UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 178

Integrating Probability- and Value Information in Judgment and Decision-Making under Risk

Cognitive Processes, Competence, and Performance

PHILIP MILLROTH

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2020

ISSN 1652-9030
ISBN 978-91-513-0924-8
urn:nbn:se:uu:diva-407865

Dissertation presented at Uppsala University to be publicly examined in Humanistiska Teatern, Engelska parken, Thunbergsvägen 3C, Uppsala, Wednesday, 20 May 2020 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Bettina Von Helversen.

Abstract

Millroth, P. 2020. Integrating Probability- and Value Information in Judgment and Decision-Making under Risk. Cognitive Processes, Competence, and Performance. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 178. 83 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-0924-8.

Many instances in human affairs involve considering the value of different outcomes and the probability (or risk) of these outcomes occuring (e.g., gambling, financial decision-making, medical decision-making, criminal behavior). The point of departure for present thesis is that descriptive theories of judgment- and decision making under risk have yet to fully utilize explanations grounded in accounts of how people integrate outcomes with their adherent probabilities. The most widely embraced accounts are positioned on opposite ends of a spectrum, holding either (*i*) that people consistently and effortlessly engage in the normative principle of multiplicatively integrating the value or utility of possible outcomes with their adherent probabilities (i.e., *weighting*), or (*ii*) that people only have the ability to engage in simple heuristics or context-dependent sampling strategies. The present thesis proposes that the field should consider positions between these extreme positions. To this end, three empirical studies were conducted in which people evaluated risky prospects in the form of numerically described monetary lotteries.

The studies show that use of weighting was robust to increases of cognitive demands, as when (*i*) other evaluations are not available as reference points (Study I), (*ii*) outcomes and probabilities are presented sequentially before the evaluation (Study II), and (*iii*) the prospect structure involves two independent outcomes (Study III). The results suggest that - even if people can turn to heuristics when they are more efficient, for specific stages in the decision process, or for very complex problems - people indeed have both the inclination and ability to weight the outcomes by their probabilities in the evaluation of individual prospects, or for a subset of decision alternatives.

In contrast to popular weighting models, however, the cognitive-modeling efforts throughout the studies speak against the notion that the weighting process can be assumed to be consistent and effortless. Instead, the cognitive process of weighting outcomes and probabilities is better characterized as an anchoring-and-adjustment strategy: people anchor on the value of the outcome and make linear adjustments downwards to account for probability. The studies show that these adjustments are often insufficient or noise-prone when the cognitive demands increase due to (*i*) properties of the task environment (Study I and Study II), or (*ii*) lack of domain-specific knowledge (i.e., numeracy and financial literacy, Study III). In conclusion, the thesis has highlighted the important, but previously neglected, nuances of human cognition in judgment and decision-making under risk - nuances found between previously conflicting standpoints. Future research exploring these nuances should make a necessary distinction between people's underlying competence and the performance they exhibit at a given moment.

Keywords: Judgment, Decision-making, Risk, Cognition, Information pro-cessing, Information integration

Philip Millroth, Department of Psychology, Box 1225, Uppsala University, SE-75142 Uppsala, Sweden.

ISSN 1652-9030
ISBN 978-91-513-0924-8
urn:nbn:se:uu:diva-407865 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-407865)

To Sofia, Simone, and Joline:

KÄRLEK!

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Millroth, P., Nilsson, H., & Juslin, P. (2018). Examining the integrity of evaluations of risky prospects using a single-stimuli design. *Decision*, 5(4):362–377.

II Millroth, P., Guath, M., & Juslin, P. (2019). Memory and decision making: Effects of sequential presentation of probabilities and outcomes in risky prospects. *Journal of Experimental Psychology: General*, 2(148):304–324.

III Millroth, P., Nilsson, H., Juslin, P., & Sundh, J. (2020). Additive weighting of outcomes in duplex gambles: Spontaneous prevalence, cognitive processing, and the effect of financial literacy. *Manuscript.*

Reprints were made with permission from the respective publishers.

For all studies included in this thesis, Philip Millroth planned and designed the experiments, analyzed the data, and wrote the manuscripts, along with contributions to all aforementioned areas from supervisor and co-authors.

Additional scientific work not included in the thesis

1. **Millroth, P**., Juslin, P., Winman, A., Nilsson, H., & Lindskog, M. (2020). Preference or ability: Exploring the relations between risk preference, personality, and cognitive abilities. *Journal of Behavioral Decision Making*. Online first publication.
2. Agren, T., **Millroth, P**., Andersson, P., Ridzén, M., & Björkstrand, J. (2019). Detailed analysis of skin conductance responses during a gambling task: Decision, anticipation, and outcomes. *Psychophysiology*, *56*(6), e13338.
3. **Millroth, P.,** Nilsson, H., & Juslin, P. (2019). The decision paradoxes motivating Prospect Theory: The prevalence of the paradoxes increases with numerical ability. *Judgment and Decision Making*, *14*(4), 513-533.
4. Sundh, J., Juslin, P., & **Millroth, P.** (2020). Appreciation for independence: does adaptation to stochastic dependence imply thinking according to stochastic principles? In Sundh, J (thesis.), *The cognitive basis of joint probability judgments: processes, ecology, and adaption*. Uppsala: Acta Universitatis Upsaliensis.
5. **Millroth, P**., Juslin, P., Eriksson, E., & Agren, T. (2017). Disentangling the effects of serotonin on risk perception: S-carriers of 5-HTTLPR are primarily concerned with the magnitude of the outcomes, not the uncertainty. *Behavioral Neuroscience*, *131*(5), 421.
6. **Millroth, P.** (2017). Descriptive statistics and Bayesian hypothesis testing show that the intervention enhances only geometric sensitivity: Comment on Dillon et al. (2017). *Science E-Letter*. Available: http://science.sciencemag.org/content/357/6346/47/tab-e-letters.
7. **Millroth, P**., & Juslin, P. (2015). Prospect evaluation as a function of numeracy and probability denominator. *Cognition*, *138*, 1-9.
8. Guath, M., **Millroth, P**., Juslin, P., & Elwin, E. (2015). Optimizing electricity consumption: A case of function learning. *Journal of Experimental Psychology: Applied*, *21*(4), 326.
9. Juslin, P., Elwin, E., Guath, M., **Millroth, P.,** & Nilsson, H. (2016). Sequential and myopic: On the use of feedback to balance cost and utility in a simulated electricity efficiency task. *Journal of Cognitive Psychology*, *28*(1), 106-128.

Contents

Abbreviations

AAM	Anchoring and Adjustment Model
ANOVA	Analysis of Variance
AW	Additive Weighting
BF	Bayes Factor
BSD	Between-Subjects Design
CE	Certainty Equivalent
CI	Credible Interval
CPT	Cumulative Prospect Theory
EUT	Expected Utility Theory
EV-C	Expected Value Computation
EVT	Expected Value Theory
FL	Financial Literacy
PNP	Precise/Non-Precise
PT	Prospect Theory
SSD	Single-Stimuli Design
STM	Short Term Memory
WSD	Within-Subjects Design

Introduction

Throughout the life span, people frequently face judgments and decisions that involve risk[1]: Should I invest money in stocks or bonds? How much money should I invest? Should I buy insurance against wild fires? Should I use my last dollars to buy a lottery ticket? All these situations involve two central attributes: the utility of a possible outcome, and the probability of that outcome occurring (i.e., the risk). For example, the lottery ticket may yield an outcome of a million dollars (the possible outcome), but the outcome is very unlikely – only one in a million lottery tickets yield this jackpot (the probability).

Research on *risk preferences* and the assumed rationality (or irrationality) of these preferences have for centuries enriched a variety of theories in the behavioral sciences (e.g., Bernoulli, 1738/1954; for reviews, see Fox, Erner, & Walters, 2015; Schoemaker, 1982), ranging from Psychology (e.g., Kahneman & Tversky, 1979; Mallpress, Fawcett, Houston, & MacNamara, 2015; Simon, 1991; Tversky & Kahneman, 1992), to Artificial Intelligence (e.g., Arthur, 1991), Political Science (Simon, 2000), and Economics (e.g., Barberis, Huang, & Santos, 2001; Camerer, Loewenstein, & Rabin, 2011). It is difficult to convey the immense impact this research program has had, but citation indexes provide a glimpse of this reality: The articles that outline the benchmark descriptive theories of judgment and decision-making under risk, prospect theory (PT: Kahneman & Tversky, 1979), and its successor - cumulative prospect theory (CPT: Tversky & Kahneman, 1992) - are the most cited articles in Economics and Psychology over the last 50 years (Simonsohn, 2014).

Given the descriptive focus and the immense impact that the field of judgment and decision-making under risk has had on adjacent fields, it is surprising that the it has yet to fulfill the central aims and purposes of Cognitive

[1] Judgment and decision making under *risk* is here defined as judgments and decisions in which the probability distribution over possible outcomes is known precisely by the judge or decision maker. This is the predominant definition used in decision theory since Knight (1921). It can be contrasted with the definition of judgment and decision making under *uncertainty:* judgments and decisions in which the probability distribution over possible outcomes cannot be known. Wakker (2010) provides arguments for the case that many results from the study of risk can be extended to that of uncertainty (for counterarguments, see Hertwig, Pleskac, & Pachur, 2019).

Psychology, namely to uncover the *cognitive processes* – the mental procedures in charge of processing all the information received from the environment in question. Judgment and decision-making inevitably recruit basic cognitive processes from systems of perception, attention, and memory. Specification of the information-processing procedures and the cognitive processes that are involved when people make judgments and decisions is necessary for a more complete explanation of any observed behavior. Without such specification, interpretations grounded in higher-level concepts such as *preferences* and *attitudes* will possibly, or perhaps even likely, be confounded with information-processing explanations that outline, for example, how information is retrieved and attended to (for reviews, see Oppenheimer & Kelso, 2015; Weber & Johnson, 2009).

One central cognitive process that is yet to be fully theoretically utilized is the process of *information integration* - a process whereby two or more distinct stimuli, or stimuli features, are combined into a whole. In other words, the processing of each individual stimulus does not by itself give rise to the integration product: it is only the combination of the stimuli that constitutes an integration, as, for example, when the representation of '5' is activated by the addition of '2' and '3' (Mudrik, Faivre, & Koch, 2014). Explanations of human behavior grounded in information-integration accounts reach back to (at least) Aristotle's discussions of how the mind perceives sums and parts (Mendell, 2017), and such accounts have over the centuries provided unified explanations of behavior in many fields of Psychology (Wundt, 1904: 1907; Anderson, 1996; Einhorn & Hogarth, 1981; Hoffman, 1960; Oppenheimer & Kelso, 2015; Slovic & Lichtenstein, 1971).

Per definition, research on judgment and decision-making under risk studies the roles of two central attributes – outcomes and the probabilities of the outcomes occurring. How are these attributes integrated? What are the characteristics of the cognitive processes underpinning the integration processes? The accumulated scientific literature over the past 60 years have resulted in contradicting information-integration accounts. The most widely adapted account assumes that people, by default, consistently and effortlessly *weight* the utility of all the outcomes with their probabilities (or probability weights) (e.g., Bell, 1985; Loomes & Sugden, 1982; Luce, 2000; Quiggin, 1993; Tversky & Kahneman, 1992). That is, probability- and value information is integrated multiplicatively – as, for example, when calculating an expected value (see Background Section). On this account, departures from normative decision theory is explained mainly, or solely, by how people respond to the key inputs of outcomes and probabilities. Another widely embraced account suggests that people are generally reluctant to engage in the weighting and addition of outcomes (e.g., Berg & Gigerenzer, 2010). Instead, people are assumed to engage in strategies that ignore much of the available information and forego the need for weighting outcomes with their probabilities. On this view, people generate preferences online based on multiple strategies that do not invoke the type of weighting described above (e.g., lexicographic heuristics or sampling from memory). Proponents of this account have even argued that models involving weighting "almost surely fails at bringing improved psychological insight" (Berg & Gigerenzer, 2010, p. 138). The present thesis will argue that both accounts fall short of explaining how people integrate probability- and value information.

The remainder of the thesis is structured as follows. The Background Section first reviews research from other areas of Psychology that have successfully applied information-integration accounts. This part serves two purposes. First, it will illuminate readers not well versed in the psychology of information integration. Second, it provides examples of the type of detailed and systematic research-pursuits regarding information integration that *should be pursued* in judgment and decision-making under risk. The Background Section thereafter reviews normative- and descriptive accounts of judgment and decision-making under risk, making it clear that there exists a number of research gaps concerning the cognitive process of integrating outcome values and their adherent probabilities. The Empirical Section then outlines the aims, methods, and results of three studies conducted with the aim of closing the identified research gaps. The thesis concludes with the General Discussion that discuss the findings in the Empirical Section in a broader perspective, highlighting how the results provide new insights about people's *competence* to integrate probabilities and outcomes. The General Discussion also provide discussions on (*i*) how the results question additional existing "truths" about risk preferences (e.g., that there may be a need to distinguish between risk preferences and risk abilities), (*ii*) practical implications at a societal level, (*iii*) limitations of the thesis, and (*iv*) proposed directions for future research.

Background

The Promised Land: The Advancement of Information-Integration Accounts in Cognitive Psychology

Throughout history, it has been repeatedly claimed that the ability to integrate information is one of the most central hallmarks of human intelligence (Descartes, 1664; Kant, 1781; James, 1890; as depicted by Mudrik et al., 2014). The necessary theoretical- and methodological frameworks to study these cognitive processes were, however, not in place until the time between 1930 and 1970. These frameworks allowed for descriptions and explanations of human behavior beyond the stimuli-response mapping proposed by behaviorism (e.g., Pavlov, 1927, Skinner, 1938) and philosophically focused approaches of early psychologists (e.g., James, 1890).

The remainder of this section will provide a brief odyssey to the development of information-integration research in the general area of judgment and decision-making, focusing on three central research movements that has shaped the field from the 1950's and onward: (**i**) the development of the *lens model* by Egon Brunswik and Kenneth Hammond in the 1950-1960's, (**ii**) the development of *information integration theory* by Norman Anderson in the 1960-1970's[2], and (**iii**) the broad research program on linear additive weighting that has, from the 1970's and onward, expanded the psychological basis of the work by Brunswik, Hammond, and Anderson. This review serves the purposes of providing frameworks that offer direction on how to best study the process of integrating probabilities and outcomes.

The Lens Model

Through the 1930-1950's, psychologists Egon Brunswik and Edward Tolman published a number of articles arguing that there is more to explaining human behavior than the stimuli-response account proposed by, for example, Pavlovian classical conditioning (e.g., Brunswik, 1952; Tolman & Brunswik, 1935, Tolman, 1948). They argued that to understand human behavior, one need to

[2] The informed reader may (correctly) object that so-called additive-weighting models proposed in the 1950's and 1960's (reviewed below under "The Lens Model") at the time also surfaced in the field of judgment and decision making under risk (e.g., Edwards, 1954; 1961), and that Norman Anderson in the 1970's conducted work on judgment- and decision making under risk. For narrative purposes, that work is presented in the section devoted to judgment- and decision making under risk.

study the structure of the environment as well as the structure of the processes within the human mind. To this end, Brunswik developed the lens-model framework (Brunswik, 1952), which holds that the purpose of judgment analysis is to understand the relationships between (*i*) the environment and the attributes (commonly called cues); (*ii*) a person's cognitive system and the cues; and (*iii*) the environment and cognitive system.

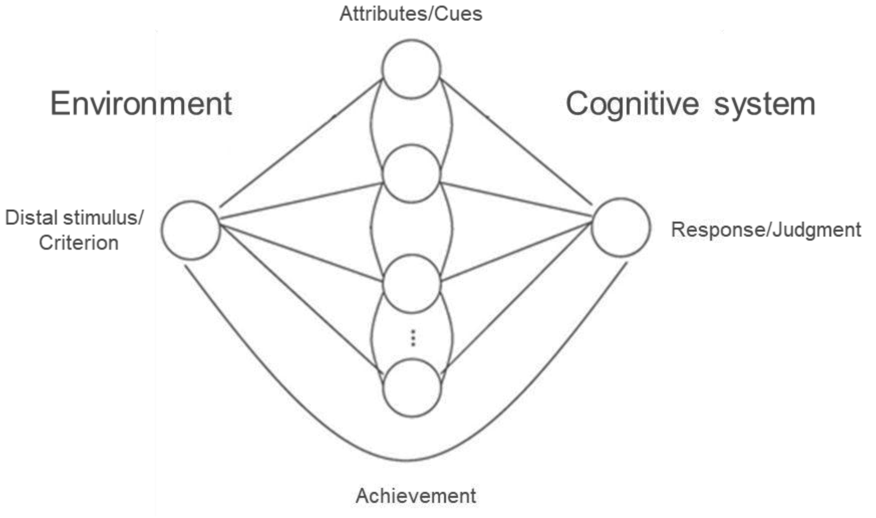

Figure 1. Brunswik's lens model (adapted from Dhami, Hertwig, & Hoffrage, 2004), showing how an agent's cognitive system is in constant interaction with the environment at hand.

Brunswik mainly studied perceptual judgments. After Brunswik's untimely passing in 1955, Brunswik's former graduate student Kenneth Hammond developed the framework to capture judgment processes in a more general sense (Hammond, 1955). In this paradigm, judgment, denoted Y_s, is modeled as a linear and additive function of a set of *n* cues, X_n, $i = 1, \ldots n$. Thus,

$$Y_j = \sum_{i=1}^{n} \beta_{s,i} X_i + \varepsilon_s, \qquad \text{(Eq. 1)}$$

where β_i represent the weights that the judge gives to cues and ε_s is the error term of the regression of Y_s on the X_is. The environmental criterion, Y_e can be modeled the same way. This model can be referred to as an *additive weighting model* (henceforth AW model) because each cue is weighted according to its importance and the weighted cue-values are then summed. The model thus correspond to that of a linear multiple regression without stated interactions. Following Hammond, the number of studies showing that AW models indeed could characterize people's judgments quickly accumulated in the early 1960's (see e.g., Björkman, 1967; Hammond, Hursch, & Todd, 1964; Wiggins & Hoffman, 1968).

To illustrate the psychological content of the model, consider the following example of the processes occurring between the organism and the cues (adapted from Shanteau, 1980). Suppose a person – Alexa - is evaluating whether to rent an apartment or not. Alexa's subjective experience of the overall utility of the apartment will likely depend on a number of cues, such as number of rooms, the condition of the bathroom and kitchen, and the distance to the work place. However, cues will likely differ in their importance. For example, Alexa may gain more utility from living close to work than having a large apartment. Thus, each cue is balanced, or *weighted*, due to its relative importance, and the weighted cue-values are summed.

Information Integration Theory

The introduction and development of the lens model coincided with the Cognitive Revolution of the 1950's and 1960's, a time during which a community of scientists developed the foundations of Cognitive Psychology. The foundation rests on the claim that the mind can depicted as a general purpose, symbol-processing system where symbols are acted on by various processes that manipulate and transforms them into other symbols that ultimately relate to properties of the external world (Eysenck & Keane, 1990). Thus, the human cognitive system can understood much like other information-processing systems (e.g., computers). Of special interest for the present thesis is the information-processing approach underlying *information integration theory* proposed by Norman Anderson (for summaries, see Anderson, 1971; 1996; 2001; 2013). Although the importance of information integration is evident in the lens model (right-side of the model depicted in Figure 1), it was first with information integration theory that it became *the* central process of empirical inquiry.

