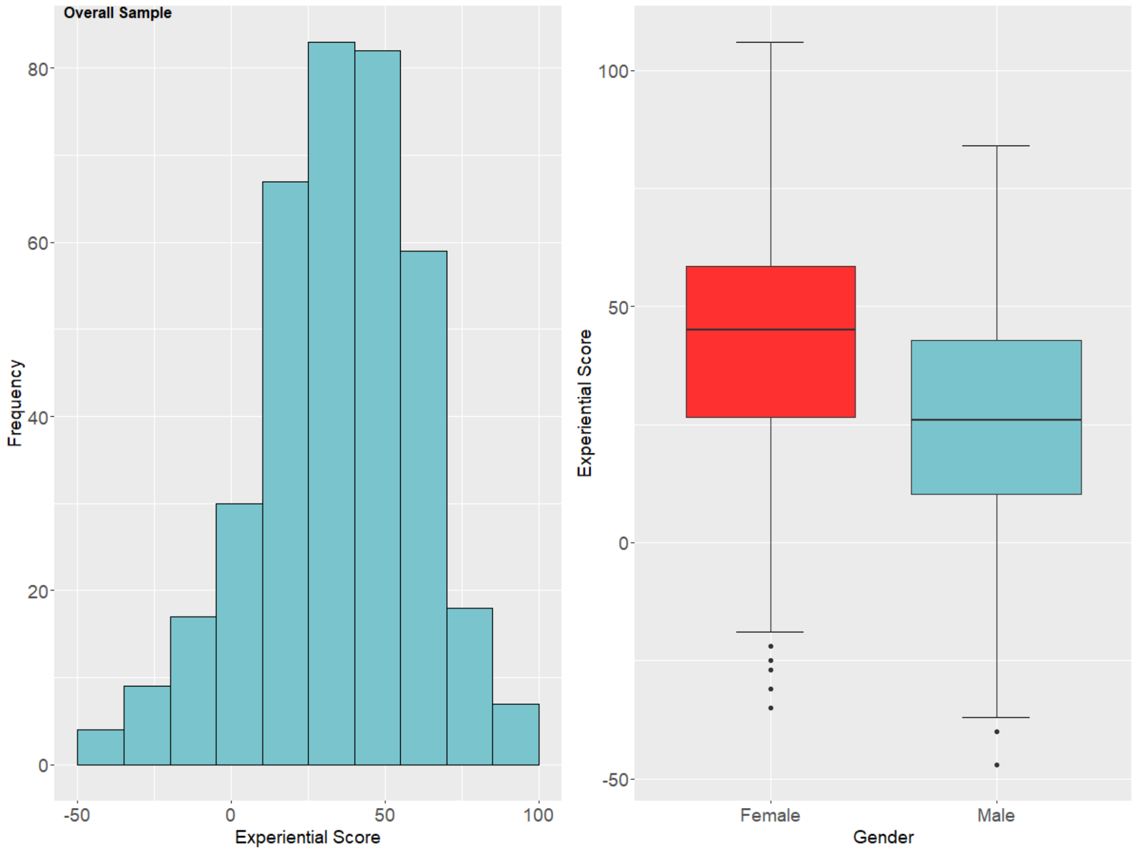
Figure 17: Rationality Score - Histogram and Boxplot



score of 48. The observed results are consistent with literature, where it is noted that men are more inclined to operate by means of the rational system (Pacini and Epstein, 1999; Epstein, 2003; Frederick, 2005).

When inspecting the distribution of the measured experiential scores (figure 18), a similar picture emerges as it is found for the rationality score. The histogram for the overall sample shows a close to normal distribution which is slightly left-skewed (-0.3), so that the distribution is heavier for the higher values in the positive region of the scale. However, the picture is reversed for the grouped boxplots, compared to the rationality score. Here, 75% of female participants score higher than 50% of men, being that Q1 for women (26.5) is almost equal to the median of men (26). Both genders cover a similar range with the difference that the female scores are shifted upwards and both genders include outliers in the negative region of the scale. It may be noted that the male sample does not report any high scoring outliers whilst the upper 25% of the female sample reaches scores of up to 106. Again, these observations conform to literature stating that

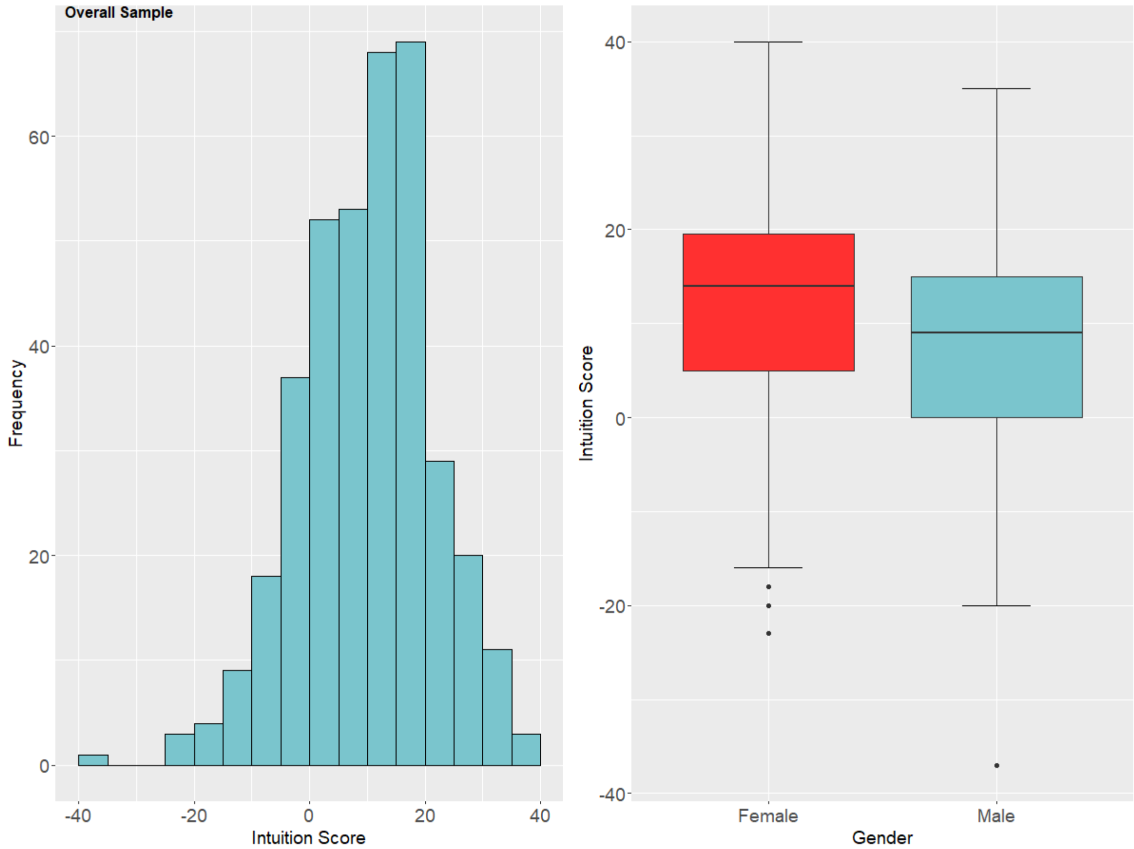
Figure 18: Experiential Score - Histogram and Boxplot



women have a higher tendency to operate in the experiential system than their male colleagues.

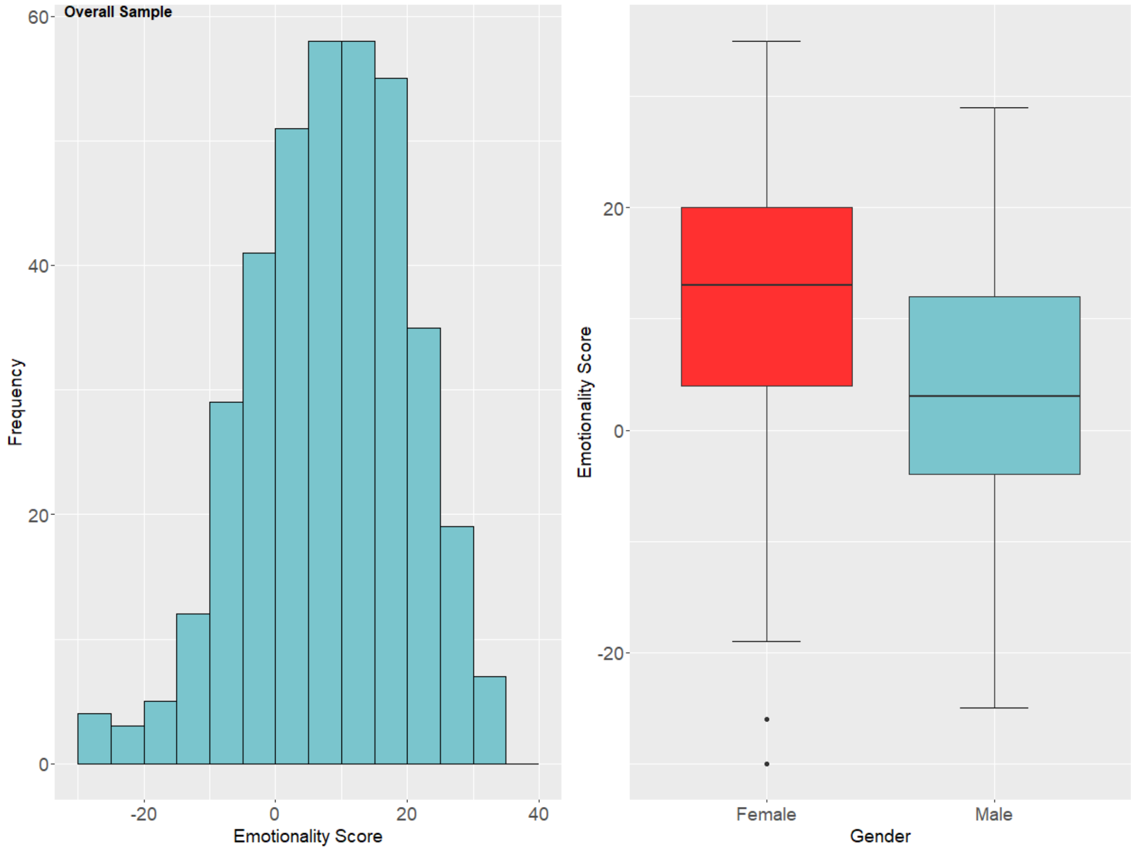
The measured intuition scores of participants also refer to the tendency to prefer the experiential system as an operating mode. As a result, the emerging pattern depicted by the histogram as well as the grouped boxplot in figure 19 is similar to the observation gained from the data on experientiality. The left tail of the distribution depicts an outlier with a considerably lower score than the remaining respondents of the sample. Although the female part of the sample also reaches the upper end of the scale, the difference between female and male respondents is less pronounced than for the previous measurements. The mean values for women (12.26) and men (7.71), as well as their medians (5 | 0), first quartile (14 | 9), and third quartile (40 | 35) show noticeable but not excessive jumps. However, the male sample contains the already mentioned extreme outlier in the negative range of the scale while the female part of the sample accounts for three outliers with closer proximity to the lower 25% of women.

Figure 19: Intuition Score - Histogram and Boxplot



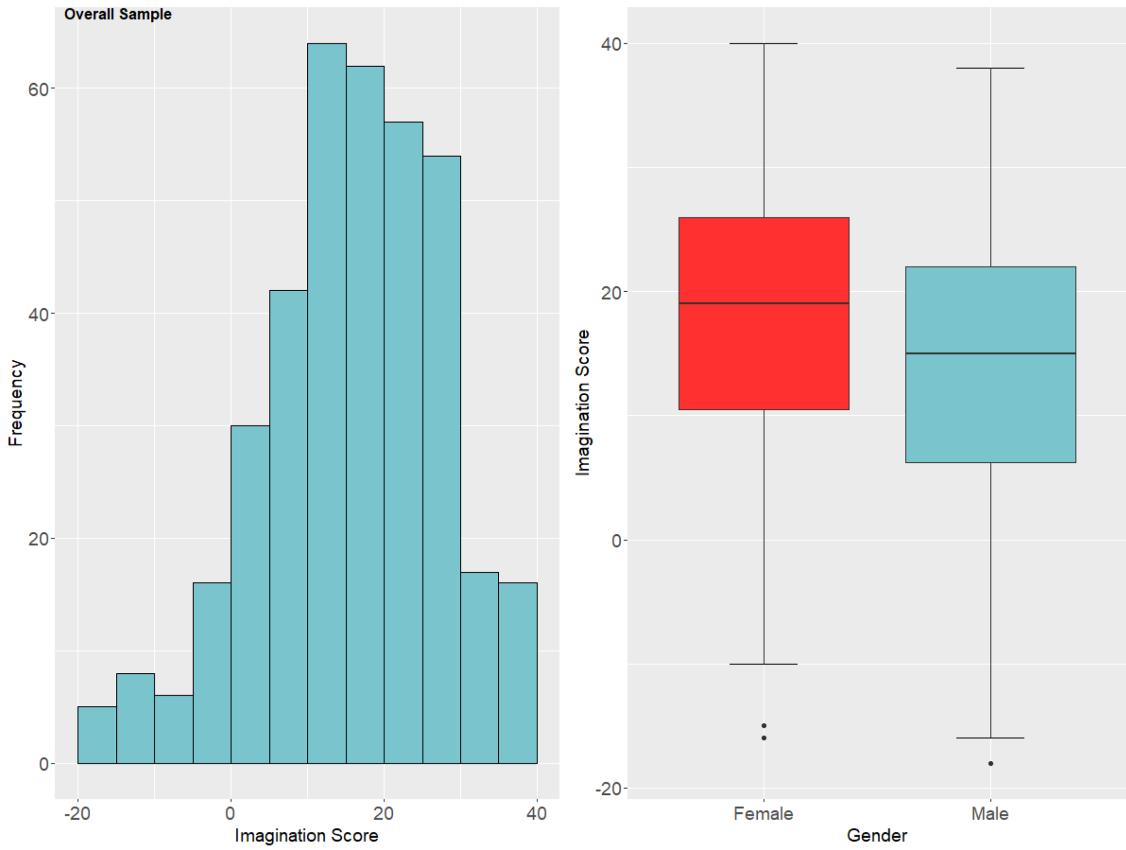
In consonance with the intuition score, the emotionality score provides another sub-measure of the experientiality of an individual. For this reason, the impression gained from the data complies with what was found for the experiential score and the intuition score. The histogram (figure 20) shows a distribution which is, again, heavier for the larger values in the positive range of the scale so that it is slightly left-tailed with a skew of -0.33. The end of the left tail shows a small increase for the lowest observed intuition scores, which hints at the presence of outliers in this area. The boxplots then reveal that these two outliers belong to the female part of the sample. Emotionality is the first variable stemming from the REI questionnaire where the male sample does not depict any outliers. Hence, the two female outliers score lower on emotionality than all male participants of the sample. However, women generally score higher on emotionality. Consequently, their corresponding boxplot is shifted into the upper region of the scale. The difference between the two groups is more pronounced compared to the intuition score. Regarding emotionality, the first quartile for women is with a value of 4 even slightly higher than the median for men, which lies at a score of 3. The implication is that 75% of women are rated as more emotional than 50% of men.

Figure 20: Emotionality Score - Histogram and Boxplot



The final measurement of the REI questionnaire yields the imagination score, which is another variable for addressing the experiential processing system. Hence, it can be expected that the gathered data behaves similarly to the previous three scores and opposite to the rationality score, where a larger fraction of men achieves higher scores than female respondents. The graph-pair displayed in figure 21 below confirms this expectation. The distribution in the histogram is skewed to the left with outliers in the left-hand tail. The more concentrated right side at first displays a stepwise decrease in frequency from the peak of observed values towards the higher end of the scale but is then followed by a sharp drop. It follows that, whilst a large fraction of the sample achieves scores between 10 and 30, only a few individuals score higher on the scale. Regarding the differences in gender, one can note that, although the expected difference is present, it is smaller than for emotionality and more similar to the results on intuition. In this sense, both genders cover a similar range of values, which is shifted upwards for the female fraction, and the median values (19 | 15) as well as quartiles (Q1: 10.50 | 6.25; Q3: 26 | 22) are not drastically different between the two groups. In addition, both genders display outliers that are positioned in a similar region. Besides the rationality score, imagination is the

Figure 21: Imagination Score - Histogram and Boxplot

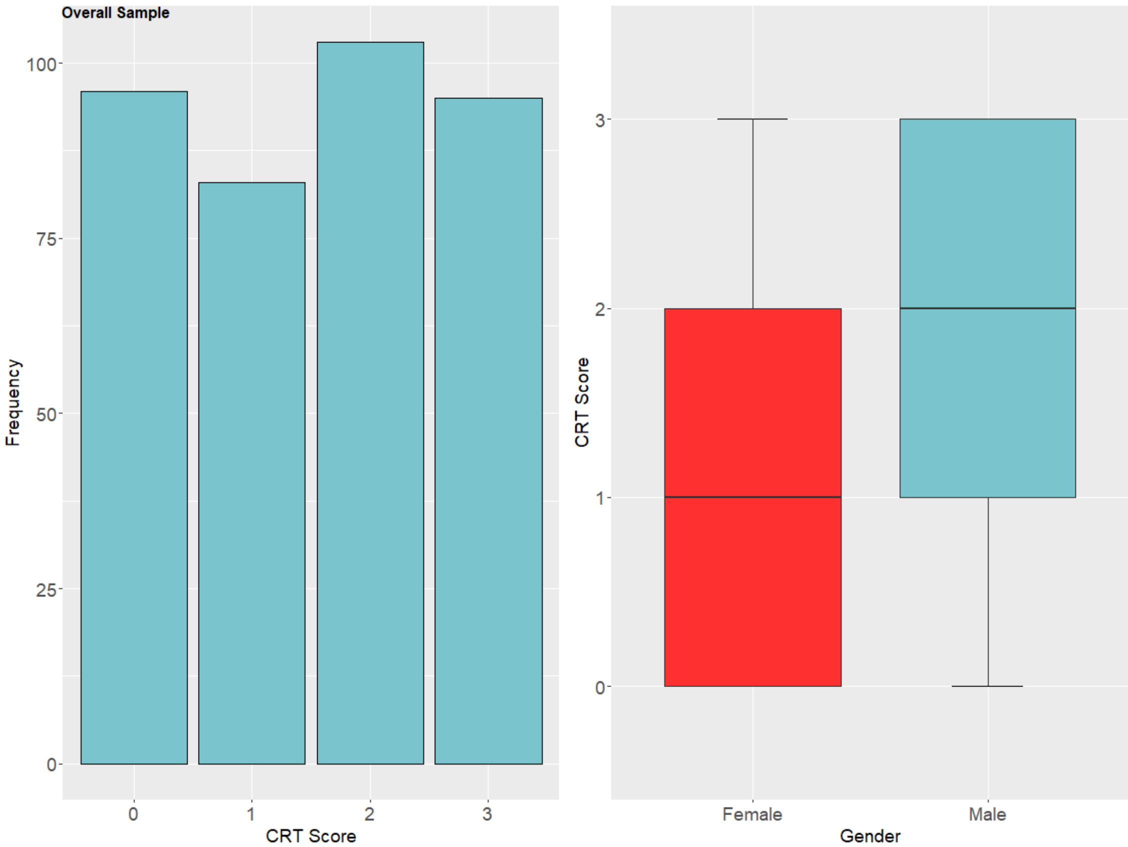


first measure of the REI questionnaire where both genders achieve similar maximum scores with women achieving a maximum of 40 while men score 38 at most.

Taken together, the five scores of the REI questionnaire support the findings in literature. Men are found to generally score higher on rationality, whilst the female fraction of the sample scores higher on the four measures that reflect the experiential system. It follows that men tend to operate in the rational system whereas women are inclined to engage in experiential processing.

The concluding psychographic characteristic is the impulsiveness of respondents, which is measured by the CRT. As it was introduced for the Berlin Numeracy Test, the created variable for this measure also describes the number of correct answers given in the three-item survey. In contrast to the scores on numeracy, the bar chart for the CRT score (figure 22) reveals a surprisingly equal distribution across all four values of the variable. This implies that the sample consists of four similarly large fractions of individuals, who are characterised by either higher or lower impulsiveness. When examining the boxplots for the genders, a distinct difference between the two groups becomes apparent. Both genders

Figure 22: Impulsiveness – Bar Chart and Boxplot



cover the full range of CRT scores but they mirror each other in the opposite direction. 75% of women in the sample achieve at most two correct answers while 50% of men respond correctly to at least two items. 75% of men score higher than 50% of women, who have a median of 1 correct answer. This gender effect may be viewed as complementing the results from the REI questionnaire where women are found to be more likely to operate in the experiential system, which is commonly described as fast, impulsive, and unconscious (Epstein et al., 1996; Epstein, 2003).

The repeating pattern for four of the psychographic characteristics found in the descriptive presentation already hints at the relatedness of some variables. In particular, experientiality may also be expressed through the measurements on intuition, emotionality, and imagination. The correlation matrix, created by the use of the R package corrplot (Wei and Simko, 2017), is displayed by figure 23 and validates this result since these four variables account for the highest correlations amongst all psychometrics. The highest correlations occur between experientiality and intuition (0.78) as well as experientiality and emotionality (0.78), followed by experientiality and imagination (0.72). It follows, that either experientiality or the group of intuition, emotionality, and imagination is used in the testing procedures but never all four simultaneously. Apart

Figure 23: Correlation Matrix Psychometrics

from this variable block, the next notable but not comparably high correlation occurs between the Berlin Numeracy score and the CRT score (0.41). This may be attributed to the circumstance that both tests are analytical and require the computation of results, whereas the variables stemming from the REI are gathered by agreeing or disagreeing to statements that address personality traits. The correlations between emotionality and the Berlin Numeracy score as well as imagination and rationality are not of significant character with p-values above the 10% significance level.

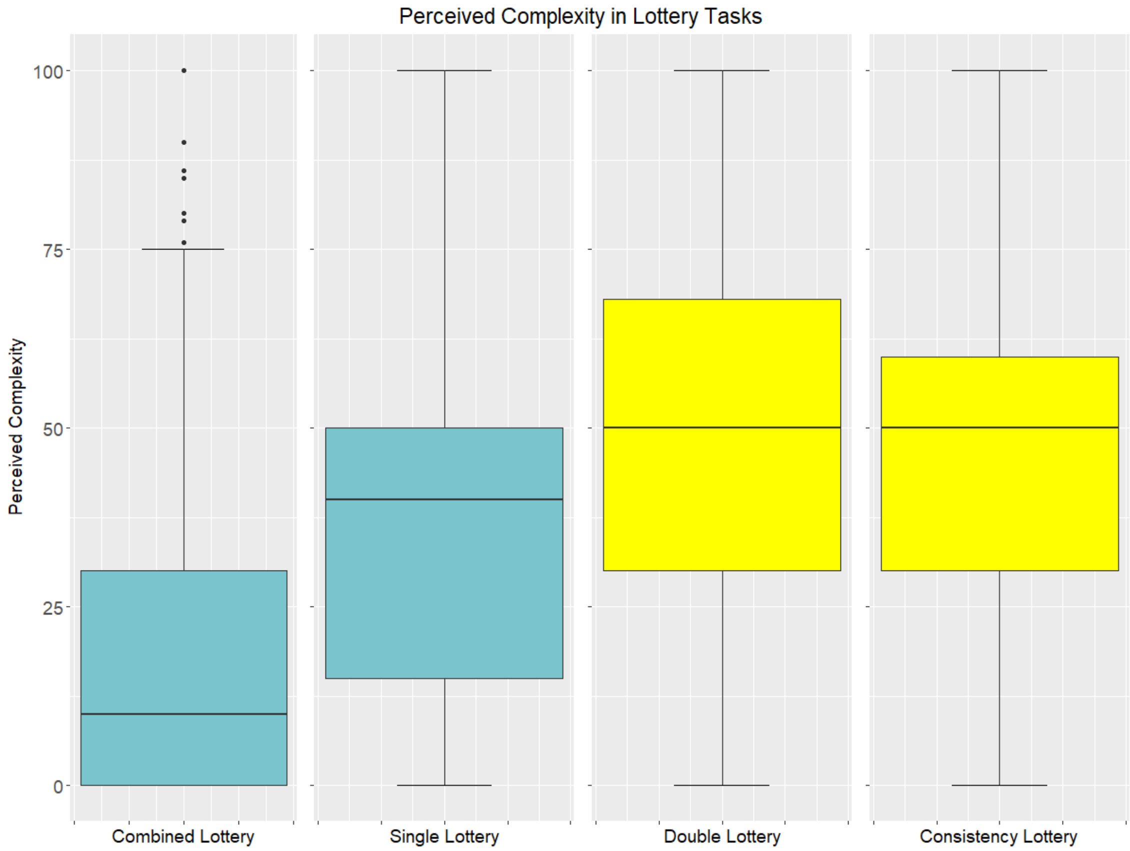
5.2 Descriptive Results: Lottery Part

Having presented some overall results of the experiment, this subchapter is dedicated to the data gathered from the four lottery tasks. The aim lies in introducing the reader to the perceived complexities, choice distributions, and times needed to complete the tasks of this part of the experiment. The table containing the descriptive statistics for the perceived complexity is given in Appendix C. Tables describing the remaining variables are to be found in Appendix D.

(1) <u>Perceived Complexity</u>

The previous subchapter already established that the perceived complexity of the lottery part is lower than the complexity respondents experience during the personnel part. Now, this perceived complexity for the overall part is further split in order to represent each of the four tasks on its own. The grouped boxplots in figure 24 show an increasing trend in perceived complexity, where the combined lottery is experienced as the least complex. For this task, 75% of respondents perceive a complexity level of 30 or less and the median is at a complexity of 10. Although all four tasks cover the full complexity range, the combined lottery does this as a result of outliers in the region above a complexity of 75. The jump to the single lottery is especially noticeable when paying attention to the median, which lies at 40 for this task. In this sense, more than half of the participants report a perceived complexity that is higher for this task than what is observed for 75% of respondents in the combined lottery. Half of the gathered data entries for the single lottery lie between a perceived complexity of 15 and 50.

Figure 24: Perceived Complexity Lottery Part

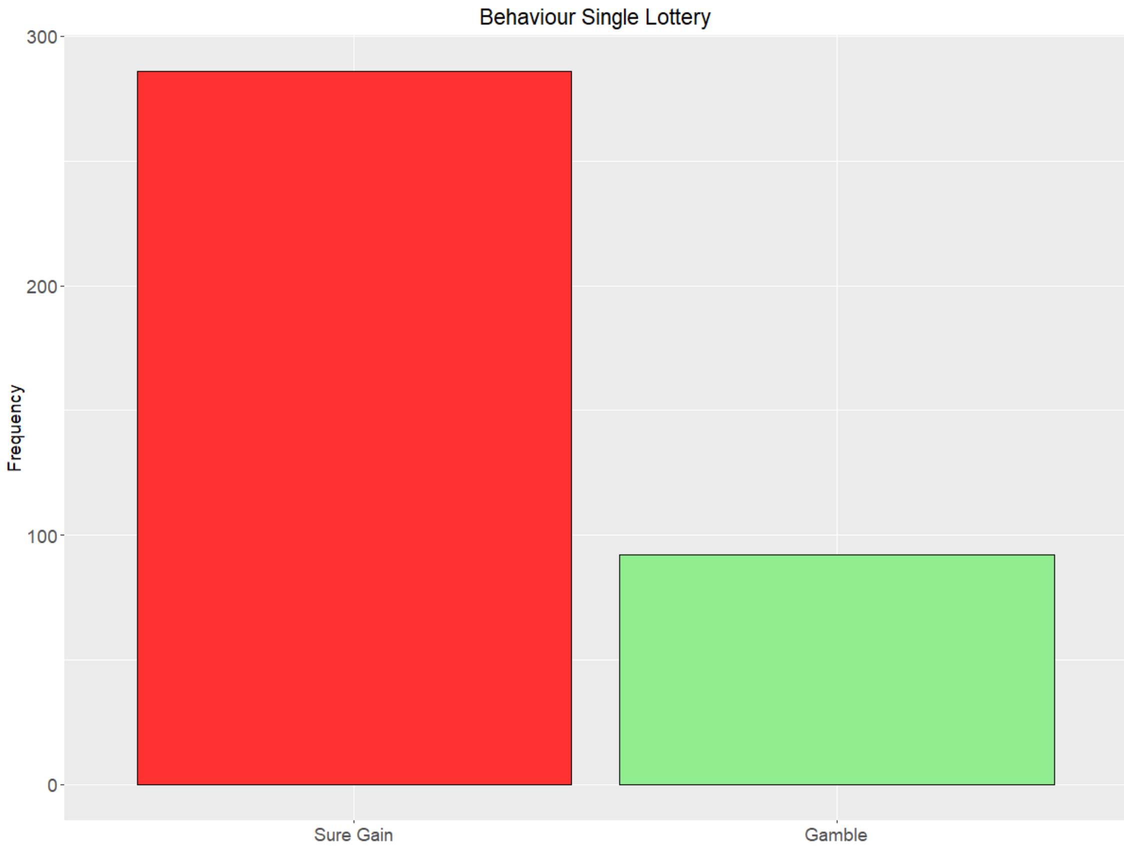


The increase in complexity with respect to the median is less pronounced when comparing the switch from the single to the double lottery as it was for the jump from the combined to the single lottery. Concerning the median, the perceived complexity increases from 40 in the single lottery to 50 in the double lottery. The median of the double lottery lies at the same level as the third quartile of the single lottery, meaning that 50% of participants describe the double lottery as more complex than 75% of respondents in the single lottery. The median (50) as well as the first quartile (30) are identical for the double lottery and the consistency lottery. This observation reflects the intention that these two tasks were designed to belong to the same objective complexity stage. Only the third quartile is higher for the double lottery (68) than it is for the consistency lottery (60). However, this may be explained by the circumstance that the consistency lottery succeeds the double lottery, so that when a respondent is confronted with a task of this type for the second time, it is experienced as less complex for the upper end of the range. In general, the development of the perceived complexity across the four tasks nicely reflects the complexity stages that the tasks were assigned to.

(2) Decision Behaviour

The single lottery task is the first one played by participants and it is free of any bracketing behaviour. The reason for this is that the task does not involve a second decision that may be broadly bracketed by integrating it into the decision set of the first decision. Instead, the task provides a baseline impression of the respondents' risk attitudes by either preferring the gamble or the sure gain in a lottery. Nevertheless, respondents will not be explicitly categorised into risk averse or risk-seeking for the reason that the prospects of the task do not exactly fit the definition of risk aversion. This is because the gamble is of higher EV than the sure gain (19.000 vs. 20.000). The definition of behaving in a risk averse manner would require the DM to choose the sure gain over a gamble of equal EV (Kahneman and Lovallo, 1993). Therefore, it is only possible to classify subjects who choose the safe gain as risk averse, but those who choose the gamble cannot be declared as being risk seeking without any doubt. However, the underlying risk attitude with respect to gambling may be observed. Such an observation may then hint at behaviour in the bracketing-relevant double lotteries, where broad bracketing involves gambling in the domain of gains whilst accepting a sure loss in the domain of losses. Figure 25 shows that 286 respondents (75.66%) choose the sure gain over the gamble, which is selected by the

Figure 25: Choice Frequency - Single Lottery Task

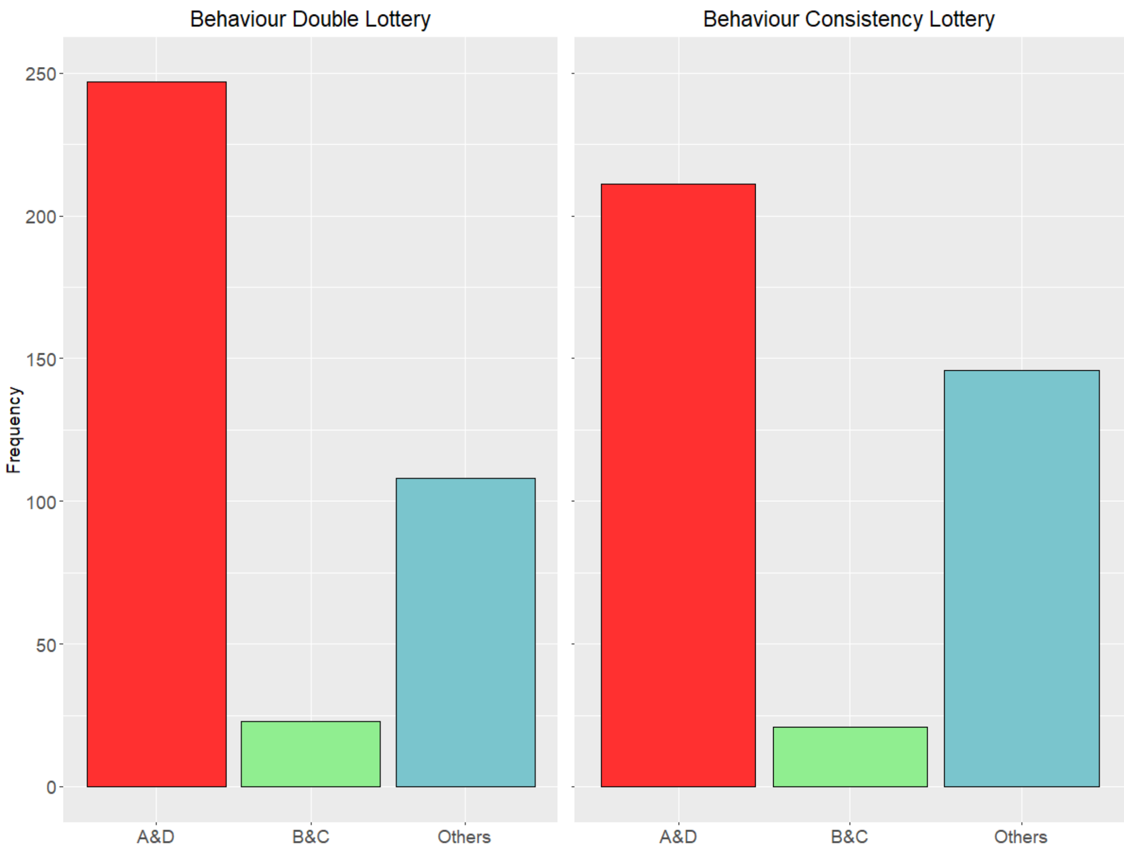


remaining 92 (24.34%) participants. The result displays a clear preference for a sure gain when confronted with a lottery, despite the circumstance that the gamble is of higher EV.

The gathered data for the following double lottery and consistency lottery tasks reveal a pattern that resembles the baseline risk preference elicited by the single lottery. The dummy variables that store the behaviour for each of the two lotteries yield the resulting choice combination of decision one and decision two. The combination A&D represents behaviour according to prospect theory, meaning that the sure gain is preferred over the gamble in the first decision while the gamble is preferred over the sure loss for the second decision. The choice-pair B&C describes broad bracketing by behaving opposite to PT predictions and choosing the gamble in the gain-domain and the sure loss in the loss-domain. All other choice combinations are summarised under the denomination "Others". Note that, as it was discussed in the theoretical part of this thesis, A&D as well as "Others" describe narrow bracketing behaviour. Narrow bracketing in this part of the study is split into these two versions in order to monitor prospect-theoretic behaviour.

The choice patterns displayed by figure 26 present an obvious dominance of the prospect-theoretic A&D combination for the double lottery (247 subjects, 65.34%) as well as the consistency lottery (211 subjects, 55.82%). The broad bracketing combination B&C is

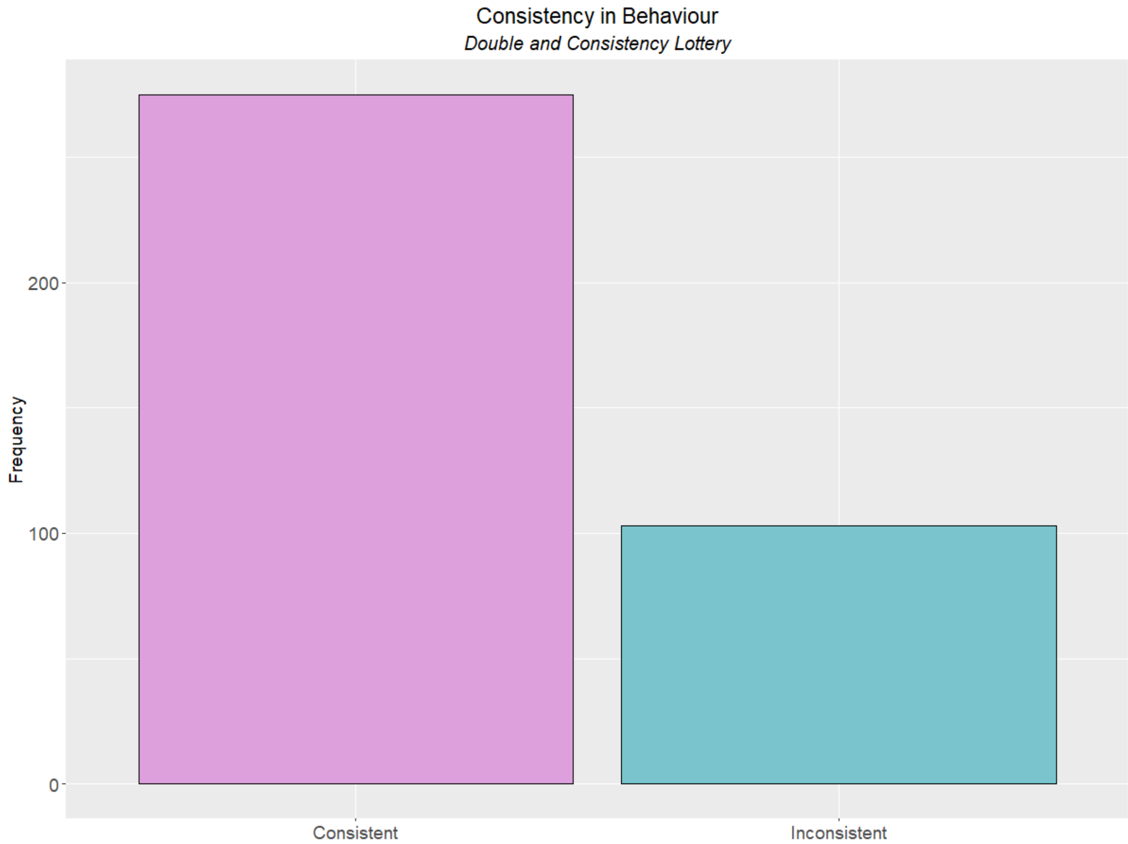
Figure 26: Choice Frequency - Double Lottery and Consistency Lottery Tasks



selected with an almost equal frequency in both tasks, namely by 23 (6.08%) subjects in the double lottery and 21 (5.56%) individuals in the consistency lottery. When including the other choice combinations, narrow bracketing behaviour is exhibited by 355 (93.92%) respondents in the double lottery and 357 (94.44%) individuals for the second task. This distribution strongly suggests that narrow bracketing is indeed the prevalent behaviour of individuals. Furthermore, the separation of the variable into A&D and "Others" demonstrates that the majority of narrow bracketing in such a task setting is accounted for by PT predictions. These findings are consistent with the research of Kahneman and Tversky (1981), who report that 73% choose A&D while 3% select the B&C combination, as well as Rabin and Weizsäcker (2009) observing a frequency of 66% for A&D and 3% for BC. Compared to these studies, the present research finds a marginally larger share for the broad bracketing B&C combination in both tasks. This may be attributed to the majority of respondents being students of business administration or economics.

When comparing both tasks, one can note that the fraction of prospect theoretic decisions decreases as the experiment proceeds while more respondents choose another combination. Nevertheless, the share of broadly bracketed decisions also decreases so

Figure 27: Consistency in the Double and Consistency Lottery Tasks



that the holistic choice frequencies do not indicate a possible learning effect between the two tasks. Respondents' consistency in their choice pattern is visualised by figure 27. The bar chart shows that the large majority, 275 individuals (72.75%), is consistent in their selection of decision alternatives. Consistent here means that the exact same choice combination is selected in the double and the consistency lottery. This result indicates that the decision-making behaviour of the sampled participants is not altered or reconsidered for similar tasks. This holds although the tasks are interrupted by the Berlin Numeracy test.

The following table 11 enables a more detailed examination of the development of choice selections as participants proceed from the double lottery to the consistency lottery. What is interesting is that, although 23 and 21 individuals bracket broadly in the two tasks considered separately of each other, only 9 (2.38%) participants consistently choose B&C in both lotteries. Four Participants switch from broad bracketing to the prospect-theoretic combination A&D, while 10 respondents divert in favour of another combination. The noticeable difference in A&D selections between the tasks can be traced to 7 respondents starting to bracket the consistency lottery broadly and 55 individuals diverting from PT in order to select another combination. In addition, 5 participants who were amongst the "Others" choice combinations in the double lottery, engage in broad bracketing for the second task by choosing B&C. While the overall B&C choice frequencies for the task appear relatively stable, the cross table reveals that 12 individuals "learned" to bracket broadly when confronted with a similar task for the second time. However, the 14 participants diverting from broad bracketing might imply that their behaviour in the first

Table 11: Behavioural Development from the Double Lottery to the Consistency Lottery Task

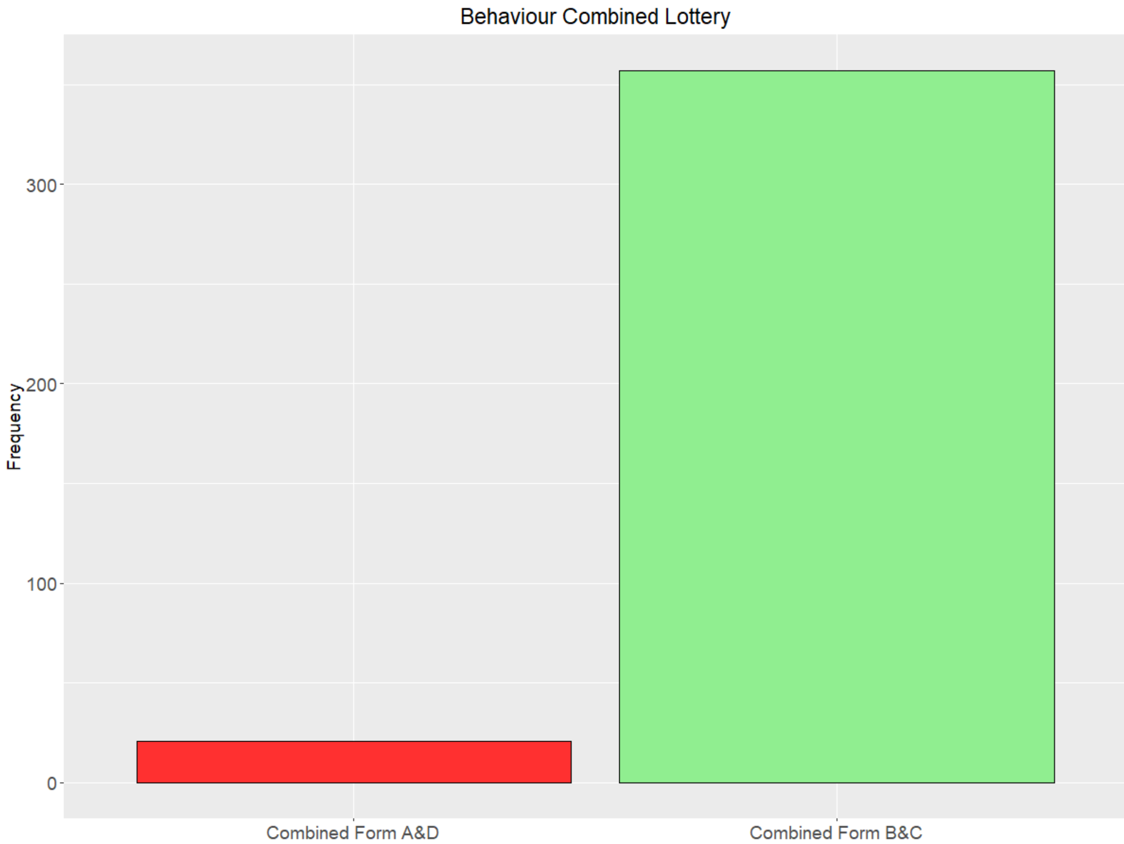
		Consistency Lottery		
	Choice Combination	A&D	B&C	Others
Double Lottery	A&D	185 (48.94)	7 (1.85)	55 (14.55)
	B&C	4 (1.06)	9 (2.38)	10 (2.65)
	Others	22 (5.82)	5 (1.32)	81 (21.43)

*relative frequencies in parentheses

task stems from a "lucky" guess. Alternatively, the respondents who "learned" and switched to B&C for the second task might be the result of a lucky guess. Overall, it becomes apparent that the fraction of individuals who is able to consistently apply broad bracketing to lotteries is, with 2.38%, even smaller than the first impression of the data in figure 26 suggests.

Subsequent to the two bracketing-relevant lotteries, the combined lottery aimed at guiding behaviour towards broad bracketing by highlighting the dominance of B&C over A&D. This is done by presenting the combination A&D and B&C in their explicitly combined form instead of their two components separated into two decision. The findings for this task (figure 28) clearly show that respondents are able to choose the dominant option when they are forced to consider the separate alternatives in an integrated manner, instead of assessing them as isolated decisions. The frequency of A&D answers is drastically reduced to 21 (5.56%) while the remaining 357 (94.44%) now choose the solution that broad bracketing would have yielded in the previous two tasks. Hence, the combined presentation form enables individuals to recognise their previous error or inattention to the circumstance that both decisions of a double lottery need to be considered simultaneously. In their study, Tversky and Kahneman (1981) achieve an even stronger

Figure 28: Choice Frequency - Combined Lottery Task



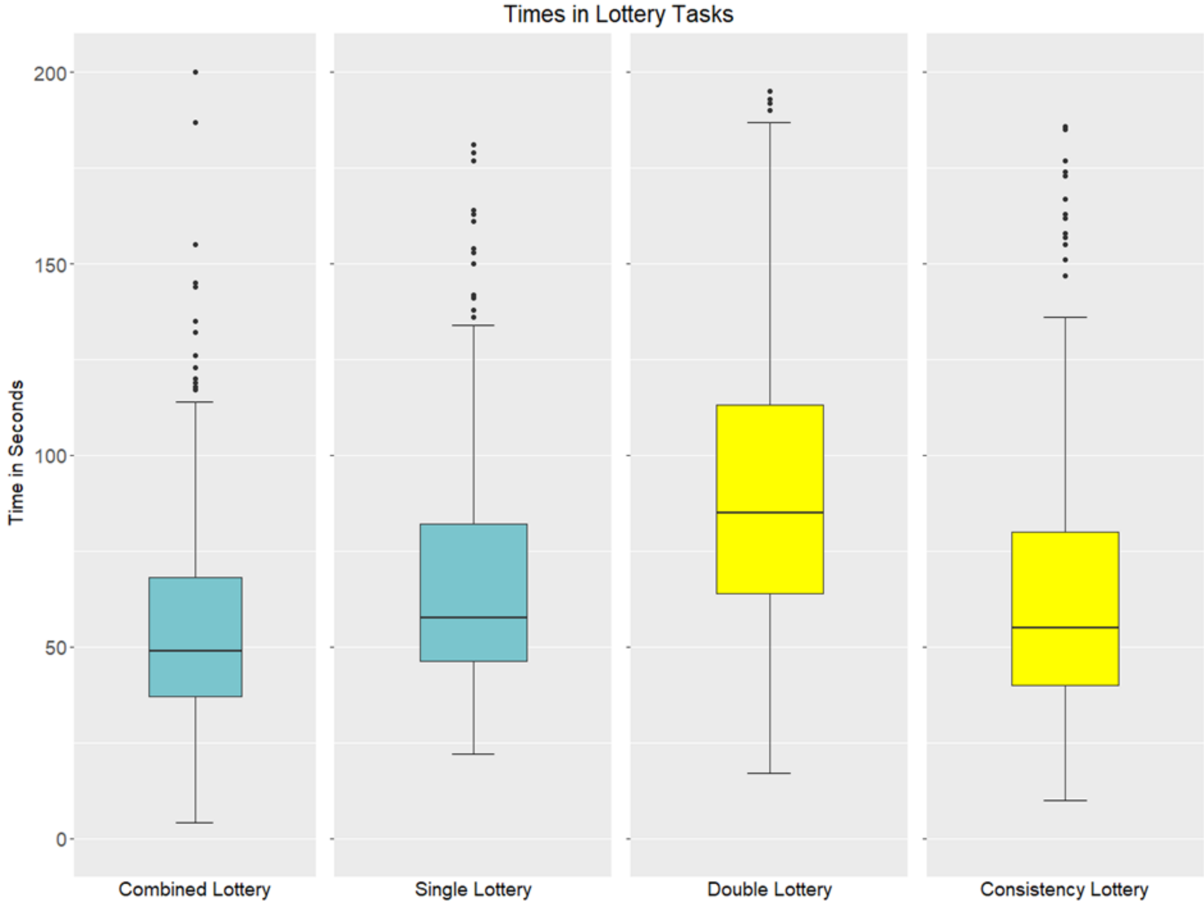
effect for a combined presentation form and report 0% A&D combinations in their sample.

(3) Times for Task Completion

Before discussing the time needed for the lottery tasks, one general statement regarding the recorded time stamps in the experiment has to be made. As already mentioned, the access to the experiment was open for one week, so that respondents were free to choose when they would wish to start the experiment. In this setting, it was not controlled for the circumstance that some individuals might participate with breaks during their sessions or even complete the experiment on a different day. Also, the sampled students were not explicitly instructed to complete the whole experiment in one session and without any breaks. As a result, the time records contain several large outliers as well as numerous entries of medium magnitude. Nevertheless, it was decided not to exclude any data entries of completed experiments. This is because we may not assume that completing the experiment in multiple sessions also implies that the tasks were granted insufficient attention from the respective participant. Moreover, the patterns found in the descriptive time distributions may still contribute to the findings of this study, but the time variable will not be used for any statistical testing.

Figure 29 then presents the boxplots for the times (seconds) needed in each task of the lottery part. The y axis of the graph was limited to 200 in order to provide an improved overview and exclude any excessive outliers from the presentation. In consequence, the graph excludes 5 entries from the combined lottery, 12 for the single lottery, 34 for the double lottery, and 14 regarding the consistency lottery. The boxplots for the combined, single, and double lottery depict an increasing trend, which matches the objective complexity stages as well as perceived complexity of these tasks. 75% of subjects need at most 1.15 minutes for the combined lottery, 1.38 in the single, 2.08 in the double, and 1.40 minutes for the consistency lottery. In this sense, a more complex task requires more time to be solved, which is a rather plausible observation. However, while the double and consistency lotteries belong to the same complexity stage and show only a minor difference in perceived complexity, the times needed for these two tasks are distinctly different. Respondents need considerably less time when facing a double lottery for the

Figure 29: Time Stamps in the Lottery Part



second time. It is likely that a large contributor to this pattern is the similarity of the two tasks, since only the prospects but not the wording is changed. This, in combination with recognising the functioning of the task, is likely to accelerate the decision-making process. Alternatively, subjects may have paid less attention to the task once they assumed it to be of great similarity to the preceding one. This assumption can be also complemented by the previous finding of 72.75% choosing the exact same combination in both lotteries. At the same time, this task reports the largest outlier of the whole experiment at a time of 79.40 hours for completing the consistency lottery. With respect to a relationship between the perceived complexity and the time stamp of a task, it is found that the correlations for all four tasks are of insignificant nature.

5.3 Descriptive Results: Personnel Selection Part

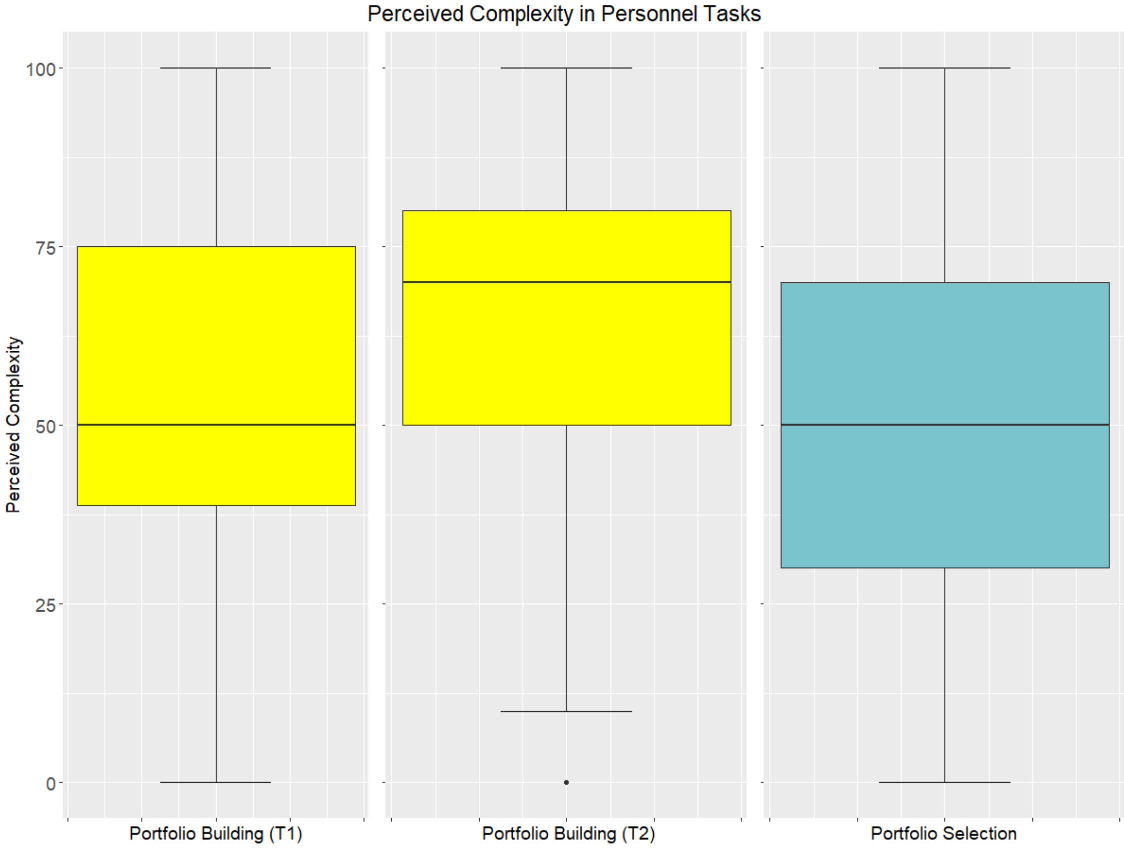
The last of the descriptive results subchapters seeks to present the findings of the main part of the experiment. This chapter first discusses the observed perceived complexities and time stamps for the three personnel selection tasks. Subsequent to these overall aspects, subchapter 5.3.1 addresses the evaluation of variety seeking behaviour as well as the results gathered in the two treatments of the portfolio building task. Subchapter 5.3.2

continues by reviewing the data concerning the portfolio selection task. The tables containing the descriptive statistics of this chapter as well as the diversification evaluation tables can be found in Appendix E. The table containing the perceived complexity is again found in Appendix C.

(1) <u>Perceived Complexity</u>

The comparison of the two parts of the experiment in chapter 5.1 already revealed that the personnel part is perceived as more complex than the lottery part. As it was done for the lotteries in the preceding subchapter, the overall complexity is further divided into the respective complexities per task. The grouped boxplot is depicted in figure 30 and illustrates a trend that is consistent with the introduced objective complexity stages. Especially the difference between the two treatments of the portfolio building task is of great interest. For this task, T1 refers to the segregate presentation form while T2 denotes the aggregate form. Both treatments span the whole range of the perceived complexity scale but T2 does this as a result of an outlier at a complexity of 0%. In general, the aggregate treatment is experienced as more complex since 75% of subjects report a higher complexity in T2 than 50% of participants in T1. The median of T1 lies at a complexity

Figure 30: Perceived Complexity Personnel Selection Part



of 50%, whereas the median of 70% for the aggregate T2 almost reaches the third quartile (75%) of T1.

These findings indicate that the manipulation of the information presentation form used to construct the two treatments has a noticeable impact on the perceived complexity of the task. Moreover, the observed complexity data coincides with the assigned objective complexity stage, where T1 is assigned to stage 4 while stage 5 refers to T2. The remaining portfolio selection task, which is played by both treatments in the same presentation form, was assigned to stage 3. Although the difference in perceived complexity to T2 is obvious, the difference to T1 of the building task is considerably less pronounced. Compared to T2, 75% perceive the selection task as less complex than 50% of subjects who played T2 of the building task. However, when comparing the selection task to the segregate T1, it is noticeable that the task also spans the full range of complexities and its median is identical to T1 (50). Nevertheless, the majority of data entries is concentrated at a slightly lower level than the perceived complexity of T1. This is shown by Q1 (30) and Q3 (70) of the portfolio selection task being located below the respective quartiles of T1 in the building task (Q1 = 38.50; Q3 = 75). Furthermore, the gathered complexities amongst subjects in the selection task are close to being symmetrically distributed across the complexity scale, which can be seen by the almost central position of the median between the first and third quartile.

The reason for the overall similarity of the T1 building task and the selection task is likely that the pre-allocation of candidates in the segregate presentation form provides sufficient decision support, so that the presentation of pre-built teams in the selection task is not experienced as an excessive additional support. The slightly more complex experience may then stem from the circumstance that T1 presents the pre-allocated candidates with their full skill item list, whereas the selection task presents the candidates of the pre-built teams with their total skill values. For participants in T2 of the building task, however, the pre-built teams are a considerable supporting measure compared to evaluating all nine candidates simultaneously and with their complete skill item list.

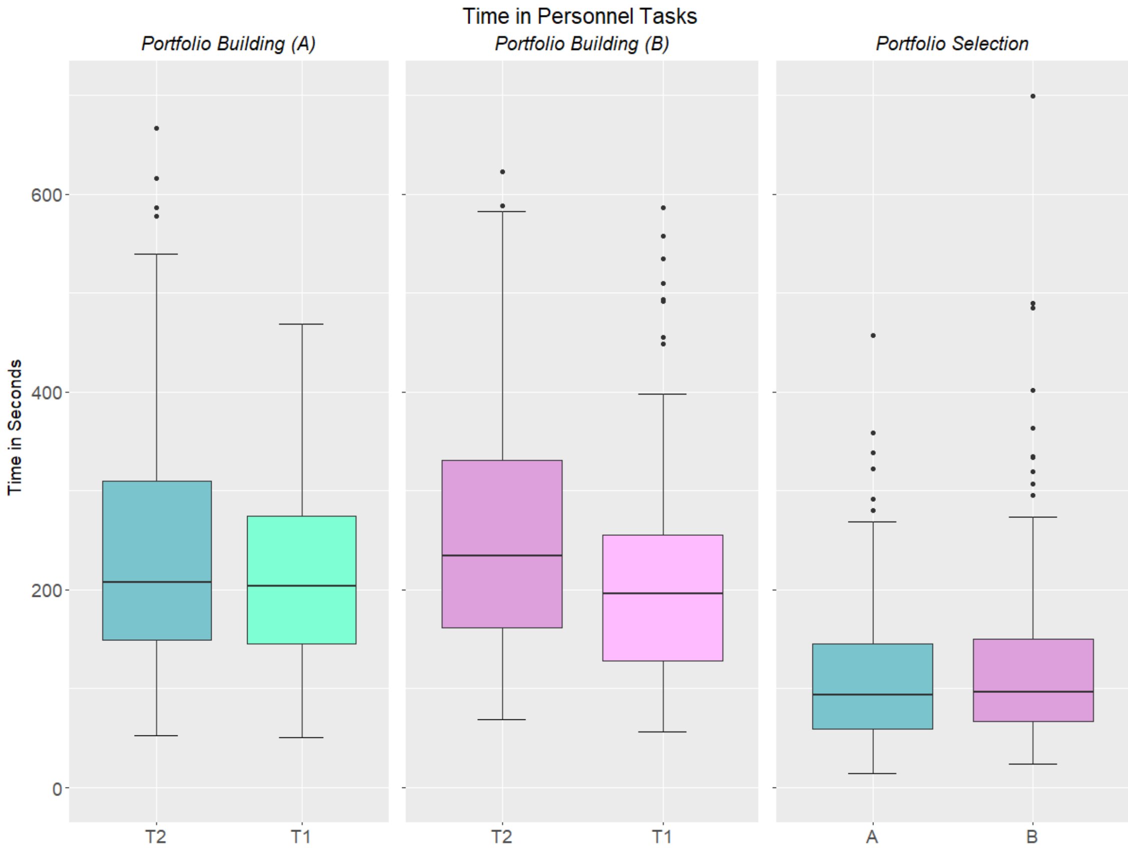
(2) <u>Times for Task Completion</u>

As already addressed in subchapter 5.2, the time records for the experiment are affected by numerous outliers as a result of not explicitly demanding subjects to complete the experiment in one session. Despite this, the overall picture of the gathered data is still able to convey some information.

The boxplots for the time needed to complete the tasks of the personnel selection part are presented by figure 31. The graphs for the portfolio building task are first split into the two candidate versions A and B before the data for the respective version is grouped according to the two information presentation forms T1 and T2. The boxplot for the portfolio selection task is grouped according to the versions A and B, being that the respondents in both treatments of a group played the same selection task. To improve the visual presentation, the y-axis of the boxplots is limited to 700 seconds. This measure results in the exclusion of 1 outlier from group A of the portfolio building task, 9 outliers for group B of the same task, and 1 data entry in the portfolio selection task.

Both groups show a similar pattern, where T1 requires less time for completion than T2. However, the treatments in group A are considerably closer together and report almost identical first quartiles, 145 for T1 and 150 for T2. The median time of players in the segregate T1 is 204 seconds while participants in T2 require 208 seconds. A notable difference is that the third quartile of T1 (275) is below Q3 of the second treatment (311.50) and T1 covers a smaller range. Regarding group B, the difference between the

Figure 31: Time Stamps in the Personnel Selection Part



treatments is more distinct as Q1 (130 in T1; 167.5 in T2), the median (201 in T1; 246 in T2), as well as Q3 (295 in T1; 344.50 in T2) of the segregate treatment lie below the aggregate presentation form. The recorded time stamps for the portfolio selection task reveal an almost identical distribution in both groups. This task requires considerably less time than any of the treatments in both groups, being that 75% of respondents of this task lie below the median of all other treatments, below the first quartile of T2 (B), and almost below 75% of subjects in both treatments of group A. Overall, the illustrated pattern conforms with the observed perceived complexities and assigned objective complexity stages of these tasks.

5.3.1 Portfolio Building Task

Having gained a first impression of how the complexity of the personnel tasks is perceived, this subchapter investigates the behaviour exhibited in the two treatments of the building task and the resulting utility. Before proceeding to these paragraphs, it is necessary to discuss how the variety seeking behaviour for the selection of candidates in a team is evaluated.

(1) Diversification: Candidate and Team Evaluation

In order to assess the variety seeking behaviour of respondents, the candidates of both groups, A and B, are either classified as generalists or specialists. The procedure that is applied for this step is the following: The four skill items (i) of each candidate (n) are weighted by their respective team compatibility (TC) value before the variance of the weighted skill items is computed for each candidate. Candidates with a high variance in their skill set can then be viewed as specialists. For the purpose of evaluating the built teams, however, the classification is further refined. For any given team constellation, the weighted maximum value (max) for each skill item achieved by this team is selected. This value is then compared to the weighted maximum possible value (best) for this skill amongst all nine candidates (m) and the corresponding ratio is calculated for each item. The mean value of the four variety ratios yields the final diversification rate for the respective team constellation. The variable created by this procedure gives the achieved skill variety of a team compared to the maximum possible skill variety and represents the measure of how well a specific participant diversified their team building decision. This process can be described through the following formula:

$$Variety\ Ratio = \frac{1}{4} * \sum_{k=1}^{4} \frac{TC_n * i_k^{max}}{TC_m * i_k^{best}}$$

A second dummy variable states whether a participant achieved a diversified team or not. For this purpose, a cut-off criterion is introduced to the ratio of each skill item in a team. All teams that yield a ratio of at least 65% for each of the four skill items are classified as diverse. The aim of applying the cut-off criterion to each skill item instead of the mean ratio is to avoid that a high ratio in one or two skills compensates for insufficient diversification in the remaining items. All 84 possible team constellations are evaluated according to this procedure, leading to 10 teams (11.90%) being classified as diverse. Table 12 presents an example of such an evaluation with the candidates 1, 2, and 3 of group A. The resulting variety ratio is 69.02% and the team would not be classified as diverse since the ratio for the items IT and Languages does not fulfil the 65% criterion. This team constellation also demonstrates why the criterion is applied to each item instead of the mean ratio, being that the mean fulfils the criterion despite insufficient diversification on two items.