Information integration theory, focused on the cognitive-system side of the lens model, posit that accounts of the cognitive process of information integration can provide simple and straightforward explanations of many human behaviors. This applies (according to Anderson) to virtually every part of psychology: from, for example, psychophysical sensations such as how sweet something is, to the formation of abstract concepts such as pride. The general idea of the framework, illustrated in Figure 2, holds that one can understand the response (R_i) to external stimuli ($\Phi_{i:}$ corresponding to proximal cues in the lens model) through a series of stages where (i) Φ_i is first transformed through a valuation function to a unique mental representation of the individual, s_i, (ii) the different s are combined through an integration function to an integrated assessment (Ψ_i), which is (iii) ultimately transformed and expressed as an external response (R).

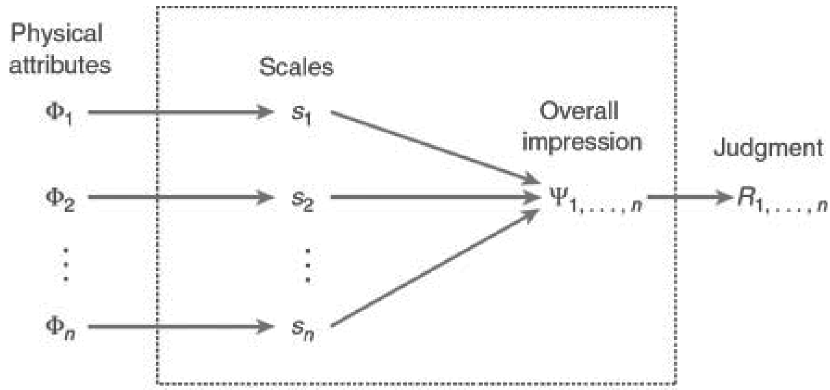

Figure 2. Depiction of the information-processing scheme proposed by information integration theory (e.g., Anderson, 1996; figure adapted from Schwartz, 2009), showing how physical attributes (from the external world) encoded into subjective counterparts and are then integrated and translated to a response in the external world.

Anderson argued that lens-modeler's reliance on comparing model fits (i.e., R^2) of multiple regressions to assert patterns of information processing is insufficient: linear regression models can typically explain most variance even when $\beta_{s,i}$ are not linear and X_is are combined through multiplication. Thus, model fits do not address information *meaning change* (stimuli interact and change one another's meaning in order to make a more consistent, unified whole) and *meaning invariance* (the informers are integrated with no change of meaning).

Anderson also argued that many experimental designs do not allow for simultaneous measurement of the valuation functions (s_i) and integration function, but simultaneous measurement of the two is necessary in order to delineate their independence. Anderson proposed the use of *factorial designs*, where the mode of integration is determined by the combination of the scientist's visual inspection (see Figure 3 for examples) and explicit testing of the interaction term (which in Eq. 1 is typically not spelled out).[3] It is only when the functions have been estimated in a factorial design that one can draw conclusions about meaning change and meaning invariance, conclusions that then can be used in settings outside of the factorial design. Using these methodological improvements, Anderson asserted that the evidence for AW models is widespread across many fields in Psychology, including how people make perceptual judgments, moral judgments, and social-attribution judgments.

The increased effort to understand integration processes did not only lead to an increased focus on the study of integration between cue-values, but on

[3] More specifically it is the interaction term when combining $si_{:n}$ and not $S_{i:n}$ that should be tested, where the subjective interval-level scales of can be derived from participant's responses in a factorial design in terms of the overall marginal means for the independent variables.

the integration process concerning weighting of cue-values as well (for a review, see Shanteau, 1980). The cue-weights in the AW model capture not only the relative importance, but also the transformation of objective stimulus to subjective counterparts (e.g., placing the digit 10 on a mental number line), the order in which stimulus information occurs, perceived credibility of the information, and range-effects induced by the task at hand (i.e., increasing or decreasing the spread of the alternative stimuli along an attribute dimension), and probability information.

This last notion – that weights also reflect probability information – is of particular interest for the present thesis. Shanteau (1980) argues that the initial work by Hammond and colleagues prematurely assumes that probability information is incorporated as cue values and thus integrated additively into the overall judgment. This assumption is, according to Shanteau (1980), premature because probabilities, by themselves, do not provide psychological meaning. For example, .99 probability is something good if it refers to a 99 per cent chance of winning a million dollars, but not so much if it refers to a 99 per cent chance of losing one's job. As Shanteau (1980) argues, additive integration of probability in some instances would need to involve the notion that probabilities are assigned weights on the scale of its adherent outcomes, while the outcomes are weighted on a different scale. For example, the weight of .99 in the scenario of winning a million dollars would have to be derived from the scale of the outcome value in order to contribute in similar strength to the other cues. There is no reason, according to Shanteau (1980), to assume that such a cognitive process is simpler and more efficiently executed than having probabilities directly consumed into the cue values as weights.

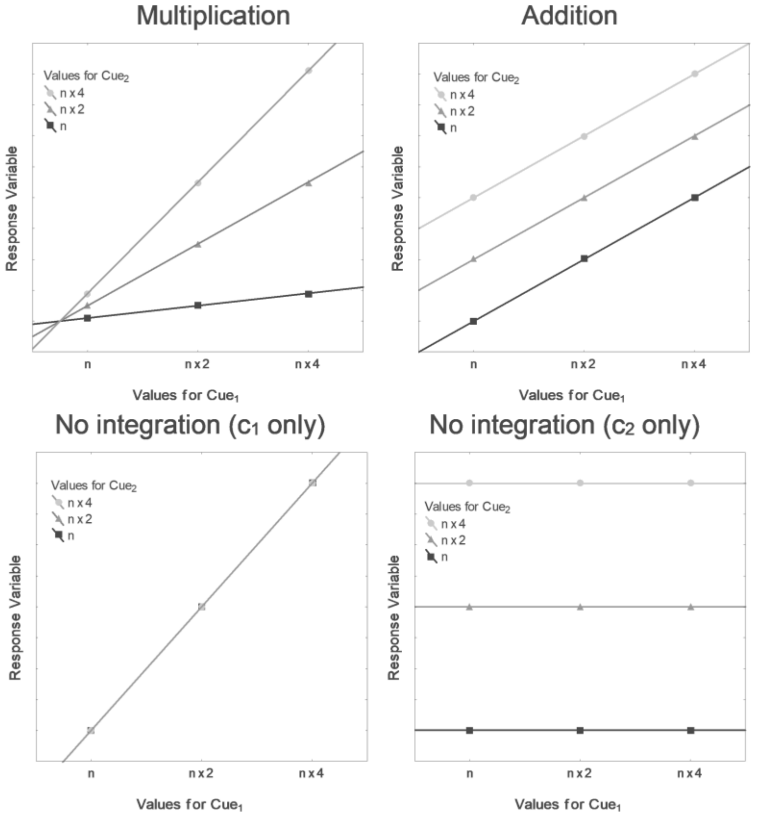

Figure 3. Illustration of the graphical patterns that result from a factorial design with one dependent variable (i.e., the response variable) and two independent variables with three levels each. In this context (and many others), the algebra of multiplication captures also the operations of division and fractioning, and the algebra of addition captures the operations of subtraction and averaging.

Expanding the Psychological Contents of Additive Weighting Models

With the integration function established as a central process, and the AW models as operating characterization of the process, psychologists of the late 20th and early 21st century turned towards understanding the psychological contents and boundary conditions of the model. The research conducted in this tradition typically study behavior in so-called multiple-cue judgment tasks in which participants are presented with values for a number of cues and thereafter asked to make a prediction of a criterion based on their observation of the cues. The tasks may also involve asking participants to attune a number of cues so to meet a specific criterion. A wealth of studies in this paradigm, over a large variety of task manipulations, suggests that people are spontaneously inclined to presume that relations in the world are linear and additive in nature (for reviews, see Brehmer & Brehmer, 1988; Glöckner & Betsch, 2008; Hammond & Stewart, 2001; Juslin, Winman, & Nilsson, 2009, Karelaia & Hogarth, 2008). Linearity entails that a certain change in a cue always leads to a certain change in the criterion. Additivity entails that a change in a cue affects the criterion in the same manner regardless of the level of other cues.

This does not mean, however, that people are unable to make judgments in nonlinear and non-additive tasks. The mastery of such environments does, however, often involve recruiting other resources than intuitive judgments, such as application of abstract rules, exemplar memory (e.g., Juslin, Karlsson, & Olsson, 2008; Karlsson, Juslin, & Olsson, 2007; von Helverson & Rieskamp, 2008; Pachur & Olsson, 2012), or heuristic extraction of important statistical properties in the environment (e.g. von Helverson & Rieskamp, 2008).

Accounts of *why* people are spontaneously inclined to engage in AW have differed. Some have argued that the behavior arises from strong "priors" for linear additive models, perhaps as result of experience (see e.g., Brehmer, 1994). The most popular view, however, holds that it arises because of cognitive limitations (e.g., Glöckner, 2007; Juslin, Nilsson, & Winman, 2009; Juslin, Nilsson, Winman, & Lindskog, 2011, Sundh & Juslin, 2017). These limitations can reside either as limitations in how neural networks evoke specific mental representations (see Glöckner, 2007), or from the sequential and capacity-constrained nature of controlled thought (e.g., Juslin et al. 2009; Juslin et al. 2011, Sundh & Juslin, 2017).

The controlled-thought account posits that people consider each cue in isolation with no memory or attention to the previous cues. For example, a physician may consider the symptoms of a patient sequentially one by one, each time updating the probability of a diagnosis. Because each adjustment does not refer to cues other than the currently attended cue, and the adjustment is the same regardless of the values of previous cues, it implements AW (even though subjects do not average, add, or multiply in any arithmetical sense of that term). This memory-free model is an idealization and probably too extreme: people evidently have some residual working ("cache") memory capacity by which they can moderate their reaction to the currently attended cue – capturing the weighting component of the AW model. Hence, it is possible that people could use this "cache" memory to intuitively moderate their reaction also in view of previously attended cue values and capture interactions between cues without resorting to the application of abstract rules. In other words, people have the capabilities to, spontaneously and effortlessly, engage in multiplicative integration processes when *weighting* cues, but it seem less likely that this capability extends to multiplicative integration of cue values (i.e., not unlike linear regression models without interaction terms).

Normative Theories of Judgment and Decision-Making under Risk

In the 17th century, Blaise Pascal, a philosopher and physicist, presented the first formalized normative theory of human decision-making under risk. Through his correspondence with a fellow mathematician, Pierre Fermat, he developed expected value theory (EVT) (Gigerenzer, Switjink, Porter, & Daston, 1990; Pascal, 1889; 1665/1991). The theory describes the overall value (*V*) of a risky prospect *x* as:

$$V_x = \sum_{i=1}^{n}(p_i) \times (v_i), \qquad \text{(Eq. 2)}$$

where i =1…n is a possible outcome, $\sum_{i=1}^{I} p_i = 1$ (summation is across the outcome space). This expected value is a probability-weighted sum of the possible outcomes, which corresponds to the long-run average return if a lottery plays out infinitely. In other words, it is an AW model, similar to that in Eq. 1 (p. 12). Modern renderings of this idea in the field of decision theory do in fact explicitly refer to the expected-value model as an AW model where probabilities take the role of weights rather than cue values (Keeney & Raiffa, 1976, see also Shanteau, 1980). To illustrate its use, consider a person buying a lottery ticket. The lottery has only one prize, $1000 (otherwise nothing). The probability of winning the $1000 is one in a thousand. Hence, the expected value of the lottery is $1. Following this formula, options and choices with risk can be calculated and compared with each other.

After becoming the dominating model of rationality, EVT was applied to understand many elements of human affairs (e.g., commercial trading and religion: Gigerenzer et al., 1990). In the 18th century, the mathematician Daniel Bernoulli (1738) revised EVT after noting the theory did not take notice of the characteristics of the decision maker (Bernoulli, 1738; as cited in Gigerenzer et al. 1990). For example, winning a specific gamble does not lend the same happiness to a rich person as to a poor person. Bernoulli (1738; as cited in Gigerenzer et al. 1990) also drew examples from the real world of trading, showing that (in his opinion) reasonable businessmen are risk aversive in that they spread their distributions over many ships in case one would sink. Thus, the new theory - expected utility theory (EUT) - took notion of expected utilities instead of expected values, and noted that (reasonable) people are risk averse; ultimately adding a free parameter to Eq. 2. The idea that people should seek, and do seek, to maximize utility continued to stimulate scientific debate in Philosophy, Economics, and Psychology throughout18th- and 19th centuries (e.g., Bentham, 1789; Mill, 1861; Marshall, 1890; for review see Edwards, 1954).

In the 20th century, attention turned to axiomization of the theory, with Von Neumann and Morgenstern (1944/2007) presenting four axioms of

EUT.[4] Even though Von Neumann and Morgenstern's (1944) axioms originated in the theory of games, they argue that these axioms are applicable to social situations and to economics as well. Von Neumann and Morgenstern (1944) theorized that probabilities can be objectively inferred in problems, as coherent with both EVT and EUT. However, as noted by Keynes (1921), Ramsey (1931), and Savage (1954), most probabilities in real life cannot be objectively inferred but are dependent on the opinions of the person who observes the problem. Savage (1954) thus argues that one should adhere subjective probabilities to Von Neumann and Morgenstern's (1944) axioms of rationality. Rather than speaking of probabilities, Savage (1954) discusses decisions in terms of "preferences of acts" and "states of the world": the "probabilities" are inferred from subjective views of the states of the world, as well as personal preferences of acts. Still, Savage's framework maintains the axioms of rationality formed by Von Neumann and Morgenstern (1944), with additive weighting of the components withstanding as a central aspect of human decision-making under risk.

Descriptive Theories of Judgment and Decision-Making under Risk

A Note on the Experimental Paradigms

The bulk of studies on cognitive- and behavioral aspects of judgment and decision-making under risk typically involve measuring people's preferences for monetary gambles. In these gambles, outcomes and probabilities are numerically stated[5]. To illustrate, consider a monetary Prospect X, simplified as [v_1, p_1, v_2]: for example, a gamble that yields either an outcome of $100, with a probability of .50; or an outcome of $0 (i.e., receiving nothing) with a probability of .50. Insights about people's risk preferences and information-processing strategies are typically elicited in one of two ways from these gambles: by *evaluations of gambles* or by *choices between gambles,* reflecting the reality that people engage in both types of processes outside of the laboratory.

---

[4] The four axioms are completeness (the decision maker has well defined preferences and can rank prospects), transitivity (the preferences are consistent over all options), independence (the preference between two prospects is not affected by a possible third prospect common to both), and continuity (for example, if a person prefers A to B and B to C, then there should be a mix of A and C which the person prefers as much as B).

[5] Other behavioral measures have also been proposed, but have not been as dominant (e.g., so-called decisions-from-experience, the balloon-analogue risk task, the Iowa gambling task). It is, however, out of scope of the present thesis to review findings from all these paradigms, but possible relations of the present thesis to these experimental paradigms will be further discussed in the General Discussion.

To exemplify choice under risk in an everyday environment, contemplate the case of purchasing insurance. The person may have a choice between an insurance with a particular feature that will protect against a low probability of a large loss (e.g., assets worth more than a million dollars), and an insurance that does not have that feature (it does not cover assets worth more than a million dollars). To exemplify evaluations under risk, consider the situation where a person is shopping from a retail site, such as eBay, that incorporate an auction-type mechanism under which buyers submit a bid for a product. A sale occurs if this bid exceeds a threshold price set in advance by the seller. Alternatively, consider a situation where a person needs to decide on how much money to allocate to a specific savings fund. This situation warrants some estimate of the funds expectancy in order to avoid allocating too much money.

In the laboratory, individuals performing the typical evaluation task are asked to report either (*i*) how much they at most are willing to pay to procure a gamble (i.e., willingness-to-pay, WTP), or (*ii*) what amount of money would make them indifferent between receiving the amount of money for sure or take the chance of procuring the risky gamble (i.e., the certainty equivalent, CE). A person may for example be willing to pay $40 to participate in the above stated gamble. In the typical choice tasks, individuals are asked to either (*i*) choose between *n* different gambles, or (ii) choose between a gamble and its expected value. For example, a participant may be asked to either play the above stated gamble or choose to receive $50 for sure.

Neo-Bernollian Accounts

In the 1950's and 1960's, the demand for a new descriptive theory mounted after a number of experiments demonstrated that people often violate the axioms of EUT (e.g., Allais, 1952; for reviews see Edwards, 1954; 1961, Schoemaker, 1982). In 1979, Daniel Kahneman and Amos Tversky presented their alternative descriptive theory of human decision-making under risk – PT. Similar to EUT, though, PT also assumes the use of AW (and not only as an *as-if* assumption: Wakker, 2010); people's violations of EUT are explained by factors other than how probability- and value information is integrated.[6]

First, PT posits that people are risk averse in choices involving sure gains, and risk seeking in choices involving sure losses, resulting in the notion that people's value function is convex for losses and concave for gains. Second, the theory posits that people perceive value and utility from a reference point rather than a final asset position. A final pillar in in PT is the notion of decision weights, a concept already implied by other researchers (e.g. Edwards, 1962;

[6] The use of AW was kept intact, with reference to a number of empirical studies that had suggested that people engage in AW (see e.g., Edwards, 1961; Tversky, 1967). The additive-weighting models in these studies were, however, not compared to simpler models (Stewart, 2009), and the cognitive nature of this process were left unexplained.

Fellner, 1961). Decision weights are not beliefs (as the case of subjective probabilities), but refer to the actual weight an outcome receives in a decision. Decision weights may be affected by the probability, but also by other factors, such as personal experiences and perception. Kahneman and Tversky (1979) show that people tend to overweight low probability events and judge highly likely events as certain. Probabilities near 0 and 1 weigh more in a decision than mid-high probabilities. A difference between the probabilities .99 and .98 (or .01 and .02) has a larger effect on the decision than a difference between .55 and .54. Thus, the probability-weighting function in PT assumes a nonlinear, s-shaped, use of probability.

A problem with the original formulation of PT was that it could not account for empirical findings showing that the same outcome is treated differently depending on whether it in the present context is a relatively attractive (henceforth-*focal*) outcome or relatively unattractive (henceforth *non-focal*) outcome (e.g., Birnbaum, 1973; Fishburn, 1978; Parducci, 1965. The rank-dependent utility models that surfaced in the 1980's (Quiggin, 1982; Schmeidler, 1989, for a review, see Lopes, 1995) provided a widely accepted account of such context effects. Rank-dependent utility models generally assume that people's attention is biased towards the focal outcomes, causing systematic overestimation/underestimation of the probabilities associated with focal/non-focal outcomes. In general, these models describe the overall value (V) of prospect x as

$$V(X) = \sum_{i=1}^{n} f(p_i) \cdot f(v_i), \qquad \text{(Eq. 3)}$$

hence similar to the expected-value formula (Eq. 2) but where $f(\cdot)$ symbolizes functions, that transform the inputs p_i and v_i into subjective counterparts. In the seminal revision of PT, Tversky and Kahneman (1992) characterize the functions as,

$$f(p_i) = \frac{{p_i}^a}{({p_i}^a + (1-p_i)^a)^{(\frac{1}{a})}}, \text{ and} \qquad \text{(Eq. 4)}$$

$$f(v_i) = {v_i}^{\frac{1}{b}} \qquad \text{(Eq. 5)}$$

where a and b are free parameters of the probability-weighting function and the value function – capturing people's risk preferences through overweighting of small probabilities, underweighting of large probabilities, and diminishing marginal returns.