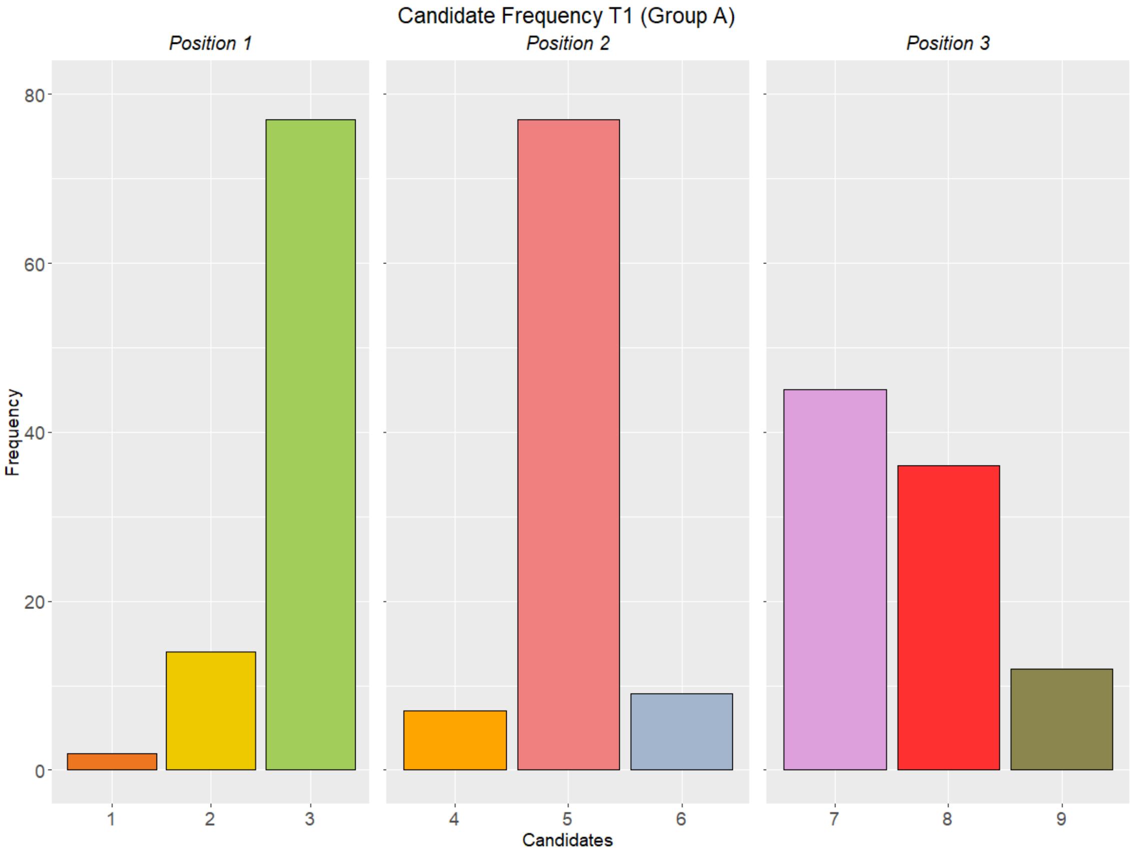
Table 12: Example - Variety Ratio for Team Evaluation

Weighted Skill Items	Selected Candidates 1	2	3	Evaluation MAX	MAX - Ratio
Analytics	8,00	45,00	33,25	45,00	100,00%
IT	14,00	24,75	24,50	24,75	41,25%
Languages	17,00	9,00	26,25	26,25	64,81%
Negotiation	11,00	20,25	21,00	21,00	70,00%
				Mean Variety Ratio:	69,02%

(2) Candidate Selection – Group A

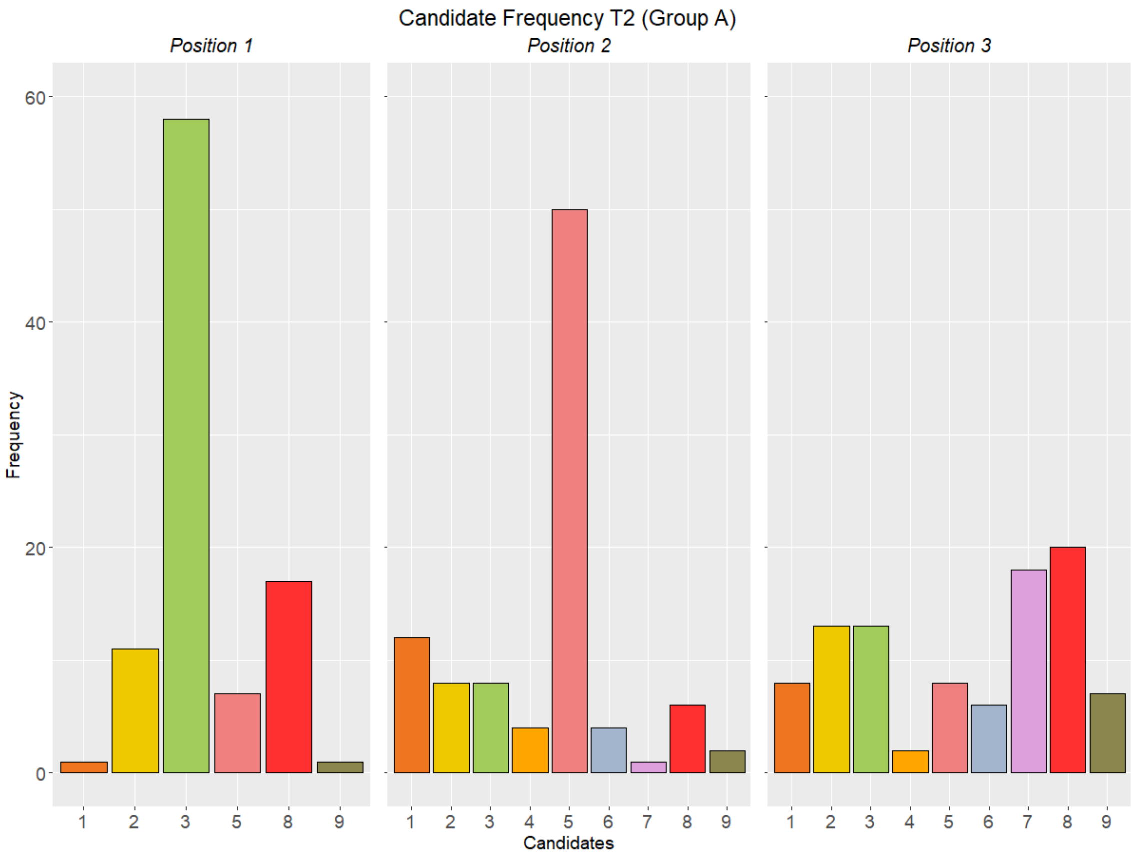
The introduced pre-allocation of candidates in treatment 1 of the task required participants to select one amongst three candidates for each position. In contrary, the aggregate treatment 2 asked for a decision between all available candidates for each position. The observed decision behaviour of subjects in T1 of group A is illustrated by figure 32. For the first and second position, the data reveals a clear dominance of candidate 3 (82.80%) and candidate 5 (82.80%) respectively. These two candidates would also be part of the utility maximizing team constellation, which includes the candidates number 3, 5, and 8. The third position has a more distributed pattern with candidate 7 as preferred choice for 48.39% of respondents. However, candidate 8 (38.71%) would have been the optimal

Figure 32: Candidate Distribution for each Position - T1 (A)



choice with respect to the achieved utility of the team. Due to the pre-allocation of candidates, it is not possible to build all 84 team constellations in T1. Despite this restriction, participants in T1 of group A are able to construct the team with the maximum diversification ratio, consisting of candidates 2, 4, and 8. It follows that the preference of candidate 7 over 8 is not a result of seeking a more diversified team but is of another source. A likely explanation can be found in the fact that candidate 7 yields the highest utility before the TC factor is taken into account. Hence, respondents favouring this candidate may have forgotten to factor in the TC value when evaluating the presented candidates. When ignoring the TC factor, candidate 8 would in fact be the worst one of this row. However, since the skill utilities need to be weighted by the TC factor, candidate 8 dominates with a utility of 120 compared to 55 (candidate 7) and 70 (candidate 9). In fact, candidate 8 yields the highest utility of all nine candidates.

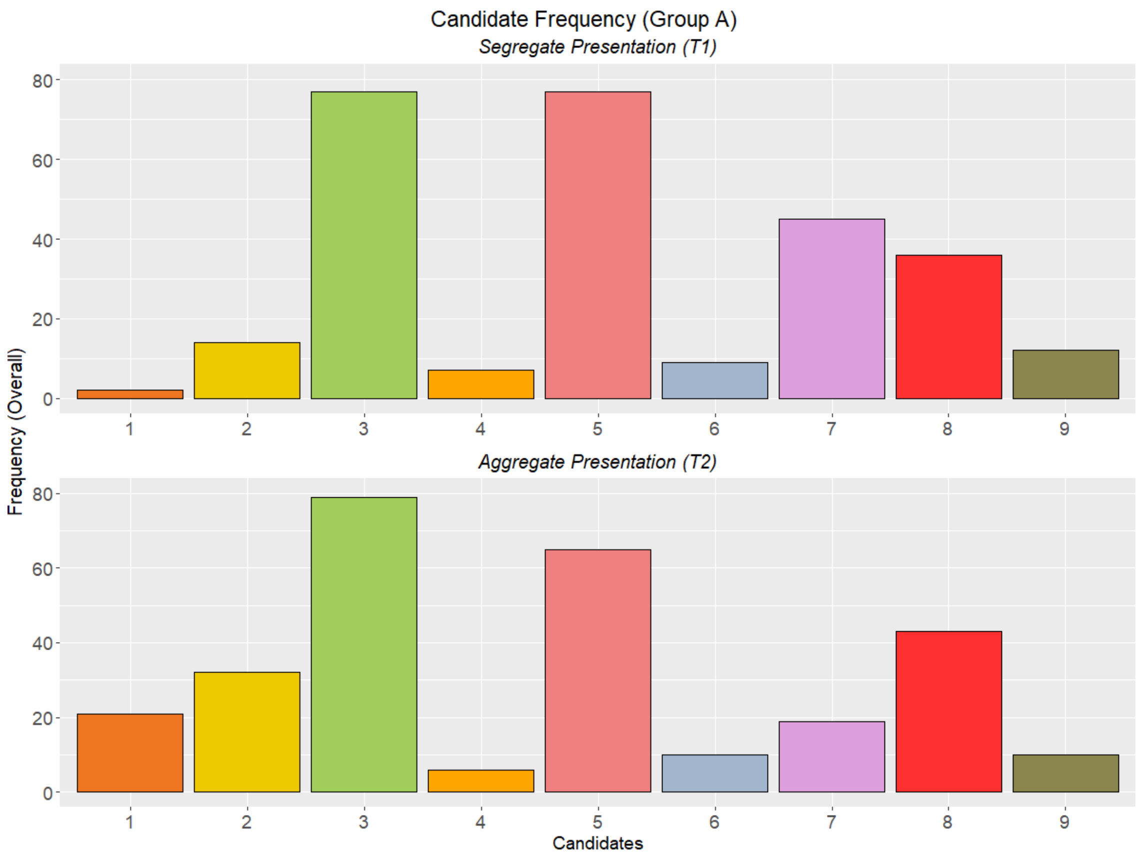
Figure 33: Candidate Distribution for each Position - T2 (A)



The choices in T2 of group A reveal a similar distribution (figure 33). For the first position, candidate 3 is preferred by the majority (61.05%) of respondents and is followed by candidate 8 (17.89%). Both candidates belong to the utility maximizing solution and the pattern implies that subjects either choose from the first or the last row of the presentation table. Interestingly, the candidates 4, 6, and 7 are never chosen for the first position. The second position is clearly dominated by candidate 5 (52.63%), which is also part of the optimal team constellation. Again, the distribution for the last position is more dispersed. This may suggest, that while the attention for position 1 is concentrated on the first or last row of the table and position 2 is focused on the middle row, subjects tend to review the whole table when deciding on their third choice. In contrary to the segregate T1, the majority (21.05%) of participants select the dominant candidate 8 followed by candidate 7 (18.95%), which yields the highest utility when the TC factor is not taken into account. Candidate 2 is the third most favoured candidate in all three positions and is selected with an identical frequency to the utility-optimising candidate 3 for positions 2 and 3. This may be explained by candidate 2 being part of the most diversified team constellation, given by the candidates number 2, 4 , and 8. The noticeable selection frequency of candidate 2 may then be the result of some participants focusing solely on diversification and not on the utility generated by the constructed team.

When considering the overall selection rates for both treatments in group A, the pattern displayed by figure 34 emerges. The segregate T1 is characterised by individuals correctly identifying candidates 3 and 5 as superior before the majority of 16.13% then exhibit an error in their decision-making by choosing candidate 7. The overall frequencies in T2 also illustrate the preference of candidates 3 and 5. Moreover, when the decision-making of participants is not restricted to one row of the presentation table, candidate 8 is correctly identified as being part of the optimal solution by 15.09%, whilst the frequency of the suboptimal candidate 7 is 6.67% . T2 also displays a noticeably higher frequency for candidates number 1 and 2. Candidate 1 is similar to candidate 7 in the sense that he yields a high utility when the TC factor is ignored while candidate 2 belongs to the team with the highest variety ratio. This comparison of the treatments indicates that, although T1 is perceived as less complex, the overall frequencies in T2 display that the three most selected candidates are those belonging to the utility-maximising team.

Figure 34: Overall Candidate Distribution - T1 versus T2 (A)

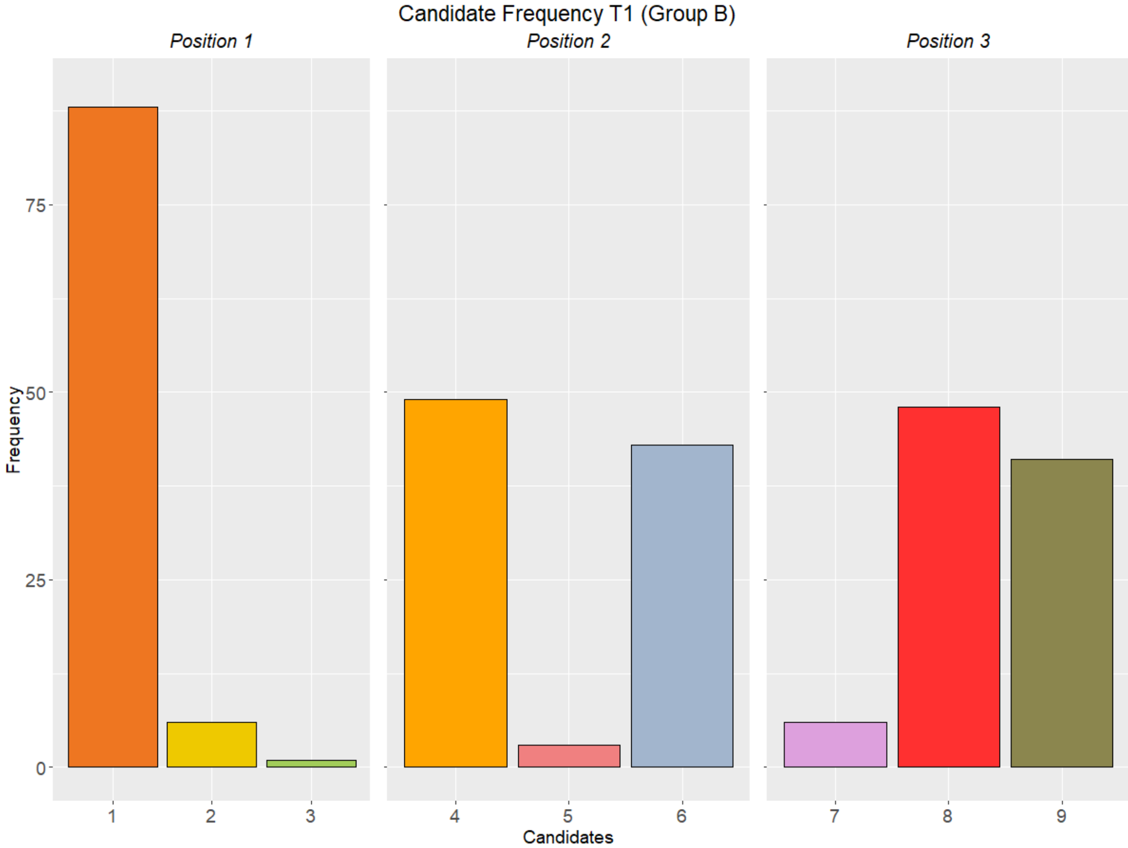


(3) <u>Candidate Selection – Group B</u>

The creation of group B involved a rearrangement as well as slight adaptation of the candidates in the presentation table, so that e.g. candidate 1 does not refer to the same skill values in both groups. For this reason, the decision behaviour for the candidate selection in the two groups needs to be presented separately. The utility-maximising team constellation for this group is given by the candidates number 1, 6, and 8.

The segregate T1 of group B reports a more dispersed candidate distribution than it is found for the treatment of group A. Due to the restrictions on the possible team constellations generated by the pre-allocation of candidates, the highest possible diversification ratio for T1 of group B can be achieved through the creation of three possible teams. The team constellations of candidate 2, 6, and 7, as well as candidates 2, 6, 8, and the team of 2, 6, and 9 all yield a variety ratio of 91.25%. The bar chart (figure 35) describes a clear dominance of candidate 1 for the first position, with 92.63% of respondents correctly identifying the candidate as most beneficial. Respondents' choices for the second and third position are of a less straightforward nature. Position two displays a preference for candidate 4 (51.58%) followed by the utility-maximising candidate 6 (45.26%). Again, a likely explanation for this behaviour is that respondents did not

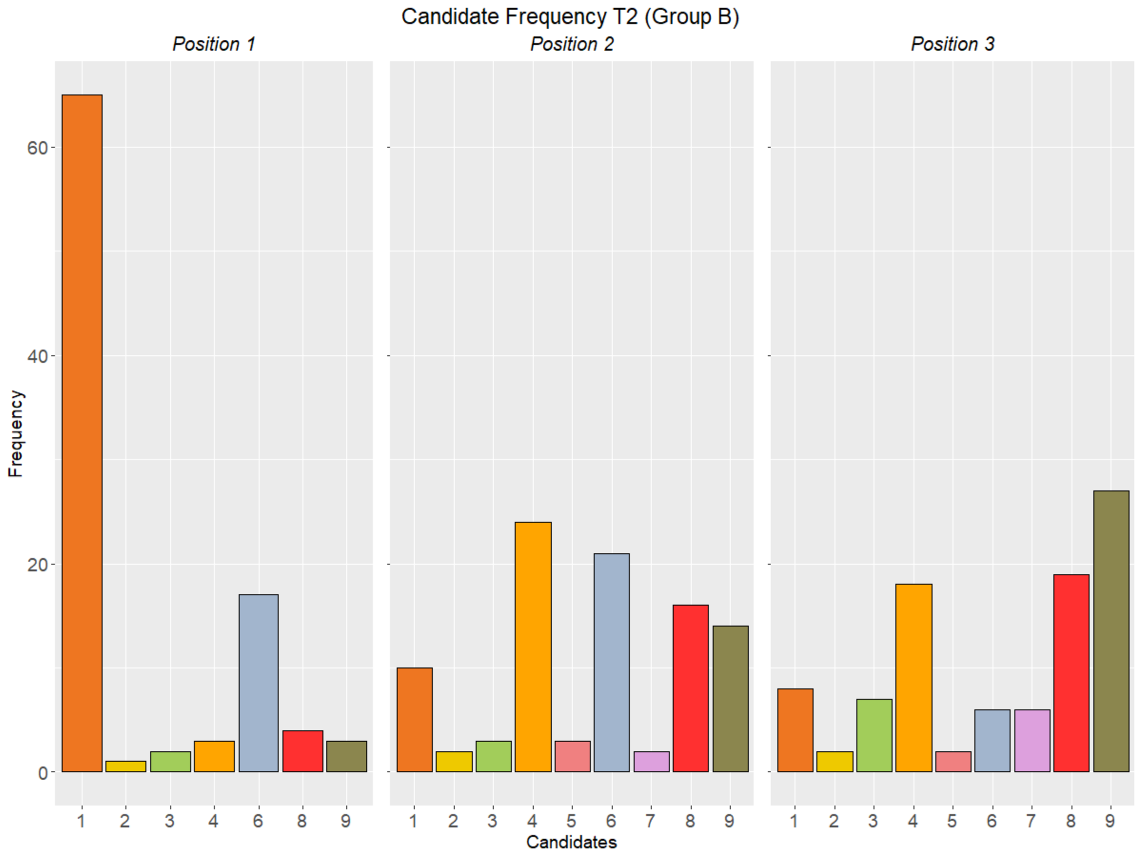
Figure 35: Candidate Distribution for each Position - T1(B)



correctly consider the TC-factor in their evaluation procedure. If the TC-factor of the candidates is neglected, candidate 4 would excel candidate 6 with a utility of 220 to 160. However, when the candidates are assessed accurately, candidate 6 yields a utility of 120 compared to 99 for candidate 4. Regarding the last position, the majority of 50.53% selected the superior candidate 8, whilst a large fraction (43.16%) erroneously prefers candidate 9. As one of the variety-maximising teams contains candidate 9, the source of this choice may be overdiversification according to the diversification bias. In this sense, participants exceedingly focus on variety and fail to consider the overall utility of the team. Alternatively, the source of the decision error may be a faulty candidate evaluation, being that candidate 9 yields the highest skill utility but at the cost of a low team compatibility, which reduces the candidate utility to the worst for position 3. Moreover, this candidate yields the lowest utility of all nine options in group B. Thus, when the TC-factor is ignored by participants, candidate 9 would dominate the other options for this position.

Turning towards T2, the bar chart (figure 36) continues to show a distinct dominance of candidate 1 for the fist position (68.42%), which is followed by the preference for another utility-maximising candidate, namely candidate 6 (17.89%). The candidates number 5

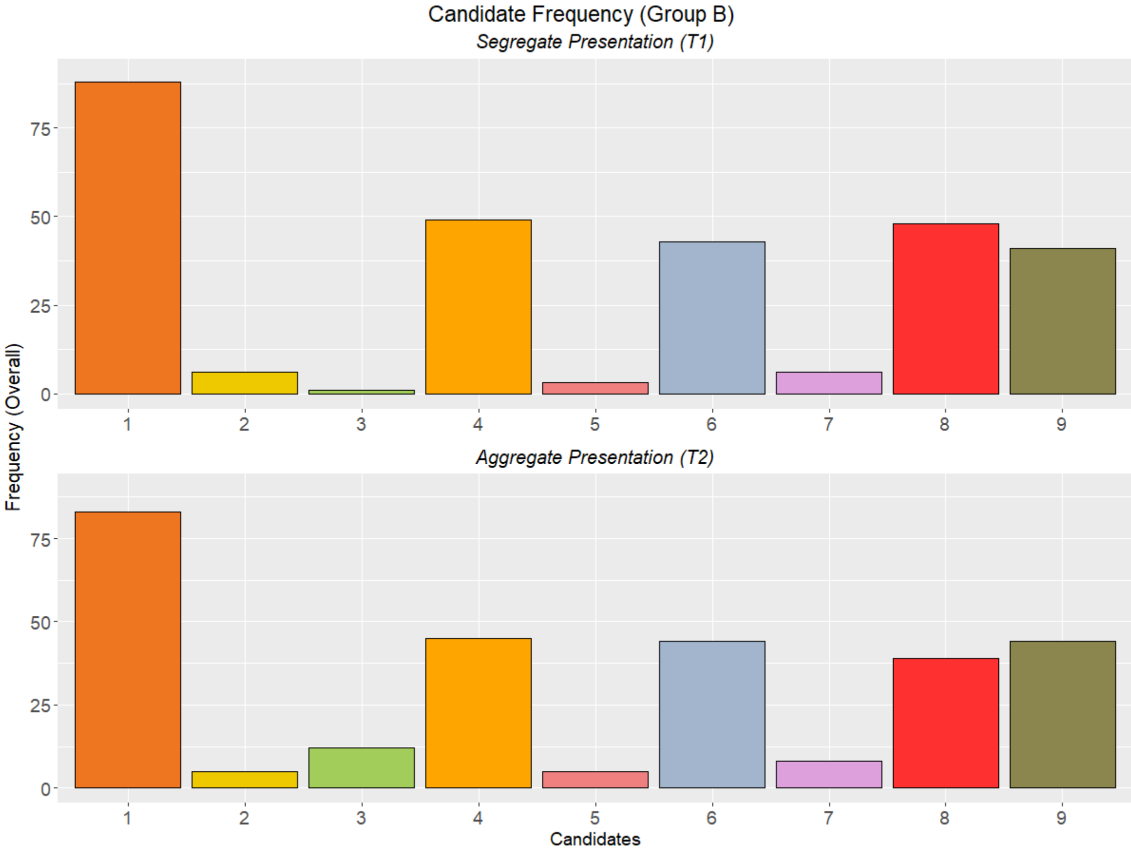
Figure 36: Candidate Distribution for each Position - T2 (B)



and 7 are never chosen to fill the first position. Similar to the findings for T1, the choices for the other positions are more distributed. The candidate favoured by the majority of participants to fill the second vacancy is candidate 4 (25.26%), which is not part of the maximum utility team. However, this candidate would be part of the team with the highest variety ratio of the aggregate treatment, which is formed by candidates 2, 4, and 6. The following most frequent selections are candidate 6 (22.11%) and 8 (16.84%), where both maximise the utility while candidate 6 additionally contributes to a high diversification. As discussed in the preceding paragraph, the notable frequency of candidate 9 (14.74%) may originate from an evaluation error, being that this candidate does not belong to the variety-maximising team of this treatment. With respect to the third position, candidate 9 gains in popularity amongst respondents so that 28.42% are likely to be oblivious to the remarkable low TC-factor. Candidate 8 (20.00%) and 4 (18.95) are the next most frequently observed choices, where candidate 8 is of utility-maximising character while candidate 4 maximises the variety ratio.

A direct comparison of the overall frequencies in both treatments (figure 37) reveals a rather similar pattern. The recorded relative selection frequencies for candidate 1 (30.88% in T1; 29.12% in T2), as well as candidate 4 (17.19% in T1; 15.79% in T2) and option

Figure 37: Overall Candidate Distribution - T1 versus T2 (B)

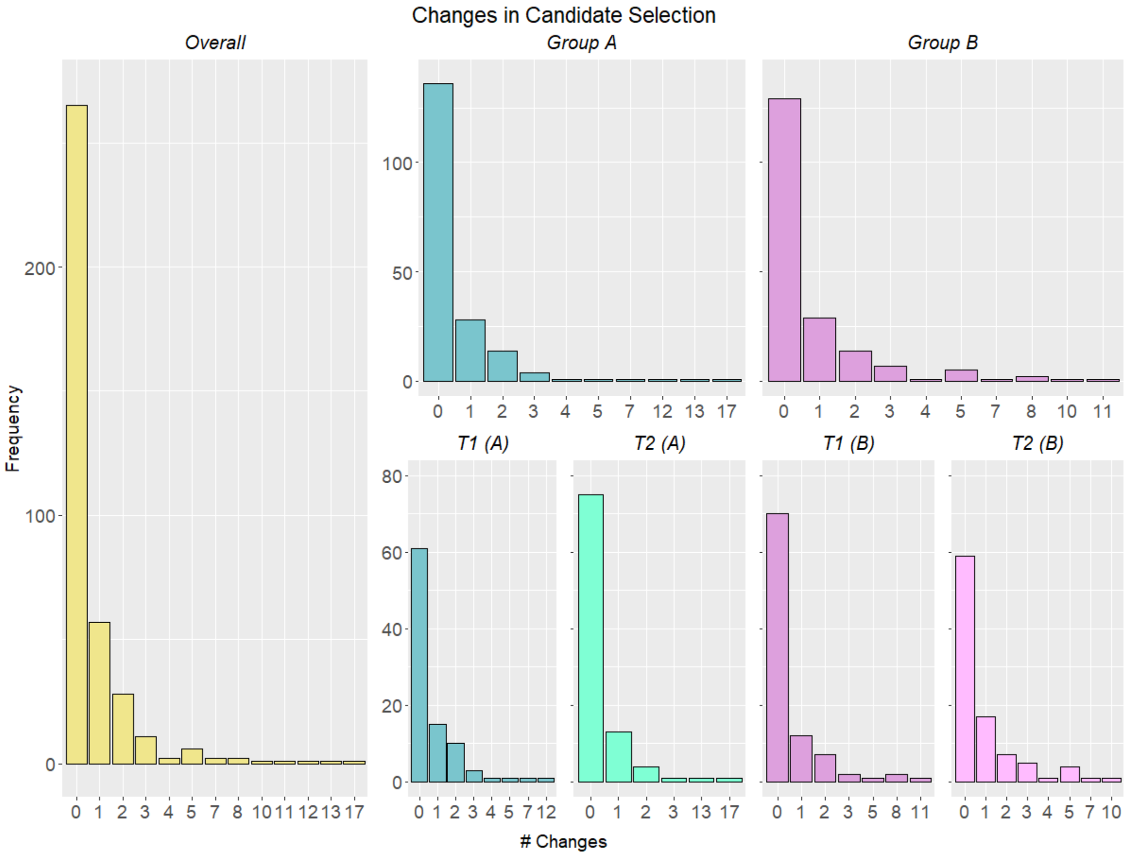


number 6 (15.09 in T1; 15.44 in T2) show no pronounced difference between the treatments. However, a difference is observed in the preferences for candidates 8 and 9. With respect to utility, more respondents in the segregate T1, correctly identify candidate 8 (16.84%) as beneficial compared to candidate 9 (14.39%). For the segregate T1, candidate 9 is also part of one of the most diversified teams (candidates 2, 6, 9), which might explain the high selection frequency. However, the frequency of candidate 2 is rather low so that diversification by means of this team constellation is not causing the frequent selection of candidate 9. In contrary, more individuals in the second treatment fall prey to the seemingly high-utility candidate 9 (15.44%), although it is not included in the maximum-variety team of T2 (given by 2, 4, 6). In consequence, the high frequency may be explained by an error in the evaluation procedure. However, 13.68% of subjects correctly incorporate the TC-factor and choose candidate 8 instead. In this sense, the more complex T2 displays a slight increase in the rate of faulty choices as candidate 9 is chosen over number 8. With respect to group A, the findings for T2 show an opposite general tendency. In the sense, the overall frequencies for the optimal candidates prevailed in T2 instead of the objectively less complex T1, which is also perceived as being of lower complexity. For both treatments of group B, the three candidates with the highest frequencies would not result in the maximum-utility team (candidates 1, 6, 8).

(4) Changes in Decision-Making

During the experiment, a variable traced any changes that participants made to the three entry fields for their candidate selection choices. The bar charts in figure 38 depict the frequencies for the total number of changes that were recorded for the three positions on the overall, group, and treatment level of the experiment. The majority of subjects (70.11%) did not change their decision on any of the three positions once it was entered. However, 15.08% modified their candidate selection once while 7.41% did so twice and 11 subjects (2.91%) undertake three adaptations. All other numbers of changes occur at a low frequency. Those between 10 and the maximum of 17 changes are performed by one participant each. The behaviour on group level is rather similar to each other, with the majority of participants not changing their selection (72.34% in A; 67.89% in B) and some modifying their decision once (14.89% in A; 15.26% in B). Group B includes more selection changes of a higher magnitude but the three highest numbers of 12, 13, and 17 changes are recorded in group A. Note that not only the number of changes was recorded but this measure was accompanied by a string variable storing the sequence of the entered

Figure 38: Number of Decision Changes in the Portfolio Building Task



candidates. Based on the inspection of this sequencing, it can be excluded that the changes of high magnitude are the result of repeatedly entering the same candidate. Continuing with the examination of the grouped bar charts for the individual treatments, one can state that the general trend of not changing the selection persists in both treatments of both groups. Nevertheless, more participants of T1 in group A changed their selection at least once (34.41%) compared to their colleagues in the aggregate T2 (21.05%). This finding is opposite to what would be expected from the less complex treatment. On the other hand, it may be explained by the pre-allocation of candidates that restricts the attention of the DM to one row of the presentation table at a time. When the DM then reaches the final row of the table and perceives his choices as a team constellation, he might be more inclined to modify single selections compared to a DM in the aggregate treatment, who has to consider the whole presentation table for each selection. However, group B yields the opposite finding, where 26.32% of individuals in T1 modify their selection at least once while this is done by 37.89% of subjects in T2. This might indicate that the rearranged candidate table of group B facilitates decisions in the segregate T1 to such an extent, that adaptations are not perceived to be necessary for the majority of participants.

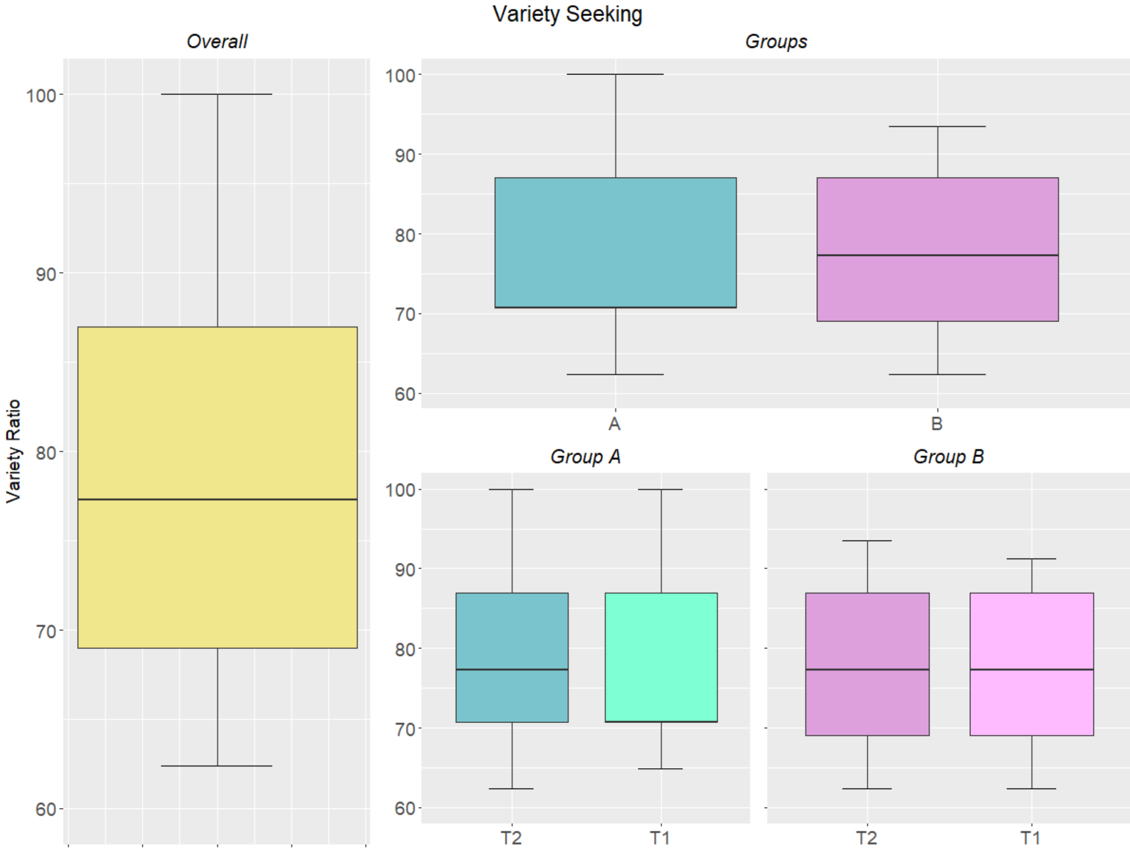
In general, the changes variable may provide insights into the bracketing behaviour and quality of the solution in the latter part of the statistical analysis.

(5) Variety Seeking Behaviour

As introduced in the introductory part to chapter 5.3, the variety seeking behaviour of subjects is measured by two variables. The first one yields the variety ratio of a built team, which describes how successful a given participant diversified their selected candidates compared to the team constellation with the maximum possible variety in skill levels. Figure 39 displays the boxplots for this ratio variable for the overall sample, as well as for the two groups and their respective treatments. The motivation behind the grouping procedure is the stepwise inspection of the patterns in the data going from a holistic level down to each treatment.

Taking an overall perspective, respondents built teams with a variety ratio ranging from 62.38% to 100%. The least diversified team in group A would have yielded a variety ratio of 50.35% (candidates 1, 7, 9), while the minimum possible degree of diversification for group B lies at 50.30%. Therefore, the teams with the lowest variety ratios as well as all other teams up to a ratio of 62.38% were not constructed by any participant. Note that the

Figure 39: Diversification in Team Building – The Variety Ratio Measurement



least diversified team does not represented a variety ratio of 0% for the reason that each buildable team has at least some degree of variety. A team of 0% variety would only be possible by hiring the same candidate for all three positions. Half of the participants achieved a variety ratio of 77.27% or higher with the top 25% of participants lying above a ratio of 86.99%. When grouping the graph according to the two versions, a noticeable difference in the median as well as range of the diversification performance is revealed. Group A covers the range up to the most diversified team, whereas group B reaches a maximum of 93.52%. At this point, it is important to remember that, in order to address the theory of the diversification bias, the teams providing the maximum variety ratios are not the ones that yield the highest utility. The maximum-utility teams in both groups are described by a variety ratio of 86.99%. Nevertheless, both utility-maximising teams are classified as diverse according to the 65% per item cut-off criterion. In consequence, group A may have been more prone to overdiversification, being that group B does not build teams above 93.52% diversification. The median of group A is equal to its first quartile at a level of 70.74%, so that 50% of respondents diversified to an equal or lower extent. For group B, although subjects may not cover the whole scale of the variety ratio, half of the participants construct teams of up to 77.27% variety. Hence, group B exhibits more diversification in their behaviour whilst at the same time mostly refraining from overdiversification. When inspecting the grouped boxplots for the treatments of each group, it becomes apparent that the difference in the median between the groups originates from the segregate T1 of group A. For this group, the pre-allocation in T1 may have distracted participants from diversifying their team so that they failed to consider the single candidate choices as a team.