Although there exists a vast amount of rank-dependent utility models, each with their own claims about central psychological processes (e.g., regret aversion, disappointment aversion, seller/byer point-of-view: for a review, see Wakker, 2010), no other model has had a stronger influence on the field than CPT. Its functional forms for probability and value have been used to explain behavior in a number of different domains in the social sciences:

for example, labor supply (Camerer, Babcock, Loewenstein, & Thaler (1997), international relations (e.g., Jervis, 1992), and conflict theory (e.g., Levy, 1996). It has also been proposed that these functional forms are evolutionary adaptive (McDermott, Fowler, & Smirnow, 2008; Mallpress et al., 2015). The ability to account for these patterns has been elevated to a benchmark that any model has to meet before it is allowed into the debate on human decision-making under risk (see, e.g., Brandstätter, Gigerenzer, & Hertwig, 2006; Erev, Ert, Plonsky, Cohen, & Cohen, 2017; see also Birnbaum, 1999; 2008).

The neo-Bernoullian account of human behavior has been subject to extensive critique. There is a long tradition in cognitive psychology of showing that preferences are often constructed and generated at the time of judgment, resulting from the interaction of the environment at hand and the cognitive limitations of the decision maker (Gigerenzer & Goldstein, 1996; Lichtenstein & Slovic, 2006; Payne, Bettman, & Johnson, 1993; Simon, 1956; Hsee, Blount, Loewenstein, & Bazerman 1999, Stewart, Reimers, & Harris, 2015). Proponents of this tradition have argued that the neo-Bernoullian approaches place unrealistic assumptions about the cognitive capacities of the human mind. It has, for example been argued that the neo-Bernoullian accounts comprise the implicit assumption that the mind is an error-free multiplication-calculator (e.g., Berg & Gigerenzer, 2010), when people in reality actually *do not* engage in weighting. It has even been suggested that additive weighting models "almost surely fail at bringing improved psychological insight" (Berg & Gigerenzer, 2010, p. 138), and that they are "implausible for most decisions because they would require computational resources […] that are beyond those available to decision makers" (Bossaerts & Murawski, 2017, p. 927).

Heuristics and Sampling

One information-processing account that came to be especially influential in the end of the 20th-century is the *fast-and-frugal heuristic account*. It entails that people do not make use of all available information and ignore trade-offs (integration) are ignored. Instead people are thought to rely on heuristic strategies (e.g., Gigerenzer, Hertwig, & Pachur, 2011; Hertwig et al. 2019; Payne et al. 1993; Thorngate, 1980). This account can be viewed as an extension of Herbert Simon's pioneering work on *bounded rationality:* the idea that human behavior is a function of the decision maker's cognitive constraints (e.g., limited memory) and the structure of the environment. (e.g., Newell & Simon, 1972; Simon, 1956; Simon, 1991). This does not necessarily imply that human behavior will be irrational; the use of heuristics may offer adaptive responses that, in many environments, can provide more "fitness" (e.g., Simon, 1976; Schurz & Hertwig, 2019).

To exemplify the account, consider how the LEX heuristic deals with the choice set below (from Payne, Bettman, & Johnson, 1988). First, the most important attribute for each alternative is determined based on highest probability of occurrence (i.e., ".75 to gain $100" for A, and ".75 to gain $200" for B). Thereafter, the payoffs of these two attributes are compared (i.e., "$100" for A, vs. "$200" for B), and the alternative with the highest payoff is chosen (B in this case). If there are ties, the second most important attribute is examined, and so on until the tie is broken.

Alternative	Probability of outcome	
	.75	.25
A	$100	$400
B	$200	$0

The fast-and-frugal heuristics account does not provide a universal calculus, as is the case of neo-Bernoullian models. Instead, it provides a set of domain-specific mechanisms, similar to the parts of a Swiss army knife, which arise as a product of the structures of the environment at hand and the structural limitations of the human mind (Gigerenzer & Selten, 2001, Hertwig et al., 2019). Many have argued that people rely on lexicographic heuristics when making choices between risky prospects (for reviews, see e.g., Cokely & Kelley, 2009; Gigerenzer et al., 2011; Pachur, Suter, & Hertwig, 2017; Payne et al., 1988; 1993; Simon; 1990; Vlaev, Chater, Stewart, & Brown, 2011).

Another account that stands in contrast to the additive-weighting account is the *sampling-account.* The idea is that people sequentially sample evaluations for each course of action over time until the strength of preference exceeds a threshold and a decision can be made in favor of one alternative. Although the sampling approach was introduced as a general framework for the cognitive process in which one, at least in principle, could implement both neo-Bernouillian and lexicographic models (e.g. Busemeyer & Townsend, 1993), recent incarnations of this tradition have stressed that there is no need for weighting. The decisions from sampling theory (Stewart, 2009; Stewart et al., 2015) assume that people make a series of binary, ordinal comparisons between attributes (i.e., between values and between probabilities) in working memory. Frequency accumulators tally the number of favorable comparisons for each option. A choice is made when the difference in tallies exceeds a threshold. Integration of probabilities and outcomes are thus captured by an additive integration function, but not in the sense that people engage in controlled judgments. As pointed out by Stewart (2009), additive integration of probability and outcome value is observed when the response variable is perceived riskiness or attractiveness rather than CEs (Mellers, Chang, Birnabaum, & Ordonez, 1992; Mellers & Chang, 1994; Joag, Mowen, & Gentry, 1990; Sokolowska & Pohorille, 2000). It is also thought that the sampling is constrained by the context at hand, leading to the typical range-frequency effects first acknowledged by Parducci (1965)[7].

At face value, the fast-and-frugal account and the sampling-account provide compelling arguments that people do not engage in controlled integration strategies captured by AW models. In contrast to the neo-Bernoullian theories, they also provide arguments of how both the structure of the environment and the cognitive limitations of the decision maker in tandem form the strategies that decision maker relies on. Yet, the accounts tend to overlook research from the multiple-cue judgment paradigm reviewed in the previous suction, as well as research providing strong evidence for the notion that people *can* – and at times *do* - engage in weighting of outcomes and probabilities. Although the latter types of studies are rare, there are – as reviewed below - a handful of them that have not only shown the general prevalence of an AW model, but that have also tried to characterize the cognitive process underlying AW.

The Cognitive Neo-Bernoullian to the Rescue?

As was hinted in the section about neo-Bernoullian accounts (e.g., CPT), early decision-making scientists strived to empirically determine (in a choice setting) the weighting of probabilities and outcome values (e.g., Tversky, 1967). The choice paradigm, however, has the drawback that it by nature involves a large space of possible models, making the disentangling of integration modes a more complex task than considered by early decision-making scientists (see, Stewart, 2009; 2011). It is harder to discount the evidence from numerous studies asking people to evaluate the subjective worth of prospects in a more constrained model space, studies in which the analyses comply with the strict measurement tradition proposed by Norman Anderson's information-integration theory. More than a handful of such studies have been conducted, all reaching the conclusion that people's behavior (even children's) is best characterized by an AW model (Anderson & Shanteau, 1970; Birnbaum, Coffey, Mellers, & Weiss, 1992; Lichtenstein & Slovic, 1971; Lopes, 1976; Millroth & Juslin, 2015; Shanteau, 1974; 1975;

[7] Stewart (2009, p. 1058): "The range principle states that the stimulus range is divided into equal-size categories (one for each category label) irrespective of the distribution of stimuli. The frequency principle states that the stimulus range is divided into categories so that each category is used equally frequently and contains an equal number of stimuli. Thus the division of the stimulus range under the frequency principle is completely dependent on the distribution of stimuli. For example, if line lengths are positively skewed, then the smaller lengths will be divided into many categories and the larger lengths into fewer categories. Effectively, under the frequency principle, the category label assigned to a particular stimulus is determined by the stimulus's rank position. The overall category assigned to a given stimulus is a weighted average of the categories given by the range and frequency principles."

Schlottmann & Anderson, 1994; Schlottmann, 2001; Slovic & Lichtenstein, 1968; Viegas, Oliveira, Garriga-Trillo, & Grieco, 2012).

These studies have also provided discussions of the nature of the cognitive processes thought to underlie the AW model, arguing the process can bedressed in terms of *serial fractionation*, or *anchoring-and-adjustment* (Licthenstein & Slovic, 1971; Carlson, 1990; Lopes & Ekberg, 1980; Lopes, 1982; Ganzach, 1996; Schlottmann, 2001). Here, the general idea is that the amount to win serves as a natural anchor and people then adjust this amount downward in order to incorporate the aspect of the probability of the win. Anderson's (1981; 1996) notion of *analogue fractionation* offers additional insights about how mental representations are used to execute these adjustments: each outcome is located on an analogue response dimension according to its value, where probability acts as a fractionation operator producing proportional downward adjustments of the value.

Thus, using anchoring-and-adjustment to calculate fractions does not require the use of the traditional algorithms for multiplication, nor the use of the multiplication table facts. Instead, this approach to multiplication only requires that one knows how to go from a number to a quantity, indicated as a position on a number line (Levin, 1981; Lopes, 1982). For example, to determine what 50% of a unit line is, one needs to be able to specify where 0.5 is on a number line going from 0.0 to 1.0. For this skill to generalize, one needs to be able to apply this skill to lines of different sizes (i.e., where is $35 on a line between $0 and $100, and be able to determine what number is represented by a particular position on a mental number line. Arguably, two basic skills are required (Levin, 1981): (1) translating from number to position on a number line, and (2) translating from position to number. Since these are basic components of knowledge about number, it should be easy to develop exercises for teaching these two skills. Research has indeed shown that even young children can adopt this mindset with only limited instructions and brief prior experience (see e.g., Levin, 1981; Dehaene, 2011; McCrink & Spelke, 2016; Siegler, Thompson, Schneider, 2011).

In agreement with findings in the more general anchoring-and adjustment literature (Tversky & Kahneman, 1975; Epley & Gilovich, 2006), the probability adjustment will typically be rather coarse and insufficient because it is cognitively demanding to make adjustments, forcing people to draw on a range of values rather than precise estimates (Epley & Gilovich, 2001; 2006). The anchoring-and-adjustment account has not yet been mathematically formalized for the context of judgment and decision-making risk. Yet, the general idea seems to have potential. For example, it has been able to explain *preference reversals*, the finding that people's preferences differ over elicitation procedures (choice vs. evaluations). The explanation holds that people resort to heuristic strategies, at least partially, when making choices, but engage in anchoring-and-adjustment when making evaluations (Ganzach, 1996; Busemeyer, Johnson, & Jessup, 2006).

Summary of the Literature Review and Highlighting of the Outstanding Research Gaps

Over the past 70 years, cognitive psychologists have highlighted the importance of studying the integration process for understanding people's judgments and decisions. Research on AW models has been especially fruitful, providing near-complete accounts of integration processes. Arguably, a satisfactory information-integration account should include **(1)** a formalized description of how people combine information from different sources in their environment to arrive at an overall assessment of the situation; (**2)** an explanation of why people adhere to this mode of integration, **(3)** an outline of the boundary conditions for when the formalized description is thought to hold, **(4)** conceptualization of the cognitive processes that facilitate or constrain the ability to engage in the described mode of integration should be reflected in the model, and **(5)** empirical testing of the first four components[8].

The AW account has provided clearly formalized models (e.g., Hammond, 1955; Juslin et al., 2008). It has also provided explanations of why people adhere to this mode of integration (cognitive constraints – e.g., Juslin et al., 2008, and task demands – e.g., Brehmer, 1994), and outlined boundary conditions for when people resort to other strategies such as exemplar memory or analytic application of learned rules (e.g. Juslin et al. 2008; von Helverson & Rieskamp, 2008). Conceptualization of the cognitive processes that facilitate or constrain the ability to engage in AW is directly reflected in the models (working-memory constrained adjustments, e.g., Juslin et al., 2008, or network configurations, e.g., Glöckner, 2006). Finally, empirical testing has been central when pursuing understanding of all components.

The state of affairs is less encouraging in the case of judgment and decision-making under risk. The field has centered on different additive-weighting models (e.g., EUT, CPT) that, although providing formalized descriptions, do not fulfill requirements for the other components. Neo-Bernoullian accounts posit that people engage in AW because it is rational (in the sense that behavior will correspond with normative theory). This rationality assumption has, however, been questioned (e.g., Gigerenzer & Goldstein, 1996, Simon, 1976; Schurz & Hertwig, 2019). It has also been argued that additive-weighting models fail to conceptualize and capture the cognitive processes people engage in (e.g., Simon, 1955; Gigerenzer et al., 2011; Payne et al., 1993; Stewart et al., 2015).

Although information-processing accounts of judgment and decision-making under risk are dearly needed, an abandonment of weighting models may be premature; research on evaluations shows that people *do* engage in weighting of outcome values and probabilities, perhaps through an intuitive

[8] The components advocated here are derived from agreed commons in influential discussions on theory development in the cognitive sciences (Anderson, 1990; Hammond & Stewart, 2001; Newell, 1994; Marr, 1982; Simon, 1996) and then adapted to the present context.

anchoring-and-adjustment process rather than by explicitly number crunching the two. However, these pioneering studied did not vary the degrees of cognitive demands: all information was simultaneously presented in within-subjects designs - conditions optimally set to allow for an AW model (see e.g., Shanteau, 1970; Kahneman & Fredericks, 2002). Moreover, the cognitive processes are not richly conceptualized, leaving many questions unanswered: What are the characteristics of the cognitive processes that facilitate the ability to engage in AW of outcomes and probabilities? Are the processes foremost intuitive or analytical? Is the ability to engage in weighting primarily constrained by information processing constraints or by lack of knowledge (or acceptance of) the normative rule involved? How can the process of weighting be formalized?

Aims and Structure of the Empirical Section

Given the conflicting claims in the literature on people's inclination and ability to weight the outcomes with their probabilities, the first aim is to explore the boundary conditions, or robustness, of this process of weighting the outcomes by their probabilities. A system, organism or design is robust if it is capable of coping well with variations in its operating environment without alteration or loss of functionality (Cavazzuti, 2013). Translated to the present context, weighting can be defined as robust as long as its multiplicative nature is not alternated, either by an algebraic shift (i.e., that people shift to additive integration) or that it collapses completely (i.e., people stop integrating, and instead rely on valuation of only one attribute). In other words, do people have a spontaneous inclination and robust ability to engage in weighting? The Empirical Section pursues the question in the context of evaluations of risky prospects because of the advantages this context offers regarding model discrimination (as compared to a choice context: Anderson, 1996).

Of course, choosing the appropriate manipulations for testing the robustness is not a trivial task: failing to document any effects could be attributed to the notion that the manipulations were not strong enough. A number of considerations are thus necessary. First, what is a reasonably strong manipulation? Second, what are the most theoretically grounded manipulations? This problem also refers to the issue of ecological validity: what are the typical manipulations of the judge's environment that not only usually lead to large effects, but also are typical of manipulations that occur in the judge's or decision maker's environment?

The present thesis focuses on examining robustness under three conditions. Study I tests the possible effects of making only an isolated evaluation instead of making evaluations in the context of many other. Similar manipulations of design type, Within-Subject Design (WSD) or Between-Subject Design (BSD), have shown to affect judgment and decision-making, with more normatively aligned behavior reported in WSDs (e.g., Birnbaum, 1999; Kahneman & Frederick, 2002). Study II tests the possible effects of sequentially presenting probabilities and outcomes. Sequential presentation has shown to affect the integration of variables in other instances (e.g., in probabilistic reasoning tasks: Shanteau, 1970). Study III tests the possible effect of if increased complexity, in the sense that prospects have two independent possible outcomes instead of one. Increased complexity has shown to affect the normative alignment of judgments and decisions (e.g., Payne, 1976; Camerer et al., 2011).

The second aim is to acquire a better understanding of the actual cognitive process of weighting. To this end, a formalized model – the anchoring- and adjustment model (AAM) - was developed (Study II) and compared with a benchmark model, CPT (Studies II and III). Study III applies cognitive modeling of the processes to delineate if the processes are of an intuitive or analytical nature. Study III also tests if the prevalence of AW among participants is best predicted by knowledge factors (financial literacy - FL, numeracy, and knowledge about expected values – EV-C) or by information-processing capacity (numerical short-term memory, STM). Apart from using these cognitive-modeling efforts to understand the cognitive processes, the thesis will also use a, in the present context, novel theoretical framework to interpret findings about the processes. This framework, first proposed by Chomsky (1965) in the context of language, makes a distinction between individual's *normative competence* and their *normative performance*: decision-making behavior produced at any moment – the performance – may be affected by various cognitive and situational limitations, leading to deviations from the normative competence (for the origins of this idea, see Chomsky, 1965). This idea has not yet been utilized in judgment and decision-making research, but offers an alternative to interpreting findings about cognitive processes in terms of dual-systems (i.e., the intuitive System 1 vs. the analytical System 2: Evans, 2008; Kahneman & Fredericks, 2002 – for a review on the problems of dual-system interpretations, see Melnikoff & Bargh, 2018). The discussion of Study I elaborates further on the idea, and the General Discussion use the framework to dissect the results from all three studies.

Empirical Work

Methods

Experimental Designs

In all three studies, participant evaluated risky prospects resulting from factorial stimulus designs created by crossing *n* levels of probability with *n* levels of value. The dependent measure was the reported CE: what amount of money that would make the participant indifferent between receiving that monetary amount for sure and playing the lottery. Before starting, participants were presented with an example of a lottery and informed about the meaning of the CE, and how to report it.

CEs were assessed by a two-step procedure that "homed in" on the CE with a resolution of $1, between $0 and the maximum possible outcome (see Tversky & Kahneman, 1992, for a similar procedure). First, they chose an approximate CE among six linearly spaced values ($0 plus the maximum outcome divided into fifths). Participants then specified their CE further by choosing among more fine-grained CEs. For example, if participants first chose $5, they would in the second step choose among five CEs from $3 to $7. This elicitation method produces less measurement error as compared to other popularly used measures – for example, willingness-to-pay - (Hey, Morone, & Schmidt, 2009).

Experiment 3 in Study II and Experiment 2 in Study III use monetary incentives: participants were made aware that one of the lotteries would be played out for real, for one of the participants (for evidence that this incentive-scheme works, see Abdellaoui, Bleichrodt, & L' Haridon, 2008).

Participants

Participants in all experiments, except Experiment 1 in Study III, were recruited through the crowdsourcing-service CrowdFlower.com (similar to Amazon's Mechanical-Turk service) and compensated with an average wage for crowdsourcing tasks. CrowdFlower.com offers selection of workers from three settings: all available workers (1), a group of more experienced and accurate contributors (2), and the smallest group of the most experienced and accurate contributors (3). All experiments used participants from Groups 2 and 3. Data was also pre-screened before analyses, checking if the CEs were positively correlated with the expected value of the prospects (implied by virtually all descriptive theories of decision-making under risk), resulting in an exclusion rate of approximately 10 per cent. The participants in Experiment 1 in Study III were recruited through public advertisement at Uppsala University and other public spaces within the city of Uppsala, and compensation was awarded in the form of a cinema voucher or course credit.

Statistical Inferences

Statistical inferences throughout the three studies were conducted using Bayesian hypothesis testing in JASP (JASP Team, 2019) and R (Morey, Rouder & Jamil, 2015). This allows for direct comparison of the Bayes factor (BF) for different models; that is, the relative likelihood of the observed data, given different models (e.g., a $BF_{10} = 2$ implies that the data are twice as likely if Model 1 is true, as compared to if Model 0 is true).

Cognitive Modeling

The present thesis apply multiple cognitive-modeling tools to answer questions about the integration process. All three studies use Bayesian factorial regressions determine the *mode* of integration. Study II and Study III use two models to explore the nature of the weighting process: whether it is best characterized as a cognitively effortless endeavor that should be unaffected by cognitive demands (CPT), or as a cognitively effortful endeavor that is easily affected by cognitive demands (the AAM). Study III use the PNP-framework to delineate whether participants classified as AW are engaging in (Brunswikian) intuitive or analytical thought-processes.

Determining the Mode of Integration

Visual inspection of the data typically goes a long way in determining the mode of integration. If the data produce the "fan-patterns" typical of multiplicative integration, it most likely means that the underlying model is of multiplicative nature; and if the data produce parallel patterns, it most likely means that the underlying model is of additive nature. There are, however, instances where visual inspections are not sufficient because the variable functions are strongly nonlinear: multiplicative models with nonlinear variable functions can close to mimic the visual pattern of responses of simpler linear models (Anderson, 1996, p. 45; Stewart, 2011). Testing the statistical interaction between the variables is thus necessary.

Bayesian factorial-regression analyses (Rouder & Morey, 2012) determine if a weighting model best described participants' responses.[9] Of specific concern for the present purposes is that nonlinearity of the underlying variable

[9] Study I, in its published form, involved the use of functional measurement (Anderson, 1996) and frequentist hypothesis testing (determining the probability of obtaining test results at least

functions increases simply as the level of response error increases (e.g., Blavatskyy, 2007). Model recoveries showed that Bayesian factorial regressions could identify an AW model even under severe cases of nonlinearity in the variable functions, thus making the error classification rate (classifying a participant as using AW when the "true" model does not involve weighting) a minor problem in the present studies.[10]

Characterizing the Process of Weighting

Two different cognitive models - mathematical formulizations of the assumed psychological processes – occurs throughout the empirical studies. The CPT-model occur in all three studies, and is defined in Eq. 3-5 (pp. 23-24). The AAM is first presented in Study II, and is applied to Study II and Study III. The AAM defines a simple prospect as:

$$V(X) = \sum_{i=1}^{I}(1 - a \cdot (1 - p_i)) \cdot v_i^b, \qquad \text{(Eq. 6)}$$

where a is a free parameter for the insufficient adjustment from P=1 and b is the standard concave value function for gains, and $a = 1$ and $b = 1$ implies that the $CE = EV$. CPT and the AAM thus posit very different cognitive mechanisms: CPT explain deviations from normative behavior in terms of people's *psychophysical functions* for probabilities and outcome values – The AAM assume that deviations stem from linear use of probability, but with a process of insufficient adjustment.

Weighting as Facilitated by Intuitive- or Analytical Thought Processes

The Precise/Non-Precise (PNP) modeling framework (Sundh, Millroth, Collsiöö, & Juslin, 2020) occurs in Study III and in the General Discussion. The psychological contents of CPT and the AAM has been covered in the Background, but the PNP framework warrants some further elaboration.

Preceding the study of intuitive- and analytic cognitive processes as properties of different systems, Egon Brunswik (1956) proposed an *operational definition* of intuitive- and analytical cognitive processes. This distinction define the processes directly from observable properties of the data, in terms of the empirical shape of the error distributions. Brunswik's original idea

---

as extreme as the results actually observed during the test, assuming that the null hypothesis – that there is no difference between tested conditions is correct). In order to achieve methodological comparability, the results were re-analyzed for the present purposes.

[10] The specifics of model-recovery simulations are described in the manuscript for Study III, and the methods there were translated to the context of the Study I and II. Previous research (Anderson, 1996) have provided evidence that erroneous classification of integration modes foremost occur in the direction AW→simpler integration models, and not the other way around. The factorial graphs of the descriptive data provided through the results-sections are clearly reflective of the models derived in from the factorial regressions, supporting the notion that Anderson's (1996) conclusion holds also in the present context.

posits that intuitive processes typically produce a ubiquitous but small random error captured by a Gaussian (i.e., normal) distribution, whereas analytic processes produce leptokurtic (spiked) error distributions with more occasional but potentially large errors (Brunswik, 1956). Brunswik considered intuitive processes reliant on probabilistic and interchangeable cues and analytic processes reliant on measurement and calculation.

The Precise-Non-Precise (PNP) model (Sundh, Millroth, Collsiöö, & Juslin, 2020) implements this idea in cognitive modeling, assuming analytic processes to involve deterministic (noise-free) application of integration rules to exact (noise-free) symbolic representations of cues, as naturally applies to deduction, calculus, or computer algorithms, as well as to unaided attempts at mental "number crunching". Analytic processes will hence most often produce estimates that are *exactly correct*, but occasionally they will be marred by errors (that are potentially large, for example if multiplying by 100 instead of 10). Intuitive processes depart from analytic procedures by always being perturbed by a random noise, either in the encoding of the cues (early noise), inconsistent and unreliable application of non-symbolic integration rules (late noise), or both (see Li et al., 2017 on these origins of noise). Sundh et al. (2020) showed that standard regression models fail to account for behavior generated by the two proposed processes (see Figure 4 below for an illustrative example of what data-patterns may look like in the PNP-framework, and Figure 2 in the manuscript for Study III for how regular regression models fail to account for analytic processes). The PNP model, on the other hand, could distinguish between the processes over a variety of tasks (perceptual area estimation, algebra tasks, estimates of student's motivation, and typical multiple-cue inference tasks), one of which was evaluations of risky prospects.

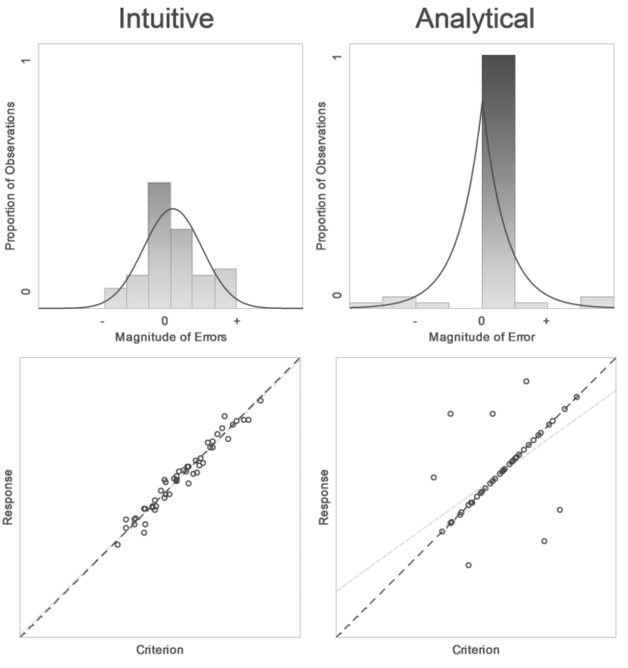

Figure 4. Illustrative examples of how responses are distributed for Brunswikian intuitive- and analytical cognitive processes in a task where participants are to estimate some known criterion on basis of a number of cues in relation to a criterion. Upper panels show distributions of responses as a function of the magnitude of the errors (deviations from model or criterion). Lower panels illustrate, with the same data, how standard regression fail to account for the notion that responses are from different statistical distributions. The thinner line represents the predictions by the regression model and the identity line (*x* = *y)* represents error-free computation of the expected values. Note that data points overlap.

For each response, an error occurs a probability λ. If *B* is a Bernoulli random variable with probability λ, each estimate y given some function $g(\mathbf{x}|\boldsymbol{\theta})$ is defined by

$$y|(B=b)=\begin{cases} g(\mathbf{x}|\boldsymbol{\theta})+N(0,\sigma^2), b=1 \\ g(\mathbf{x}|\boldsymbol{\theta})+N(0,\tau^2),\ b=0 \end{cases} \qquad \text{(Eq. 7)}$$

For a fully intuitive process, $\lambda = 1$ and ubiquitous Gaussian noise perturb the output of the model (corresponding to the default assumption in most statistical modeling). For an analytic process λ is presumably a small but non-zero number (no one is perfect, after all). For technical reasons, outlined in Sundh et al. (2020), it is prudent to introduce a tolerance τ. The parameter τ specifies the precision of responses when errors are not presumed to have occurred. Therefore, τ should not be understood as an estimate of error, despite delineating a Gaussian distribution, but as a small tolerance for what counts as correct. For further details, see manuscript of Study III.

Study I

Background

Study I was motivated by the fact that the design type (within-subject design, WSD or between-subjects design, BSD) has shown to affect the extent to which people align their judgments and decisions with normative principles of decision theory. When the decision problems are laid out transparently with access to relevant comparisons and anchors – as is typically the case in WSDs - people often recognize and endorse certain normative principles, like transitivity, dominance, and the extension rule, while they in BSDs more often violate normative principles (Kahneman & Frederick, 2002).

The purpose was to compare the results from a WSD and a BSD (in the extreme case of a single-stimuli design, SSD – see Method) to determine if three of the most central normative intuitions of EUT are dependent on design type. First, is the evaluation of a risky prospect linear in its probability? Second, is it concave in its objective outcome (e.g., in $)? Third, and most directly related to the overall theme of this thesis, is probability- and value information integrated through weighting?

Method

Thirty-six prospects were created by crossing six levels of probability (.01; .20, .40, .60, .80; .99) with six levels of value ($15; 30; 45; 60; 75; 90). The expected values of the prospects ranged from 0.15 to 89.1 (M = 26.25, SD = 23.81). The 20 participants in the WSD sample evaluated all 36 prospects (presentation order randomized across participants). The 720 participants in the SSD sample evaluated only one of the 36 prospects, making the total number of data points equivalent in the two samples (720 = 36 x 20).

Results

There was no difference in the average reported CEs for the two samples (M = 19.0, 95% CI = 14.8 – 23.2, for the SSD sample, versus M = 17.10, 95% CI = 12.8 – 21.4 for the WSD sample). Figure 5 presents the interaction plots between probability and outcome, revealing the bilinear interaction pattern that is typical of multiplicative integration for both the WSD sample and the SSD sample. In both samples, this finding was supported by a test of the interaction term in factorial regression (WSD: BF >1,000; SSD: BF = 228). The statistical interaction was confirmed for 18 of 20 participants in the WSD sample.

Fitting CPT to the aggregated responses from the two samples showed that WSD sample had a "more linear" probability-weighting function (.798, 95% confidence intervals = .575 – 1.02) than the SSD sample (.401, 95%

confidence intervals =.320 – .483), but there was no difference for the value function (WSD = .904, 95% confidence intervals = .825 – .983; SSD = 1.00, 95% confidence intervals =.947 – 1.06). Figure 6 illustrates the shape of the variable functions implied by CPT. As will be discussed in the General Discussion, this does not necessarily entail that participants in SSD sample invoked nonlinear use of probability. A more probable explanation holds that they made adjustments that were less sufficient (but still linear). The difference between the modeling results for the WSD and the SSD were not a result of fitting individuals versus average group data or that random responses differed between the samples (see published article for details).

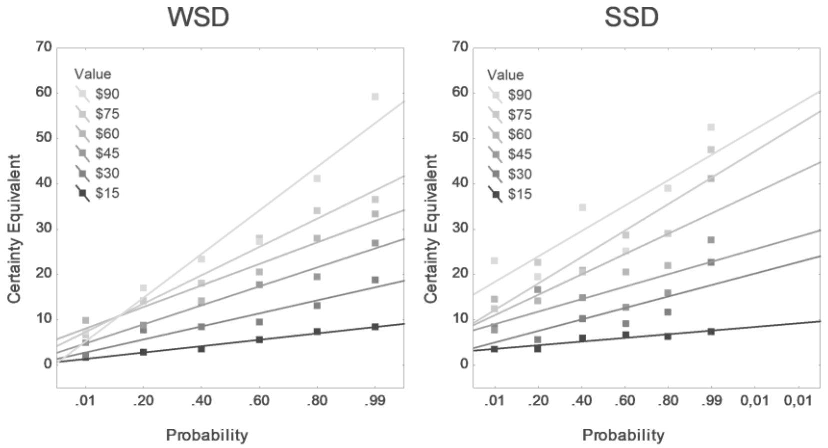

Figure 5. The mean reported CE (y-axis) for each value (marked by color) and probability level (x-axis) along with linear fits (lines) for the within-subject design (WSD, left side) sample and the single-stimuli design (SSD, right) sample.

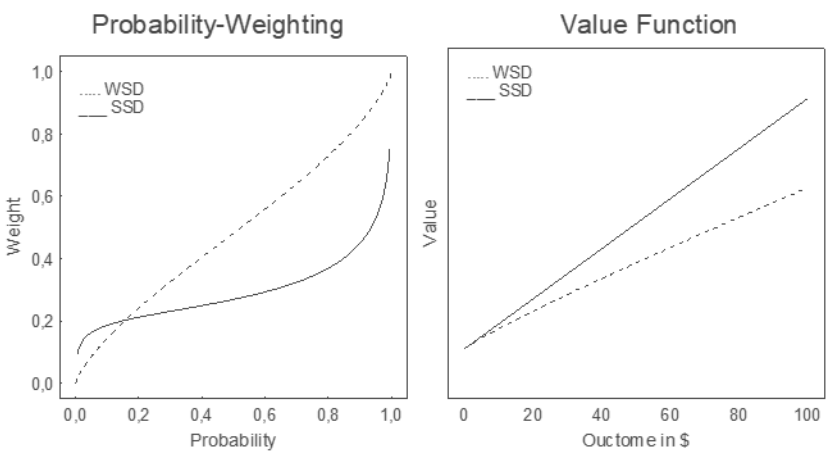

Figure 6. Illustrates the shapes of the probability-weighting function (left panel) and value function (right panel) implied by the best-fitting parameters of cumulative prospect theory (CPT) for both samples (marked line shows results for the within-subject design, WSD, sample, and the dotted line shows the results for the single-stimuli design, SSD, sample.

Discussion

The results suggest that most people are spontaneously inclined to integrate outcomes and probabilities through weighting, and that this inclination is robust to the increased cognitive demands that surface when other evaluations are not present as comparative anchors. Although people across both samples seemed to grasp this normative intuition, there were differences in how they accounted for the symbolic probability; responses from the WSD sample were more aligned with normative theory than the responses from SSD sample.

The standard way to account for this duality in normative abilities is by distinguishing between fast intuitive System 1 processes and slow analytical System 2 processes, where the latter processes, which carry the intuitions related to normative models, become seemingly more engaged in a WSD (Kahneman & Fredericks, 2002). From this view, the more normative behavior in the WSD could be interpreted as a methodological artefact in the sense that the design introduces demand characteristics. This idea holds that participants may believe that the experimenter is asking for a specific characteristic (in this case, weighting, and evaluations that are linear in probability) and they adapt these characteristics even though they would not do so when unobserved.

As a complement – or alternative – to the dual-systems view (Kahneman & Fredericks, 2002), one can argue that people carry many beliefs that potentially informs the evaluation of a decision option. Together, these beliefs define people's *normative competence*. Some of these beliefs may refer to properties of individual objects or situations (e.g., a set of dead oily birds is not appealing), while other beliefs may refer to relations between objects or situations (e.g., a set of 200,000 birds is a hundred times larger than a set of 2,000 birds). Although these beliefs may be similar in other respects, they may differ in their requirement for contextual information for implementation in a specific situation. Therefore, it may depend on the situation whether this normative competence is materialized in *normative performance* or not. To illustrate the idea in the present context, consider the possibility that it is only when people are asked to explore the full probability scale, from 0 to 1, that they evoke the belief that the use of probability should be linear. This idea of normative competence- and performance is akin to the influential distinction between language competence and language performance introduced by Noam Chomsky (1965). The General Discussion provides additional elaboration of the idea and how this theoretical framework, compared to a dual-systems framework, can fruitfully account for the results in this thesis.

Study 2

Background

Simon (1955) originally suggested that memory limits constitute one of the foremost bounds of human rationality. Prominent accounts of decision-making behavior indeed contain mechanisms related to memory sampling, such as the availability heuristic (Tversky & Kahneman, 1973, 1974), the decision-by-sampling framework (Stewart, Chater, & Brown, 2006), and query theory (Johnson, Häubl, & Keinan, 2007). Manipulating the temporality of presentation mode, presenting information simultaneously or sequentially, is one of the more common methods to impose increased demands on memory. Research on temporal presentation mode suggest that sequential presentation may be detrimental to performance, affecting behavior through various routes: (for a review, see Hogarth & Einhorn, 1992): **i**) it can lead to primacy and/or recency effects, **ii**) it can affect the integration function, and **iii**) it can lead to less in-depth cognitive processing.

Lopes and Ekberg (1980) offer a telling example in the context of judgment and decision-making under risk: participants' response time was higher when probabilities were presented first, as compared to when outcome-values were presented first. Lopes and Ekberg (1980) argue that this suggest that sequential presentation with probability first is more difficult than sequential presentation with outcome first. They also interpret the findings as evidence that people engage in either an anchoring-and-adjustment strategy or that sequential presentation lead participants to an additive combination of probability and outcome value.

Relating to this finding, Shanteau (1980) also show, in a Bayesian belief-revision task, that sequential presentation can lead to a change in the integration mode. For simultaneously presented stimuli, the study participants integrated the probabilities multiplicatively - but for sequentially presented stimuli, participants integrated the information additively. Shanteau (1970) offers no explanation for this phenomenon, but other lines of research suggest that people tend to retreat to the default of linear additive integration of the information, when the task is too complex or people fail to recall content-specific rules (see, e.g., Nilsson, Winman, Juslin, & Hansson, 2009; Juslin et al., 2009; Juslin, Lindskog, & Mayerhofer, 2015).

Study II involves four experiments that test the effects of presenting simple lotteries with simultaneous presentation of outcomes and probabilities, as compared to presenting them sequentially. In Experiments 1, 2, and 3, the participants memorize outcomes/probabilities in a separate phase of the experiment and the focus is more on long-term memory than on STM. Experiment 4 tests if the effects in the first experiments hold also when emphasizing the use of STM (by investigating immediately sequential presentation of the two components).

Method

Thirty-six prospects were created by crossing six levels of probability (.01; .20, .40, .60, .80; .99) with six levels of value ($15; 30; 45; 60; 75; 90). The expected values of the prospects ranged from 0.15 to 89.1 (*M* = 26.25, *SD* = 23.81). Prospects were thus the same as in Study I.