The second dummy variable reports whether a respondent built a diverse team with respect to the 65% criterion, which was explained in the beginning of chapter 5.3. The bar chart below (figure 40) illustrates the proportion of diverse and non-diverse teams for the overall sample, as well as groups and their treatments. In general, 243 (64.29%) subjects did not diversify to such an extent that fulfils the introduced criterion. Regarding the group level, the share of non-diverse teams is almost identical for both groups (62.76% in A; 65.79% in B) so that the slight changes of the presented candidates did not influence the diversification behaviour. This picture continues for the inspection of the individual treatments of the groups, where only minor differences can be found. This consistent finding implies that the presentation form of the treatments does not affect

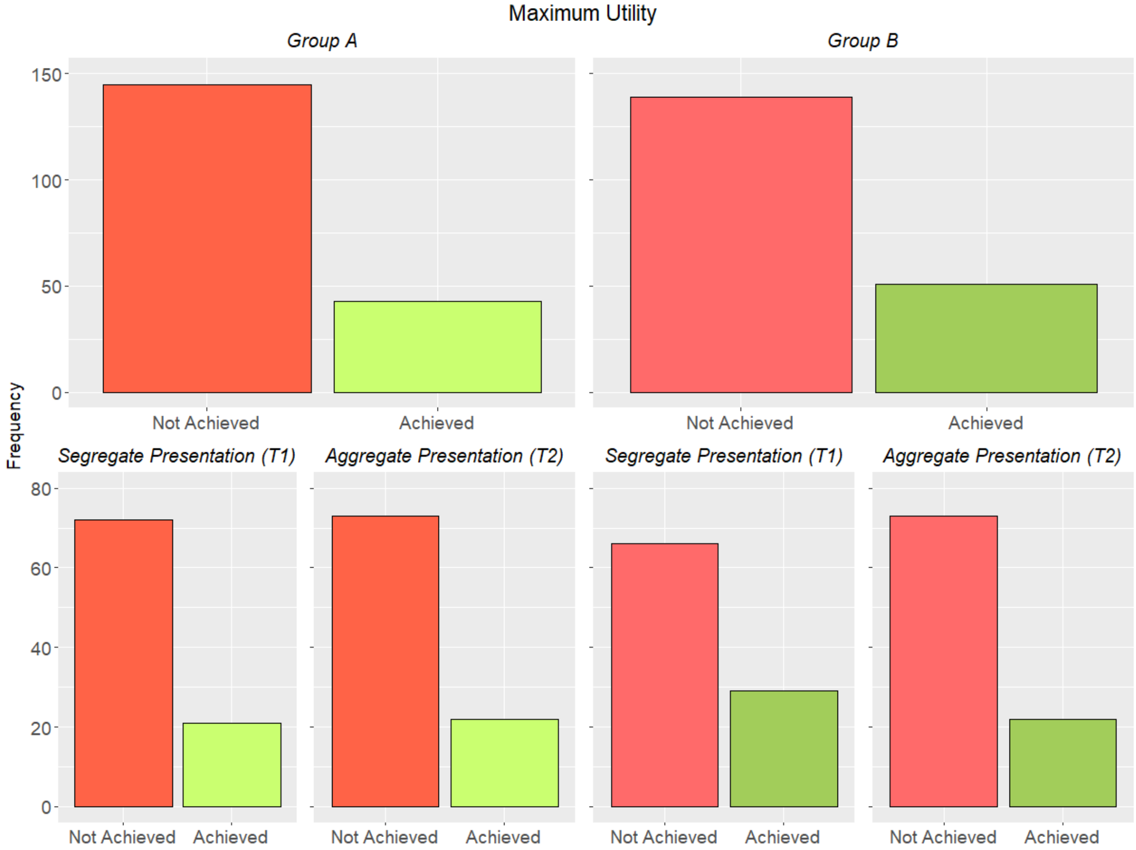
Figure 40: Diversification in Team Building - The Classification as Diverse

whether a diverse team is built. However, the boxplot for the variety ratio variable above (figure 39) visualised that the treatments for group A influence the degree to which a built team is diversified. For that measure, the segregate treatment is found to induce a lower diversification rate. Nevertheless, the treatments of group B did not record such a difference, which may be the result of the adaptations in the candidate table. Taken together, neither the groups nor the presentation forms seem to impact participants' ability to build a team that classifies as diverse. However, the presentation form may influence the general level of observed variety ratios in dependence of the arrangement of candidates in the presentation table.

(6) Utility Measurements for the Constructed Teams

There are three variables that describe the utility that subjects achieve in their portfolio building task. The first two are dummy variables, where one states whether a participant reached the maximum utility whereas the second dummy assigns participants to one of four utility categories defined below. This categorization is introduced in order to provide an enhanced overview of the distribution across the sample. The last variable reports the utility achieved by respondents in a metric form.

Figure 41: Frequency of the Optimal Team – Portfolio Building Task



Regarding the maximum utility, figure 41 portrays that the majority of respondents do not find the optimal team constellation with only 43 (22.87%) subjects of group A and 51 (26.84%) of group B reaching the maximum value. This finding also indicates that the adapted candidates on the group level do not constitute a major influence in the likelihood that the optimal solution is found. The individual treatments convey a similar impression, with no notable difference between the treatments of group A and only a minor difference in group B. For the second group, playing the segregate T1 increases the frequency with which the optimal team constellation is created from 23.16% in T2 to 30.53% in T1. This is consistent with the difference in perceived and objective complexity for the treatments in the sense that the less complex T1 facilitates finding the optimal solution.

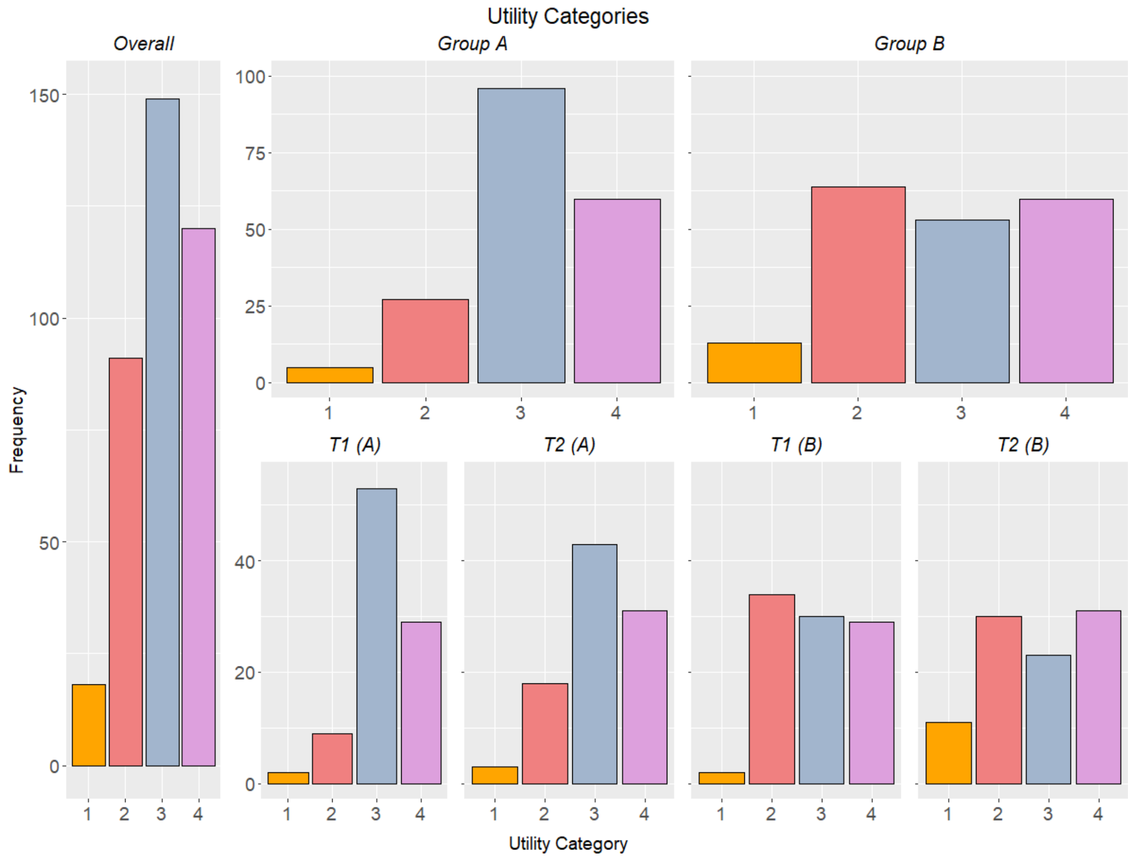
For the purpose of providing a more detailed impression of the distribution of participants' generated utilities, four utility categories are created. The categorisation also offers the opportunity to not solely focus on the influence of the groups and treatments on the ability to find the optimal solution but to investigate the general trend in achieved utilities. The categories are defined relative to the maximum possible utility in the form of the following intervals:

- Category 1: [0%; 70 %]
- Category 2:]70%; 80%]
- Category 3:]80%; 95%]
- Category 4:]95%; 100%]

The top category is not restricted to the optimal solution for two reasons: first, identifying the optimal team constellation is already described by a separate variable. More importantly, the participants of the experiment are not perfectly rational DMs. In consequence, one may not expect them to always find the optimal solution and utility category 4 broadens the perspective to those who achieve a high performance instead.

The distribution for the overall sample (figure 42) reveals that the majority of respondents (39.42%) is situated in category 3 but a substantial fraction of 31.75% belong to the top utility category 4. The trend of most participants scoring within the range of category 3 is even more pronounced for group A, where 51.06% of subjects are assigned to this category. Group B is of a more dispersed nature with utility category 2 being the most frequent (33.68%) but closely followed by category 4 (31.58%) and 3 (27.89). In comparison to group A, group B also includes an increased share of category 1

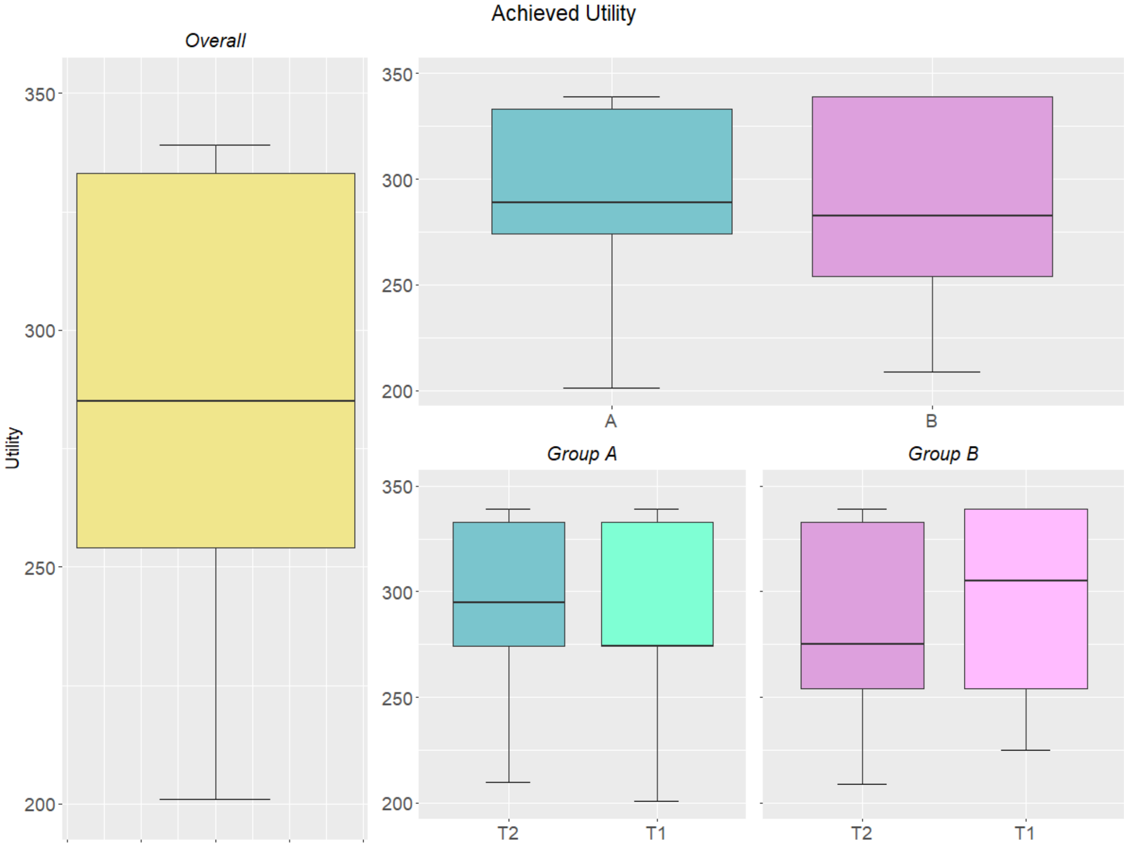
Figure 42: Distribution of Utility Categories – Portfolio Building Task



participants (2.66% in A; 6.84% in B). Proceeding to the treatment level, one may note a difference between the treatments of group A. Participants in the segregate T1 are most likely to belong to category 3 (56.99%), followed by the top category (31.18%) and only a few participants in the lower utility categories. Although subjects in the more complex T2 are also most frequently classified into category 3 (45.26%), the gap to category 4 is considerably smaller (32.63%) and the frequency of category 2 increases from 9.68% in T1 to 18.95% in T2. The implication is that the more complex presentation form leads to a slightly increased frequency of top-utilities but at the same time leads a greater fraction of individuals to build teams of lower utility. On the other hand, the less complex T1 appears to facilitate reaching one of the higher utility categories with most individuals scoring in category 3. While the treatments did not show a considerable effect on finding the optimal solution, they suggest a difference in the general trend of achieved utilities. In this sense, decision support in the form of the segregate presentation facilitates a higher performance for group A. The treatments of group B display a less pronounced difference. As in group A, slightly more individuals of T2 (32.63%) belong to category 4 compared to T1 (30.53). Nevertheless, the share of respondents in category 3 and 2 is noticeably higher for T1 (31.58% in C3, 35.79% in C2) than for T2 (24.21% in C3; 31.58% in C2). Moreover, the more complex T2 reports a distinct increase in the share of category 1 respondents (11.58%) compared to T1 (2.11%). Thus, the trend for group B reflects the finding on group A, although the general pattern in the distribution is of different nature for this group. Overall, the decision support implemented in the segregate treatment appears to support a higher performance in decision-making.

The final variable that addresses the utility of the constructed teams directly records the achieved utilities gathered during the experiment. The corresponding boxplots in figure 43 again describe the overall sample, as well as the two groups, and the individual treatments. The achieved utility of participants ranges between 201 and a maximum of 339, which is given by the optimal team constellation in both groups and all treatments. The distribution is slightly skewed to the left (-0.1) so that the majority of data entries are located in the middle to upper section of the scale. In fact, 50% of subjects range between a utility of 254 (Q1) and 333 (Q3) and the median of the whole sample lies at 285. Proceeding to the group level, the first differences between the groups become visible. What was already outlined by the utility categories is that group B is more dispersed than group A. While 50% of observed utilities in A are found in the range of 274 up to 333,

Figure 43: Achieved Team Utility – Portfolio Building Task



half of the participants playing group B range between a utility of 254 and the maximum of 339. For the achieved utility, the treatments of both groups display recognizable differences. The treatments of group A are described by identical first and third quartiles (Q1 = 274; Q2 = 333) but the median of T2 lies at a utility of 295, whilst this measure for T1 coincides with its Q1. This means that 50% of individuals in T2 achieve a utility of at least 295, whereas half of respondents in T1 achieve at least 274. Since the median and Q1 coincide, 25% of all subjects achieve exactly 274. Therefore, the observations in T2 are concentrated at a higher utility level compared to T1. However, one needs to be aware of the circumstance that the pre-allocation of candidates in T1 does not allow for all 84 team constellations and consequently also restricts the variety in achievable utilities. Group B of the experiment also reports identical first quartiles (254) but the third quartile of T1 lies at the maximum utility of 339, while T2 is located below at 333. In comparison to group A, the median of T1 (305) in group B is considerably above T2 (275). It follows that half of the respondents in the less complex treatment are located in the upper range of utilities while the upper half of their colleagues in T2 cover a broader area. This finding supports the suggestion that the segregate T1 enhances the performance of participants in group B. However, the findings for group A delivered the opposite picture so that the upper half of T1 players are more dispersed while those in T2 are concentrated in an area of higher utility. This observation may be attributed to the different candidate tables of the two groups. In this sense, the rearranged candidate table in group B may facilitate the recognition of well-performing candidates. Nevertheless, the groups did not affect the rate at which the optimal team was identified, which was illustrated by figure 41. Overall, the emerging picture is that the presentation form does not necessarily facilitate building an optimal team but does impact the proportion of respondents who achieve a higher quality solution. When comparing the mean achieved utility (332.99) of the group of participants who bracketed broadly and thus achieve a diverse team with the mean (270.17) of the narrow bracketing group, the difference between the two measures is significant (W = 32039, p-value < 2.2e-16) and can be interpreted as the cost of narrow bracketing. For the collected sample, this cost of narrow bracketing is an average loss in utility of 62.82.

5.3.2 Portfolio Selection Task

The second task that participants face in the personnel selection part of the experiment is the portfolio selection task. Since this task is played in a within-subjects design, the results will only be presented at the overall and group level. The purpose of the portfolio selection task is comparable to the combined lottery task, which concluded the lottery part of the experiment. The aim of that lottery lies in providing decision support by explicitly combining the choice alternatives and thereby highlighting the inferiority of the most frequent choice pair A&D. Similarly, the portfolio selection task introduces a greater degree of decision support, compared to the pre-allocation in T1 or no support in T2. This is achieved by presenting three already formed teams to choose from along with reducing the skill item list of each team member to stating the sum of their skill values. This also implies that, like for the single lotteries, only one decision is asked from participants so that the task is free from bracketing behaviour. Furthermore, the setting of choosing a pre-built team without having information on the distribution of the individual skill values amongst the members excludes variety seeking from this task. Consequently, participants are not subjected to a diversification bias so that their attention is directed towards the resulting team utility. This subchapter will proceed by first presenting the observed selection behaviour of respondents followed by elaborating on the corresponding utility measurements.

(1) Team Selection

Figure 44: Distribution of the Team Selection Choices

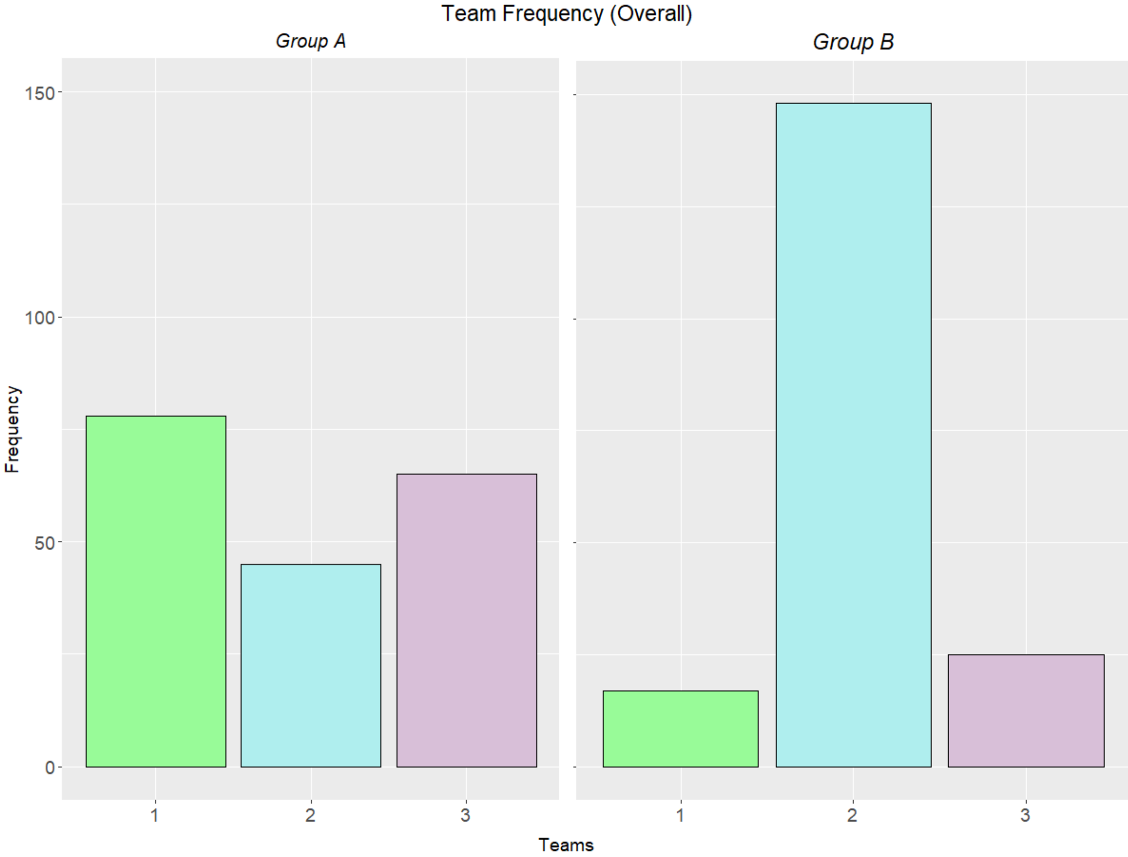
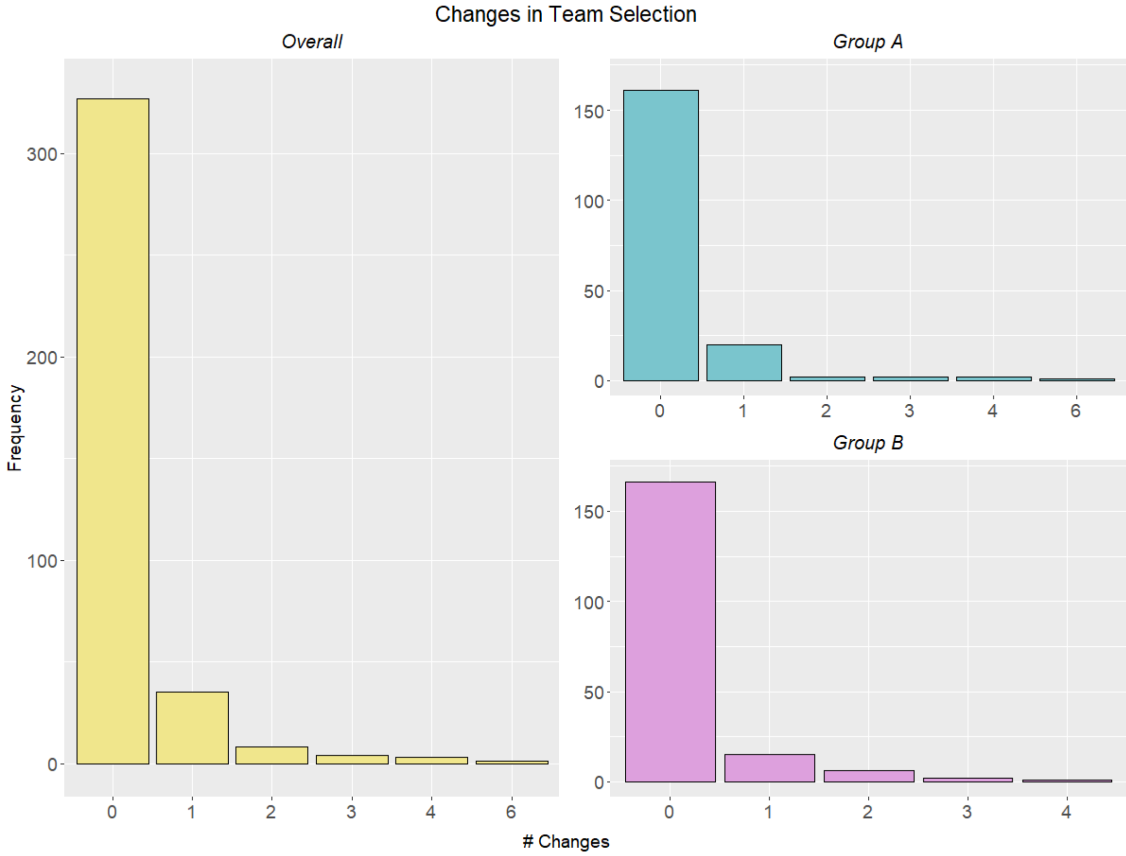


Figure 44 presents the observed selection frequencies in the portfolio selection task for both groups. The behaviour in both groups shows a clear difference in its pattern, where group A reports well distributed team frequencies whilst group B is characterised by the distinct dominance of one team. Regarding group A, the most frequently selected team is team 1 (41.49%), which yields the highest utility (314) of the three options and thus represents the optimal solution for this group. This is followed by a high preference for team 3 (34.57%) and team 2 (23.94%) being the least frequent choice. However, providing a utility of 294, team 2 yields a higher utility compared to team 3 (276) and should therefore be observed more frequently. The distinct dominance of team 2 in group B suggests that the large majority of 77.89% of respondents are able to identify the optimal solution, which is given by the generated utility of 312 from team 2. A considerably lower fraction (13.16%) select team 3 and only 8.95% choose team 1. This sequence coincides with the utility order of the teams, where team 3 yields a utility of 291 and team 1 provides 273. Comparing these results to the frequencies of finding the optimal solution in the portfolio building task, it can be noted that the supporting design of the present task enhances the likelihood that participants identify the dominant team constellation. The results of the portfolio building task showed that 22.87% of group A and 26.84% of group B select the candidates that together yield the maximum utility. Hence, the portfolio selection task increases the success rate to 41.89% (group A) and 77.89% (group B). The gained picture is consistent in the sense that group B already performed slightly better for the tasks with less decision support and again outperforms group A when given enhanced supporting measures. This supports the emerging impression that the adapted candidate table of group B facilitates the recognition of the superior candidates.

(2) Changes in Decision-Making

As it was implemented for the portfolio building task, a variable traced any modifications that subjects performed on their entry field for their team selection choice. Figure 45 visualises the number of changes that participants made to their decision for the overall sample as well as on the group level. Consistent with the findings for the building task, the large majority of 327 (86.51%) participants do not adapt their decision once it is entered into the entry field. A considerably smaller fraction of 9.26% changes their team selection once while only 2.12% do so twice. The remaining observed number of changes are performed by only a few participants and the maximum recorded number is one

Figure 45: Number of Decision Changes in the Portfolio Selection Task



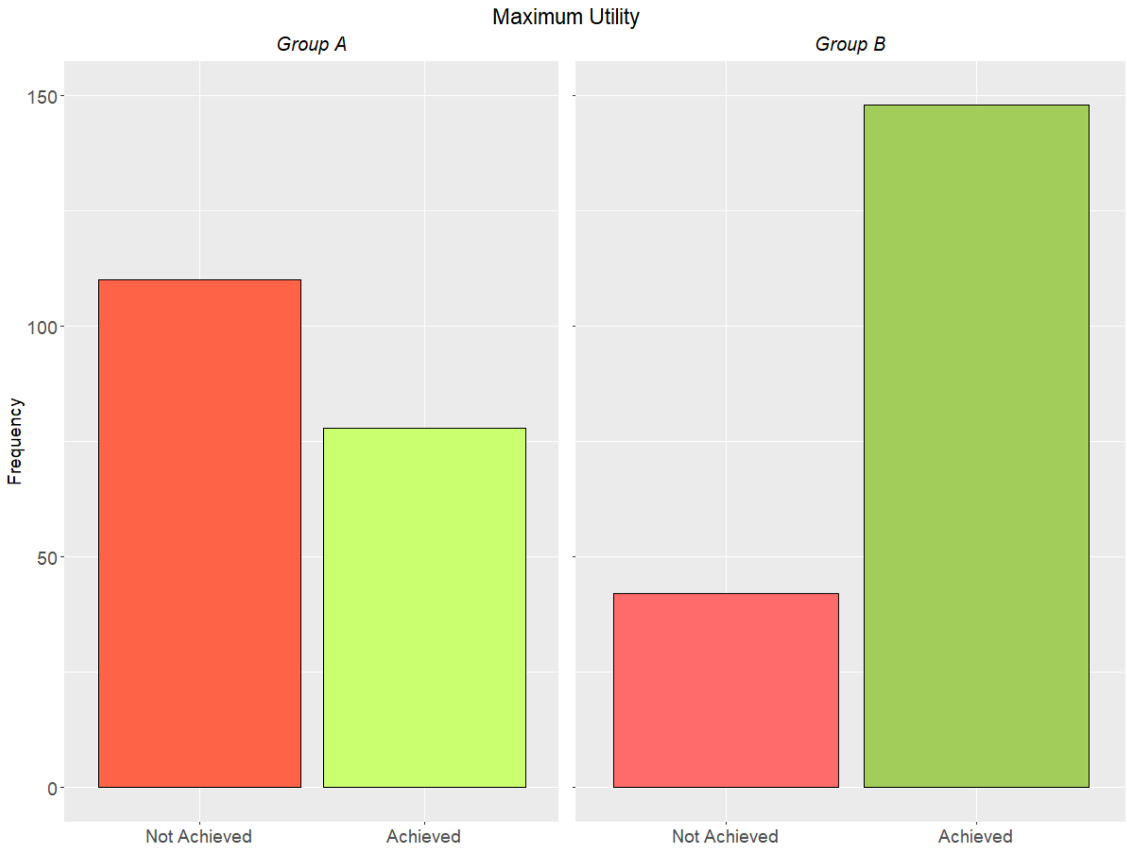
individual changing their selection 6 times. At the group level, both groups convey a similar picture with 85.64% of subjects not modifying their choice in group A and 87.37% keeping their first entry in group B. The circumstance that slightly less subjects reassess their choice in group B may again hint at this group offering a small advantage in the arrangement of candidates. In comparison to the findings concerning the portfolio building task, it becomes clear that the maximum number of changes as well as the fraction of respondents who change their selection at least once is noticeably reduced in the portfolio selection task. Remember, that for the building task 72.34% of subjects in group A and 67.89% of group B did not perform any adaptations to their choice. Furthermore, the maximum number of changes in the portfolio building task was 17 in group A and 11 in group B. The resulting implication is that the reduced complexity due to incorporating stronger decision support into the task design, causes subjects to be more confident of their first selection. Even in cases where changes are undertaken, the number of such alterations is considerably smaller and the maximum of observed changes for the portfolio selection task is 6 (in A) and 4 (in B).

(3) Utility Measurement for the Selected Teams

Since participants choose between three pre-built teams, they may only achieve one of three possible utilities. Consequently, the bar chart presented in the paragraph on the selection frequencies (figure 44) also represents the distribution of achieved utilities, with team 1 being the optimal choice in group A and team 2 yielding the maximum utility for group B.

Overall, 59.79% of respondents select the dominant team but the groups show a clear difference in the success rate of their players. Reflecting the selection distribution discussed before, figure 46 illustrates the proportions of participants who recognise the utility-maximising team. The majority of 110 individuals (58.51%) in group A reaches a solution of inferior quality whilst only 42 subjects (22.11%) select a dominated team in group B. Hence, the two groups provide an opposite picture, which fortifies the impression that group B consistently performs better than group A. This is contrary to the intention of creating both groups as equal as possible for comparability. However, compared to the success rates observed for the portfolio building task in chapter 5.3.1, both groups show a noticeable improvement in their ability to achieve an optimal solution. Taken together, the reduced objective complexity of this task seems to have a profound

Figure 46: Frequency of the Optimal Team – Portfolio Selection Task



impact on the performance of participants. This is especially interesting when one takes the perceived complexities into account. Remember that the segregate T1 of the building task is experienced as only marginally more complex than the portfolio selection task, so that one would expect similar results in the frequency of successfully solving the task. Only 22.58% of subjects in group A and 30.53% in group B achieve the maximum possible utility in T1 of the portfolio building task. However, 34.41% of subjects who played T1 of the building task (group A) and 77.89% of those assigned to group B now select the optimal solution.

5.4 Broad Bracketing Across Parts

Having presented the findings on the four lottery tasks as well as the data gathered from the personnel selection tasks, this subchapter concludes the descriptive section of the results chapter. Its purpose is to summarise the observed bracketing behaviour across the two experimental parts before chapter 5.5 continues with the results on statistical testing.

Regarding the lottery part of the experiment, the double lottery as well as the consistency lottery are the two tasks were bracketing behaviour is a relevant factor. Broad bracketing for both lotteries is represented by behaving opposite to the predictions of PT. This means choosing the gamble over the sure gain in the domain of gains and the sure loss over the gamble for a negative EV lottery. For both tasks, this behaviour is observed when the choice combination of B for the first decision and C for the second one is selected. Only 6.08% of respondents paired these two options in the double lottery while 5.56% did so in the consistency lottery. The majority of 65.34% in the double lottery and 55.82% in the consistency lottery behave in compliance with PT and select A&D. The remaining fraction decided for another choice combination.

Advancing to the personnel selection part, the two treatments of the portfolio building task involve bracketing behaviour. For these tasks, broad bracketing refers to diversifying the skill distribution of the selected candidates, so that the resulting team scores highly with respect to each of the four skill items. In compliance with the theory on the diversification bias (Read and Loewenstein, 1995), broad bracketing in this part is subjected to the possibility of choosing an excessive degree of variety. In this sense, the team that provides the maximum variety ratio is not the team that yields the highest utility. Therefore, the measure for broad bracketing is not whether a participant achieved the maximum utility but whether a team that classifies as diverse is built. Since the utility-maximising team is also diverse according to the introduced 65% criterion, the diverse variable also accounts for all respondents who find the optimal solution. For that reason, subjects who identify the optimal team constellation are broad bracketers in the sense that they diversify their team but do not let the diversification bias distract them from the overall objective of maximising utility.