Participants were randomly allocated to one of three conditions: the Simultaneous Condition, the Outcome-First Condition, and the Probability-First Condition. Participants in the Simultaneous Condition evaluated the 36 prospects with all information about probabilities and outcome values available to them on the screen. Participants in the sequential conditions were informed that they would make evaluations about six different lotteries, but before evaluating these, they would in a first phase only be able to observe the probabilities/outcomes of the lotteries. The task in this phase was to memorize them. In Experiment 1 tested participants' memorization by having participants choosing among alternatives in a list, thus relying in recognition memory. In Experiment 2 and 3 participants had to freely recall the outcome values/probabilities.

After the participants had succeeded in reproducing the correct outcomes/probabilities, they were informed that they would be given the adherent outcomes/probabilities of the lotteries and then evaluate the lotteries. When presented, the lotteries in the Simultaneous Condition were numbered 1-36 and the lotteries in the sequential conditions were numbered 1-6. In Experiment 4, participants did not engage in a memorization task before the evaluation phase. They were instead directly presented with the probability/outcome of the prospect, and when they had observed this information, they pressed "Next". Participants were then shown the adherent probability/outcome value and asked to report their CEs. Experiment 2 and 3involved a follow-up test following the evaluation phase, and here were participants once again instructed to recall the outcome values/probabilities for the six lotteries.

In Experiments 1 and 2, the Probability-First Condition were told to memorize both the probability of Outcome A and Outcome B, even though Outcome A was the only outcome that produced a positive outcome. In the Outcome-First Condition, the participants also had to memorize the values for both outcomes, but the value for Outcome B was always $0, effectively allowing the participants to focus on memorizing only one outcome. In Experiment 3, participants in the Probability-First Condition only had to remember the probability for the outcome with the positive outcome, Outcome A. Experiment 3 also involved real monetary incentives (participants were made aware that one subject would be randomly selected to play out one of the gain questions with the actual payment: see Abdellaoui et al., 2008). Experiment 3 also tested the possible effect of providing participants with a more embodied representation (i.e., pie charts and pictures of dollar bills), since this has shown to reduce other decision-making biases (Carlin, 1990).

Results

Experiment 1

The results showed that the CEs were between 47 % and 80 % higher in the sequential conditions (Outcome-First: $M = 27.4$, $SD = 8.30$; Probability-First: $M = 22.3$, $SD = 5.98$) than in the Simultaneous Condition ($M = 15.2$, $SD = 11.4$). There were interactions with fan patterns typical of weighting in the Simultaneous and the Outcome-First Conditions, but almost flat functions in the Probability-First Condition, suggesting that the probabilities were ignored (see Figure 7). However, Experiment 1 did not disclose to what extent these effects were mediated by the neglect of known and retrievable information or an inability to retrieve information in the evaluation phase.

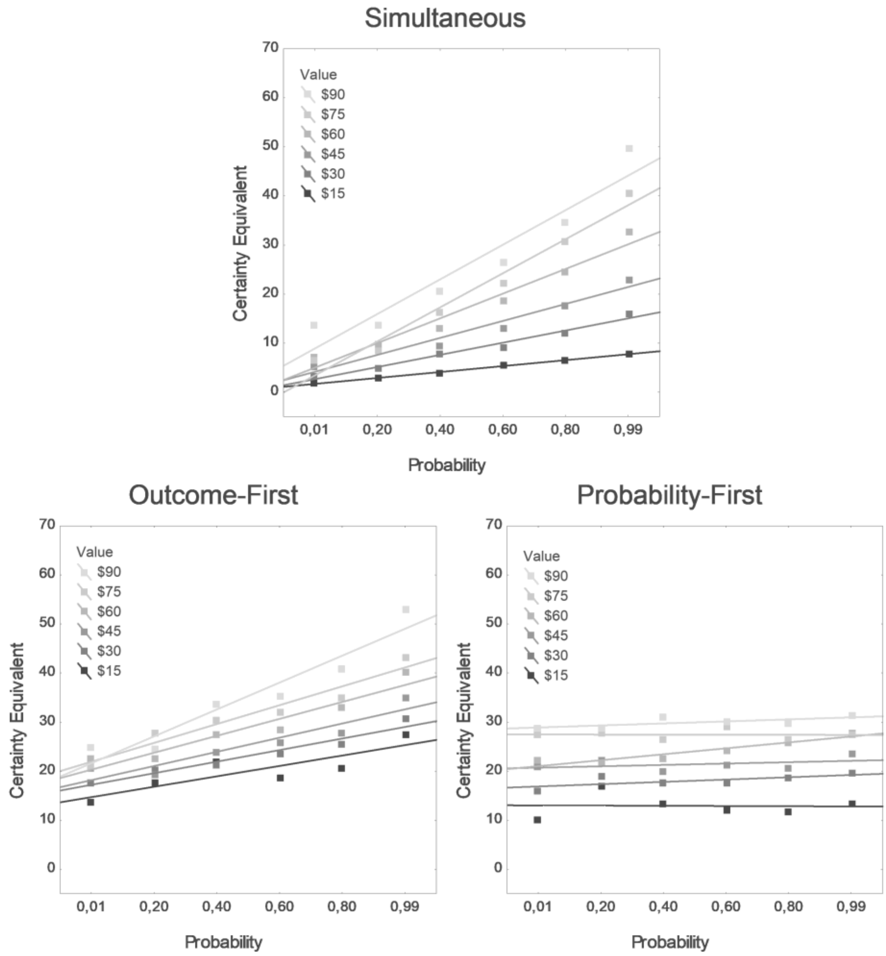

Figure 7. Average CEs (y-axis) as a function of probability (x-axis) and the outcome value (lines) for each condition in Experiment 1.

Experiment 2

Similar to Experiment 1, the CEs were higher in the sequential conditions than in the simultaneous condition (*M* = 19.5, *SD* = 13.8 in the Simultaneous Condition vs. *M* = 33.80, *SD* = 12.1 in the Outcome-First Condition, and *M* = 37.7, *SD* = 10.9 in the Probability-First Condition). Of main interest is the extent to which these patterns persist also when considering only the participants with perfect posttest recall of the outcomes and the probabilities.

The percentage (%) of participants successfully remembering the correct outcome values/probabilities after the evaluation phase was 42.9 for the Probability-First Condition and 44.4 for the participants in the Outcome-First Condition. As evident in Figure 8, participants in the Probability-First Condition generated, on average, higher CEs (*M* = 38.9, SD = 14.9) than participants in in the Simultaneous Condition (*M* = 19.5. *SD* = 13.8). The Outcome First

Condition did not differ on average from the Simultaneous Condition (M = 19.6, SD = 13.6). Factorial-regression analysis supported a weighting model for 78.8% of participants in the Simultaneous Condition, 58.3% for the Outcome-First Condition (with perfect memory), and 50% for the Probability-First Condition (with perfect memory).

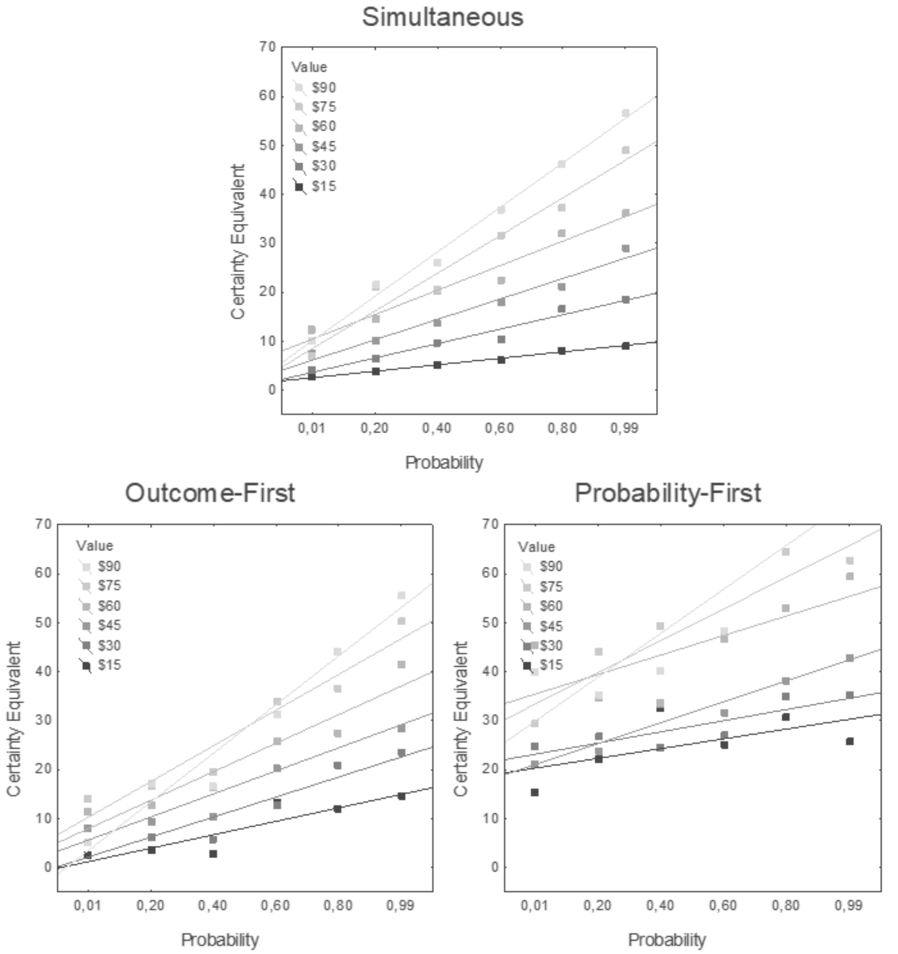

Figure 8. Average CEs (y-axis) as a function of probability (x-axis) and outcome value (lines) for each condition in Experiment 2 (all participants).

Experiment 3

The two embodiment conditions were collapsed with the regular conditions because there were no evidence of an effect of embodiment.

Seventy per cent of participants in the Probability-First Condition and 61.5% of the participants in the Outcome-First Condition successfully remembered the correct outcome values/probabilities after the evaluation phase. Participants in the sequential conditions reported, on average, approximately 33% higher CEs than the participants in the Simultaneous Condition (M = 21.1, *SD* = 13.7 in the Simultaneous Condition vs. M = 28.3, *SD* = 11.2 in the Outcome-First Condition, and M = 27.7, *SD* = 10.9 in the Probability-First Conditions. Figure 9 graphs the behavioral data for each condition, showing the typical fan-patterns associated with weighting of outcomes and probabilities. Factorial-regression analysis supported a weighting model for 85% of the participants in the Simultaneous Condition, for 66.7% of the in the participants Outcome-First Condition (with perfect memory), and for 64.4% of the participants in the Probability-First Condition (with perfect memory).

A more detailed analysis on these data was pursued to further explore the consequences of the higher CEs observed in the sequential conditions. As illustrated in Figure 10, the participants produce CEs that amount to a "between-subjects" violation of stochastic dominance, assessing higher CEs for lotteries with a lower probability of the same outcome. In Figure 10, participants indicate that they would pay more ($17) for a .01 probability of winning $45 in the Probability First Condition than for a .40 probability of winning $45 in the Simultaneous Condition ($15), despite perfect posttest recall of the probabilities.

The systematicity of these violations of stochastic dominance were explored by comparing the proportions of items in the sequential conditions that had higher CEs than in the Simultaneous Condition, despite having a lower probability of the same outcome. As a benchmark, it was first assessed if the Simultaneous Condition had any such items. They did not, in accordance with normative theory. There were 15 possible instances where the sequential conditions could report higher CEs despite a lower chance of winning the same outcome (5 for the .99 level, 4 for the .80 levels, and so on). The proportion of the 15 items that violated stochastic dominance in this sense was .347 with a 95 percent CI between .152 and .587 for the Probability-First Condition and .347 with a 95 percent CI between .152 and .587 for the Outcome First Condition. The example in Figure 10 is therefore not an isolated finding but it occurs in about one third of all relevant items.

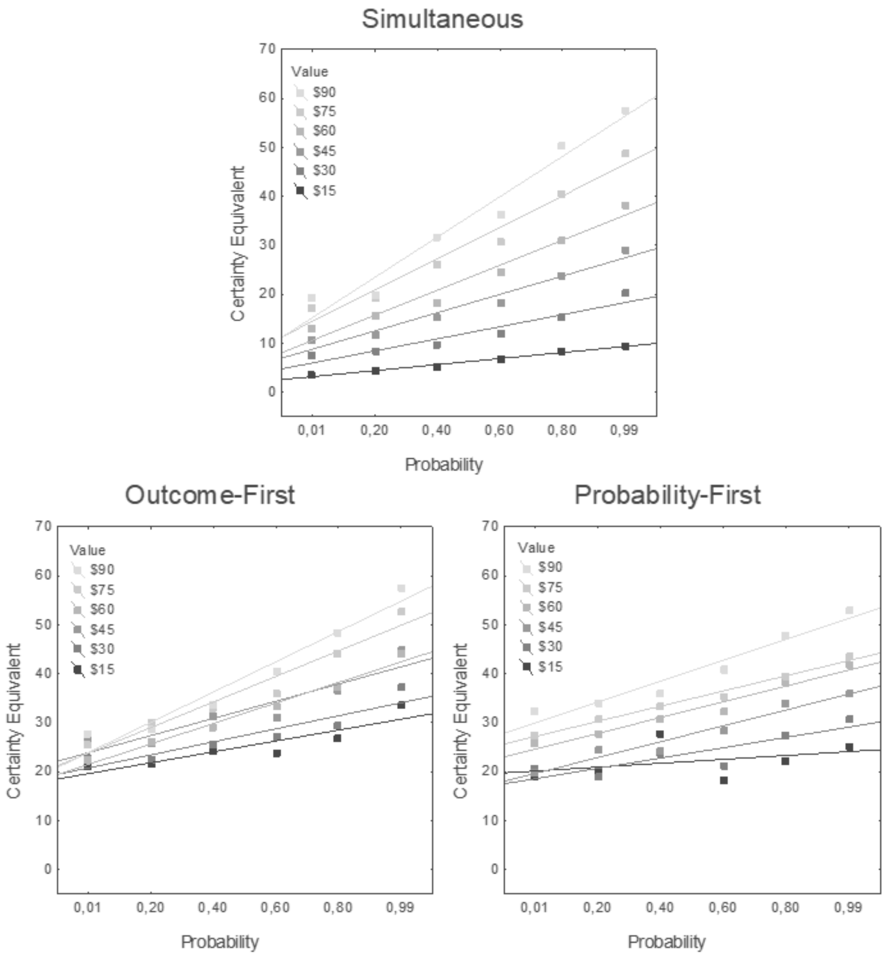

Figure 9. Average CEs (y-axis) as a function of probability (x-axis) and the outcome value (lines) for each condition in Experiment 3 (all participants).

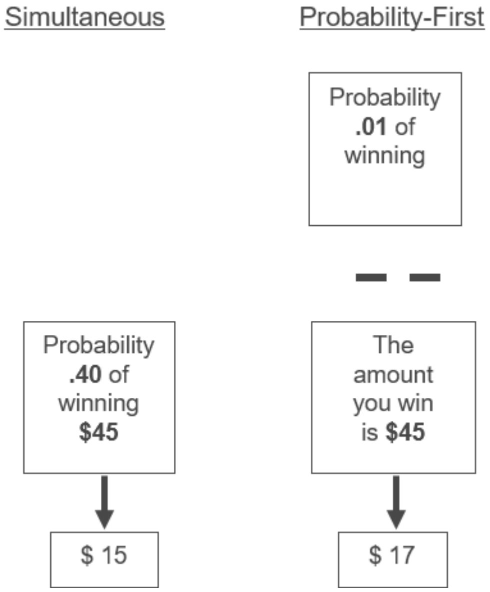

Figure 10. Example of data from two conditions in Experiment 3 reported below. In the Simultaneous Condition (left side of the figure) the outcome value and the probability is simultaneously available. In the Probability-First Condition (rights side of the figure) the participants first memorize the probability of events to asymptote in a first phase and later receive the outcome value in a later phase where they assess the Certainty Equivalents (CEs) for the lotteries. The mean CEs for these lotteries are provided at the bottom of the bottom of the figure.

Experiment 4

The experiment involved STM rather than memorization and long-term memory. Participants in the sequential conditions reported on average 29% higher CEs than those in the Simultaneous Condition (M = $18.21, *SD* = 12.8 in the Simultaneous Condition vs. M = $25.05, *SD* = 11.1 in the Outcome First Condition, and M = $21.58, *SD* = 8.55 in the Probability-First Condition). Figure 11 graphs the behavioral data for each condition, showing the typical fan-patterns associated with weighting of outcomes and probabilities. Factorial-regression analysis supported a weighting model for 80.1% of the participants in Simultaneous Condition, for 65.8% of the participants in the Outcome-First Condition, and for 61.9% of the participants in the Probability-First Condition. The proportion of the 15 items that violated stochastic dominance was .225 with a 95 percent CI between .073 and .456 for the Probability-First Condition. For the Outcome-First Condition the proportion of the 15 items that violated stochastic dominance was .347 with a 95 percent CI between .152 and .587. The substantially higher CEs accordingly arise robustly also with the immediately sequential presentation.

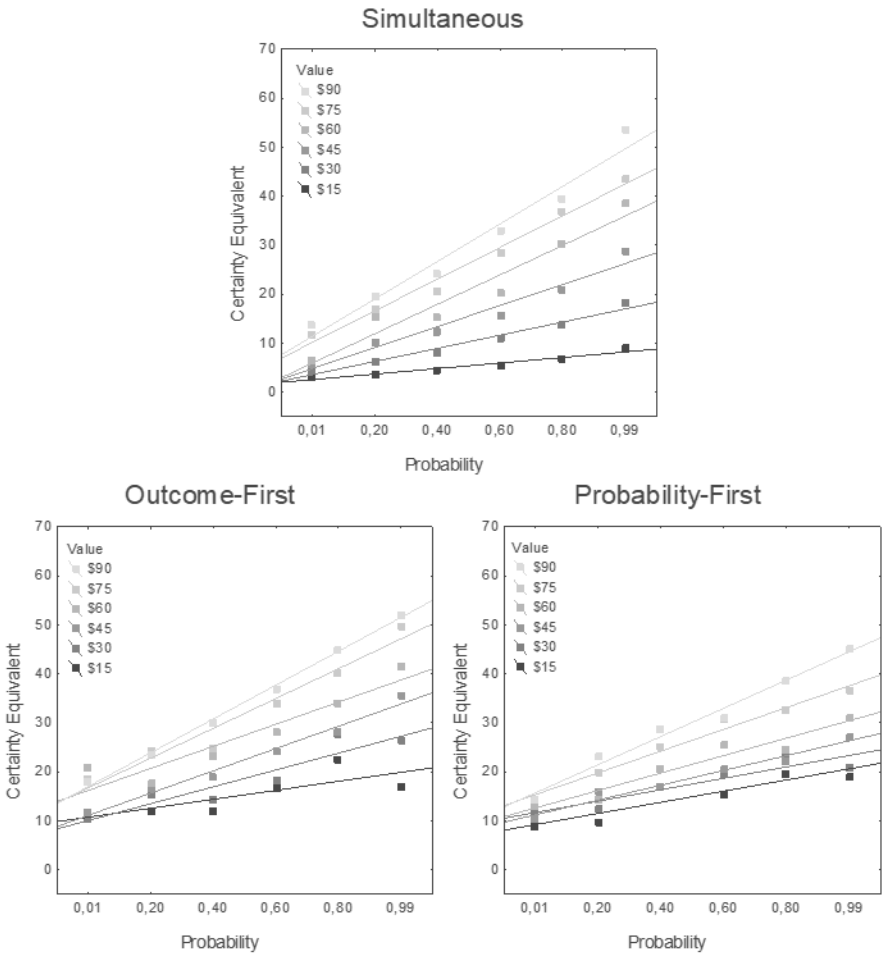

Figure 11. Average CEs (y-axis) as a function of probability (x-axis) and the outcome value (lines) for each condition in Experiment 4.

Discussion

The results across the four experiments suggests, again, that people are spontaneously inclined to integrate outcomes and probabilities through weighting, and that this tendency is robust to increased cognitive demands. The results also showed, however, that participants reported substantially higher CEs when they evaluated options sequentially, even when offered monetary incentives, and after controlling for the participant's recall of the information. What psychological mechanism could explain the main effect of format that is repeatedly evident in these experiments?

To elucidate this question, CPT and AAM were fitted to the data for the participants with perfect recall from Experiment 3. Model fit and best fitting parameters for each of the two models are summarized in Table 1 and the overall fit of the models to the data from the condition with simultaneous presentation is visually illustrated in Figure 12. Three conclusions can be obtained from Table 1. First, there is generally poorer quantitative fit for both models in the sequential rather than the simultaneous conditions, seemingly because of generally more "regressive" functions in the sequential conditions, presumably because of the larger memory demands in these situations.[11] Second, the AAM provides consistently better fit to the data in all three conditions than CPT. Figure 12 indeed suggests that the residuals from the predictions made by CPT are systematic already in the condition with simultaneous presentation. Finally, the best-fitting parameters of CPT are almost identical for all three conditions, not providing any account of *why* the CEs are higher in the sequential conditions. By contrast, the parameter for insufficient adjustment in the AAM (*a*) suggest systematically less adjustment from the outcome in the sequential conditions.

Although the experiments reported in this article were not designed to provide strong tests of these mechanisms, these results suggest that the insufficient adjustment captured by the AAM is at least one candidate mechanism when accounting for the data. Hence, it is plausible that memory demands can affect people's CEs from two routes: either as a result of remembering the wrong probability or outcome for the present context, or as a result of the adjustment process becoming even more effortful when having to recall probabilities and outcomes and store them in working memory when making the adjustments.

Table 1. *The model fit (coefficient of determination r^2 and mean square error (MSE) of prediction and best-fitting parameters for CPT and AAM.*

		Model fit		Parameters	
Model	Condition	r^2	*MSE*	*a*	*b*
CPT	Simultaneous	.85	27.2	.567	.959
	Outcome-First	.46	96.7	.544	.998
	Probability-First	.37	99.9	.548	.992
AAM	Simultaneous	.98	4.19	.712	.889
	Outcome-First	.67	40.5	.603	.912
	Probability-First	.67	38.2	.582	.904

[11] Remember that only about 60-70% of the participants with perfect memory-recall were classified as engaging in AW: it is plausible that the other participants did forget, or felt unsure about, some outcomes and probabilities, and hence did not make use of them at the time of the evaluation. Indeed, a model simply assuming responses that are regressive to the mean could account for a large amount of the variance (see published article).

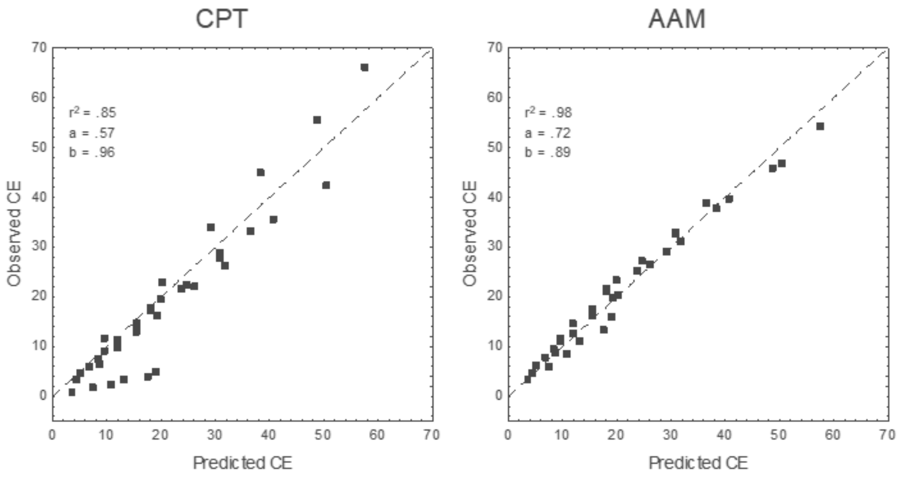

Figure 12. Visual illustration of the quantitative fit of CPT and AAM. The observed CEs are plotted against the prediction CEs for the condition with simultaneous presentation of outcome and probability, where the legends report the coefficient of determination r^2 and the best-fitting parameters.

Study 3

Background

Study I- and II suggest that people are spontaneously inclined to engage the normative intuition that outcome values should be weighted by their adherent probabilities, and that this inclination is robust to increased cognitive demands. Study III tests if this inclination extends beyond prospects of the simple structure used in Study I and Study II. To this end, the participants in Study III evaluate so-called duplex-gambles: prospects that have two independent outcome stages. Study III also explores the nature of the cognitive processes that underlie this ability by testing if AW involves analytic calculation or intuitive integration processes, and by testing if the inclination toward AW is best predicted by information-processing abilities or by lack of knowledge (or endorsement) of AW as a normative principle.

To the extent that probabilities are treated as cue-weights rather than exogenous cues[12] in the integration, research on multiple-cue judgments suggests that people should have the ability and the inclination for AW of the outcomes in situations that involve several outcomes. Some research suggest this to be the case (e.g., Anderson & Shanteau, 1970; Birnbaum et al. 1992; Shanteau, 1974; 1975; Schlottman, 2001; Viegas et al., 2012). Other research suggest, however, that people resort to simpler integration models that involve weighting only to some or extent, or to no extent at all (e.g., Levin, Johnson, Russo, & Delden, 1985; Joag et al., 1990; Mellers et al.,1992; Mellers & Chang, 1994; Mullet, 1992; Sokolowska & Pohorille, 2000; Shanteau, 1974; 1975; Sjöberg, 1968; Svenson, 1983).

The present study pursues the issue of pinpointing the characteristics of the cognitive processes that underlie AW from two stances. First, using the PNP-framework (Sundh et al., 2020; described in "Cognitive Modeling" of the present thesis), testing if the prevalence of spontaneous inclination to engage in AW is contingent on analytical thought processes or intuitive thought processes. For the purposes of investigating whether people rely on intuitive or analytical thought-processes when evaluating compound lotteries we re-analyzed data from 759 participants that had all assessed *prospects of the structure* [v_1, p_1, v_2], where v_2 *equals zero* (datasets from Millroth & Juslin, 2015; Study I; Study II). The results from that analysis showed that 20.1 per cent of these participants relied on analytical processes, 64.2 per cent relied on intuitive processes, and 15.7 per cent could not be classified

[12] On the meaning of exogenous cues: An exogenous variable in the context of regression analysis is a variable that is not affected by other variables. Weights are not exogenous variables; they invoke *exogenous change.*

(classification criteria described under "Cognitive Modeling").[13] That people rely on intuitive processes when evaluating simple prospect suggest that they may apply such strategies also when evaluating complex prospects. Second, it is tested if the ability to uphold AW is constrained by information-processing constraints (operationalized as short-term memory, STM, capacity), or by lack of knowledge (or acceptance) of the normative principle (operationalized as financial literacy, FL, and knowledge about expected values, EV-C).

Method

Both experiments involved a 3x3x3x3 factorial design, with probability for the first outcome, probability for the second outcome value, value for the first outcome, and value for the second outcome as independent within-subjects variables. Probability levels for the first- and second outcome were .10, .50, and .90 in both experiments. Value levels for the first- and second outcome were 100, 500, and 900 SEK in Experiment 1, and 10, 50, and 90$ in Experiment 2 – at the time of data collection corresponding more or less to the same amount of money. Probability- and outcome levels were factorially crossed to create 81 monetary lotteries with mathematical expectations ranging from 20 to 1620 (M = 500, SD = 362) in Experiment 1, and from two to 162 (M = 50, SD = 36.2) in Experiment 2.

After the prospect evaluations in Experiment 1, participants performed a task capturing numerical STM (see e.g., Hansson, Juslin, & Winman, 2008). In the task, sequences of integers were presented to participants for 5 seconds, and participants were thereafter allowed 15 seconds to enter their response with the computer keyboard. On a starting level, participants were shown four sequences of four integers each sequence. If at least three of the four sequences were reproduced correctly, participants reached the next level where the sequence length was extended by one. The test was terminated when participants were unable to exactly reproduce at least 3 out of 4 sequences on a given level. Thereafter, participants answered five questions that have been widely used in the literature to measures people's FL (Lusardi & Mitchell, 2009). Lastly, people answered three questions aiming at determining whether they had a specific understanding of expected values per se (expected value calculation, EV-C).

Participants in Experiment 2 also answered the five FL-questions (Lusardi & Mitchell, 2009), and the three questions capturing EV-C. The STM-measure was dropped in Experiment 2 because there was no effect of it in Experiment 1. Instead, Experiment 2 involved a measure of participants' levels of

[13] In Study I, 10% of the participants were classified as analytical; 20% as unclassified; and 70% as intuitive. In Study II, 14.2% of the participants were classified as analytical; 10.4% as unclassified, and 75.4% as intuitive.

numeracy - differences in how and how much people think about, and meaningfully understand, a numerical decision problem. This because there was an effect of FL in Experiment 1 for which it could not be ruled out that it was due to numeracy rather than FL.

Results

Experiment 1

Eighty per cent of participants were best classified by an AW model, suggesting that most participants had no trouble engaging the normative integration rule. The CE-plots of the aggregated data in Figure 13 reflect the results of the factorial regressions, showing that participants engaging in AW exhibit the fan-patterns typical of weighting of probabilities and outcomes. Their mean CE was 424 (95% credible interval, CI = 362; 486). The aggregated plots for participants *not engaging* in AW showed signs of both weighting (fan-patterns) and addition (lines being parallel), as expected given the occurrence of both types of models among participants. Their mean CE was 312 (95% CI = 280; 345).

An intriguing aspect of these data patterns is that participants *not engaging* in AW exhibit data-patterns at the aggregated level suggestive of the notion that AW could describe them on average even though no single subject were classified as using AW. Indeed, regression analyses of the aggregated data favored the AW model with a BF of 18.0 (R^2 = .886).

Both CPT and AAM provided good explanations of the data (median R^2 for CPT = .907; AAM = .961) – validating the use of them as basic models for the PNP-modeling. Importantly, the PNP-models were able to classify participants as both analytical and intuitive, offering proof-of-concept. Figure 14 provide a telling example of how responses from an analytical participant and an intuitive participants differ: the responses of the former generally aligns exactly with the EV, while the latter's responses seldom align with the actual EV but are instead normally distributed around the best fitting function.

The results of the PNP-modeling convey three findings. First, very few participants were classified as engaging analytical thought-processes where they seemingly number-crunched the expected values (6.25%); most participants were classified as engaging intuitive thought processes (93.75%). Second, most intuitive participants exhibited linear functions of probability and value (see Table 2). Third, the AAM model described behavior for more participants than CPT did: 14 participants were best described by the AAM; two by CPT, and the models could not be distinguished for 16 participants that had close to linear use of probability and outcome value, essentially engaging in noise-prone expected value calculation.

The use of probability is generally linear for AAM-participants – CPT predicts nonlinear use of probability. Indeed, if one derives the probability-weighting parameter for CPT for these participants it suggest nonlinear use of probability (median estimate = .759), suggesting that CPT model idiosyncrasies in the data rather than the underlying psychological properties. Also intriguing is the finding that the participants classified by the AAM exhibited a more concave value function than participants equally well described by the AAM as CPT: there is no a-priori reason to assume that AAM-participants should be more risk averse. Another intriguing prediction of the AAM in the present context – compared to CPT -is that the average CE can be higher than the average EV of 500 (it is, however, not a prerequisite of the model in the PNP framework), and this was actually the case for 13 per cent of the participants engaging in AW.

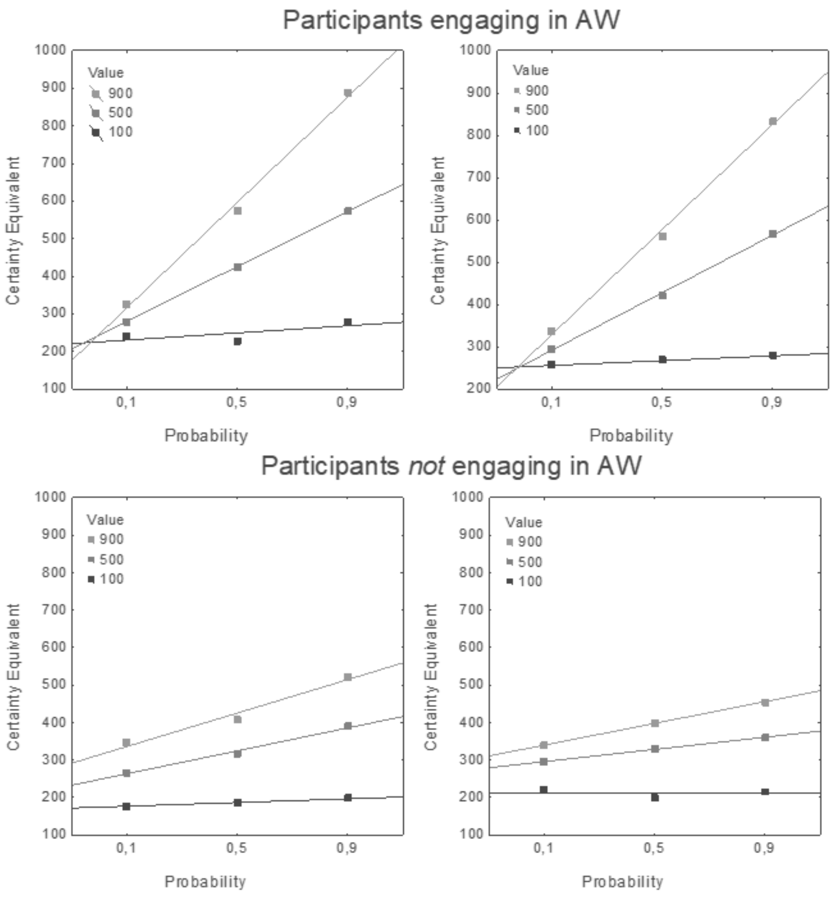

Figure 13. The mean reported CE (y-axis) for each value and probability level (x-axis) for each independent gamble-stage (left side = first stage, right side = second stage) for participants in Experiment 1 engaging in CM (top panels) and not engaging in CM (bottom panels). Horizontal bars indicate 95% confidence intervals.

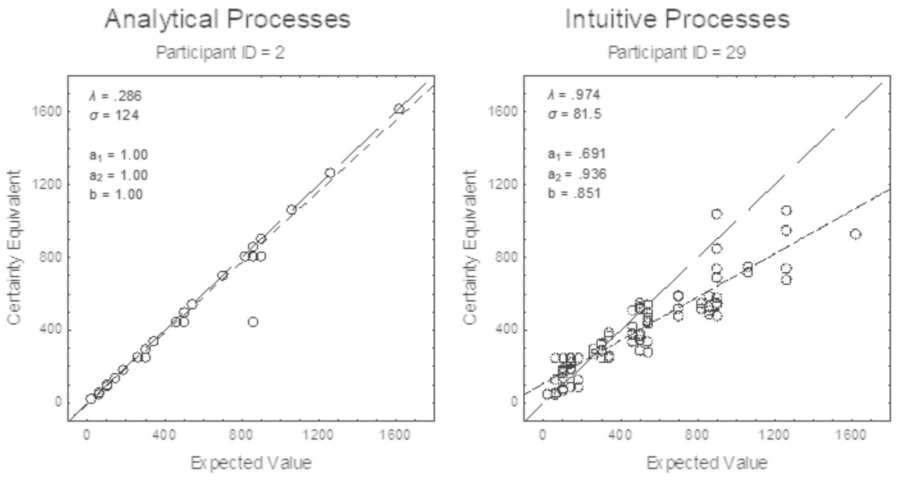

Figure 14. Illustrative examples of participants in Experiment 1 engaging mainly in analytical thought-processes (left panel) and intuitive processes (right panel). A line representing the predictions of a linear regression model (thinner line) and a reference line $x = y$ (thicker line) is included in the graph.

Table 2. *Summary of PNP-modeling, showing the median parameter estimates for each model (95% credible intervals are presented in parentheses).*

	Parameter Estimates				
Model	Lambda	Sigma	a_1/r	b	a_2
AAM	.941 (.919; .952)	99.1 (76.6; 180)	.960 (.719; 1.00)	.949 (.910; .965)	.975 (.925; .990)
CPT	.921 (.573; >1.00)	127 (12.9; 241)	.883 (.841; .926)	.882 (.881;.883)	-
Expected Value	.855 (.675; .903)	120 (109; 150)	1.00 (.997; 1.00)	1.00 (.987; 1.00)	1.00 (.987; 1.00)

Note. Parameter Estimates for "Both" is from the AAM, but could as well have been from the CPT-model; they are more or less virtually equivalent in these cases (1.00/1.00)

In an initial effort to explore the effects of FL, STM, and EV-C, three Bayesian ANCOVAs were computed, with FL, STM, and EV-C as the dependent variable, AW model (yes or no) as the independent variable, and covariates were composed of the variables not used as the dependent variable[14]. For example, STM and EV-C constituted the covariates when FL constituted the dependent variable. Table 3 summarizes the results, with the only effect being that the levels of FL differed depending on whether an AW model was

[14] Appendix B in the manuscript for Study III provide the results from logistic- and factorial regressions that ultimately provide the same conclusion as the results from the Bayesian ANCOVAs; presenting the ANCOVAs do, however, provide additional pedagogical efficiency in communicating the results.

used or not (M = 4.13 for the participants engaging in AW, vs. 2.64 for the others).

Logically, analytical number-crunching of the expected value requires actual knowledge about how to calculate expected values. Descriptive data indeed showed that the participants classified as analytical scored on average higher than the participants classified as intuitive on EV-C (3.00 vs. 1.83), as well as FL (3.50 vs. 3.13) and STM (9.38 vs. 8.04), but it was not possible to assert these patterns with inference testing because there were only two analytical participants.

Table 3. *Results of three Bayesian ANCOVAs with AW model (yes/no) as independent variable. The evidence for the presence of an effect of aw model on the dependent variable (FL, STM, EV-C) is combined across models that include the effect. P(incl) convey the prior inclusion probability, P(incl|data convey the posterior inclusion probability, and BF_{incl} convey the BF – the change from prior inclusion odds to posterior inclusion odds.*

Dependent Variable	P(incl)	P(incl|data)	BF_{incl}
FL	.500	.903	9.36
STM	.500	.287	.402
EV-C	.500	.431	.757

Note. STM and EV-C constituted covariates for the first model; FL and EV-C for the second model; FL and STM for the third model.