For the overall sample, 135 individuals (35.71%) are found to be broad bracketers by building a diverse team while 94 (24.87%) of these subjects also achieve the optimal solution. Compared to the lottery tasks, these findings represent a considerable increase in the proportion of broadly bracketing individuals. It follows the implication that the lottery and personnel parts of the experiment are sufficiently different in their design to provoke different decision-making behaviour in a noticeable fraction of the sample. One possible explanation may be given by the circumstance that the lottery tasks are of a rather analytical nature and refer to the area of gambling in an obvious manner. In contrary, the personnel part is set in the context of personnel selection, where the gambling component of not knowing which skill set will be required by future projects is less obvious. Furthermore, although the tasks of the lottery part are phrased to fit the context of personnel selection, the personnel part has been given a more detailed storyline. Hence, the context in which a decision is set seems to noticeably impact the bracketing behaviour

Table 13: Broad Bracketing Across Experimental Contexts

		BB in Lottery Part			
		Double Lottery	Consistency Lottery	Neither	Both
BB in Personnel Part	No	9 (2.38)	8 (2.12)	221 (58.47)	5 (1.32)
	Yes	5 (1.32)	4 (1.06)	122 (32.28)	4 (1.06)

*relative frequencies in parentheses

of a DM. This means that a single individual may switch between being a broad or narrow bracketer for different contexts.

Table 13 provides an overview of the observed broad bracketing behaviour across the two parts of the experiment. As it is to be expected from the examination of the experimental parts in isolation, the majority of participants (58.47%) bracket the lottery as well as the personnel part narrowly. This supports the argument that narrow bracketing is the default mode in decision-making. When comparing the rows of the personnel part for the double and consistency lottery, two observations emerge: First, from all participants who only bracket the double (14 subjects) or consistency lottery (12 subjects) broadly, approximately one third also exhibits broad bracketing behaviour for the personnel part. Secondly, this proportion does not hold for the 9 participants who are broad bracketers in both lottery tasks. Here, almost half of the broad bracketers for both lotteries are found to bracket the personnel part broadly as well. While 144 (38.10%) individuals bracket broadly in at least one task in one of the two parts but not in the other, an even smaller fraction of 13 (3.44%) subjects is able to transfer this behaviour to at least one task in both contexts. Consistent broad bracketing for all tasks of both contexts implemented in the experiment is only achieved by 4 (1.06%) respondents. This small fraction highlights that broad bracketing itself is a demanding task, which may be further hindered by changing decision-making contexts. However, changing contexts represent the environment that a DM encounters in his everyday life.

The most interesting observation that addresses the influence of context can be found amongst the respondents who bracket neither of the lotteries broadly. A large share of

32.28% of the sample fail to bracket the lotteries broadly but then switches to do so for the personnel part. In comparison, only 5.82% of subjects exhibit broad bracketing for at least one of the lotteries but then fails to adapt this to the personnel part. The difference becomes even more remarkable when one takes into account that the lottery part is of lower objective and perceived complexity than both treatments of the personnel part. This means that despite the personnel selection tasks being perceived as more complex, more subjects apply broad bracketing and diversify their team constellation. This further stresses that context seems to make a difference and that, with regard to the sampled respondents, the context of personnel selection supports broad bracketing behaviour while lotteries impede it.

5.5 Hypotheses Testing

Subsequent to the descriptive statistics of the relevant variables, the following part aims at presenting the results of the statistical testing and is divided into two subchapters. The first is concerned with the main hypotheses, whereas the second subchapter presents the results of the task-specific hypotheses. In this sense, chapter 5.5.2 offers a more profound perspective on the findings that relate to the experimental tasks, whilst chapter 5.5.1 seeks to investigate the relationships presented in the research framework at a more holistic level. Each of the two subchapters conclude with a table that summarises which hypotheses are rejected as a result of the gathered data.

5.5.1 Main Hypotheses

The following subchapter will cover the results regarding the first 15 hypotheses. For this purpose, the undertaken regressions on different dependent variables will be presented and discussed. In accordance with the flow of the research framework, first a look will be taken at the drivers of perceived complexity, followed by three models aiming to shed light on possible effects leading to a broad bracketing behaviour. An analysis of the influence of broad bracketing by means of diversification on the solution quality will be succeeded by a table showing the main hypotheses and if they have to be rejected or not.

(1) Perceived Complexity

Starting off with the perceived complexity, the research framework shows two major influences, namely the objective task complexity on one hand and the psychographics of a participant on the other. While the influence of the objective complexity of a task is trivial to explain, the psychographics, except for the impulsiveness, are hypothesised to influence the perceived complexity of an individual in different directions. While the numeracy and the rationality of a participant are believed to decrease his perceived complexity, the other psychographics are assumed to have the opposite effect with respect to the same task.

Since the experientiality is also expressed by three other psychometrics of the REI questionnaire, it is necessary to determine which variables are to be included in the following regression models in order to avoid multicollinearity. As the three variables of emotionality, imagination, and intuition provide more insight on the characteristics of a participant than the single variable of experientiality, it is decided to use them for the further analysis.

As it is depicted in figure 14 and 15 (chapter 5.1), the task complexity of the lottery part is lower compared to the personnel part and is also perceived as such. To analyse the drivers of perceived complexity, three different models are presented. In the first model ("Overall"), it is not accounted for the difference in task complexity across the two parts of the experiment. Besides the first model, the second model ("Lottery Part") includes only the reported complexity for the tasks of the lottery part and the third model ("Personnel Part") does the same with only the complexities reported in the personnel tasks. This gives the opportunity to take an overall perspective and compare it with a context-specific analysis. It was decided to use such an approach instead of a dummy variable for the two different parts in order to be able to not only see the influence of the context, but also how the single covariates change for different perspectives.

Table 14 presents the abovementioned three linear mixed effects models, which were built with different datasets that are all of a long format. A linear mixed effects approach with the participant-specific label as a random variable is chosen, for the reason that it allows to take out the differences between the individual entries of each participant and assess the psychographics separately. This analysis is conducted using the package lme4 (Bates et al., 2015) in the statistical computing environment R (R Core Team, 2020), and all regression tables are created with the stargazer package (Hlavac, 2018). The stargazer package also provides the calculation of p-values for random effects models.

The participant-specific label explains 220 of the predictors' variance in the overall model, 250 in the lottery model and 398 in the personnel model. These values show that the individual differences of the respondents have a considerable influence on the individually perceived complexity. The marginal and conditional R^2, as proposed by

Table 14: Regression Models - Perceived Complexity

Hierarchical Regressions on Perceived Complexity

	Dependent variable:		
	Perceived Complexity		
	Overall	Lottery Part	Personnel Part
Intercept	20.458*** (2.506)	20.169*** (2.659)	50.325*** (3.191)
Berlin Numeracy	-0.070 (0.701)	-0.504 (0.761)	0.829 (0.946)
Rationality	-0.170** (0.074)	-0.163** (0.080)	-0.184* (0.100)
Emotionality	0.098 (0.084)	0.098 (0.091)	0.099 (0.114)
Imagination	0.042 (0.082)	0.110 (0.089)	-0.090 (0.111)
Intuition	-0.041 (0.090)	-0.068 (0.098)	0.010 (0.122)
Stage (1)	18.788*** (1.430)	18.788*** (1.343)	
Stage (2)	26.964*** (1.238)	26.964*** (1.163)	
Stage (3)	29.336*** (1.430)		
Stage (4)	33.815*** (1.797)		5.489*** (1.518)
Stage (5)	44.257*** (1.790)		13.921*** (1.511)
Observations	2,268	1,512	756
Log Likelihood	-10,247.560	-6,811.445	-3,416.374
Akaike Inf. Crit.	20,521.110	13,642.890	6,852.747
Marginal R^2 / Conditional R^2	0.206 / 0.494	0.180 / 0.527	0.061 / 0.648
Note:			*p<0.1; **p<0.05; ***p<0.01

Values in parentheses are standard errors

Nakagawa and Schielzeth (2013), are two measures to judge model quality for linear mixed effects models. While the marginal R^2 describes the amount of variance in the dependent variable explained through only the fixed factors, the conditional R^2 also considers the random variables. Both R^2 values are calculated using the “tab_model” function of the package sjPlot (Lüdecke, 2020). The largest marginal R^2 (0.206) can be found in the overall model, which is not surprising since this model has the most fixed effect variables. However, at the same time this model also has the lowest conditional R^2 (0.494). This means that compared to the part-specific models, the effect of the random variable is weakest but still not ignorable. The personnel model has the biggest discrepancy between the marginal (0.061) and conditional (0.648) R^2. This leads to the assumption that, especially in the personnel part, the random effect explains a major part of the observed variance. Insofar it can be assumed, that the individuals themselves have

the highest influence on their perceived complexity and only a small part of their perceived complexity can accurately be predicted by the fixed effects stated in the model.

The participant-specific label explains 220 of the predictors' variance in the overall model, 250 in the lottery model and 398 in the personnel model. These values show that the individual differences of the respondents have a considerable influence on the individually perceived complexity. The marginal and conditional R^2, as proposed by Nakagawa and Schielzeth (2013), are two measures to judge model quality for linear mixed effects models. While the marginal R^2 describes the amount of variance in the dependent variable explained through only the fixed factors, the conditional R^2 also considers the random variables. Both R^2 values are calculated using the "tab_model" function of the package sjPlot (Lüdecke, 2020). The largest marginal R^2 (0.206) can be found in the overall model, which is not surprising since this model has the most fixed effect variables. However, at the same time this model also has the lowest conditional R^2 (0.494). This means that compared to the part-specific models, the effect of the random variable is weakest but still not ignorable. The personnel model has the biggest discrepancy between the marginal (0.061) and conditional (0.648) R^2. This leads to the assumption that, especially in the personnel part, the random effect explains a major part of the observed variance. Insofar it can be assumed, that the individuals themselves have the highest influence on their perceived complexity and only a small part of their perceived complexity can accurately be predicted by the fixed effects stated in the model.

Other than that, the linear mixed effects model and its coefficients can be interpreted in the same fashion as a linear regression, and the values in parentheses are standard errors. The intercept of the overall analysis as well as the lottery part incorporates the lowest objective complexity, the combined lottery, as base group. Since the combined lottery does not belong to the second experimental part, the intercept of the personnel part has the portfolio selection task (stage 3) as base group. All intercepts are highly significant and the difference between the lottery and the personnel part serves as a good example to once again highlight the difference in complexities across the two parts, being that the intercept of the personnel part more than doubles the one found for the lottery part.

The "Berlin Numeracy" variable is metric and describes how many of the four asked examples were answered correctly by participants. Although it is insignificant in all three models, it points in the hypothesised direction in the first two models but does not do so for the third one. The reason for the switch may lie in the nature of the tasks themselves. Remember, that numeracy describes the ability to process risk probabilities in a statistical manner and act on that information. In the Lottery part, where explicit probabilities are stated and understanding them is required to act on this information, people who score highly in numeracy also perceive the task as less complex. On the other hand, in the personnel part the different context of the problem as well as the design of the tasks does not highlight the underlying risk probabilities explicitly. Therefore, although numeracy itself could be a complexity reducing characteristic, participants may fail to detect the underlying risks and as a result perceive higher complexity. Altogether, this variable cannot be taken as a good measure for perceived complexity.

The “Rationality” variable is also metric and highly significant in all three models. This means that an individual with a higher tendency to operate in the rational processing system perceives a task as less complex compared to a less rational person. The three variables that represent experiential processing, namely “Emotionality”, “Imagination”, and “Intuition” are all insignificant. For this reason, it is not possible to formulate a general statement on how they influence the perceived complexity. This observation further strengthens the impression that the two contexts of the experiment are distinctly different. In this sense, imagination does not facilitate the processing of lotteries but reduces the perceived complexity in a personnel selection context. A possible explanation may lie in imagination facilitating the mental representation of different team constellations while this cannot be applied to the analytical gamble in a lottery task. Surprisingly, intuition shows an opposite effect by reducing the perceived complexity for lotteries while increasing it for the personnel selection. It may be argued that the design of lotteries as decisions between two options encourages participants to act more intuitively and as a consequence, to experience a reduced complexity. In contrary, the large candidate tables in the personnel tasks impede a quick intuitive processing, which may cause intuitive respondents to perceive a higher complexity.

Performing the same analysis with experientiality as a substitute for the three characteristics, an insignificant result can be observed. This may be the result of the opposite effects of the individual variables, so that an aggregate measurement yields an insignificant relationship. Thus, according to what was expected, the approach of examining the three variables that together refer to experiential processing provides an enhanced impression. Moreover, when comparing the new results with the abovementioned models, the AIC and also BIC clearly state that the model containing "Emotionality", "Imagination", and "Intuition" is preferable. Before finally committing to these models, it is necessary to control the variance inflation factors (vif's), which are all in the range of 1 – 1.18. This signifies that multicollinearity is no issue in the inspected models.

"Stage" is an ordered variable containing the ranking of the objective complexity of each task in the experiment. The highly significant coefficients for all values suggest, that compared to the base group the jump in objective complexity is a main driver of perceived complexity. In order to check if the differences between the separate stage coefficients are significant, the base group for the models is changed and the new coefficients are reviewed. By doing this analysis it is found that the differences between the coefficients are significant for all stage comparisons except for stage 2 and stage 3. Hence, it can be stated that stage 2 is perceived significantly more complex compared to stage 1 or stage 3 is perceived significantly less complex compared to stage 4. On the other hand, it is not possible to interpret stage 3 as significantly more complex than stage 2 or vice versa.

According to the analysis, the anticipated perception of a lower complexity with increasing numeracy (H1a) as well as the hypothesised perceived complexity increasing effect of intuition (H4a) and imagination (H6a) cannot be observed for all parts of the experiment. The emotionality (H5a) characteristic shows the assumed behaviour of increasing perceived complexity with increasing emotionality, but as for the aforementioned variables the results are insignificant. Testing the experientiality (H3a) in a separate analysis also yields no significant result. Therefore, the respective hypotheses need to be rejected. In contrary, H2a is not rejected since, as depicted by table 14 above, the hypothesised relation between the rationality and the complexity an individual perceives is negative in all three models.

(2) Bracketing Behaviour: Lottery Part

Continuing the analysis, the bracketing behaviour of subjects will be inspected. With respect to the research framework, the impulsiveness as well as the other proposed psychographics of a DM, and the perceived complexity of the task at hand should influence the bracketing behaviour. As a result of the experimental design, it is necessary to split the analysis of the bracketing behaviour into a lottery, a personnel, and an overall perspective. The reason for this is, that the various parts require contrasting measures of what represents broad bracketing. In addition, diverse datasets and different testing methods are needed in order to appropriately address the separate parts.

Starting with the lottery part, the measure for broad bracketing behaviour is the same as it is used in literature (Tversky an Kahneman, 1981; Rabin and Weizsäcker, 2009), namely whether a subject chooses the dominant B&C combination or not. Being that this measure refers to a dichotomous variable, a logistic regression is required for testing. As the experiment entails two very similar double lotteries, it is decided to take the mean value of the two perceived complexities reported by each individual for the subsequent analysis. In addition, the score reached in the CRT (Frederick, 2005) serves as a measure for the impulsiveness and is used as metric variable. In addition, the risk attitude of the subjects, measured through the single lottery, is included in the regression as well as all the psychographic characteristics.

Table 15 summarizes the acquired results with the coefficients being logit values and the figures in parentheses stating the corresponding standard errors. These coefficient values are obtained as standard output for generalized linear models in R. Although it is not intuitive to interpret these values, the sign of the coefficients can be understood as the independent variable having a positive or negative influence on the predictor. Hence, logits can be used for predictions of the dependent variable in the same way as in an OLS regression. Nevertheless, the result should then be converted into a probability to enhance the understandability of the prediction. Being that logits are not as understandable as odds or probabilities, table 16 shows the odds (e^{logit}) as well as the respective probabilities ($\frac{odds}{1+odds}$) for every coefficient. What can easily be derived from this table is that the probability to bracket broadly in the lottery parts is around 1%, assuming every other coefficient is zero. This probability can be especially altered by three covariates of our model, namely the perceived complexity, the impulsiveness as well as the risk attitude of the subjects. All three variables have an enhancing effect on bracketing behaviour leading participants to bracket broadly with a higher probability, the higher these values are.

Table 15: Regression Model - Bracketing Behaviour Lottery Part

Logistic Regression on Broad Bracketing in the Lottery Part

	Dependent variable:
	Broad Bracketing
Intercept	-4.630*** (0.642)
P.C. Double Lotteries	0.016** (0.007)
Risk Seeking	0.638* (0.329)
CRT Score	0.441*** (0.168)
Berlin Numeracy	-0.049 (0.135)
Rationality	0.001 (0.013)
Emotionality	0.017 (0.015)
Intuition	0.007 (0.015)
Imagination	-0.011 (0.016)
Observations	756
Log Likelihood	-158.359
Akaike Inf. Crit.	334.718
Pseudo-R^2 (Nagelkerke)	0.069
Note:	*p<0.1; **p<0.05; ***p<0.01

Values in parentheses are standard errors

Table 16: Conversion Table - Bracketing Behaviour Lottery Part

Coefficient	Odds	Probability
Intercept	0.010	0.010
P.C. Double Lotteries	1.016	0.504
Risk Attitude	1.892	0.654
CRT Score	1.555	0.609
Berlin Numeracy	0.952	0.488
Rationality	1.001	0.500
Emotionality	1.017	0.504
Intuition	1.007	0.502
Imagination	0.989	0.497

For the perceived complexity, this finding is counterintuitive at first and contradicts the hypothesised relation of an increasing probability to bracket broadly with decreasing perceived complexity (H8). The reason for the observed behaviour may originate from the extent to which a respondent understands the task at hand. A person understanding what is required and how to find an optimal solution might experience the task as more difficult compared to a person who does not fully understand the instruction of taking both decisions into consideration simultaneously. Such an individual that performs two single lotteries instead of one double lottery would then perceive the task as being of lower complexity. Consequently, a higher perceived complexity may enhance the likelihood that broad bracketing is observed for it constituting a measure of whether participants understand that the two decisions in the task are interrelated. Moving on, it can be seen that increasing the CRT score also increases the probability of a person to bracket broadly (H7). This finding is highly significant and means that a less impulsive person is more likely to bracket broadly compared to an impulsive individual. This observation supports the hypothesised influence and direction of impulsiveness on bracketing behaviour.

The third significant result is the risk attitude, which is measured by means of the single lottery task. If an individual chooses the sure gain over the higher EV gamble, they exhibit risk averse behaviour. The weakly significant coefficient found in the regression table indicates that less risk aversion in the single lottery increases the likelihood to bracket the double lotteries broadly. Therefore, such an individual may show a lower tendency to behave according to PT and instead selecting the gamble for the first decision of the double lottery. Regarding the psychometrics, no significant direct influence on bracketing behaviour is recognized. Neither an increase in the likelihood to bracket broadly can be observed with increasing numeracy (H1b) or rationality (H2b) score, nor a decreasing effect in the same can be detected for increasing intuition (H4b), emotionality (H5b), or imagination (H6b). However, the coefficients of the rationality as well as the imagination characteristic point in the hypothesised direction. Again, the experientiality (H3b) is tested as a substitute in a separate regression and is found to be insignificant. Therefore, for the lottery part analysis, H1b-H6b must all be rejected. Despite the model having a small pseudo-R^2 value (0.069), it provides an impression of what measures might influence the bracketing behaviour towards broad bracketing in the lottery part.

(3) Bracketing Behaviour: Personnel Part

Regarding the personnel part, a different measure of broad bracketing is required. As introduced in chapter 5.3.1, a broad bracketing individual is determined by the degree of diversification in the skill items of the built team given by the variety ratio. Broad bracketers are then defined as those subjects whose team fulfils the 65% cut-off criterion for each item of the four skill items.

Having determined the measure for broad bracketing as achieving a diverse team, a linear regression is used in order to analyse what drives the diversification behaviour of respondents, represented by the metric variety ratio. The covariates displayed in the results table 17 are the same as in the regression in the lottery part, but the score of correctly answered control questions is included, while the risk attitude is excluded. For the reason that the preliminary analysis shows no significance, the variable documenting the number of changes that respondents perform on their candidate selection, is excluded as well. As before, the figures in parentheses represent the standard errors.

In the resulting regression, no significant effect of impulsiveness on the degree of diversification (H7) can be observed. This finding is contrasting to the analysis of bracketing behaviour in the lottery part, where the impulsiveness is found to be highly significant. A possible explanation for the insignificance of impulsiveness may be found in the nature of the task. The context of lotteries facilitates an impulsive decision since respondents are only asked to make two decisions with two options each. Participants are not confronted with any additional information that would require them to process the task with an increased level of attention. In contrary, the personnel selection context provides respondents with information on how to elaborate a given candidate as well as a candidate table that contains nine options. Thus, the information load that subjects encounter in this part may lead them to contemplate longer and repress their first intuitive answer. However, although the coefficient of impulsiveness is not of significant character, it points into the direction that would have been hypothesised. In this sense, an individual that is described by a lower impulsiveness is more likely to achieve a high variety ratio. Consequently, such an individual is also more likely to build a team that classifies as diverse according to the 65% criterion and by doing so, qualifying as broad bracketer.

Table 17: Regression Model - Bracketing Behaviour Personnel Part

Linear Regression on Broad Bracketing in the Personnel Part

	Dependent variable:
	Degree of Diversification
Intercept	70.616*** (2.171)
CRT Score	0.260 (0.481)
P.C. Building Task	0.044** (0.019)
Control Score	1.958*** (0.625)
Berlin Numeracy	-0.192 (0.418)
Rationality	0.075* (0.041)
Emotionality	-0.096** (0.047)
Intuition	0.036 (0.050)
Imagination	-0.022 (0.046)
Observations	378
R^2	0.067
Adjusted R^2	0.047
Residual Std. Error	9.302 (df = 369)
F Statistic	3.314*** (df = 8; 369)
Note:	*p<0.1; **p<0.05; ***p<0.01

Values in parentheses are standard errors

When looking at the mean time spent in the double lottery tasks compared to the personnel selection task, a longer time spent on the personnel task can be observed. This further supports the argument that subjects take more time to form their decision and also reflects that the personnel part is of higher objective and perceived complexity. This perceived complexity again is highly significant in a positive direction, showing the same counterintuitive behaviour as was recognized in the lottery part (H8). According to the results, a higher perceived complexity induces a higher degree of diversification. As it was argued for the preceding model, this finding might stem from the circumstance that subjects who understand what the task at hand requires in order to find an optimal solution, rate it as being of a higher complexity. And since such participants understand what is needed to give a good answer their team constellations are more diversified. The highly significant effect of the control score strengthens this argument – the more control questions are answered correctly, the more diversified the chosen teams are. Opposite to the analysis of the lottery part, the rationality (H2b) score is weakly significant but also shows a positive influence on variety seeking behaviour. Subsequently, a more rational individual is more likely to diversify and by doing this is also more likely to find a team constellation that classifies as diverse. The characteristic of emotionality (H5b), on the other hand, decreases the degree of diversification the more developed this feature is for an individual. The numeracy (H1b), intuition (H4b), and the imagination (H6b) psychometrics are found to be insignificant and must therefore be rejected. As before, the (adjusted) R^2 is of a smaller magnitude with the model explaining approximately 5% of the variation in the data. The variance inflation factor is inconspicuous for all covariates.

(4) <u>Bracketing Behaviour: Overall Perspective</u>

At last, the bracketing behaviour shall be examined not separated into the different experimental contexts but from a holistic perspective. For this reason, the dependent variable "Broad Bracketing" is designed to be either 1, if a subject brackets broadly in the respective task or 0 otherwise. Bracketing broadly in the personnel task is again measured by achieving a team constellation that complies to the introduced 65% criterion and thus classifies as diverse.

Since there is one observation for every task, resulting in multiple observations for each participant, a random effects model (Appendix F) must be conducted. For this model, the task represents the random factor. This approach leads to the finding, that the task itself explains a considerable amount of the variation of the model. Hence, it can be concluded, that the task has an influence on the bracketing behaviour which also indicates, that there is a substantial difference between the two contexts. In order to get a better insight of how big this influence of context on bracketing behaviour is, a logistic regression with the task tag as a dummy variable is conducted. This can be done, since the random effects model showed, that the tasks actually can be viewed independently of each other, which is a prerequisite for a standard logistic regression. By comparing the logistic regression (including the dummy variable) with the random effects model, it can be seen that the coefficients as well as the p-values stay almost equal. Since it provides a better insight of how the different tasks with their respective contexts influence the bracketing behaviour, the logistic regression will be used for further analysis. The result of this regression with the coefficients being logits and the values in parentheses being standard errors, can be seen in table 18. Table 19 further provides the conversion of those figures into log odds and percentages.

Table 18: Regression Model - Bracketing Behaviour Overall

Logistic Regression on Broad Bracketing Overall

	Dependent variable:
	Broad Bracketing
Intercept	-1.692*** (0.351)
Perceived Complexity	0.010*** (0.004)
CRT Score	0.266*** (0.089)
Berlin Numeracy	-0.00001 (0.076)
Rationality	0.010 (0.008)
Emotionality	0.002 (0.009)
Imagination	0.002 (0.008)
Intuition	0.002 (0.009)
Consistency L.	-2.326*** (0.253)
Double L.	-2.257*** (0.244)
Observations	1,134
Log Likelihood	-410.559
Akaike Inf. Crit.	841.118
Pseudo-R^2 (Nagelkerke)	0.290
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 19: Conversion Table - Bracketing Behaviour Overall

Values in parentheses are standard errors

Coefficient	Odds	Probability
Intercept	0.184	0.156
Perceived Complexity	1.010	0.503
CRT Score	1.305	0.566
Berlin Numeracy	1.000	0.500
Rationality	1010	0.503
Emotionality	1.002	0.501
Imagination	1.002	0.501
Intuition	1.002	0.500
Consistency L.	0.098	0.089
Double L.	0.105	0.095

The covariates displayed in the regression are the same as used above but extended by two dummy variables – “Consistency L.” and “Double L.” – indicating the respective task compared to the base group given by the portfolio building task. When inspecting the intercept of the model and using the conversion table (table 19), it can be seen that the probability of broad bracketing behaviour in the personnel part of this experiment, holding all the covariates at 0, is about 15.6%. By comparing this probability with the one obtained in the analysis of bracketing behaviour in the lottery part (table 16), it can be seen that the probability to bracket broadly in the personnel task is about 14.6 percentage points higher. This result is further confirmed by the highly significant task dummy variables at the bottom of the regression table 18, showing a strong negative influence of the lottery tasks on broad bracketing behaviour. This observation is another supporting indicator for the possible influence that context has on the behaviour of a DM.

The psychographics, apart from the CRT score, are all insignificant, thus showing no direct influence on the bracketing behaviour. The CRT score, on the other hand, is highly significant, which indicates a higher probability to bracket broadly with decreasing impulsiveness (H7). This is in line with the other two regressions on bracketing behaviour, although the result on impulsiveness in the personnel part is of insignificant character. The same relationship is observed for the perceived complexity (H8) as it was found in the part-specific models, which endorses the finding of an increased probability to bracket broadly with increasing complexity. The model on overall bracketing behaviour has the highest goodness of fit value of all three models, with a pseudo-R^2 of around 29%. Like in the models before, the covariates do not show any signs of multicollinearity.

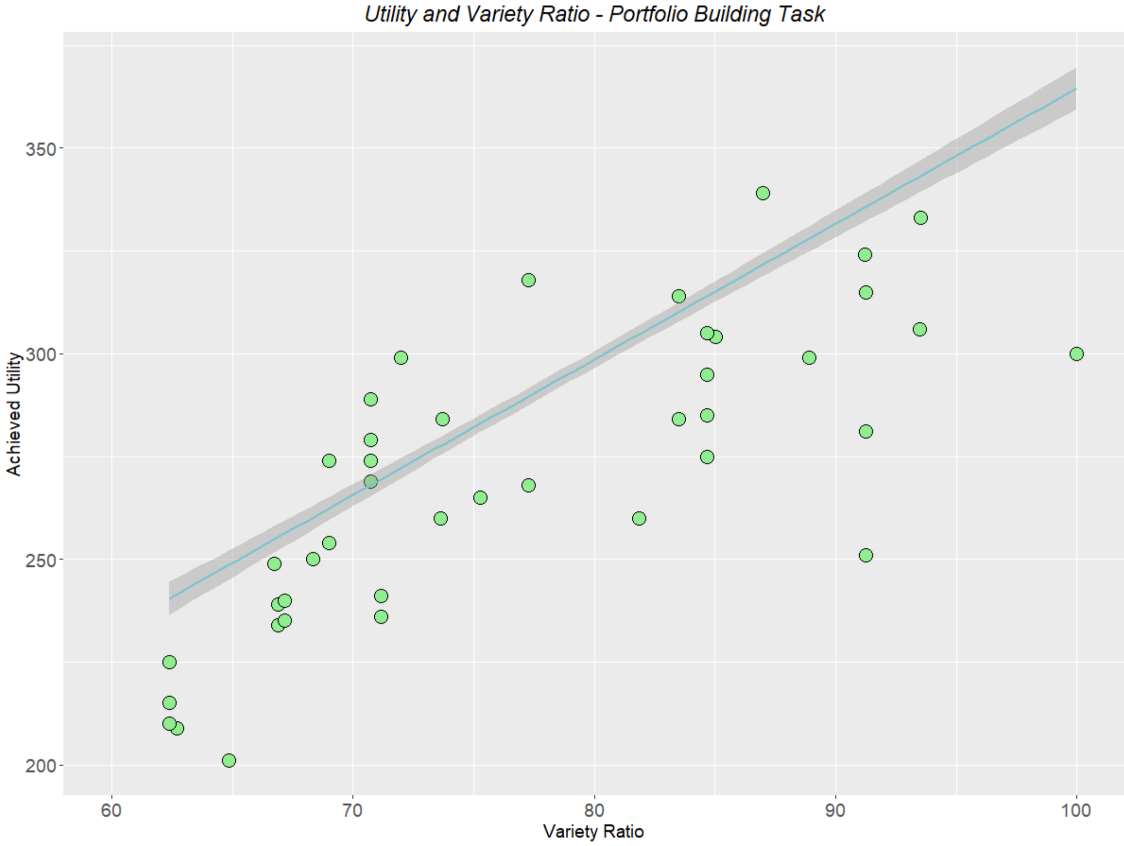
Summarizing the analysis of the bracketing behaviour, the influence of rationality (H2b) and emotionality (H5b) on bracketing behaviour can only be statistically identified for the model of the personnel part. These two hypotheses must therefore be rejected for the lottery part as well as for the overall analysis. Apart from that, an individuals’ numeracy (H1b), intuitiveness (H4b), and imagination (H6b) as well as the experientiality (H3b), which was always tested separately, do not significantly influence the bracketing behaviour. Thus, these hypotheses must be rejected for all models. Regarding an individuals’ impulsiveness, H7 is not rejected for the lottery and overall part of the analysis but must be rejected when looking at the personnel part. On the other hand, H8 needs to be rejected with respect to all three models, since exactly the opposite of the hypothesised influence of the perceived complexity can be observed.

(5) Solution Quality

The last step of the research framework is determining the influence of the applied bracketing behaviour on the outcome quality. Regarding the lottery part, no specific analysis is needed for this purpose since broad bracketing is denoted as selecting the dominant B&C combination. Consequently, broad bracketing achieves the optimal solution whilst any other form of behaviour is of inferior quality. For the personnel part, on the other hand, it is possible to examine this interdependence between bracketing behaviour and solution quality. This is because bracketing behaviour is measured by the degree of diversification and due to the diversification bias, the formation of a diverse team does not necessarily yield the optimal solution. For this aspect of the thesis, several regression analyses have been conducted. The difference in explanatory power of the models that include the variety ratio as independent variable compared to the ones without this covariate is substantial. This finding shows the importance of diversification on solution quality but also displays it as being the only important factor. That is why the analysis on solution quality is split into two parts. First it is investigated how the degree of diversification directly influences the solution quality by means of calculating the respective correlation. Afterwards it will be assessed if there is an influence of the psychographics on the solution quality when the effect of diversification is removed. This will be achieved by regressing the psychographics on the residuals of the OLS regression (adj. R^2 = 70.71%) between the achieved utility as dependent and diversification as independent variables.

By checking the correlation between the achieved utility in the personnel part and the degree of diversification, the assumption expressed before is confirmed. The two measures are strongly correlated with a highly significant correlation of 0.8414. Figure 47 illustrates this relation in the form of a scatterplot. The blue line is fitted through an OLS regression between the two variables. The graph does not display a degree of diversification smaller than 62,38%, since subjects built no team composition that would yield a smaller value. The variety ratios of all built team constellations are displayed, and

Figure 47: Scatterplot - Achieved Utility and Variety Ratio



there are no notable gaps in the scatterplot. The fitted regression line tends to overestimate the solution quality, especially at the beginning and at the end.

Table 20 shows the regression of the psychographics on the residuals of the achieved utility model. What can be observed here is, that even when the effect of diversification on the achieved utility is removed, the CRT score and therefore the impulsiveness of a DM remains significant. Also, the understanding of the task at hand, measured through the number of correctly answered control questions has a significant positive effect on the achieved utility. This is in line with the finding that a high control score increases the likelihood of an individual bracketing broadly. At last, the objective complexity of the task itself, defined by the aggregated and segregated presentation modes of the two treatments, reveals a significant influence on the achieved utility. It follows that the less complex segregate treatment increases the mean achieved utility. The influence of the other tested psychographics is insignificant. This means that, except for the CRT, no direct influence of an individuals' characteristics on the solution quality can be identified. Taking the two steps of the analysis together it is possible to discuss what is proposed in H9.

Table 20: Regression Model - Achieved Utility

Linear Regression on Residuals of Achieved Utility

	Dependent variable:
	Residuals of Achieved Utility
Intercept	-12.724*** (4.244)
CRT Score	2.357** (1.031)
Control Score	2.847** (1.339)
Segregated Presentation	4.121** (2.072)
Berlin Numeracy	-0.492 (0.897)
Rationality	-0.044 (0.088)
Emotionality	0.107 (0.100)
Imagination	0.144 (0.098)
Intuition	-0.108 (0.108)
Observations	378
R^2	0.046
Adjusted R^2	0.025
Residual Std. Error	19.936 (df = 369)
F Statistic	2.207** (df = 8; 369)
Note:	*p<0.1; **p<0.05; ***p<0.01

Values in parentheses are standard errors

In the first phase of this two-step analysis, a strong correlation between diversification and the outcome quality is found. Hence, a higher degree of diversification directly improves the solution quality. Again, it needs to be noted, that this linear relationship holds only up to a certain point, before the diversification bias takes effect. The model displayed in table 20 constitutes the second part of the analysis by showing a significant effect of the treatment, even if the effect of diversification on utility is removed. This means, that the presentation form itself also has a significant influence on the outcome quality of a decision. In consequence of the findings yielded by the underlying data, H9 is not rejected.

Summarizing the first part of the statistical testing procedure, table 21 on the following page displays the main hypotheses and whether they are rejected or not.

Table 21: Overview Main Hypotheses

Hypothesis	Statement			Reject (x) | Do not reject (o)		
				Overall	Lottery Part	Personnel Part
H1a	High numeracy low P.C.	→	low P.C.	x	x	x
H1b	High numeracy	→	likely to bracket broadly	x	x	x
H2a	High rationality	→	low P.C.	o	o	o
H2b	High rationality	→	likely to bracket broadly	x	x	o
H3a	High experientiality	→	high P.C.	x	x	x
H3b	High experientiality	→	likely to bracket narrowly	x	x	x
H4a	High intuitiveness	→	high P.C.	x	x	x
H4b	High intuitiveness	→	likely to bracket narrowly	x	x	x
H5a	High emotionality	→	high P.C.	x	x	x
H5b	High emotionality	→	likely to bracket narrowly	x	x	o
H6a	High imagination	→	high P.C.	x	x	x
H6b	High imagination	→	likely to bracket narrowly	x	x	x
H7	High impulsiveness	→	likely to bracket narrowly	o	o	x
H8	High P.C.	→	likely to bracket narrowly	x	x	x
H9	Broad bracketing	→	better solution quality	-	-	o

5.5.2 Task-Specific Hypotheses

After presenting the analysis of the main hypotheses and the created models, this subchapter highlights the analysis of the task-specific hypotheses. As before, the chapter will be closed with an overview of the respective hypotheses stating whether they are rejected or not.

(1) Lottery Part

The assumption of H10 is in line with the already existing literature stating that the dominant choice combination (B&C) will be chosen less often compared to the prospect theoretic choice combination (A&D). In order to test if this holds true for the underlying experiment, it needs to be assessed whether a significant deviation from a uniform distribution of answers is observed. A χ-square test represents the appropriate instrument, since the number of observations for all choice combinations is sufficiently large. Note that, since the experiment incorporates two double lotteries, this hypothesis needs to be tested for both tasks. In figure 26 (chapter 5.2) the frequencies of the different combinations for both lotteries are displayed. The χ-square test confirms what was hypothesised for both, the double (χ-square = 202.97, p-value < 2.2e-16) and the consistency lottery (χ-square = 148.02, p-value < 2.2e-16). By looking at the frequencies of the different choice combinations, the highly significant deviation from a uniform distribution can be interpreted as a clear sign of participants choosing the A&D combination more frequently than B&C or any other choice combination. For this reason, H10 is not rejected.

In H11 the consistency between the first and the second double lottery is inspected by hypothesising that the behaviour of DMs is consistent, especially for those who select the superior choice combination. Again, a χ-square test is the statistical test used to check whether the distributions of both lotteries significantly deviate from each other. For this purpose, we first test whether the participants select the exact same combination in both lotteries, or if they deviate from their initial choice. In figure 27 (Chapter 5.2), the overall consistency in behaviour between the lotteries is shown and the χ-square test (χ-square = 78.265, p-value < 2.2e-16) confirms the statistical significance of the deviation from a

Table 22: Distribution of B&C Combination in Lottery Tasks

		Consistency Lottery		
	Choice Combination	B&C	Not B&C	Sum
Double Lottery	B&C	9 (2.38)	14 (3.70)	23 (6.08)
	Not B&C	12 (3.17)	343 (90.74)	355 (93.92)
	Sum	21 (5.56)	357 (94.44)	378 (100)

*relative frequencies in parentheses

uniform distribution. This means, that significantly more individuals (approximately $^2/_3$ of all participants) are consistent as opposed to being inconsistent. In order to investigate, whether those individuals who bracket broadly by selecting the dominant choice combination in the first lottery also did so in the second one, another χ-square test is conducted. Table 22 shows the distribution of participants explicitly selecting the combination B&C. For the χ-square test (χ-square = 6.125, df = 1, p-value = 0.01333) the two values used are 23, which is the number of people choosing B&C in the first double lottery and 9, the number of participants who stayed consistent in the second lottery. In this case the significant result paired with a look at the values expresses the opposite of the finding for overall consistency. This signifies, that whilst holistic behaviour in the lotteries is consistent, respondents who broadly bracket the first lottery tend to deviate from the superior option for the second task.

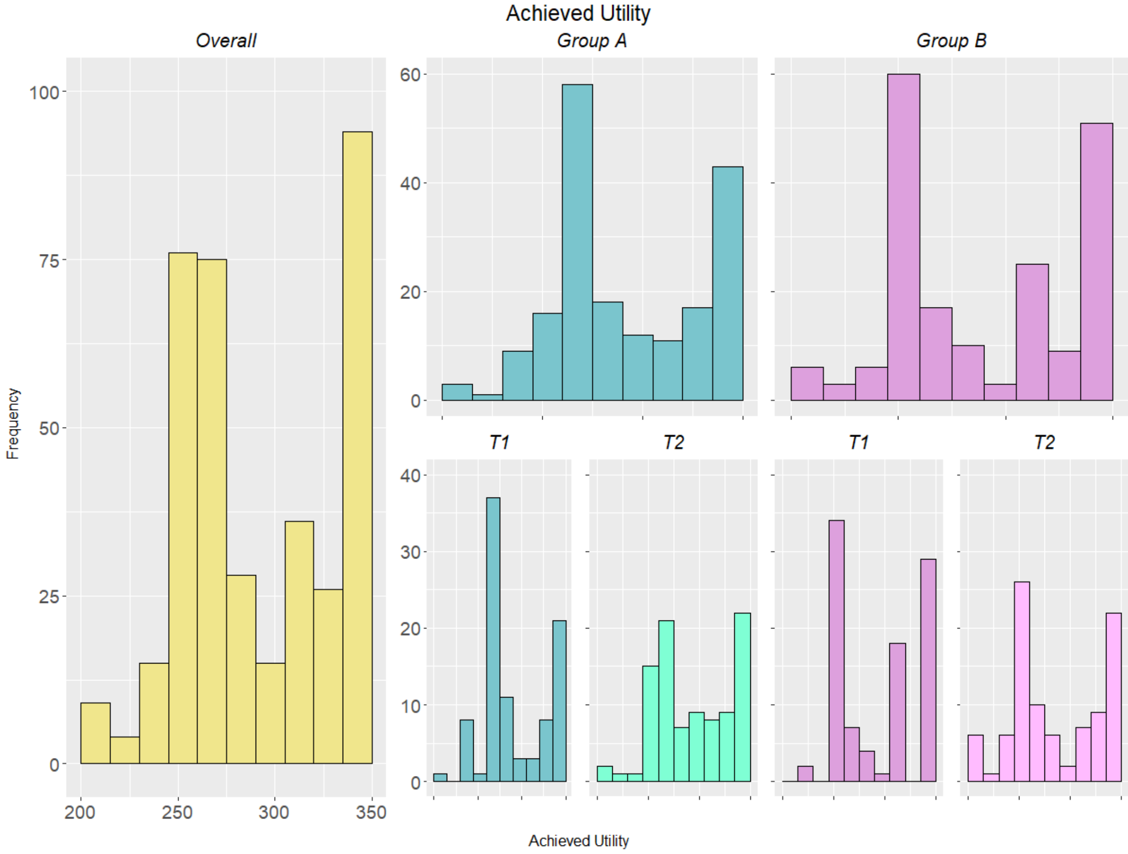
The last task-specific hypothesis linked to the lottery part is H12, stating the assumption that participants will be able to recognize the dominant choice when they are presented with a combined presentation form of the task. In figure 28 (chapter 5.2) the relevant frequencies are depicted which already indicated a strong effect of the presentation mode. When testing this suggestion with a χ-square test (χ-square = 298.67, p-value < 2.2e-16) the highly significant result supports the aforementioned assumption.

(2) Personnel Part

The first task-specific hypothesis concerning the personnel part, H8a, assumes a difference in the frequency of the optimal team constellation due to the different presentation modes of the treatments. Again, a χ-square test is conducted, comparing the number of subjects who choose the utility-optimising team constellation in the different presentation modes. While 23.76% of participants in the aggregated presentation mode (T2) choose the optimal solution, slightly more individuals (26.6%) did so for the segregated form (T1). The resulting insignificant χ-square test indicates that there is no notable difference in the frequency with which the optimal solution is selected across the two treatments. This result is also found when the treatments of the two groups of the experiment (A and B) are tested separately.

The next hypothesis, H8b, addresses the contrast in achieved team utility between the two presentation modes. In figure 43 (chapter 5.3.1) it can be seen that there is a slight discrepancy in the overall comparison as well as in group A, but a bigger difference when it comes to group B. To test if a significant difference in mean values between the two treatments can be found, it is necessary to determine whether a parametric or a non-parametric test is suitable. As a first step, the distributions for the achieved utility in both

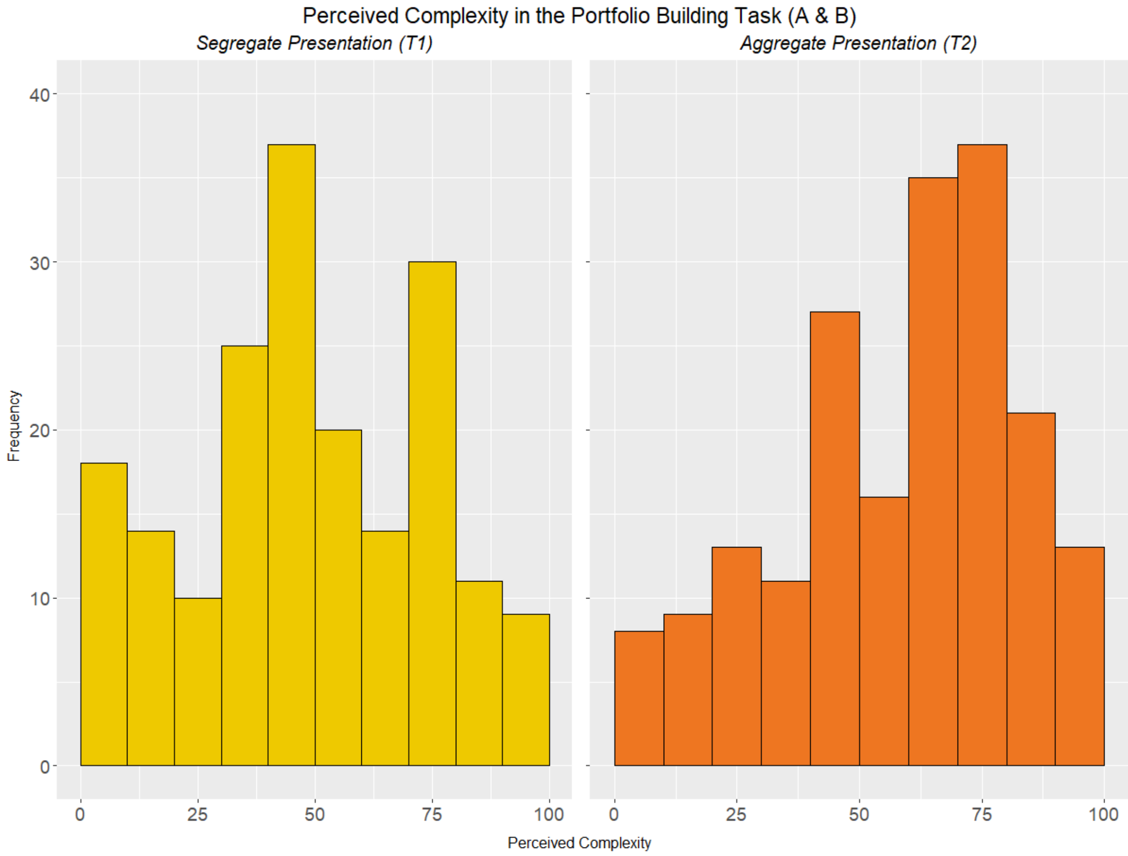
Figure 48: Achieved Utility in the Portfolio Building Task - Histograms



groups with their corresponding treatments are inspected (figure 48). Since all histograms do not show any similarities to a normal distribution, non-parametric tests are required for the comparison. Accordingly, a Mann-Whitney-U-test is conducted, which compares the overall means of the independent samples of achieved utilities across treatments. The result of this test indicates that there is no statistically significant difference in achieved utility across different treatments when taking an overall perspective. By repeating the same procedure for the two groups, very similar results for group A and for group B are obtained. This means, that the hypothesis needs to be rejected when examining the data for the treatments at an overall level, as well as when considering the group level of the treatments.

H13 draws attention to the difference in perceived complexity between the two treatments by specifying that the aggregate presentation form (T2) is experienced as the more difficult task. In order to determine whether a parametric or non-parametric analysis must be conducted, the histograms of the separate treatments are inspected (figure 49). The histograms reveal a shape of the underlying data that is close to a normal distribution, so that the samples are tested for homoskedasticity. A first descriptive comparison of

Figure 49: Perceived Complexity for Treatments - Histograms



variances hints towards variance homogeneity and the conducted insignificant Levene's test provides the statistical confirmation for this assumption.

Going forth, a parametric approach is taken by comparing the perceived complexities with a Student's t-test. The result is highly significant (t = 3.8756, p-value = 0.0001255), meaning that there is a deviation different from zero between the mean values of the two treatments. By comparing the mean perceived complexity in the aggregated configuration of the experiment (62.51) with the segregated one (52.53), it can be concluded that the data supports the hypothesis. Continuing the analysis, the two groups A and B are tested according to the same procedure. In group A (t = 3.502, p-value = 0.0005806) a similar, highly significant result can be found, while the statistical result for group B (t = 2.0335, p-value = 0.04341) is not as strong but still significant at the 0.05 level. Consequently, neither the analysis on an overall perspective nor a group level approach provides a statistical basis that would justify rejecting H13.

Now that it was possible to show the statistically relevant difference in perceived complexity for both treatments, H13a and H13b intensify the analysis of the different complexity levels. According to H13a, participants who bracket broadly in the aggregate presentation mode are expected to show a higher average numeracy score, compared to the broad bracketers in the less complex segregate T1. Likewise, H13b assumes the average rationality score of broad bracketers in the more complex T2 to be higher than the average score for broad bracketers in T1. An identical testing process is conducted for both hypotheses. First, the data is split into participants who bracket broadly and the ones who bracket narrowly. Remember that the criterion to distinguish between broad and narrow bracketers in the personnel part is not achieving the best possible outcome but forming a diversified team constellation. By choosing this criterion attention is not only placed on those participants who achieve the best possible solution but also include those who bracket broadly but fall prey to the diversification bias. In the next step of the analysis, it is tested whether a parametric or non-parametric approach is suitable. Since the criterions for a parametric approach as discussed above are not met, the new dataset is tested with a Wilcoxon rank-sum test, which compares the numeracy (H13a) and rationality (H13b) scores in the aggregated and segregated treatment. The results for both hypotheses are insignificant, which means, that the average numeracy and also rationality score of broad bracketers does not vary across treatments.

Two implications follow from this finding: Either the complexity difference of the treatments is not sufficient to sort out the participants with a lower numeracy or rationality score who still managed to bracket broadly. Alternatively, there may be a threshold for numeracy and rationality from which on a person is enabled to bracket broadly when confronted with such a problem, regardless of the presentation form used in the underlying experiment. Ultimately, both, H13a and H13b must be rejected. However, note that the regression analysis on bracketing behaviour for the personnel part (chapter 5.5.1) yields a significant positive influence of rationality on broad bracketing, so that a general relation is found. Combining these findings, it is possible further support the second possible explanation: Rational people tend to bracket broadly, but it is not necessary to be even more rational in order to bracket broadly in the second task, meaning that once a benchmark rationality is achieved, broad bracketing becomes likelier for similar tasks of different complexity.

For H14 the difference in the degree of diversification between the two presentation forms is tested with the non-parametric Wilcoxon rank-sum test, since the conditions for a parametric method are not given. The insignificant result of this test shows a similar finding as for H13a and H13b, namely that there is no statistically significant difference in the degree of diversification of the chosen teams across treatments. Therefore, H14 must be rejected.

The last two task-specific hypotheses related to the personnel part are concerned with the comparison of the optimal solutions found in the aggregated presentation form (H15) and the segregated presentation form (H16), compared to the portfolio selection task. Due to the substantial difference in objective complexity between the two treatments and the portfolio selection task it is hypothesised that compared to both configurations, more participants will find the optimal solution in the portfolio selection task. Table 23 shows the number of people who achieved the best possible solution in the different tasks and treatments.

Table 23: Frequency of Optimal Solutions in the Personnel Part

		Task		
		Portfolio Building	Portfolio Selection	Sum
Presentation Form	Segregate Presentation (T1)	50	106	156
	Aggregate Presentation (T2)	44	120	164
	Sum	94	226	320

*relative frequencies in parentheses

The testing approach is the same for both hypotheses, namely a χ-square test to evaluate whether there is a deviation from a uniform distribution of optimal solutions found across the tasks. Before the analysis of the hypotheses, the first undertaken χ-square test examines whether a statistical difference in the frequency of optimal solutions can be found between the portfolio building task, without consideration of the treatments, and the portfolio selection task. Unsurprisingly the result is highly significant (χ-square = 54.45, p-value = 1.595e-13) and can be interpreted insofar as that the frequency of optimal solutions recorded for the portfolio selection task is significantly higher compared to the portfolio building task. The same procedure but with consideration of the separate treatments is repeated to check H15 (χ-square = 35.22, p-value = 2.946e-09) and H16 (χ-square = 20.103, p-value = 7.34e-06) leading to the same result with an equal interpretation. Therefore, H15 and H16 both are not rejected.

Table 24 displays which task-specific hypotheses can or cannot be rejected.

Table 24: Overview Task-Specific Hypotheses

Hypothesis	Statement	Reject (x) Do not reject (o)
H10	For double lottery: less "B&C" than "A&D"	o
H11	Broad bracketing in double lottery → broad bracketing in consistency lottery	x
H12	For combined lottery: more individuals choose dominant option	o
H8a	Optimal solution is found more often in the segregated T1 than in the aggregated T2	x
H8b	Avg. team utility is higher in the segregated T1 than in the aggregated T2	x
H13	P.C. is lower in teh segregated T1 than in the aggregated T2	o
H13a	Numeracy score of broad bracketers in the aggregated T2 is higher than the numeracy score of broad bracketers in the segregated T1	x
H13b	Rationality of broad bracketers in the aggregated T2 is higher than the rationality of broad bracketers in the segregated T1	x
H14	Variety ratio of teams is higher in the aggregated T2 than in the segregated T1	x
H15	More subjects find optimal solution in portfolio selection compared to the aggregate T2	o
H16	More subjects find optimal solution in portfolio selection compared to the segregate T1	o

6 Discussion

After presenting the main findings of the experiment, the following chapter aims at discussing these in more detail. As before, the flow of the research framework will serve as a guide for the discussion. Thus, findings regarding the perceived complexity of participants will be discussed first and insights about which measures drive perceived complexity will be given. Subsequently, the bracketing behaviour, which is the main focus of this thesis, will be highlighted. The hypothesised drivers of bracketing behaviour will be examined before the centre of discussion moves on to the context of a decision problem as main driver of bracketing behaviour. The importance and possible influence of the decision problem's context on an individual's behaviour found in this thesis is examined. Additionally, managerial implications, derived from the findings for the bracketing behaviour, are presented.

One of the main purposes of this thesis is to establish a relationship between complexity and bracketing behaviour. In contrast to other scholars, the approach of the underlying thesis is not exclusively based on the objective task complexity, but rather on the perceived complexity that an individual faces when confronted with a decision problem. This approach is chosen, since this thesis suggests that different individuals will not perceive the same complexity for the same task. Therefore, subjects of the conducted experiment are requested to submit a value between 0 (simple) and 100 (difficult), rating their perceived complexities for the respective tasks This provides the possibility to retrieve an insight of how difficult the different tasks, which are of different objective complexity, are perceived by participants. Figure 15 (Chapter 5.1) perfectly illustrates the assumption made, namely that every task reports a perceived complexity ranging from 0% to 100%, independently of the objective task complexity. As the prevalent belief is that complexity drives bracketing behaviour, the question raised is what drives perceived complexity.

According to existing literature, cognitive capacity is the main source of coping with complexity and therefore a possible reason for suboptimal bracketing behaviour (Read and Loewenstein, 1995; Koch and Nafziger, 2016; Stracke et al, 2015). In order to measure the cognitive ability of their participants, previous studies frequently utilize self-reported math grades. In this thesis, the self-reported math grades are replaced by the Berlin numeracy score as an instrument to capture the cognitive ability of participants. Unfortunately, this measure did not show the effect it was hoped to have. Contrary to what was expected, the collected data provides no support for the numeracy score representing a good predictor of perceived complexity. The insignificance of the variable would imply that the complexity perceived by an individual is independent of cognitive ability but driven by other factors. A possible explanation for this rather surprising statement could lie in the circumstance that the chosen measure for cognitive capacity may not actually be a good measure of cognitive capacity. Then the finding would be reduced to the statement that an individual's numeracy attribute does not have a significant impact on perceived complexity, but no proposition is made for cognitive ability. In addition, the numeracy variable also proved to be insignificant with respect to exhibited bracketing behaviour.

Three potential explanations can be formulated for why this score is found to be insignificant. One reason could be, that the cognitive ability is of no relevance for bracketing behaviour and does not lead to a predefined tendency for a specific bracketing behaviour. Alternatively, the same circumstance as for perceived complexity may hold and numeracy does not represent cognitive ability to a sufficient degree so that only numeracy is insignificant, and no statement is made for cognitive ability. Since numeracy only measures a specific aspect of cognitive ability, this observation means that it may be needed in order to solve a task but is not relevant for assessing a task's complexity. The third reason for the failure of the variable might lie in the modus that this experiment was executed in. Since the experiment needed to be conducted online, it was not possible to control for the usage of calculators or other unauthorized aids. If candidates disregarded the instructions, which clearly stated not to use calculators, the collected data for cognitive ability as well as its relation between elicited cognitive ability and observed behaviour could be distorted. Since participants of this study were granted bonus points for their courses, an incentive to perform well is introduced to the experiment. On one hand, the performance-dependent course credits serve as a motivator to dedicate an increased amount of attention to the experiment while, on the other hand, it may provide an incentive to cheat. Therefore, the second explanation appears to be adequate.

In addition, the REI is used to test for additional psychometric traits that may influence how complex different tasks are perceived. The rationality of a subject is the only psychographic captured through this questionnaire that proves to be a consistent predictor for complexity – the higher the measured rationality of a subject is, the lower the perceived complexity. This influence of rationality holds true for both parts of the experiment and is therefore observed in different contexts. All other measured psychographics appear to be insignificant, leading to the impression of rationality being the only trait of importance for the determination of changes in perceived complexity. The term "changes" is selected to emphasize, that rationality alone is not able to predict perceived complexity because the decision problem with its respective environment influences the complexity significantly. This is shown by the stages variable in the regression table for perceived complexity (table 14), being that this variable not only incorporates the different objective complexity levels but also the switch in decision context (lottery vs. personnel part). Interpreting these findings, it can be said that the decision problem itself is the biggest and also unsurprising driver for complexity, while the rationality of an individual moderates the transition from objective complexity to the subjectively perceived complexity. Therefore, more rational individuals are likely to achieve better results when confronted with decision problems, since experiencing a reduced complexity should facilitate the efficient processing and subsequent evaluation of the necessary information. Other tested psychographics have no significant influence on the perceived complexity.

The perceived complexity is then assumed to be a main influencing factor for the exhibited bracketing behaviour. This assumption is confirmed by the conducted statistical testing procedure, but the influence is reversed to what is hypothesised. Increasing perceived complexity is found to also increase the probability to bracket broadly in two out of three models. This is counterintuitive at first and stands in contrast to the common assumption that increasing the demand on cognitive capacity will decrease the probability to bracket broadly (Stracke et al., 2015). We argue that this finding may be explained by how well subjects understand a given task. In order to apprehend a decision problem in its whole difficulty, a high cognitive ability is required. In this sense, an individual of lower cognitive ability would state a lower perceived complexity for the reason that he does not fully understand the extent of the task. Thus, the link to cognitive ability can be made in such manner that the perceived complexity also acts as an indicator of a higher cognitive ability, which is needed to grasp the problem at hand. An additional explanation could be found in the effort provided by the subjects in order to fulfil the required task. If a high amount of effort is supplied, the subject may be more inclined to perceive the task as more complex compared to an individual who do not invest a considerable amount of effort for working on the task.

Interestingly, this finding can be observed for the lottery part as well as for the personnel part. Consequently, it can be assumed that the detected effect is independent of the context of the decision problem. Stracke et al. (2015) report in this regard, that the effect of cognitive ability on bracketing behaviour can only be observed when the difference in complexity between tasks is sufficiently large. This statement is supported by the findings of the underlying thesis through the following observation: Although the perceived complexity significantly increases the probability of bracketing broadly a significant increase in broad bracketing behaviour can only be observed when comparing the portfolio building task to the portfolio selection task. For the comparison of the two presentation forms of the portfolio building task, however, the difference in perceived complexity is not sufficiently large, so that no significant increase in the degree of broad bracketing can be found.

These findings imply, that decreasing the task complexity by means of decision support can only have a significant effect on the resulting behaviour, if the difference in complexity is large enough. Stracke et al. (2015) find similar results in their study, where they conclude that only substantial decreases in complexity can improve the observed bracketing behaviour. Nevertheless, for the two treatments of the portfolio task, it is found that decision support significantly increases the solution quality although it does not induce major differences in bracketing behaviour. This result is stated in table 20 (Chapter 5.5.1), where the presentation form is found to be significant after the effect of diversification, the proxy for bracketing behaviour, on the achieved utility is removed. It follows that regardless of the bracketing behaviour, decision support mechanisms lead to an increasing solution quality. In addition, an increased understanding of the task at hand as well as a reduced impulsiveness ultimately lead to a higher solution quality when the effect of broad bracketing by means of diversification is removed. The insignificant other psychographics paired with the mentioned significant effects offer the following implication: An individual of low impulsiveness, who understands the task at hand, or receives the appropriate decision support has an advantage in achieving a high-quality solution. Nevertheless, the biggest influence on the solution quality still is explained through the degree of diversification and therefore the bracketing behaviour.

Moving forward, the conducted testing procedure showed that there are context dependent differences regarding the influence of individual psychometrics on bracketing behaviour. While for the personnel selection context the emotionality as well as the rationality have a direct effect on bracketing behaviour, in the lottery context merely the impulsiveness is found to have a significant impact. Similar to the results reported in literature, the more rational a subject is, the greater is the likelihood that such a subject brackets broadly. This stresses that choice bracketing is a rational procedure (Rabin and Weizsäcker, 2009). On one hand, rationality decreases the perceived complexity, which is a comprehensible relation. However, a reduced perceived complexity then negatively impacts broad bracketing behaviour, which is explained by the presented suggestion that understanding the task to its full extent leads subjects to report a higher value. At the same time, pronounced rationality directly increases the likelihood to bracket broadly. Thus, a two-stage process of assessing the complexity is suggested. First, any given task is defined by its objective complexity level, resulting from its context, number of decision options, information load, etc. This leads to higher benchmark complexity values for difficult tasks and lower ones for less complex tasks. The second step is executed by the rationality of the DM itself, in the sense that an increasing rationality leads to decreasing perceived complexity, starting from the already set benchmark complexity.

Emotionality, in contrast to rationality, decreases the probability to bracket broadly in the personnel part of this experiment, which again supports the assumption of bracketing as being a rational procedure. It must be noted though, that these characteristics do not induce or impede bracketing behaviour for the lottery context. There, only a subject's impulsiveness directly influences the bracketing behaviour. The more impulsive a DM is, the higher is the probability to bracket narrowly.

This statement, on the other hand, can only be made for the lottery part, but not for the personnel part. The results, especially the analysis of times spent on each task, support the reasoning that decisions are made rather quickly and intuitively compared to the personnel part. The reason for this may be found in the higher complexity of the personnel tasks as well as the circumstance that the lottery tasks represent a decision between two alternatives, which may allow for impulsive behaviour more easily than the large candidate tables of the personnel tasks. It follows that decisions in contrasting contexts are influenced by different psychographic traits. This means that the context itself can be defined to be a driver of bracketing behaviour. This finding is rather unexplored by other scholars, since most studies focus on decision problems in only one context. Supporting this argument, the analysis of broad bracketing at an overall level (table 18), shows a highly significant difference for broad bracketing behaviour between the various contexts. While broad bracketing as a default mode for bracketing behaviour (holding all other covariates at 0, the probability to bracket broadly lies around 1%) is rather unlikely in the lottery context, the probability of broad bracketing being the default mode is significantly higher for the personnel context (approximately 15.6%). In addition to this finding, table 13 (Chapter 5.4) shows, that participants who managed to bracket broadly in one part of the experiment, do not repeat this behaviour for the other context in approximately 60% of cases. Following this argumentation, it can be stated that because of a context dependency, it is not possible to deduce a universal default mode of bracketing behaviour for an individual that would then be applied to every problem. Therefore, it is suggested that it is only possible to detect a tendency for bracketing behaviour in a specific context. With respect to the hiring process in an organisation, this finding transforms to the implication that a principal must seek to find a match between an agent's context specific bracketing mode and the context of the position that the agent is supposed to operate in. As a result, standardized tests only assessing behaviour with respect to one decision-making context are not advisable for an organisation-wide hiring criterium. Opposed to that, division or even job-specific standardized assessments of agents would enhance the probability of finding a match between a broad bracketing applicant and the future context in which this agent will operate.

In this thesis, no evidence can be found of demographic-specific influences on bracketing behaviour or the solution quality. Although it is clearly stated in the descriptive statistics, that men tend to be more rational compared to women, in the subsequent analysis no gender difference can be observed in terms of bracketing behaviour or solution quality. For principals, this implicates that the individual psychometrics instead of the demographics should be the main focus when deciding which agent to hire.

Limitations and Directions for Future Research

A major limitation of this study is caused by the pandemic of 2020, which required the transition to an online-experiment. Conducting the study in this form entails a loss of control that may, to some degree, result in noise in the measured variables. In particular, it was not possible to control subject's behaviour with respect to them communicating with each other, using search engines, pen and paper, or calculators as aids for completing the tasks. Additionally, participants might have paid less attention while reading the instructions than they would have in a controlled lab-setting. Furthermore, the online-design did not allow for an opportunity to ask questions if the instructions were not sufficiently clear. In order to reduce the possible distortion of the collected data, participants were explicitly informed about what represents a forbidden aid. Furthermore, subjects were given the possibility to read the instructions for the personnel part at any time while working on the respective tasks. Aside from this, respondents had access to the e-mail address of both authors as a possibility to establish contact. Nevertheless, the comparison of the collected results to a follow-up study that is executed in a lab-setting would be an interesting and potentially valuable direction for future research. Findings for the personnel context might be of particular interest, being that one may evaluate whether a large fraction of respondents used a calculator in the present study and by doing so potentially disrupted the relation between cognitive ability and bracketing behaviour. Such a use of unauthorised aids might then be the cause of the numeracy characteristic, which represents the cognitive capacity, having no significant influence in any of the models.

Another potential limitation addresses the used sample, which mainly consists of students of business administration and economics. Although this implies that participants of the study are not real DMs, the introduced incentive of gaining course credits represents a real-world consequence for behaviour. Consequently, participants are motivated to make their decisions similar to a real DM. Nevertheless, a study that uses real DMs would be able to validate the findings of the present research. Furthermore, opposite to what was previously intended, the switch to an online-experiment also lead to the decision to reduce the number of sampled courses to two. While the achieved sample size is satisfactory, both courses belong to a bachelor's programme and are directed at students that are still at the beginning of their studies. Although experience shows that these courses are also attended by more advanced students, the change of the execution mode excludes master students as well as bachelor students that have already specialised in specific fields. Future studies might use a broader sample in order to investigate the effect of specialising in a particular field. This is of interest because the present study finds the context of the decision problem to be of great importance for the exhibited bracketing behaviour. Thus, the specialisation of students in a particular context might be an additional influencing individual trait. In this sense, a matching of specialisation and decision problem context could reduce the perceived complexity and facilitate broad bracketing for a task in this context.

As already noted, the experiment was accessible for one week and had no time constraints in place, which lead to the loss of control for the time tracking variables. The data shows that some participants started and ended the experiment on different days. While such extreme outliers can be easily identified, the sample also contains numerous middle-sized outliers, where subjects seem to have taken smaller breaks in between tasks. This behaviour was not explicitly forbidden by the instructions and renders the time variables unsuitable for the statistical analysis. This is because for those outliers, it cannot be distinguished whether the respective participants actually needed the recorded time for working on a task or merely took a break. A lab experiment or the implementation of a time restriction into an online-setting would provide reliable measures and constitutes a suggestion for future studies.

Another limitation may be found in the design of the question that requests the perceived complexity of respondents. Participants were free to enter any integer value between 0% and 100%, which were assigned the label simple and difficult respectively. The collected data reveals entries of 0% and 100% perceived complexity for all tasks of the five objective complexity stages. This may be provoked by the wording of the question, since the notion that a specific task is "simple" may be interpreted differently by different individuals. A better reference point for the 0% extreme point might be "I do not need to think about what I want to answer" while the 100% mark could be labelled as "I need extensive time and effort to give an answer". Alternatively, future research could validate the scales in order to be able to compare the gathered entries. Furthermore, the newly suggested scale might also incorporate the impulsiveness of an individual, so that the connection to the results of the CRT are of interest. In any case, one needs to keep in mind that the perceived complexity remains a subjective measure, meaning that the understanding of 0% and 100% complexity may not be generalised.