Experiment 2

Twenty-seven per cent of the participants were classified as AW. The CE-plots of the aggregated data in Figure 15 reflect the results of the factorial regressions, showing that participants engaging in AW exhibit the fan-patterns typical of multiplicative integration. Their mean CE was 30.7 (95% CI = 23.4; 37.9. The aggregated plots for participants *not engaging* in AW showed signs of both weighting of cues (fan-patterns) and neglect of probability information (lines being parallel), as would be expected with those heterogenic patterns in the data. Their mean CE was 32.9 (95% CI = 26.7; 39.1. In contrast to Experiment 1, there was no evidence that the aggregated responses from participants not engaging in AW could be explained by an AW model ($BF_{\text{against best model}}$ = 32.1, R^2 = .728).

The results align with the results from of Experiment 1 in several ways. First, a minority of participants were classified as engaging in analytical thought-processes where they seemingly number-crunched the expected values (27%; see Figure 16 for illustrative examples of participants engaging in both types of processes). Second, the AAM model best described behavior for more participants than CPT (16 participants were best described by the AAM; 1 by CPT, and the models could not be distinguished for 8 participants that had close to linear use of probability and outcome value). Third, most participants exhibited linear, or close to linear, functions for probability adjustments (see Table 4).

The participants best described by the AAM exhibited a more concave value function compared to participants essentially calculating a noisy EV. An intriguing possibility is that the more concave value function is an effect of a difference in symbolic number mapping. People best captured by the AAM may be people with lower levels of numeracy, a group that typically exhibit more inexact number mapping. The results from a Bayesian Mann-Whitney Test with numeracy as the dependent variable support this notion (M = 2.19 vs. 2.88 BF = 5.05)[15].

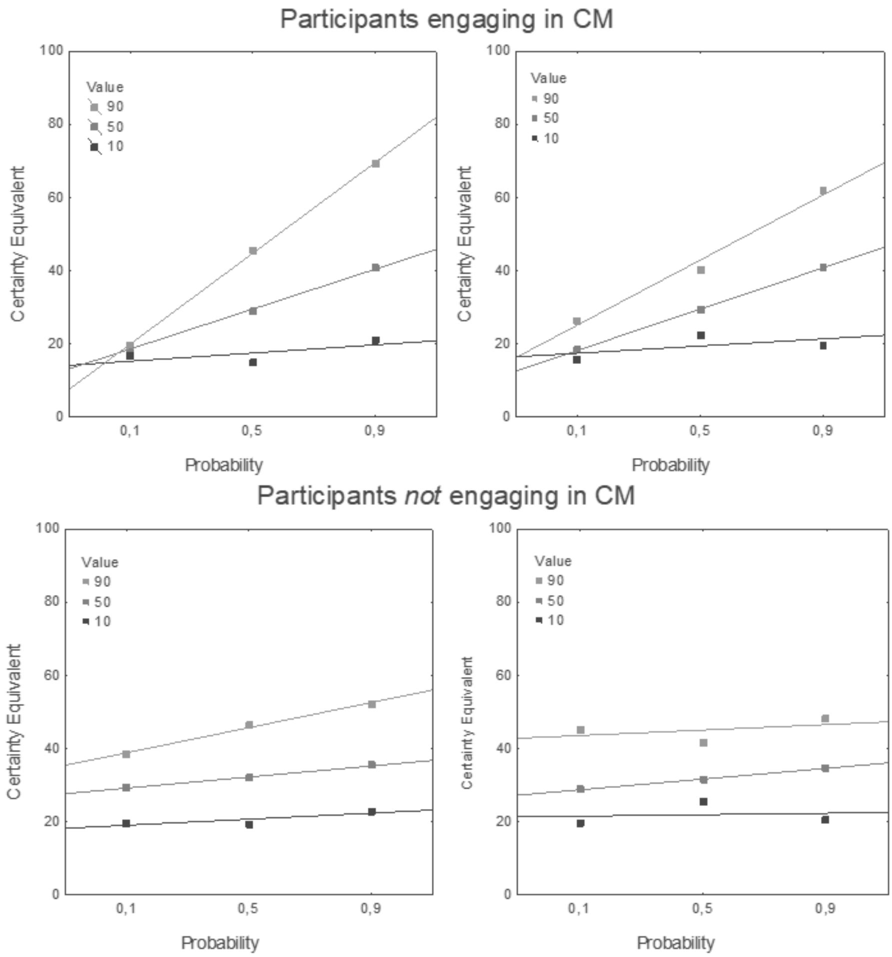

Figure 15. The mean reported CE (y-axis) for each value and probability level (x-axis) for each independent gamble-stage (left side = first stage, right side = second stage) for participants in Experiment 2 engaging in CM (top panels) and not engaging in CM (bottom panels). Horizontal bars indicate 95% confidence intervals.

[15] The participant best explained by CPT was excluded from the analysis.

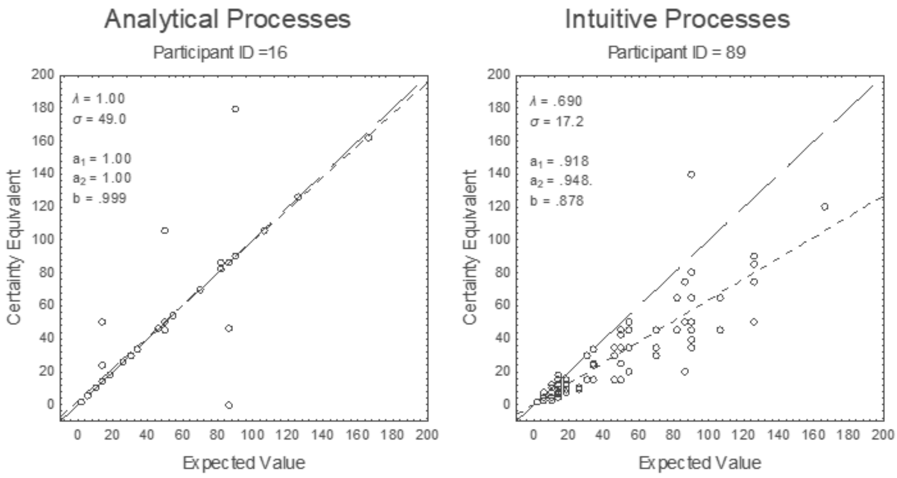

Figure 16. Illustrative examples of participants in Experiment 2 engaging in analytical thought-processes (top panels) and intuitive processes (bottom panels).

Table 4. *Summary of PNP-modeling, showing the median parameter estimates for each model (95% credible intervals are presented in parentheses).*

	Parameter Estimates				
Model	Lambda	Sigma	a_1/r	b	a_2
AAM	.589	17.9	1.00	.855	.924
	(.460; .694)	(13.0; 24.0)	(.923; 1.00)	(.635; .878)	(.827; .950)
CPT	.312	20.9	1.00	.318	1.00
Both	.532	30.0	1.00	.980	1.00
	(.241; .686)	(23.4; 41.3)	(.998; 1.00)	(.923; 1.00)	(.984; 1.00)

Note. Parameter Estimates for "Both" is from the AAM.

As in Experiment 1, three Bayesian ANCOVAs were conducted, but now with FL, numeracy, and EV-C as the dependent variables, AW model (yes or no) as the independent variable, and covariates were composed of the variables not used as the dependent variable. Table 5 summarizes the results, showing that there was only an effect of FL: the levels of FL differed depending on whether participants used an AW model or not (M = 4.11 for the participants engaging in AW, vs. 3.06 for those *not engaging* in AW). Aggregating evidence for the effect of FL across the two experiments yields a BF of 51.4[16]. Bayesian Mann-Whitney Tests showed that participants classified as analytical scored higher than participants classified as intuitive for EV-C (M = 2.71 vs. 1.27; BF = 11.1) and numeracy (M = 3.00 vs. 1.91; BF

[16] Wagenmakers (2007a) proposed aggregating BFs across studies through multiplication. Another proposed method is to analyze the effect when assuming that the responses are from the same sample (Wagenmakers, 2007b). Applying this procedure to the present results yielded a BF > 1,000.

= 5.32), but that there was no difference in levels of FL (M = 4.00 vs. 4.18; BF = 1.16).

Table 5. *Results of three Bayesian ANCOVAs with AW model (yes/no) as independent variable. The evidence for the presence of an effect of aw model on the dependent variable (FL, Numeracy, EV-C) is combined across models that include the effect. P(incl) convey the prior inclusion probability, P(incl|data convey the posterior inclusion probability, and BF_{incl} convey the BF – the change from prior inclusion odds to posterior inclusion odds.*

Dependent Variable	P(incl)	P(incl|data)	BF_{incl}
FL	.500	.846	5.49
Numeracy	.500	.253	.339
EV-C	.500	.209	.265

Note. Numeracy and EV-C constituted covariates for the first model; FL and EV-C for the second model; FL and numeracy for the third model.

Discussion

The results showed that people are often inclined to engage in AW when evaluating duplex gambles, and that they often do so by engaging in an intuitive process often well described by the AAM. The results also showed, however, that this inclination was dependent on FL: more financially literate individuals were more likely to engage in AW.

The results of Experiment 2 differed notably from that of Experiment 1, the primarily difference being that only 27 per cent of the participants were classified as engaging AW– compared to 80 per cent of the participants in Experiment 1. A straightforward explanation holds that the difference is a direct consequence of the variance in the cognitive skills between the samples (i.e., FL, EV-C). Data supports this explanation: A logistic regression with AW (yes/no) as the dependent variable, Experiment (1/2) as the fixed factor, FL and EV-C as crossed covariates, yielded a BIC-difference of 9 in favor of the hypothesis that Experiment could *not* predict use of AW. Table B1 and Table B4 provide the mean and median estimates of FL and EV-C for both experiments. The FL estimates of the sample in Experiment 2 also correspond to recent research that have applied FL in an online setting using over 2,000 participants (see Krische, 2015 – who also show that these estimates reflect estimates of the general U.S. population). Moreover, if subject inattentiveness that arise from lacking motivation was the driving factor, then the use of AW should have been predicted also by numeracy because there are no a priori reasons to suspect that participants suddenly would become motivated for that particular task. Finally, more random-responses, or sticking to the same response, would have yielded more evidence for NI reasoning, since the null-model is included in the NI category)– but there was only a minor drive towards more NI reasoning (14% in Experiment 2 compared to 0% in Experiment 1). Another difference between the two experiments is that the explained variance for linear- and nonlinear AW models (the linear regressions, AAM, CPT) were generally lower in Experiment 2. One possibility is that different aspects of the decision-making process are affected by different components: while FL may determine the type of type of process applied (AW model or not), numeracy may determine how well this process is executed. Exploratory correlation analysis supported the notion that numeracy was related to the form of the variable functions, while FL was not (see Appendix B in the manuscript). Some of the discrepancy in explained variance of the linear AW model between Experiment 1 and 2 may thus interpreted in this light: participants in the online sample are less numerate, and less numerate people exhibit more nonlinear variable functions (Millroth & Juslin, 2015; Traczyk & Fulawka, 2016).

General Discussion

Summary and Conclusions

The thesis aimed at exploring the robustness of the process of weighting probabilities and outcomes, in the context of evaluations of risky prospects. Robustness was tested under three conditions. Study I tested the possible effect of making only an isolated evaluation compared to making evaluations in the context of many other evaluations. Study II tested the possible effect of sequentially presenting probabilities and outcomes. Study III tested the possible effect of increasing the complexity of prospects (i.e., duplex gambles). The results across the three studies show that people are often spontaneously inclined to engage in the normative principle of weighting the value or utility of possible outcomes with their adherent probabilities, even when cognitive demands are increased.

A second aim of the thesis was to acquire a better understanding of the actual cognitive process of weighting. To this end, a formalized model – the anchoring-and-adjustment model (AAM) - was developed (Study II and Study III). The AAM holds that the outcome value serves as an anchor from which people make downward linear adjustments to take into account that the probability of the outcome is less than one. This adjustment may be coarse and insufficient because people's cognitive constraints force them to draw on a range of values rather than precise estimates. The AAM was compared with a benchmark model, cumulative prospect theory (CPT), whose explanatory power lay in the proposed psychophysical functions for probability and utility (Studies II and III). Study III applied cognitive modeling of the processes to delineate if the processes posited by the AAM are of intuitive or analytical nature. Study III also tested if the prevalence of additive weighting (AW) among participants is best predicted by knowledge factors (financial literacy – FL, numeracy, and knowledge about expected values – EV-C) or by information-processing capacity (numerical short-term memory, STM).

The results showed that the AAM described people's behavior better than CPT in Study II and Study III. For the present purposes, also the results of Experiment 1 were re-analyzed using the PNP-framework, showing that the AAM provided a better explanation of the data, and that analytic processes seemed more prevalent for participants in the WSD sample.[17] Study III showed

[17] Parameters were fitted for three levels of tau (as in Study III), but even the AAM models that did not involve rounding outperformed all CPT models (BIC-differences between 5 and 15). Best fitting AAM

that the processes underlying the AAM were dependent on cultural knowledge (i.e., FL) rather than information-processing constraints (i.e., STM). All three studies reported behavior that are problematic for benchmark descriptive models where the psychological content of the parameters captures risk preferences and diminishing marginal utility. More specifically, there is no clear a-priori reason suggesting that the number of available anchors (Study I), sequential presentation of probabilities and outcomes (Study II), or the size of the model space (Study III) should affect people's diminishing marginal utility or perception of risk per se.

Theoretical Implications

As argued in the Introduction- and Background sections, the research field of judgment- and decision-making under risk has yet to provide satisfactory accounts of how people integrate probabilities and outcomes. This likely, at least partly, a result of the polemic debates between proponents of theories holding that people behave engage in AW (e.g., CPT), and proponents of theories holding that people do not engage in AW. The present thesis makes this issue a key focus for empirical investigation and provide a novel middle ground between the conflicting accounts. The results suggest that even if people can turn to heuristics when they are more efficient, for certain stages in the decision process, or for very complex problems, they indeed have both the inclination and ability to weight the outcomes by their probabilities in the evaluation of individual prospects, or for a subset of more promising decision alternatives. This weighting is, however, not consistent and effortless. Adjustments for probability are often insufficient or noise-prone – especially when the cognitive demands increased as a property of the task environment (Study I and Study II), and when people lacked domain-specific knowledge (i.e., numeracy and FL, Study III).

Competence and Performance

Study I introduced the idea that judgment- and decision-making research should make distinctions between *normative competence* and *normative performance.* This idea posits that the decision-making behavior produced at any moment – the performance – is affected by various cognitive and situational limitations, leading to deviations from the normative competence (for the origins of this idea, see Chomsky, 1965). The idea offers a complement – or alternative –to the popular dual-systems view which holds that specific systems are responsible for upholding behavior that align with normative theory (System 1 vs. System 2, or intuitive system vs. analytics system;

---

parameters for the WSD sample: lambda = .861, sigma = 2.91, a = .825, b .860. Best fitting AAM parameters for the SSD sample: lambda = .964; sigma = 4.35; a = .737; b = .856.

Kahneman & Fredericks, 2002; Evans, 2008). The results of the present thesis indeed support the idea that most people have the normative competence to engage in AW, whether this competence reflected in normative performance depends on a variety of circumstantial and contextual factors. Whether this competence/performance distinction can be neatly mapped onto two systems or not remains to be charted in future research. As shown in Study III, however, people can engage the normative integration-rule through both intuitive- and analytic cognitive processes. This conclusion is in line with the large literature on multiple-cue judgments (Hoffman, von Helversen, & Rieskamp, 2014; 2016; Juslin et al., 2008; Pachur & Olsson, 2012; Platzer & Bröder, 2013; von Helverson & Rieskamp, 2009).

The results also showed that people often made insufficient adjustments or adjustments that were error-prone, and that domain-specific knowledge determine the degree of engagement in AW. An intriguing possibility is that interventions addressed at increasing numeracy and financial literacy could lead to people's competence being more strongly reflected in their performance (for support of this hypothesis, see Garcia-Retamero, Sobkow, Petrova, Garriodo, & Traczyk, 2019; Lührmann, Serra-Garcia, & Winter, 2018; Schoemaker, 1979; Sobkow, Fulawka, Tomczak, Zjawiony, & Traczyk, 2019, for counter-evidence see Montgomery & Abelbratt, 1982; Williams & Connolly, 2006).

The results may carry implications regarding strategy selection in a choice setting. People may switch to simpler models in choice tasks because these tasks often introduce more than four attributes in total over all alternatives (corresponding to the two non-zero monetary outcomes and their adherent probabilities that are involved in the compound lotteries of the present study). For example, consider the problem below that Kahneman and Tversky's (1979) used as part of their famous illustration of the Certainty Effect:

"Problem 2: Choose between

C: 2,500 with probability .33	D: 2,400 with probability .34,
0 with probability .67	0 with probability.66",

where the amount of calculation for an AW agent is in essence similar to the situation of evaluating duplex lotteries (except that the products are compared instead of added). It is not far-fetched to assume that people will turn to strategies that they feel more confidence in, or strategies that they applied in previous similar situations with favorable results. Of course, it may also be that people switch to heuristics because they do not *endorse* the normative rule in these cases. This idea holds that people may choose to endorse heuristics *even if* they have the competence to engage in AW. Because heuristics offer specific advantages. For example, they may decrease error variance in prediction (Gigerenzer & Brighton, 2009), provide cost-efficient solutions (e.g., Payne et al., 1993), and can result in consequences that are more favorable for people in specific environments (Schurz & Hertwig, 2019). Recent research has indeed shown that perceived cognitive-effort predicts the use of normative rules (Traczyk et al., 2018).

From where originates the normative competence of AW? The broader literature in Cognitive Psychology suggest two plausible explanations. One possibility is that the process of weighting is so intrinsic to the evaluation of risky prospects that evolution has provided innate support of this competence. The fact that both animals and children spontaneously engage in this process supports this claim. From this point of view, the remarkable fact is that adult humans have acquired the ability to use purely symbolic representations of risk to inform their weighting of the possible outcomes. Another explanation, albeit possibly overlapping, is that the ability to engage in weighting is extremely important to many human affairs and that the process thus becomes automatized already early on in people's lives. To illustrate, consider the difference between judging some physical quality (e.g., the area of a square) and judging the amount one is willing to pay for a prospect. The ability to judge physical qualities seems to develop in tandem with a more analytical mindset from an age of around eight years, before which children integrate the variables by additive integration. Still, children as young as five years old seemingly use the weighting rule when integrating probabilities and outcomes, suggesting they have developed an intuitive understanding of the concept well before they are apt to have formed an analytical strategy of integrating the variables (for a review, see Schlottmann, 2001). Processes can presumably – when performed repeatedly – become automatized so that the integration and retrieval processes are elicited directly by the input when in a specific circumstance (e.g., if you are used to approximate the area of a plank, you might naturally gravitate to retrieving the product of the height and length when confronted with such a task). However, it may also be that there are specific semantic triggers causing people to rely on automatized reasoning schemes (e.g., relative quantifiers).