Concerning the presentation of the candidate table (table 8) in the aggregate T2, another design limitation may be found. Although the table incorporates no explicit pre-allocation of candidates, the 3x3 design may result in an implicit pre-allocation. In this sense, the arrangement of three rows with three candidates per row may cause respondents to feel obligated to select one candidate from each row. When recalling the distribution of selected candidates for group A (figure 33) and group B (figure 36), it is in fact observable that the most frequent candidate for position 1 is selected from the first row, whereas position 2 and 3 are most frequently filled from the second and third row respectively. However, it has to be noted that building the utility maximising team would require respondents to choose one candidate per row. Nevertheless, the arrangement of the candidate table for this treatment should be adapted for future research in order to avoid the formation of an implicit pre-allocation or choice restriction.

With respect to the selected psychometrics, the Berlin Numeracy variable entails a limitation. This measurement is shown to be of insignificant influence for the perceived complexity as well as the bracketing behaviour and achieved utility. This is a surprising result because numeracy was intended to represent the cognitive ability of an individual for this study. The choice for this measurement is based on the circumstance that, especially the lottery part of the experiment, requires the correct interpretation of risks. However, the underlying risky component in the personnel part may be less clear so that numeracy is of reduced importance for this context. While the result that the numeracy of a DM does not significantly influence perceived complexity, bracketing, or utility is acceptable, it is difficult to imagine that the same applies to cognitive ability. This leads to the suggestion that future research may implement a different measure of cognitive ability. The insignificance of the numeracy attribute as a measure of cognitive ability may also indicate, that there is a default mode for bracketing behaviour that is independent of an individual's cognitive capacity. As already noted in the paragraph on the online-setting of the experiment, the use of a calculator or discussing the tasks with colleagues might cause noise in the relation between numeracy and bracketing behaviour. In this sense, even an individual of low cognitive ability could easily achieve an optimal solution when making use of a calculator. The only pre-requisite would then be, that such an individual understands how the candidates in the personnel context need to be evaluated.

Another possible limitation is that the introduced framework addresses seven psychographic characteristics, the objective task complexity, as well as the presentation form but it cannot be ruled out that the results are dependent of other variables. Thus, it shall be encouraged to extend the proposed framework in order to investigate further possible determinants of perceived complexity and bracketing behaviour. A suggestion for an interesting relationship is the consideration of the cognitive style of a DM. A suitable instrument is given by the Cognitive Style Index of Allison and Hayes (1996), which places subjects on an analytic-intuitive scale. Concerning the extension of the framework, future research may also further investigate the role of the decision-making context for choice bracketing. Especially, the inconsistency of exhibited behaviour across different contexts deserves additional attention. It is this inconsistency that prevents the classification of a specific individual into either a broad or narrow bracketer based on their psychometric traits. It follows that a DM has no universal default mode for bracketing behaviour but needs to be classified for specific contexts or possibly groups of contexts. Additional attention should also be given to the interaction of decision support and choice bracketing. This is because the appropriate level and method of support can facilitate broad bracketing for individuals that would otherwise bracket the specific context narrowly as well as reduce the cognitive demand placed on already broad bracketing DMs. Thus, their cognitive resources are not depleted and may be used for the upcoming tasks. Another related direction for research can be found in examining the possibility of learning effects across several tasks of the same context. Findings on such effects could yield important implications on the possibility of training narrow bracketing agents and thereby optimising the decision quality in a given organisation.

Concluding this thesis, it can be claimed that as long as decision-making is of relevance, so will be choice bracketing. Since it is rather unlikely that our private and professional life will be governed by fully automated decision-making in the near future, choice bracketing may not lose its appeal to research. In a world that is increasingly competitive and interconnected, the consequences of a single decision spread through a vast network of decisions and cross-effects. Organisations in such an environment need to rely on agents that are able to process and act on a broader picture, rather than making their decisions in isolation. Hence, the study of what drives bracketing behaviour may grow increasingly important for the hiring process in organisations seeking to gain an advantage in decision-making.

References

Abeler, J., & Marklein, F. (2017). Fungibility, Labels, and Consumption. *Journal of the European Economic Association*, *15*(1), 99-127.

Ainslie, G., & Haslam, N. (1992). Self-control. In G. F. Loewenstein & J. Elster (eds.), *Choice Over Time (pp. 177-212)*. New York: Russel Sage Foundation.

Allinson, C. W., & Hayes, J. (1996). The cognitive style index: A measure of intuition-analysis for organizational research. *Journal of Management Studies*, 33(1), 119-135.

Andersen, S., Harrison, G., Lau, M., & Rutström, E. (2008). Eliciting Risk and Time Preferences. *Econometrica*, *76*(3), 583-618.

Andreoni, J., Gravert, C., Kuhn, M. A., Saccardo, S., & Yang, Y. (2018). Arbitrage Or Narrow Bracketing? On Using Money to Measure Intertemporal Preferences. *NBER Working Paper Series*. https://doi.org/10.3386/w25232

Baddeley, A. (1992). Working Memory. Science, 255(5044), 556-559.

Barberis, N., Huang, M., & Thaler, R. (2006). Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case for Narrow Framing. *American Economic Review*, *96*(4), 1069-1090.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48. doi:10.18637/jss.v067.i01

Benartzi, S., & Thaler, R. (1995). Myopic Loss Aversion and the Equity Premium Puzzle. *The Quarterly Journal of Economics*, *110*(1), 73-92.

Benartzi, S., & Thaler, R. (2007). Heuristics and Biases in Retirement Savings Behavior. *Journal of Economic Perspectives*, *21*(3), 81-104.

Benartzi, S., & Thaler, R. H. (1999). Risk aversion or myopia? Choices in repeated gambles and retirement investments. *Management Science*, *45*(3), 364-381.

Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2013). Who is ‘behavioral’? Cognitive ability and anomalous preferences. *Journal of the European Economic Association*, *11*(6), 1231-1255.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica, 22*(1), 23-36.

Bertrand, M., & Morse, A. (2011). Information Disclosure, Cognitive Biases, and Payday *Borrowing. Journal of Finance, 66*(6), 1865-1893.

Blanco, M., Engelmann, D., Koch, A. K., & Normann, H. T. (2010). Belief elicitation in experiments: is there a hedging problem?. *Experimental Economics, 13*(4), 412-438.

Bland, J. (2019). How many games are we playing? An experimental analysis of choice bracketing in games. *Journal of Behavioral and Experimental Economics, 80,* 80-91.

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences, 106*(19), 7745-7750.

Camerer, C., Babcock, L., Loewenstein, G., & Thaler, R. (1997). Labor supply of New York City cabdrivers: One day at a time. *The Quarterly Journal of Economics, 112*(2), 407-441.

Chang, E. H., Kirgios, E. L., Rai, A., & Milkman, K. L. (2020). The isolated choice effect and its implications for gender diversity in organizations. *Management Science*, 66(6), 2752-2761.

Chen, D., Schonger, M., & Wickens, C. (2016). OTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88-97.

Cokely, E., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making, 7*(1), 25-47.

de Finetti, B. (1974). *Theory of Probability: A critical introductory treatment: 1*, New York: Wiley.

Ellis, A., & Freeman, D. (2020). *Revealing Choice Bracketing*. https://arxiv.org/abs/2006.14869

Engin, A., & Vetschera, R. (2017). Information representation in decision making: The impact of cognitive style and depletion effects. *Decision Support Systems, 103*, 94-103.

Epstein, S. (2003). Cognitive-Experiential Self-Theory of Personality. In I.B. Weiner (Ed.), *Handbook of Psychology* (pp. 159-184).Hoboken, New Jersey: John Wiley & Sons.

Epstein, S. (2014). *Cognitive-Experiential Theory: An Integrative Theory of Personality*. Oxford University Press.

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual Differences in Intuitive–Experiential and Analytical–Rational Thinking Styles. *Journal of Personality and Social Psychology, 71*(2), 390-405.

Fellner, G., & Sutter, M. (2009). Causes, Consequences, and Cures of Myopic Loss Aversion – An Experimental Investigation. *Economic Journal, 119*(537), 900-916.

Fels, M. (2020). Mental accounting, access motives, and overinsurance. *The Scandinavian Journal of Economics*, *122*(2), 675-701.

Felső, F., & Soetevent, A. (2014). Broad and narrow bracketing in gift certificate spending. *European Economic Review, 66*, 284-302.

Finucane, M., Alhakami, A., Slovic, P., & Johnson, S. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making, 13*(1), 1-17.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives, 19*(4), 25-42.

Fudenberg, D., & Levine, D. (2006). A Dual-Self Model of Impulse Control. *American Economic Review, 96*(5), 1449-1476.

Gneezy, U., & Potters, J. (1997). An experiment on risk taking and evaluation periods. The *Quarterly Journal of Economics, 112*(2), 631-646.

Haisley, E., Mostafa, R., & Loewenstein, G. (2008). Myopic risk-seeking: The impact of narrow decision bracketing on lottery play. *Journal of Risk and Uncertainty, 37*(1), 57-75.

Hardin, A., & Looney, C. (2012). Myopic loss aversion: Demystifying the key factors influencing decision problem framing. *Organizational Behavior and Human Decision Processes*, *117*(2), 311-331.

Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, *32*(4), 1199-1228.

Heath, C., & Soll, J. (1996). Mental Budgeting and Consumer Decisions. *Journal of Consumer Research, 23*(1), 40-52.

Herrnstein, R. J. & Prelec, D. (1992a) Melioration. In G. F. Loewenstein & J. Elster (eds.), *Choice Over Time (pp. 235-264)*. New York: Russel Sage Foundation.

Herrnstein, R. J. & Prelec, D. (1992b) A Theory of Addiction. In G. F. Loewenstein & J. Elster (eds.), *Choice Over Time (pp. 331-360)*. New York: Russel Sage Foundation.

Herrnstein, R. J. (1990). Behavior, reinforcement and utility. *Psychological Science*, *1*(4), 217-224.

Herrnstein, R., Loewenstein, G., Prelec, D., & Vaughan, W. (1993). Utility maximization and melioration: Internalities in individual choice. *Journal of Behavioral Decision Making, 6*(3), 149-185.

Heyman, G. M. (1996). Resolving the contradictions of addiction. *Behavioral and Brain Sciences*, *19*(4), 561-574.

Hilgers, D., & Wibral, M. (2014). How malleable are choice brackets? The case of myopic loss aversion. *University of Bonn, Mimeo*.

Hlavac, M. (2018). *Stargazer: well-Formatted Regression and Summary Statistics Tables*. R package version 5.2.1. https://CRAN.R-project.org/package=stargazer

Hsee, C., Blount, S., Loewenstein, G., & Bazerman, M. (1999). Preference Reversals Between Joint and Separate Evaluations of Options: A Review and Theoretical Analysis. *Psychological Bulletin, 125*(5), 576-590.

Hsiaw, A. (2018). Goal bracketing and self-control. *Games and Economic Behavior*, *111*, 100-121.

Kahneman, D. & Lovallo, D. (1993). Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking. *Management Science*, *39*(1), 17-31.

Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *American Economic Review, 93*(5), 1449-1475.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263-292.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*(4), 341-350.

Kahneman, D., Knetsch, J., & Thaler, R. (1990). Experimental Tests of the Endowment Effect and the Coase Theorem. The *Journal of Political Economy, 98*(6), 1325-1348.

Koch, A. K., & Nafziger, J. (2020). Motivational goal bracketing: An experiment. *Journal of Economic Theory*, *185*. https://doi.org/10.1016/j.jet.2019.104949

Koch, A., & Nafziger, J. (2016a). Correlates of Narrow Bracketing. *IDEAS Working Paper Series from RePEc,* Institut for Økonomi, Aarhus Universitet.

Koch, A., & Nafziger, J. (2016b). Goals and bracketing under mental accounting. *Journal of Economic Theory*, *162*, 305-351.

Koch, A., & Nafziger, J. (2019). Correlates of Narrow Bracketing. *Scandinavian Journal of Economics*, *121*(4), 1441-1472.

Lipkus, I., Samsa, G., & Rimer, B. (2001). General Performance on a Numeracy Scale among Highly Educated Samples. *Medical Decision Making, 21*(1), 37-44.

Loewenstein, G., & Prelec, D. (1993). Preferences for Sequences of Outcomes. *Psychological Review, 100*(1), 91-108.

Lüdecke, D. (2020). *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.8.4. https://CRAN.R-project.org/package=sjPlot

Markowitz, H. (1952). The Utility of Wealth. *Journal of Political Economy, 60*(2), 151-158.

Martin, V. (2017). When to quit: Narrow bracketing and reference dependence in taxi drivers. *Journal of Economic Behavior and Organization*, *144,* 166-187.

Miller, G. (1994). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review, 101*(2), 343-352.

Mittelman, M., Andrade, E., Chattopadhyay, A., & Brendl, C. (2014). The Offer Framing Effect: Choosing Single versus Bundled Offerings Affects Variety Seeking. *Journal of Consumer Research*, *41*(4), 953-964.

Moher, E., & Koehler, D. (2010). Bracketing effects on risk tolerance: Generalizability and underlying mechanisms. *Judgment and Decision Making, 5*(5), 339-346.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2): 133-142. DOI: 10.1111/j.2041-210x.2012.00261.x

Pacini, R., & Epstein, S. (1999). The Relation of Rational and Experiential Information Processing Styles to Personality, Basic Beliefs, and the Ratio-Bias Phenomenon. *Journal of Personality and Social Psychology, 76*(6), 972-987.

Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance, 16*(2), 366-387.

Pocheptsova, A., Amir, O., Dhar, R., & Baumeister, R. F. (2009). Deciding without resources: Resource depletion and choice in context. *Journal of Marketing Research*, *46*(3), 344-355.

Pöppel, E. (2008). *Zum Entscheiden geboren: Hirnforschung für Manager*. München: Carl Hanser Verlag GmbH & Co. KG.

Prelec, D., & Loewenstein, G. (1991). Decision making over time and under uncertainty: A common approach. *Management Science, 37*(7), 770-786.

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rabin, M., & Weizsäcker, G. (2009). Narrow Bracketing and Dominated Choices. *American Economic Review, 99*(4), 1508-1543.

Rachlin, H. (1995). Self-control: Beyond Commitment. *Behavioral and Brain Sciences, 18*(1), 109-121.

Read, D., & Loewenstein, G. (1995). Diversification Bias: Explaining the Discrepancy in Variety Seeking Between Combined and Separated Choices. *Journal of Experimental Psychology: Applied, 1*(1), 34-49.

Read, D., Antonides, G., Van Den Ouden, L., & Trienekens, H. (2001). Which Is Better: Simultaneous or Sequential Choice? *Organizational Behavior and Human Decision Processes, 84*(1), 54-70.

Read, D., Loewenstein, G., & Rabin, M. (1999). Choice Bracketing. *Journal of Risk and Uncertainty*, *19*(1), 171-197.

Redelmeier, D. A., & Tversky, A. (1992). On the framing of multiple prospects. *Psychological Science*, *3*(3), 191-193.

Sabini, J., & Silver, M. (1982). *Moralities of Everyday Life*. Oxford: Oxford University Press.

Samuleson, P. (1963). Risk and Uncertainty: A Fallacy of Large Numbers. *Scientia, 98*, 108-113.

Simon, H. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99-118.

Simon, H. A. (1957). Models of man: social and rational. New York: Wiley.

Simonson, I. (1990). The Effect of Purchase Quantity and Timing on Variety-Seeking Behavior. *Journal of Marketing Research, 27*(2), 150-162.

Simonson, I., & Winer, R. (1992). The Influence of Purchase Quantity and Display Format on Consumer Preference for Variety. Journal of Consumer Research, 19(1), 133-138.

Smith, E., & DeCoster, J. (2000). Dual-Process Models in Social and Cognitive Psychology: Conceptual Integration and Links to Underlying Memory Systems. *Personality and Social Psychology Review, 4*(2), 108-131.

Stanovich, K., & West, R. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*(5), 645-665.

Stracke, R., Kerschbamer, R., & Sunde, U. (2015). *Prevalence and Determinants of Choice Bracketing-Experimental Evidence*. Annual Conference 2015: Economic Development - Theory and Policy, Verein für Socialpolitik / German Economic Association, Muenster, Germany. http://hdl.handle.net/10419/113092

Stracke, R., Kerschbamer, R., & Sunde, U. (2017). Coping with complexity – Experimental evidence for narrow bracketing in multi-stage contests. *European Economic Review, 98*, 264-281.

Thaler, R. (1985). Mental Accounting and Consumer Choice. *Marketing Science*, *4*(3), 199-214.

Thaler, R. H. (1999) Mental Accounting Matters. *Journal of Behavioural Decision Making, 12*(3), 183-206.

Thaler, R., Tversky, A., Kahneman, D., & Schwartz, A. (1997). The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test. *The Quarterly Journal of Economics, 112*(2), 647-661.

Tversky, A. (1996). Contrasting rational and psychological principles of choice. In R. J. Zeckhauser, R. L. Keeney, & J. K. Sebenius (Eds.), Wise choices: Decisions, games, and negotiations (pp. 5–21). Boston, MA: Harvard Business School Press.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207-232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. S*cience*, *185*(4157), 1124-1131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. S*cience*, 211(4481), 453-458.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*(4), 293-315.

Venkatraman, S., Aloysius, J., & Davis, F. (2006). Multiple prospect framing and decision behavior: The mediational roles of perceived riskiness and perceived ambiguity. *Organizational Behavior and Human Decision Processes*, *101*(1), 59-73.

Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton Univ. Press.

Webb, E., & Shu, S. (2017). Is broad bracketing always better? How broad decision framing leads to more optimal preferences over repeated gambles. *Judgment and Decision Making*, *12*(4), 382-395.

Wei, T. & Simko, V. (2017). R package „corrplot“: Visualization of a Correlation Matrix. Version 0.84. https://github.com/taiyun/corrplot

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag

Abstract – English

For the past decades, research has concerned itself with the grouping of several choices into decision sets and the effects of such behaviour. As the suggested effects of choice bracketing have far-reaching implications on the quality of the decisions, understanding its determinants requires particular attention. Scholars investigate how the characteristics of a decision problem and those of the decision maker are related to bracketing behaviour. This thesis contributes by investigating the combination of both aspects into the mediating factor of perceived complexity as well as the direct impact of psychographics on bracketing behaviour. For this purpose, a framework that links objective task complexity, individual characteristics, and perceived complexity to bracketing behaviour is proposed. An online-experiment implements decision problems for two distinct decision-making contexts and decision support in the form of information presentation. The experiment yields 378 valid data entries of primarily business and economics students. Results show, that the main drivers of perceived complexity are objective task complexity as well as the rationality of an individual. While perceived complexity significantly impacts bracketing behaviour in both examined contexts, distinct differences are found for the psychometric measurements. The findings indicate that the context of the decision problem is a significant driver for bracketing behaviour. Mixed effects for decision support are reported.

Keywords: Choice bracketing, perceived complexity, psychographics, information presentation

Abstract – Deutsch

Im Laufe der vergangenen Jahrzehnte hat sich die Forschung mit dem Gruppieren einzelner Entscheidungen zu Entscheidungsbündeln, sowie den Effekten eines solchen Verhaltens befasst. Da die von der Literatur behandelten Effekte des „Choice Bracketings“ bedeutende Auswirkungen auf die Güte des Entscheidungsresultats haben, muss dem Verständnis dieser Einflussfaktoren besondere Aufmerksamkeit gewidmet werden. Bestehende Forschung untersucht hierzu einerseits den Zusammenhang zwischen den Merkmalen eines Entscheidungsproblems und ausgewiesenem „Bracketing“-Verhalten und andererseits jenen zwischen persönlichen Eigenschaften und dem Verhalten. Der Beitrag dieser Arbeit liegt in der Verbindung beider Aspekte zur Einflussgröße der wahrgenommenen Komplexität sowie in der Behandlung der psychometrischen Eigenschaften als direkte Einflussgrößen. Zu diesem Zweck wird ein theoretisches Modell vorgestellt, welches die objektive Aufgabenkomplexität, die individuellen Eigenschaften und die subjektive Komplexität mit dem resultierenden Verhalten verknüpft. Die Durchführung eines Online-Experiments inkludiert Problemstellungen aus zwei unterschiedlichen Entscheidungskontexten sowie die Adaption der Informationspräsentation als Unterstützungsmechanismus für den Entscheidungsprozess. Das Experiment liefert 378 gültige Teilnahmen, welche vorwiegend von Studierenden der Betriebs- und Volkswirtschaftslehre stammen. Die Ergebnisse weisen die objektive Aufgabenkomplexität sowie die Rationalität des Entscheidungsträgers als hauptsächliche Einflussfaktoren für die wahrgenommene Komplexität aus. Die wahrgenommene Komplexität wiederum übt in beiden betrachteten Entscheidungsfeldern einen signifikanten Einfluss auf das „Bracketing“-Verhalten aus, während für die psychometrischen Eigenschaften unterschiedliche Einflüsse ermittelt werden. Die Resultate weisen insbesondere auf den Entscheidungskontext als wesentlichen Einflussfaktor für das beobachtete Verhalten hin. Für die Präsentationsform als Entscheidungshilfe werden gemischte Effekte berichtet.

Keywords: Choice bracketing, perceived complexity, psychographics, information presentation

Appendices

Appendix A: Experiment Instructions and Tasks

Page1: Rules and Role

Welcome and thank you for participating in this experiment!

Rules for this experiment:

- No calculator
- No pen and paper
- No group work
- No communication/influence from a third party
- No search machines

Please assume the following situation:

You are the head of a medium-sized department at your company. You will face different situations that require your decision during this experiment. Because you get paid depending on how well your department performs, you want that all processes and decisions made in your department are efficient.

Next

Page 2: Single Lottery

You want to take on a new project for your department and have to choose between the following two options:

Choose between:

○ A) A sure profit of 19.000

○ B) A 20 percent chance to gain 100.000, and 80 percent chance to gain nothing

I would rate the complexity/difficulty of this task as (0% simple – 100% difficult). Type your answer without the percentage sign:

Next

Page 3: Double Lottery

Your department is hiring and you have to decide which candidate you want in your team.

You have two open positions and are presented with two candidates for each position. Both candidates for the second position are no ideal fit but you have to fill in the position and choose one of them.

Please read both decisions carefully before making your choices for each one.

Decision (1): Choose between:

○ A) A candidate with a sure profit of 24.000 generated

○ B) A candidate with a 25 percent chance to generate a profit of 100.000, and 75 percent chance to generate nothing

Decision (2): Choose between:

○ C) A candidate with a sure loss of 75.000 generated

○ D) A candidate with a 75 percent chance to cause a loss of 100.000, and 25 percent chance to lose nothing

I would rate the complexity/difficulty of this task as (0% simple – 100% difficult). Type your answer without the percentage sign:

Next

Page 4: Berlin Numeracy Test (Cokely et al., 2012)

Please answer the questions below. Do not use a calculator.

Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in the choir 100 are men. Out of the 500 inhabitants that are not in the choir 300 are men.What is the probability that a randomly drawn man is a member of the choir? (please indicate the probability in percent).

Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)?

Imagine we are throwing a loaded die (6 sides). The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws, how many times would the die show the number 6?

In a forest 20% of mushrooms are red, 50% brown and 30% white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with a probability of 5%. What is the probability that a poisonous mushroom in the forest is red?

Next

Page 5: Consistency Lottery

Your department is hiring and you have to decide which candidate you want in your team.

You have two open positions and are presented with two candidates for each position. Both candidates for the second position are no ideal fit but you have to fill in the position and choose one of them.

Please read both decisions carefully before making your choices for each one.

Decision (1): Choose between:

- ○ A) A candidate with a sure profit of 19.000 generated
- ○ B) A candidate with a 20 percent chance to generate a profit of 100.000, and 80 percent chance to generate nothing

Decision (2): Choose between:

- ○ C) A candidate with a sure loss of 80.000 generated
- ○ D) A candidate with a 80 percent chance to cause a loss of 100.000, and 20 percent chance to lose nothing

I would rate the complexity/difficulty of this task as (0% simple – 100% difficult). Type your answer without the percentage sign:

Next

Page 6: Combined Lottery

Your department is hiring and you have to decide which candidate you want in your team.

You have one open position and are presented with two candidates. Since you need to fill in the position, you have to choose one of them.

Choose between:

- A) A candidate with a 25 percent chance of generating a profit of 24.000 and a 75 percent chance of generating losses of 76.000
- B) A candidate with a 25 percent chance of generating a profit of 25.000 and a 75 percent chance of losing 75.000

I would rate the complexity/difficulty of this task as (0% simple – 100% difficult). Type your answer without the percentage sign:

Next

Page 7: Cognitive Reflection Test (Frederick, 2005)

Please answer the following questions.

A bat and a ball cost 22 euro in total. The bat costs 20 euro more than the ball. How much does the ball cost? (Please give your answer in cents.):

If it takes 5 machines 5 minutes to make 5 widgets, how many minutes would it take 100 machines to make 100 widgets?

In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how many days would it take for the patch to cover half of the lake?

Next

Page 8: Introduction Personnel Part

Assume the following situation for the next two tasks:

Your department is hiring three new employees. All three positions have the same requirements.
The HR department already sorted all applications and presents you with the final nine candidates. Because all positions have the same requirements, each candidate could be hired for each position.

To help with your decision, the HR department rated the candidates on the four most important skills and their team compatibility factor (TC) based on CVs, cover letters and interviews. The skill items are ranked on a scale from 0% to 100% (0% - low ability, 100% - high ability). This gives the **position-fit** (PF) of a candidate.

The team compatibility (TC) is also rated on a scale of 0% to 100% and represents how well a candidate works in a team. This gives the **team-fit** (TF) of a candidate.

The **overall-fit** (OF) of a candidate is the team-fit multiplied with the position-fit.

OF = PF * TF

Info: You can find these instructions again at the end of the following pages.

Next

Page 9: Control Questions

Please select the statement that correctly compares candidate 1 to candidate 2.

	Candidate 1	Candidate 2
Sum Skill Scores	70	80
Team Compatibility	65	65

Candidate 1 is ... Candidate 2:

- ○ better than
- ○ worse than
- ○ equally good as

	Candidate 1	Candidate 2
Sum Skill Scores	100	50
Team Compatibility	50	100

Candidate 1 is ... Candidate 2:

- ○ better than
- ○ worse than
- ○ equally good as

	Candidate 1	Candidate 2
Sum Skill Scores	150	100
Team Compatibility	50	80

Candidate 1 is ... Candidate 2:

- ○ better than
- ○ worse than
- ○ equally good as

Next

Assume the following situation for the next two tasks:

Your department is allowed to create three new positions. All three positions have the same requirements.
The HR department already sorted all applications and presents you with the final nine candidates. Because all positions have the same requirements, each candidate could be hired for each position.

To help with your decision, the HR department rated the candidates on the four most important skills and their team compatibility factor (TC) based on CVs, cover letters and interviews. The skill items are ranked on a scale from 0% to 100% (0% - low ability, 100% - high ability). This gives the **position-fit** (PF) of a candidate.

The team compatibility (TC) is also rated on a scale of 0% to 100% and represents how well a candidate works in a team. This gives the **team-fit** (TF) of a candidate.

The **overall-fit** (OF) of a candidate is the team-fit multiplied with the position-fit.

OF = PF * TF

Treatment 1 - Segregated Form: Group A

Page 10: Portfolio Building Task

The HR department not only rated the candidates' position-fit and team-fit but also pre-allocated three candidates to each position. This way, you only have to choose between three candidates for each position. Please take a look at all the candidates presented to you before making your final decision.

Position 1:

Skill	Person 1	Person 2	Person 3
Analytics	40	100	95
IT	70	55	70
Languages	85	20	75
Negotiation	55	45	60
Team Compatibility	20	45	35

Please choose your candidate for position 1 out of the candidates 1-3 (e.g. 1):

Position 2:

Skill	Person 4	Person 5	Person 6
Analytics	65	45	20
IT	10	55	90
Languages	90	50	40
Negotiation	15	40	50
Team Compatibility	45	60	40

Please choose your candidate for position 2 out of the candidates 4-6 (e.g. 4):

Position 3:

Skill	Person 7	Person 8	Person 9
Analytics	70	10	85
IT	30	80	50
Languages	65	30	10
Negotiation	55	40	55
Team Compatibility	25	75	35

Please choose your candidate for position 3 out of the candidates 7-9 (e.g. 7):

I would rate the complexity/difficulty of this task as (0% simple – 100% difficult). Type your answer without the percentage sign:

Next

Assume the following situation for the next two tasks:

Your department is allowed to create three new positions. All three positions have the same requirements.
The HR department already sorted all applications and presents you with the final nine candidates. Because all positions have the same requirements, each candidate could be hired for each position.

To help with your decision, the HR department rated the candidates on the four most important skills and their team compatibility factor (TC) based on CVs, cover letters and interviews. The skill items are ranked on a scale from 0% to 100% (0% - low ability, 100% - high ability). This gives the **position-fit** (PF) of a candidate.

The team compatibility (TC) is also rated on a scale of 0% to 100% and represents how well a candidate works in a team. This gives the **team-fit** (TF) of a candidate.

The **overall-fit** (OF) of a candidate is the team-fit multiplied with the position-fit.

OF = PF * TF

Page 11: Portfolio Selection Task

A consultant already built three teams for you to choose from. You are presented with the position-fit and team-fit of each candidate. You cannot transfer candidates between teams.

Please choose the team that you want to hire.

Team 1	Person A	Person B	Person C
Sum Skill Scores	190	200	180
Team Compatibility	50	60	55

Team 2	Person D	Person E	Person F
Sum Skill Scores	240	165	225
Team Compatibility	30	80	40

Team 3	Person G	Person H	Person I
Sum Skill Scores	220	200	280
Team Compatibility	40	45	35

Please choose your preferred team to hire (e.g. 1):

I would rate the complexity/difficulty of this task as (0% simple – 100% difficult). Type your answer without the percentage sign:

Next

Assume the following situation for the next two tasks:

Your department is allowed to create three new positions. All three positions have the same requirements.
The HR department already sorted all applications and presents you with the final nine candidates. Because all positions have the same requirements, each candidate could be hired for each position.

To help with your decision, the HR department rated the candidates on the four most important skills and their team compatibility factor (TC) based on CVs, cover letters and interviews. The skill items are ranked on a scale from 0% to 100% (0% - low ability, 100% - high ability). This gives the **position-fit** (PF) of a candidate.

The team compatibility (TC) is also rated on a scale of 0% to 100% and represents how well a candidate works in a team. This gives the **team-fit** (TF) of a candidate.

The **overall-fit** (OF) of a candidate is the team-fit multiplied with the position-fit.

OF = PF * TF

Treatment 1 - Segregated Form: Group B

Page 10: Portfolio Building Task

The HR department not only rated the candidates' position-fit and team-fit but also pre-allocated three candidates to each position. This way, you only have to choose between three candidates for each position. Please take a look at all the candidates presented to you before making your final decision.

Position 1:

Skill	Person 1	Person 2	Person 3
Analytics	95	65	70
IT	70	10	30
Languages	75	90	65
Negotiation	60	15	75
Team Compatibility	35	45	25

Please choose your candidate for position 1 out of the candidates 1-3 (e.g. 1):

Position 2:

Skill	Person 4	Person 5	Person 6
Analytics	100	85	10
IT	55	50	80
Languages	20	10	30
Negotiation	45	55	40
Team Compatibility	45	35	75

Please choose your candidate for position 2 out of the candidates 4-6 (e.g. 4):

Position 3:

Skill	Person 7	Person 8	Person 9
Analytics	20	45	40
IT	90	55	70
Languages	40	50	85
Negotiation	50	40	55
Team Compatibility	40	60	20

Please choose your candidate for position 3 out of the candidates 7-9 (e.g. 7):

I would rate the complexity/difficulty of this task as (0% simple – 100% difficult). Type your answer without the percentage sign:

Next

Assume the following situation for the next two tasks:

Your department is allowed to create three new positions. All three positions have the same requirements.
The HR department already sorted all applications and presents you with the final nine candidates. Because all positions have the same requirements, each candidate could be hired for each position.

To help with your decision, the HR department rated the candidates on the four most important skills and their team compatibility factor (TC) based on CVs, cover letters and interviews. The skill items are ranked on a scale from 0% to 100% (0% - low ability, 100% - high ability). This gives the **position-fit** (PF) of a candidate.

The team compatibility (TC) is also rated on a scale of 0% to 100% and represents how well a candidate works in a team. This gives the **team-fit** (TF) of a candidate.

The **overall-fit** (OF) of a candidate is the team-fit multiplied with the position-fit.

OF = PF * TF

Page 11: Portfolio Selection Task

A consultant already built three teams for you to choose from. You are presented with the position-fit and team-fit of each candidate. You cannot transfer candidates between teams.

Please choose the team that you want to hire.

Team 1	Person A	Person B	Person C
Sum Skill Scores	190	220	200
Team Compatibility	50	40	45

Team 2	Person D	Person E	Person F
Sum Skill Scores	200	280	220
Team Compatibility	70	30	40

Team 3	Person G	Person H	Person I
Sum Skill Scores	160	180	240
Team Compatibility	75	55	30

Please choose your preferred team to hire (e.g. 1):

I would rate the complexity/difficulty of this task as (0% simple – 100% difficult). Type your answer without the percentage sign:

Next

Assume the following situation for the next two tasks:

Your department is allowed to create three new positions. All three positions have the same requirements.
The HR department already sorted all applications and presents you with the final nine candidates. Because all positions have the same requirements, each candidate could be hired for each position.

To help with your decision, the HR department rated the candidates on the four most important skills and their team compatibility factor (TC) based on CVs, cover letters and interviews. The skill items are ranked on a scale from 0% to 100% (0% - low ability, 100% - high ability). This gives the **position-fit** (PF) of a candidate.

The team compatibility (TC) is also rated on a scale of 0% to 100% and represents how well a candidate works in a team. This gives the **team-fit** (TF) of a candidate.

The **overall-fit** (OF) of a candidate is the team-fit multiplied with the position-fit.

OF = PF * TF

Treatment 2 – Aggregated Form: Group A

Page 10: Portfolio Building Task

Please choose one candidate for each of the three positions.
Please make sure to choose 3 different ones!

Skill	Person 1	Person 2	Person 3
Analytics	40	100	95
IT	70	55	70
Languages	85	20	75
Negotiation	55	45	60
Team Compatibility	20	45	35
	Person 4	**Person 5**	**Person 6**
Analytics	65	45	20
IT	10	55	90
Languages	90	50	40
Negotiation	15	40	50
Team Compatibility	45	60	40
	Person 7	**Person 8**	**Person 9**
Analytics	70	10	85
IT	30	80	50
Languages	65	30	10
Negotiation	55	40	55
Team Compatibility	25	75	35

Please choose your first candidate: (e.g. 1):

Please choose your second candidate: (e.g. 2):

Please choose your third candidate: (e.g. 3):

I would rate the complexity/difficulty of this task as (0% simple – 100% difficult). Type your answer without the percentage sign:

Next

Assume the following situation for the next two tasks:

Your department is allowed to create three new positions. All three positions have the same requirements.
The HR department already sorted all applications and presents you with the final nine candidates. Because all positions have the same requirements, each candidate could be hired for each position.

To help with your decision, the HR department rated the candidates on the four most important skills and their team compatibility factor (TC) based on CVs, cover letters and interviews. The skill items are ranked on a scale from 0% to 100% (0% - low ability, 100% - high ability). This gives the **position-fit** (PF) of a candidate.

The team compatibility (TC) is also rated on a scale of 0% to 100% and represents how well a candidate works in a team. This gives the **team-fit** (TF) of a candidate.

The **overall-fit** (OF) of a candidate is the team-fit multiplied with the position-fit.

OF = PF * TF

Page 11: Portfolio Selection Task

A consultant already built three teams for you to choose from. You are presented with the position-fit and team-fit of each candidate. You cannot transfer candidates between teams.

Please choose the team that you want to hire.

Team 1	Person A	Person B	Person C
Sum Skill Scores	190	200	180
Team Compatibility	50	60	55

Team 2	Person D	Person E	Person F
Sum Skill Scores	240	165	225
Team Compatibility	30	80	40

Team 3	Person G	Person H	Person I
Sum Skill Scores	220	200	280
Team Compatibility	40	45	35

Please choose your preferred team to hire: (e.g. 1):

I would rate the complexity/difficulty of this task as (0% simple – 100% difficult). Type your answer without the percentage sign:

Next

Assume the following situation for the next two tasks:

Your department is allowed to create three new positions. All three positions have the same requirements.
The HR department already sorted all applications and presents you with the final nine candidates. Because all positions have the same requirements, each candidate could be hired for each position.

To help with your decision, the HR department rated the candidates on the four most important skills and their team compatibility factor (TC) based on CVs, cover letters and interviews. The skill items are ranked on a scale from 0% to 100% (0% - low ability, 100% - high ability). This gives the **position-fit** (PF) of a candidate.

The team compatibility (TC) is also rated on a scale of 0% to 100% and represents how well a candidate works in a team. This gives the **team-fit** (TF) of a candidate.

The **overall-fit** (OF) of a candidate is the team-fit multiplied with the position-fit.

OF = PF * TF

Treatment 2 – Aggregated Form: Group B

Page 10: Portfolio Building Task

Please choose one candidate for each of the three positions.
Please make sure to choose 3 different ones!

Skill	Person 1	Person 2	Person 3
Analytics	95	65	70
IT	70	10	30
Languages	75	90	65
Negotiation	60	15	75
Team Compatibility	35	45	25
	Person 4	**Person 5**	**Person 6**
Analytics	100	85	10
IT	55	50	80
Languages	20	10	30
Negotiation	45	55	40
Team Compatibility	45	35	75
	Person 7	**Person 8**	**Person 9**
Analytics	20	45	40
IT	90	55	70
Languages	40	50	85
Negotiation	50	40	55
Team Compatibility	40	60	20

Please choose your first candidate: (e.g. 1):

Please choose your second candidate: (e.g. 2):

Please choose your third candidate: (e.g. 3):

I would rate the complexity/difficulty of this task as (0% simple – 100% difficult). Type your answer without the percentage sign:

Next

Assume the following situation for the next two tasks:

Your department is allowed to create three new positions. All three positions have the same requirements.
The HR department already sorted all applications and presents you with the final nine candidates. Because all positions have the same requirements, each candidate could be hired for each position.

To help with your decision, the HR department rated the candidates on the four most important skills and their team compatibility factor (TC) based on CVs, cover letters and interviews. The skill items are ranked on a scale from 0% to 100% (0% - low ability, 100% - high ability). This gives the **position-fit** (PF) of a candidate.

The team compatibility (TC) is also rated on a scale of 0% to 100% and represents how well a candidate works in a team. This gives the **team-fit** (TF) of a candidate.

The **overall-fit** (OF) of a candidate is the team-fit multiplied with the position-fit.

OF = PF * TF

Page 11: Portfolio Selection Task

A consultant already built three teams for you to choose from. You are presented with the position-fit and team-fit of each candidate. You cannot transfer candidates between teams.

Please choose the team that you want to hire.

Team 1	Person A	Person B	Person C
Sum Skill Scores	190	220	200
Team Compatibility	50	40	45

Team 2	Person D	Person E	Person F
Sum Skill Scores	200	280	220
Team Compatibility	70	30	40

Team 3	Person G	Person H	Person I
Sum Skill Scores	160	180	240
Team Compatibility	75	55	30

Please choose your preferred team to hire: (e.g. 1):

I would rate the complexity/difficulty of this task as (0% simple – 100% difficult). Type your answer without the percentage sign:

Next

Assume the following situation for the next two tasks:

Your department is allowed to create three new positions. All three positions have the same requirements.
The HR department already sorted all applications and presents you with the final nine candidates. Because all positions have the same requirements, each candidate could be hired for each position.

To help with your decision, the HR department rated the candidates on the four most important skills and their team compatibility factor (TC) based on CVs, cover letters and interviews. The skill items are ranked on a scale from 0% to 100% (0% - low ability, 100% - high ability). This gives the **position-fit** (PF) of a candidate.

The team compatibility (TC) is also rated on a scale of 0% to 100% and represents how well a candidate works in a team. This gives the **team-fit** (TF) of a candidate.

The **overall-fit** (OF) of a candidate is the team-fit multiplied with the position-fit.

OF = PF * TF

Page 12: Rational Experiential Inventory (Epstein et al., 1996)

The drop-down menus include the following statements:

----
very strongly agree
strongly agree
moderately agree
slightly agree
neither agree nor disagree
slightly disagree
moderately disagree
strongly disagree
very strongly disagree

Read the given arguments below and choose the best fitting statement for you personally.

I enjoy problems that require hard thinking.

I am not very good in solving problems that require careful logical analysis.

I enjoy intellectual challenges.

I prefer complex to simple problems.

I don't like to have to do a lot of thinking.

Reasoning things out carefully is not one of my strong points.

I am not a very analytical thinker.

I try to avoid situations that require thinking in depth about something.

I am much better at figuring things out logically than most people.

I have a logical mind.

Using logic usually works well for me in figuring out problems in my life.

Knowing the answer without understanding the reason behind it is good enough for me.

I enjoy reading things that evoke visual images.

I enjoy imagining things.

I can clearly picture or remember some sculpture or natural object (not alive) that I think is very beautiful.

I identify strongly with characters in movies or books I read.

I tend to describe things by using images, metaphors or creative comparisons.

Art is really important to me.

Sometimes I like to just sit back and watch things happen.

I have favorite poems and paintings that mean a lot to me.

When I travel or drive somewhere, I always watch the landscape and scenery.

I almost never think in visual images.

My emotions don't make much difference in my life.

Emotions don't really mean much: they come and go.

When I have a strong emotional experience, the effect stays with me for a long time.

When I'm sad, it is often a very strong feeling.

Things that make me feel emotional don't seem to affect other people as much.

Everyday experiences often evoke strong feelings in me.

I'd rather be upset sometimes and happy sometimes, than always feel calm.

I don't react emotionally to scary movies or books as much as other people do.

My anger is often very intense.

When I'm happy, the feeling is usually more like contentment than like exhilaration or excitement.

I like to rely on my intuitive impressions.

I often go by my instincts when deciding on a course of action.

I don't think it is a good idea to rely on one's intuition for important decisions.

I trust my initial feelings about people.

I tend to use my heart as a guide for my actions.

I enjoy learning by doing something, instead of figuring it out first.

I can often tell how people feel without them having to say anything.

I generally don't depend on my feelings to help me make decisions.

For me, descriptions of actual people's experiences are more convincing than discussions about the "facts".

I'm not a very spontaneous person.

Next

Page 13: Demographics Questionnaire

Please answer the following questions accurately.

What is your country of citizenship?

What is your age?

What is your gender?

○ Male ○ Female ○ Diverse

What is your field of study?

What is your mother language?

How would you rate your knowledge of English on a scale from 1 (poor) to 5 (fluent) ?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

Next

Page 14: Final Page

Thank you

Thank you very much for your participation !

You have completed the experiment. This is the last screen.

Your achieved bonuspoints will be added in Moodle before the next examdate.

You may close the tab now.

Appendix B: Instruction Sheet for the Learning Platform

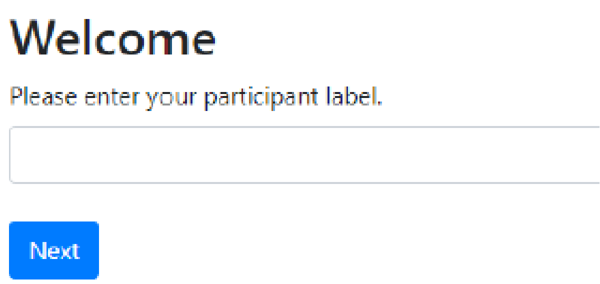
Thank you for your interest in the online experiment

In the following you will receive basic instructions for this online experiment and at the very bottom you will find the participation link.

General Instructions:

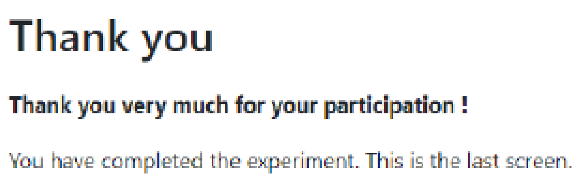
- Please answer the questions honestly and to your best knowledge
- Please do not rush through the questions just to get the bonus points (we will see the time stamps)
- Do the experiment on your own and without consulting others
- **DO NOT** use: the internet, a calculator, paper & pencil for side calculations since we are interested in your intuitive considerations

When you click the link you will be redirected to the experiment. There you have to type in a participant label. This is your **student ID** exactly as it is written on Moodle/your u:card. Once you type in your student ID, the experiment starts. Your student ID is important to us, so that we can assign your bonus points correctly. If you encounter any difficulties with your participant label, please write to:



„e-mail address“

Only when you see the “Thank You!” screen you have reached the end of the experiment and will get the bonus points for your participation. Once you see this page you can close your browser/tab.



The bonus points depend on your performance and will be a minimum of 1,5 points and a maximum of 2 points. If we notice that you just clicked through the experiment (time stamps), you will receive 0 points.

Experiment Link:

„Experiment Link“

Appendix C: Descriptive Statistics for the Holistic Results

Table 25: Descriptive Statistics - Perceived Complexity

	Descriptive Statistics Perceived Complexity							
	Lottery Part	Personnel Part	Stage 0	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Mean	36.66	52.68	18.48	37.27	45.45	47.82	52.53	62.51
Sd	26.75	25.95	24.06	25.55	23.96	25.53	25.71	24.32
Min	0	0	0	0	0	0	0	0
Max	100	100	100	100	100	100	100	100
Q1	10	30	0	15	30	30	38.75	50
Median	35	50	10	40	50	50	50	70
Q3	60	73.5	30	50	60	70	75	80
MAD	37.06	29.65	14.83	29.65	29.65	29.65	29.65	29.65
Range	100	100	100	100	100	100	100	100
Skew	0.23	-0.22	1.38	0.27	-0.05	-0.10	-0.15	-0.6
Kurtosis	-1.02	-0.79	1.11	-0.87	-0.88	-0.84	-0.75	-0.33
n	1512	756	378	378	756	378	188	190

	Descriptive Statistics Perceived Complexity						
	Lottery Part				Personnel Part		
	Single L.	Double L.	Consistency L.	Combined L.	P.-Building (T1)	P.-Building (T2)	P.-Selection
Mean	37.27	46.60	44.30	18.48	52.53	62.51	47.82
Sd	25.55	24.08	23.82	24.06	25.71	24.32	25.53
Min	0	0	0	0	0	0	0
Max	100	100	100	100	100	100	100
Q1	15	30	30	0	38.50	50	30
Median	40	50	50	10	50	70	50
Q3	50	68	60	30	75	80	70
MAD	29.65	29.65	29.65	14.83	29.65	29.65	29.65
Range	100	100	100	100	100	100	100
Skew	0.27	-0.05	-0.04	1.38	-0.15	-0.60	-0.10
Kurtosis	-0.87	-0.86	-0.91	1.11	-0.75	-0.33	-0.84
n	378	378	378	378	188	190	378

Table 26: Descriptive Statistics - Psychometrics

Descriptive Statistics Psychometrics												
	Berlin Numeracy			Rationality			Experientiality			Intuition		
	Overall	Female	Male	Overall	Female	Male	Overall	Female	Male	Overall	Female	Male
Mean	1.49	1.35	1.71	17.43	15.38	20.54	35.24	41.89	25.17	10.45	12.26	7.71
Sd	1.28	1.20	1.35	12.52	13.27	10.58	26.69	25.42	25.46	11.38	11.17	11.18
Min	0	0	0	-22	-22	-16	-47	-35	-47	-37	-23	-37
Max	4	4	4	48	46	48	106	106	84	40	40	35
Q1	0	0	1	11	5.50	15	18	26.5	10.25	3	5	0
Median	1	1	1.5	18	15	20	36	45	26	11	14	9
Q3	2	2	3	27	27	27.75	54	58.5	42.75	18	19.5	15
MAD	1.48	1.48	2.22	11.86	14.83	8.90	26.69	23.72	23.72	10.38	11.86	11.86
Range	4	4	4	70	68	64	153	141	131	77	63	72
Skew	0.54	0.65	0.34	-0.39	-0.30	-0.25	-0.3	-0.36	-0.28	-0.36	-0.37	-0.39
Kurtosis	-0.75	-0.49	-1.08	0.03	-0.39	0.67	0.13	0.37	-0.05	0.46	0.21	0.86
n	377	227	150	377	227	150	377	227	150	377	227	150

Descriptive Statistics Psychometrics									
	Emotionality			Imagination			Implusiveness (CRT)		
	Overall	Female	Male	Overall	Female	Male	Overall	Female	Male
Mean	8.55	11.67	3.83	16.24	17.97	13.63	1.52	1.35	1.79
Sd	12.06	11.39	11.55	11.59	11.16	11.78	1.13	1.12	1.08
Min	-30	-30	-25	-18	-16	-18	0	0	0
Max	35	35	29	40	40	38	3	3	3
Q1	1	4	-4	9	10.50	6.25	0	0	1
Median	9	13	3	17	19	15	2	1	2
Q3	18	20	12	25	26	22	3	2	3
MAD	13.34	10.38	11.86	11.86	11.86	11.86	1.48	1.48	1.48
Range	65	65	54	58	56	56	3	3	3
Skew	-0.33	-0.51	-0.14	-0.45	-0.40	-0.48	-0.06	0.16	-0.39
Kurtosis	-0.11	0.39	-0.30	0	-0.04	-0.14	-1.38	-1.36	-1.16
n	377	227	150	377	227	150	377	227	150

Table 27: Frequencies - Berlin Numeracy and CRT

	Frequencies Psychometrics		
	Overall	Female	Male
Berlin Numeracy - # Correct Answers			
0	100 (26.53)	66 (29.07)	34 (22.67)
1	113 (29.97)	72 (31.72)	41 (27.33)
2	81 (21.49)	49 (21.59)	32 (21.33)
3	45 (11.94)	24 (10.57)	21 (14.00)
4	38 (10.08)	16 (7.05)	22 (14.67)
Implusiveness (CRT) - # Correct Answers			
0	96 (25.46)	70 (30.84)	26 (17.33)
1	83 (22.02)	54 (23.79)	29 (19.33)
2	103 (27.32)	57 (25.11)	46 (30.67)
3	95 (25.20)	46 (20.26)	49 (32.67)

*relative frequencies in parentheses

Table 28: Descriptive Statistics - Time in Experimental Parts

	Descriptive Statistics Time (in Minutes)			
	Whole Experiment	Lottery Part	Personnel Part	Surveys
Mean	75.19	22.53	9.76	36.45
Sd	404.80	255.57	5.74	296.45
Min	12.78	1.30	2.30	7.10
Max	5792.75	4767.03	59.2	5777.63
Q1	25.85	3.67	6.38	13.86
Median	33.10	4.62	8.40	17.98
Q3	42.85	6.01	11.55	23.99
MAD	12.50	1.63	2.30	7.40
Range	5779.97	4765.73	56.9	5770.53
Skew	12.09	17.36	3.23	19.22
Kurtosis	153.23	313.58	18.82	369.14
n	378	378	378	378

Appendix D: Descriptive Statistics for the Lottery Part

Table 29: Choices in the Single Lottery Task

	Choice Single Lottery	
	Sure Gain	Gamble
Count	286 (75.66)	92 (24.34)

*relative frequencies in parentheses

Table 30: Choices in the Double Lotteries Tasks

	Choices Double Lottery		Choices Consistency Lottery	
	Sure Outcome	Gamble	Sure Outcome	Gamble
Decision 1	278 (73.54)	100 (26.46)	243 (64.29)	135 (35.71)
Decision 2	54 (14.29)	324 (85.71)	53 (14.02)	325 (85.98)
Choice Combinations:				
A&D	247 (65.34)		211 (55.82)	
B&C	23 (6.08)		21 (5.56)	
Others	108 (28.57)		146 (38.62)	
Cross Table Double and Consistency Lottery:				
	A&D	B&C	Others	
A&D	185 (48.94)	7 (1.85)	55 (14.55)	
B&C	4 (1.06)	9 (2.38)	10 (2.65)	
Others	22 (5.82)	5 (1.32)	81 (21.43)	

*relative frequencies in parentheses

Table 31: Choices in the Combined Lottery Task

	Choice Combined Lottery	
	Combined Form A&D	Combined Form B&C
Count	21 (5.56)	357 (94.44)

*relative frequencies in parentheses

Table 32: Descriptive Statistics - Times for the Lottery Tasks

	Descriptive Statistics Time (in Seconds) - Lottery Part			
	Lottery Part			
	Single L.	Double L.	Consistency L.	Combined L.
Mean	99.93	137.91	1054.35	59.58
Sd	455.87	531.08	15322.97	45.84
Min	22	17	10	4
Max	8750	10323	285844	513
Q1	47	68	40	38
Median	60	90	56	49.5
Q3	83	124.75	84	69
MAD	23.72	39.29	28.91	21.5
Range	8728	10306	285834	509
Skew	18.16	18.67	17.42	5.54
Kurtosis	340.2	354.89	315.05	43.44
n	378	378	378	378

Appendix E: Descriptive Statistics for the Personnel Part

Table 33: Descriptive Statistics - Times for the Personnel Tasks

Descriptive Statistics Time (in Seconds) - Personnel Selection Part	P.-Building (T1)			P.-Building (T2)			P.-Selection		
	Overall	A	B	Overall	A	B	Overall	A	B
Mean	236.71	217.03	255.98	276.07	249.92	302.22	119.71	110.67	128.66
Sd	175.07	101.7	223.82	192.28	156	220.5	97	72.52	115.78
Min	51	51	56	53	53	69	14	14	24
Max	1670	469	1670	1531	978	1531	1097	457	1097
Q1	132.75	145	130	156.25	150	167.5	62	59.5	67
Median	203	204	201	224	208	246	95	94	97
Q3	281.5	275	295	321.75	311.5	344.5	146.75	145.25	151.25
MAD	106.75	93.4	115.64	117.13	105.26	134.92	56.34	60.05	56.34
Range	1619	418	1614	1478	925	1462	1083	443	1073
Skew	3.95	0.55	3.49	2.63	1.75	2.69	4.06	1.47	4.26
Kurtosis	25.03	-0.36	16.42	10.87	4.12	10.15	30.32	2.93	27.5
n	188	93	95	190	95	95	378	188	190

Portfolio Building Task:

Table 34: Diversification - Candidate Evaluation (Group A)

	Candidate Skill Values - Group A								
Skill Items	1	2	3	4	5	6	7	8	9
Analytics	40	100	95	65	45	20	70	10	85
IT	70	55	70	10	55	90	30	80	50
Languages	85	20	75	90	50	40	65	30	10
Negotiation	55	45	60	15	40	50	55	40	55
TC	20	45	35	45	60	40	25	75	35

	Weighted Skill Values - Group A								
Skill Items	1	2	3	4	5	6	7	8	9
Analytics	8	45	33,25	29,25	27	8	17,5	7,5	29,75
IT	14	24,75	24,5	4,5	33	36	7,5	60	17,5
Languages	17	9	26,25	40,5	30	16	16,25	22,5	3,5
Negotiation	11	20,25	21	6,75	24	20	13,75	30	19,25
Sd	3,87	15,04	5,15	17,52	3,87	11,78	4,45	22,08	10,79
Variance	15,00	226,13	26,54	307,13	15,00	138,67	19,79	487,50	116,38

Table 35: Diversification - Candidate Evaluation (Group B)

	Candidate Skill Values - Group B								
Skill Items	1	2	3	4	5	6	7	8	9
Analytics	95	65	70	100	85	10	20	45	40
IT	70	10	30	55	50	80	90	55	70
Languages	75	90	65	20	10	30	40	50	85
Negotiation	60	15	75	45	55	40	50	40	55
TC	35	45	25	45	35	75	40	60	20

	Weighted Skill Values - Group B								
Skill Items	1	2	3	4	5	6	7	8	9
Analytics	33,25	29,25	17,5	45	29,75	7,5	8	27	8
IT	24,5	4,5	7,5	24,75	17,5	60	36	33	14
Languages	26,25	40,5	16,25	9	3,5	22,5	16	30	17
Negotiation	21	6,75	18,75	20,25	19,25	30	20	24	11
Sd	5,15	17,52	5,10	15,04	10,79	22,08	11,78	3,87	3,87
Variance	26,54	307,13	26,04	226,13	116,38	487,50	138,67	15,00	15,00

Table 36: Portfolio Building Task - Candidate Frequencies

Portfolio Building Group A								
	Segregate Presentation (T1)				Aggregate Presentation (T2)			
Candidates	Postition 1	Position 2	Position 3	Overall	Position 1	Position 2	Position 3	Overall
1	2 (2.15)	-	-	2 (0.72)	1 (1.05)	12 (12.63)	8 (8.42)	21 (7.37)
2	14 (15.05)	-	-	14 (5.02)	11 (11.58)	8 (8.42)	13 (13.68)	32 (11.23)
3	77 (82.80)	-	-	77 (27.60)	58 (61.05)	8 (8.42)	13 (13.68)	79 (27.72)
4	-	7 (7.53)	-	7 (2.51)	0 (0.00)	4 (4.21)	2 (2.11)	6 (2.11)
5	-	77 (82.80)	-	77 (27.60)	7 (7.37)	50 (52.63)	8 (8.42)	65 (22.81)
6	-	9 (9.68)	-	9 (3.23)	0 (0.00)	4 (4.21)	6 (6.32)	10 (3.51)
7	-	-	45 (48.39)	45 (16.13)	0 (0.00)	1 (1.05)	18 (18.95)	19 (6.67)
8	-	-	36 (38.71)	36 (12.90)	17 (17.89)	6 (6.32)	20 (21.05)	43 (15.09)
9	-	-	12 (12.90)	12 (4.30)	1 (1.05)	2 (2.11)	7 (7.37)	10 (3.51)
Modus	C3	C5	C7	C3, C5	C3	C5	C8	C3

*relative frequencies in parentheses

Portfolio Building Group B								
	Segregate Presentation (T1)				Aggregate Presentation (T2)			
Candidates	Postition 1	Position 2	Position 3	Overall	Position 1	Position 2	Position 3	Overall
1	88 (92.63)	-	-	88 (30.88)	65 (68.42)	10 (10.53)	8 (8.42)	83 (29.12)
2	6 (6.32)	-	-	6 (2.11)	1 (1.05)	21 (2.11)	2 (2.11)	5 (1.75)
3	1 (1.05)	-	-	1 (0.35)	2 (2.11)	3 (3.16)	7 (7.37)	12 (4.21)
4	-	49 (51.58)	-	49 (17.19)	3 (3.16)	24 (25.26)	18 (18.95)	45 (15.79)
5	-	3 (3.16)	-	3 (1.05)	0 (0.00)	3 (3.16)	2 (2.11)	5 (1.75)
6	-	43 (45.26)	-	43 (15.09)	17 (17.89)	21 (22.11)	6 (6.32)	44 (15.44)
7	-	-	6 (6.32)	6 (2.11)	0 (0.00)	2 (2.11)	6 (6.32)	8 (2.81)
8	-	-	48 (50.53)	48 (16.84)	4 (4.21)	16 (16.84)	19 (20.00)	39 (13.68)
9	-	-	41 (43.16)	41 (14.39)	3 (3.16)	14 (14.74)	27 (28.42)	44 (15.44)
Modus	C1	C4	C8	C1	C1	C4	C9	C1

*relative frequencies in parentheses

Table 37: Portfolio Building Task – Descriptive Statistics: Changes in Behaviour

	Changes in Candidate Selection						
			Group A			Group B	
	Overall	Overall	Segregate Presentation (T1)	Aggregate Presentation (T2)	Overall	Segregate Presentation (T1)	Aggregate Presentation (T2)
Mean	0.73	0.67	0.77	0.57	0.79	0.67	0.92
Sd	1.85	1.97	1.68	2.22	1.73	1.75	1.71
Min	0	0	0	0	0	0	0
Max	17	17	12	17	11	11	10
Q1	0	0	0	0	0	0	0
Median	0	0	0	0	0	0	0
Q3	1	1	1	0	1	1	1
MAD	0	0	0	0	0	0	0
Range	17	17	12	17	11	11	10
Skew	4.71	5.61	4.04	6.09	3.31	3.89	2.7
Kurtosis	28.03	36.7	21.09	38.6	12.68	16.59	8.69
n	378	188	93	95	190	95	95
# Changes							
0	265 (70.11)	136 (72.34)	61 (65.59)	75 (78.95)	129 (67.89)	70 (73.68)	59 (62.11)
1	57 (15.08)	28 (14.89)	15 (16.13)	13 (13.68)	29 (15.26)	12 (12.63)	17 (17.89)
2	28 (7.41)	14 (7.45)	10 (10.75)	4 (4.21)	14 (7.37)	7 (7.37)	7 (7.37)
3	11 (2.91)	4 (2.13)	3 (1.60)	1 (1.05)	7 (3.68)	2 (2.11)	5 (5.26)
4	2 (0.53)	1 (0.53)	1 (1.08)	-	1 (0.53)	-	1 (1.05)
5	6 (1.59)	1 (0.53)	1 (1.08)	-	5 (2.63)	1 (1.05)	4 (4.21)
7	2 (0.53)	1 (0.53)	1 (1.08)	-	1 (0.53)	-	1 (1.05)
8	2 (0.53)	-	-	-	2 (1.05)	2 (2.11)	-
10	1 (0.26)	-	-	-	1 (0.53)	-	1 (1.05)
11	1 (0.26)	-	-	-	1 (0.53)	1 (1.05)	-
12	1 (0.26)	1 (0.53)	1 (1.08)	-	-	-	-
13	1 (0.26)	1 (0.53)	-	1 (1.05)	-	-	-
17	1 (0.26)	1 (0.53)	-	1 (1.05)	-	-	-

*relative frequencies in parentheses

Table 38: Portfolio Building Task – Descriptive Statistics Variety Seeking Behaviour

		Variety Seeking					
			Group A			Group B	
Built Team	Overall	Overall	Segregate Presentation (T1)	Aggregate Presentation (T2)	Overall	Segregate Presentation (T1)	Aggregate Presentation (T2)
Diverse	135 (35.71)	70 (37.23)	34 (36.56)	36 (37.89)	65 (34.21)	34 (35.79)	31 (32.61)
Non-Diverse	243 (64.29)	118 (62.76)	59 (63.44)	59 (62.11)	125 (65.79)	61 (64.21)	64 (67.37)
Variety Ratio - Ratio to max. Achievable Variety in Team Composition:							
Mean	78.18	78.25	77.76	78.73	78.11	78.46	77.77
Sd	9.53	9.78	9.92	9.68	9.29	8.43	10.12
Min	62.38	62.38	64.86	62.38	62.38	62.38	62.38
Max	100	100	100	100	93.52	91.25	93.52
Q1	69.02	70.74	70.74	70.74	69.02	69.02	69.02
Median	77.27	70.74	70.74	77.27	77.27	77.27	77.27
Q3	86.99	86.99	86.99	86.99	86.99	86.99	86.99
MAD	12.23	5.29	0.65	12.23	12.23	12.23	12.23
Range	37.62	37.62	35.14	37.62	31.13	28.87	31.13
Skew	0.23	0.43	0.61	0.25	0	-0.08	0.08
Kurtosis	-1.4	-1.33	-1.25	-1.4	-1.53	-1.57	-1.59
n	378	188	93	95	190	95	95

*relative frequencies in parentheses

Table 39: Portfolio Building Task - Descriptive Statistics: Utility

Descriptive Statistics Utility							
			Group A			Group B	
Max. Utility	Overall	Overall	Segregate Presentation (T1)	Aggregate Presentation (T2)	Overall	Segregate Presentation (T1)	Aggregate Presentation (T2)
Achieved	94 (24.87)	43 (22.87)	21 (22.58)	22 (23.16)	51 (26.84)	29 (30.53)	22 (23.16)
Not Achieved	284 (75.13)	145 (77.13)	72 (77.42)	73 (76.84)	139 (73.16)	66 (69.47)	73 (76.84)
Utility Categories							
1 [0%;70%]	18 (4.76)	5 (2.66)	2 (2.15)	3 (3.16)	13 (6.84)	2 (2.11)	11 (11.58)
2]71%; 80%]	91 (24.07)	27 (14.36)	9 (9.68)	18 (18.95)	64 (33.68)	34 (35.79)	30 (31.58)
3]81%; 95%]	149 (39.42)	96 (51.06)	53 (56.99)	43 (45.26)	53 (27.89)	30 (31.58)	23 (24.21)
4]95%; 100%	120 (31.75)	60 (31.91)	29 (31.18)	31 (32.63)	60 (31.58)	29 (30.53)	31 (32.63)
Achieved Utility							
Mean	292.61	294.60	293.24	295.94	290.63	294.83	286.43
Sd	37.36	34.45	33.81	35.20	40.01	37.89	41.81
Min	201	201	201	210	209	225	209
Max	339	339	339	339	339	339	339
Q1	254	274	274	274	254	254	254
Median	285	289	274	295	282.50	305	275
Q3	333	333	333	333	339	339	333
MAD	45.96	40.03	22.24	56.34	48.18	50.41	57.82
Range	138	138	138	129	130	114	130
Skew	-0.1	-0.09	0.03	-0.21	-0.06	-0.06	0
Kurtosis	-1.23	-0.93	-0.85	-1.02	-1,48	-1.69	-1.43
n	378	188	93	95	190	95	95

*relative frequencies in parentheses

Portfolio Selection Task:

Table 40: Portfolio Selection Task - Team Frequencies

	Portfolio Selection	
Teams	Group A	Group B
1	78 (41.49)	17 (8.95)
2	45 (23.94)	148 (77.89)
3	65 (34.57)	25 (13.16)
Modus	T1	T2

*relative frequencies in parentheses

Table 41: Portfolio Selection Task - Descriptive Statistics: Changes in Behaviour

	Changes in Team Selection		
	Overall	Group A	Group B
Mean	0.21	0.23	0.19
Sd	0.67	0.74	0.59
Min	0	0	0
Max	6	6	4
Q1	0	0	0
Median	0	0	0
Q3	0	0	0
MAD	0	0	0
Range	6	6	4
Skew	4.4	4.62	3.62
Kurtosis	24.08	25.56	14.54
n	378	188	190
# Changes			
0	327 (86.51)	161 (85.64)	166 (87.37)
1	35 (9.26)	20 (10.64)	15 (7.89)
2	8 (2.12)	2 (1.06)	6 (3.16)
3	4 (1.06)	2 (1.06)	2 (1.05)
4	3 (0.79)	2 (1.06)	1 (0.53)
6	1 (0.26)	1 (0.53)	-

*relative frequencies in parentheses

Table 42: Portfolio Selection Task - Descriptive Statistics: Utility

	Descriptive Statistics Utility		
Max. Utility	Overall	Group A	Group B
Achieved	226 (59.79)	78 (41.49)	148 (77.89)
Not Achieved	152 (40.21)	110 (58.51)	42 (22.11)
Achieved Utility			
Mean	300.94	296.07	305.75
Sd	15.43	16.59	12.48
Min	273	276	273
Max	314	314	312
Q1	291	276	312
Median	312	294	312
Q3	312	314	312
MAD	2.97	26.69	0.00
Range	41	38	39
Skew	-0.75	-0.09	-1.73
Kurtosis	-1.12	-1.69	1.54
n	378	188	190

*relative frequencies in parentheses

Appendix F: Hierarchical Logistic Regression on Broad Bracketing Overall

Table 43: Hierarchical Logistic Regression on Broad Bracketing Overall

Hierarchical Logistic Regression on Broad Bracketing Overall

	Dependent variable:
	Broad Bracketing
Intercept	-3.225*** (0.713)
Perceived Complexity	0.010*** (0.004)
CRT Score	0.265*** (0.089)
Berlin Numeracy	-0.0001 (0.075)
Rationality	0.010 (0.008)
Emotionality	0.002 (0.009)
Imagination	0.002 (0.008)
Intuition	0.002 (0.009)
Observations	1,134
Log Likelihood	-417.583
Akaike Inf. Crit.	853.165
Bayesian Inf. Crit.	898.467
Note:	*p<0.1; **p<0.05; ***p<0.01

Values in parentheses are standard errors