It seems reasonable that probability has come to act as such a trigger because weighting - in contrast to AI, PW, and NI – may offer adaptive benefits in environments involving risk. Research on animal-cognition has indeed suggested that various animals – from bees to rhesus macaques – implement the normative integration rule when making decisions involving risk (for a review, see Farashahi, Donahue, Hayeden, Lee, & Soltani, 2019, which also discuss evidence that people and primates abandon this strategy when risks are not known or cannot be approximately inferred). Probability theory is, however, a rather recent invention in the lens of evolution (Cardano, 1663; Huyggens, 1657; Fermat & Pascal, 1654) – not widely applied in society before the 19th century (see Gigerenzer et al., 1990). That this automatized reasoning-scheme has been able to, rather quickly, transgress the barriers of numeric comprehension is impressive. It does imply, however, that it should be easier to disrupt this automatized reasoning scheme for numerical setups compared to setups where probabilities are experienced as frequencies of outcomes (for evidence that manipulation of numerical setups can lead to disruption, see Sundh et al., 2020).

Preference or Ability?

The notion of stable risk preferences lay at the core of normative and descriptive theories of decision-making under risk (e.g., von Neumann & Morgenstern, 1944/2007; Tversky & Kahneman, 1992). It does seem, indeed, that there is something equivalent of the g-factor for general intelligence in the risk-domain (Frey, Pedroni, Mata, Rieskamp, & Hertwig, 2017). Even though it has been found that the probability-weighting function and value function of CPT can vary considerable across contexts (e.g., Pedroni et al., 2017), these parameters – elicited from behavior in gambling tasks (such as monetary lotteries) – are generally held to capture risk preferences, stable or not (Glimcher & Fehr, 2014). A soaring wave of recent research has, however, questioned using monetary lotteries to elicit people's risk preferences, arguing that these lottery-preferences are contaminated by their demands on numerical and cognitive abilities and only weakly relate to any general disposition towards risk (for reviews, see Mata, Frey, Richter, Schupp, & Hertwig, 2018; Millroth, Juslin, Winman, Nilsson, & Lindskog, 2020). The present thesis align with this literature. Again, it seems far-fetched to believe that the manipulations conducted in the experiments of this thesis would have such profound effects on people's appetite for risk per se. Currently, the best way to capture any general disposition towards risk seem instead to be by measuring people's self-reported risk preferences (e.g., "Do you prefer to take risks?": Mata et al., 2018; Zhang Highhouse, & Nye, 2019).

Still, even if risk preferences revealed through behavioral tasks mainly define relatively advanced demands on people's ability to process and understand numbers and abstractions such as "probability", rather than risk preferences per se, it does not mean that behavioral task should be should abandoned. Such action would be premature for (at least) two reasons. First, behavioral tasks and self-reports may capture two complementary tools that people have at their disposal when encountering environments that involve risk and uncertainty. On one side is the general *risk preference*: a general disposition formed through human evolution (risk aversion and risk seeking can both offer adaptive advantages: e.g. Mallpress et al., 2015), and that can be informed by key events in an individual's life (e.g., losing a fortune on the stock market). This preference can be stored by biological systems that offer immediate affective responses used to guide behavior (Glimcher & Fehr, 2014). One the other side is the *risk ability*: the ability to form and use context-specific representations about outcomes and probabilities when making evaluations and choices – either through analytical- intuitive thought processes. In part, risk preferences and risk abilities may stem from different processes: Neuropsychological findings have recently implied that people approach risk taking in one of two ways: by relying on a rule-based integration of outcomes and probabilities, or by a habitual processing of rewards (see e.g., O'Doherty, Cockburn, & Pauli, 2017). Risk propensity is linked to reward sensitivity (e.g., Blum et al., 2000; Yarosh et al., 2014), which does not share the neural underpinnings used for the encoding and weighting of uncertainty (e.g., Bach & Dolan, 2012; Li, Vanni-Mercier, Isnard, Mauguiere, & Dreher, 2016; O'Neill, & Schultz, 2010). Note that this account do not necessarily correspond with the idea of dual-system theory (e.g., Evans, 2008).

To illustrate, consider the case of a drug-addict: there is definitely a probability of being severely injured from accidental drug-overdose or from consumption of "dirty" drugs – but in the mind of the addict these components do not play a role in the decision to consume the drug at hand; the urge to consume is all there is. To illustrate the opposite end of the spectrum, consider choosing between stock funds offered by the bank. Most banks typically give an estimate of expected returns from the funds, and the risk associated with the fund. In this case, people may engage in cognitive processes that are analytical in nature (e.g., by estimating an expected value) or intuitive in nature (as the type of weighting proposed in this thesis). The drug addict can very well be a well-calibrated stock trader – and such "paradoxical" conflicts does not need necessarily reside from the notion of competing "systems". The extent to which risk preferences and risk abilities interact seems an especially promising direction for future research (for promising advances, see Mishra, Barclay, & Sparks, 2017; Williams & Connolly, 2006).

Practical Implications

The theoretical implications may also translate to practical implications. Many view the malleability of people's preferences as an acknowledgement of human irrationality – a notion that has spread outside of academia with the exponential increase of popular-science books (e.g., Kahneman, 2011). Conveying the message that humans are cognitively unable of many tasks may lead to a collective self-fulfilling prophecy (Merton, 1948). Such a message is also at the risk of being used to justify policy-interventions that are to "correct" for people's cognitive limitations, interventions that can come with the (perhaps unintentional) sacrifice of people's autonomy and sense of competence (for a discussion of "nudges" – see Lembcke, Engelbrecht, Brendel, & Kolbe, 2019). For example, it been estimated that increasing normative economic performance can increase expected economical returns up to 10% in organizations and governments (Lovallo and Sibony, 2010), ultimately yielding many positive benefits for society (no matter one's personal political position).

However, pursuing these returns by means of overcoming people "stupidity" seem, in many cases, not only unethical, but also less promising than the alternatives. Sure, people do at times exhibit "biases" in the sense that they depart from normative theory. Given the impact of the decision environment (e.g., Simon, 1956), it is plausible that the decision maker should not be foremost to blame (see also Gigerenzer & Goldstein, 1996, for "rational" explanations of so-called biases). Importantly in this regard, the present thesis has indeed shown that people can connect with their underlying competence when the environment is properly structured. Future research should examine if normative performance can be increased through restructuring the environment in a manner that allows people to connect with their normative competence instead of achieving this goal through deceptive means (for promising examples in this direction, see Hertwig & Grüne-Yanoff, 2017). Study I and II both illustrate that one straightforward application is to structure the environment so that information is presented simultaneously and comparative anchors are available.

The theoretical implications of how scientists should study the elusive concept of risk-preference may also carry practical implications. Conclusions about people's risk preferences have, historically, been used to make inferences about many behaviors associated with poor life-outcomes (e.g., excessive drug-use, gambling disorders, unprotected sex, health-plan choices, and criminal behavior - see Frey et al., 2017; Meertens & Lion, 2008; Weber & Blais, 2002). The results in this thesis suggest that the cognitive processes involved when integrating probability- and value information constitute part of people's risk competence - future research should thus examine the extent to which these processes affect real-life decisions.

Limitations

Even though published evidence generally shows data to be of equal or better quality for online data (for reviews, see Gosling & Mason, 2015; Hauser & Schwarz, 2016), subject inattentiveness remain an issue (see Cheung, Burns, Sinclair, & Sliter, 2017). The experiments here did, however, implement several of the precautions that Hauser and Schwarz (2016) argue increase the probability to collect high-quality data. This includes recruitment of experienced and accurate contributors, screening the data for "impossible" vales (i.e., checking if the CEs were positively correlated with the expected value of the prospects). It is also important to note that, if anything – more "quality" data would have only strengthened the conclusion that people have the normative competence to engage in AW. It remains, however, a real possibility that adjustments for probability would not be as insufficient in a controlled lab setting.

As with any experiment, it is hard to determine to what degree the results are representative of the general population. The samples of online crowdsources align with the general population of the United States on important demographics such as age, income, and educational level, but it cannot be ruled out that crowdsource workers are special in other aspects (after all, many people are not inclined to spend time on tasks for a small monetary compensation). That core findings overlap across samples (crowdsource workers in the U.S and campus-recruited participants in the lab) suggests, however, that the findings are robust to idiosyncrasies of the samples.

The stimuli used in the experiments were limited in scope. Even though it was a deliberate decision to focus only on numerically described gambles, it means that the conclusions cannot be directly translated to, for example, the context of decisions-from-experience (where people infer probabilities from sampling outcomes). Moreover, the designs of the experiments were constrained to a very limited stimuli-space; prospects that involved probabilities evenly distributed between zero and one, and outcome values that can be considered "pocket money". The present thesis has argued that instable mental number lines affects the extent that adjustments are insufficient; it follows from that the average effects of insufficient adjustments will be even stronger with design using larger values and exclusively low-end probabilities. Although this is a limitation in regards of external reliability, it is arguably a strength in regards to the validity of the results: cognitive science should develop theories grounded in a simple drosophila, even when the aim is to understand how complex information-processing systems interact with the world outside of the laboratory (e.g., Minsky, 1960; Marr, 1982). In other words, moving towards more ecologically valid environments is appropriate first when numerous predictions have been tested and the theory updated (or reversed) accordingly.

Moreover, the AAM has yet to be evaluated using process-tracing measures. Algebraic process models are only complete when combined with process measures (Lopes, 1995). Furthermore, even though the AAM generally fared well in the present studies, these studies were not designed to delineate between models. Finally, humans are emotional systems (Damasio, 1996; Loewenstein, Weber, Hsee, & Welch, 2001; Minsky, 2006) but the present thesis does not convey to what extent emotions, such as fear or surprise, affects adjustments (but for evidence that emotion-regulative neurotransmitters can affect the processing, see Millroth, Ågren, Eriksson, & Juslin, 2018).

Personal Remarks: The Future Ahead

Some have accused the field of judgment and decision-making under risk of suffering from the "toothbrush problem" (Mischel, 2005; 2008): the tendency to treat other people's theories like dirty toothbrushes, staying away from them at all costs. Some have even argued that the polemic debates have only amounted to only modest theoretical advances over the past 50 years (Taleb, 2018). Although such critiques may be too harsh, they, nevertheless, contain elements of truth. I am personally guilty of thinking in these terms as well; I approached this project with the hope of uncovering a new set of models that would differ from both the typical additive-weighting models and the heuristic- and sampling accounts. I am grateful though that the evidence pointed to another direction, and it is my hope that the present thesis will help shift attention towards the *nuances* of decision-making under risk: humans are neither cognitive cripples nor normatively calculating agents, and do not engage in their environments with only one type of tools (e.g., heuristics, sampling, AW). This much is clear. Now we need to turn towards more careful mapping of how the structure of the environment interacts with constraints of the decision maker to evoke those specific nuances – an endeavor proposed long ago (e.g., Brunswik, 1955, Simon, 1956) but that, nevertheless, still awaits to be fully utilized.

Acknowledgments

Many before me have described the path towards acquiring a PhD-degree as at times being “effortful”, “mentally demanding”, or “stressful”: But I think a paragraph from the late (great) writer Hunter S. Thompson provides the most vivid analogy to describe the occasional feeling of despair. In Thompson’s case, he details the feeling of a working as a traveling sportswriter struggling with how to properly conclude a reportage. It could just as well have been about undergoing six years of PhD-studies:

“Should we end the bugger with that?
Why not? Let the sportswriters take it from here.
And when things get nervous, there’s always that
smack-filled \$7-a night motel room down on the seawall
in Galveston.” - Hunter S. Thompson (1973).

This section is devoted to those who not only kept the seawall in Galveston a mere glimpse in the horizon, but who also contributed to making these years the most enjoyable of my life thus far. Sorry if these pages are full of grammatical errors, but it was nice to finally be excused from supervisors and reviewers.

Peter: I am amazed and thankful that you never made me feel stupid during these years. You have provided ” tools of science” that have changed not only my stance on academic issues, but also how I confront issues outside of academia.

Håkan: you have done much more than what is required of a co-supervisor, and I am definitely a better scientist for it. It has also been nice to discuss matters that are more serious than science (e.g., how to handle the emotional states of “kvalserietider”).

Thomas Å: Our collaborations has been crucial for developing my independence as a researcher, and I look forward to new exciting projects together!

Anders, Ronald, and Tomas F: The input from all of you at the half- and full time seminar really improved the final product. Special thanks also to Anders for “showing me the ropes” as a young and naïve undergraduate student,

To my fellow PhD-colleagues in the Cognition- and Perception Group (past and present): **Joakim**: I could not have wished for a better “cellmate”! You have advanced my horizons on many issues, not only science-related ones. **August:** Your friendship makes going to work even more fun, and I am (actually) thankful that you constantly act as “Reviewer C” when I present ideas and stances. **Elina:** when everyone else jumped ship for various dubious reasons (graduating, parental leave…) you were still around to keep the ship a float – both on a social level and a professional level. **Mona** and **Ebba:** we have collaborated on projects that have been essential to my growth as a researcher and person! **Britt:** your dark and nitty sense of humor can really light up a day - nice paradox, huh?!

In the first years, I also shared room with two other lovely jailbirds - **Kahl** and **Josh:** sharing office with you in the first years was, well, "special"… In a good way (to the extent that I joined Lokalgruppen to avoid separation)!

Marcus L and **Ronald (again):** PhDs need not only formal supervisors, but also role-models who have "recently" (approx. 10 years) finished their PhD-studies and are now excelling as young and promising researchers – you guys are my go-to examples!

To my other colleagues at the department (past and present): thank you so much for all the lovely discussions and friendly encounters over the years! The corona-virus has made it painfully clear just how much I value these encounters. So, again, thanks: **Andreas, Annika L, Anna, Anton, Cecilia, Christer, Elin, Emma, Erik, Eva, Goncalo, Henry, Ingrid, Janna, Johan, Johanna, Jonas, Jörgen, Karin B, Kim, Kirsti, Laura, Leo, Linn, Maja, Matilda, Mathias, Olof, Pär , Robin, Sheila, Sven, Terje, Tommie, Örjan.** To those colleagues not mentioned: I did not really forget you, the circumstances (text limits) just did not allow me to have this memory reflected in performance!

Finally, yet most importantly, there are a number of individuals outside of academia whose support and influence over the years has been essential for me to arrive at this point in my life! **Magnus A:** for providing essential "life-lessons". **Members of HG IF:** not for having anything to do with me succeeding as a researcher (in that aspect your effect is rather the contrary!) but for providing long-lasting friendships. **Erik** and **Johan:** For fostering an appetite for competition that have helped me to succeed in academia, and for being Bröder (in blood and in the Stiko-Per sense). **Nieces, nephews, and in-laws of various kinds:** I could not have wished for better relatives, and I am so thankful to be part of such close-knit families! **Jeanette** and **Kjell:** Your support has been invaluable in so many ways; I am really at a loss of words here. **Mom** and **Dad:** I have been extremely fortunate to have parents that have always helped me accomplish my dreams and goals. I also suspect that you "nudged" me towards this career path, so thanks for that!

Simone and **Joline:** "Min bästa" = Ni! Ni har också tvingat mig att tänka på annat än forskning och för det är jag så himla tacksam! ☺

Sofia: You, and your love, have constantly reminded me that the most important aspects of life cannot be captured by numerical description! You are the love of my life – BillMills 4ever!

Puh! Six years of encounters summed up to a couple of pages. Should we end the bugger with that? I say we do. Thank you for your time and attention!

References

Abdellaoui, M., Bleichrodt, H., & l'Haridon, O. (2008). A tractable method to measure utility and loss aversion under prospect theory. *Journal of Risk and Uncertainty*, *36*(3), 245.

Anderson, N. H., & Shanteau, J. C. (1970). Information integration in risky decision making. *Journal of Experimental Psychology, 84*(3), 441

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Anderson, N. H. (2013). Unified psychology based on three laws of information integration. *Review of General Psychology, 17*(2), 125-132.

Anderson, N. H. (1996). *A functional theory of cognition*. Psychology Press.

Anderson, J. R. (1990). *The adaptive character of thought*. Psychology Press.

Bach, D. R., & Dolan, R. J. (2012). Knowing how much you don't know: a neural organization of uncertainty estimates. *Nature Reviews Neuroscience, 13*(8), 572-586.

Barberis, N., Huang, M., & Santos, T. (2001). Prospect theory and asset prices. *The Quarterly Journal of Economics, 116*(1), 1-53.

Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations Research*, *33*(1), 1-27.

Bentham, J. (1789). *An introduction to the principles of morals*. London: Athlone.

Berg, N., & Gigerenzer, G. (2010). As-if behavioral economics: Neoclassical economics in disguise? *History of Economic Ideas*, pp. 133-165.

Bernoulli, D. 1738/1954. *Exposition of a new theory on the measurement of risk*, translated by Louise Sommer (pp. 23-36).

Bernoulli, N. (1709). Specimina artis conjectandi ad quaestiones juris adplicatae. *Acta Eruditorum*, pp. 159-70.

Birnbaum, M. H. (1973). The devil rides again: Correlation as an index of fit. *Psychological Bulletin, 79*(4), p. 239.

Birnbaum, M. H. (1999). How to show that 9> 221: Collect judgments in a between-subjects design. *Psychological Methods, 4*(3), 243.

Birnbaum, M. H. (2008). New paradoxes of risky decision making. *Psychological Review, 115*(2), 463.

Birnbaum, M. H., Coffey, G., Mellers, B. A., & Weiss, R. (1992). Utility measurement: Configural-weight theory and the judge's point of view. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 331.

Björkman, M. (1967). Stimulus—event learning and event learning as concurrent processes. *Organizational Behavior and Human Performance, 2*(3), 219-236.

Blavatskyy, P. R. (2007). Stochastic expected utility theory. *Journal of Risk and Uncertainty, 34*(3), 259-286.

Blum, K., Braverman, E. R., Holder, J. M., Lubar, J. F., Monastra, V. J., Miller, D., ... & Comings, D. E. (2000). The reward deficiency syndrome: a biogenetic model for the diagnosis and treatment of impulsive, addictive and compulsive behaviors. *Journal of Psychoactive Drugs, 32,* pp. 1-112.

Bossaerts, P., & Murawski, C. (2017). Computational complexity and human decision-making. *Trends in Cognitive Sciences*, *21*(12), 917-929.

Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review, 10*(1), 3-47.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: making choices without trade-offs. *Psychological Review, 113*(2), 409.

Brehmer, A., & Brehmer, B. (1988). From thirty years of policy capturing. *Human judgment: The SJT view*, p. 75.

Brehmer, B. (1994). The psychology of linear judgement models. *Acta Psychologica, 87*(2-3), 137-154.

Broniatowski, D. A., & Reyna, V. F. (2017). A formal model of fuzzy-trace theory: Variations on framing effects and the Allais Paradox. *Decision*. Advance online publication.

Brunswik, E. (1952). The conceptual framework of psychology. *Psychological Bulletin, 49*(6), 654-656.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Univ of California Press.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100*(3), 432.

Busemeyer, J. R., Johnson, J. G., & Jessup, R. K. (2006). Preferences constructed from dynamic micro-processing mechanisms. *The Construction of Preference*, pp. 220-234.

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.

Camerer, C. F., Loewenstein, G., & Rabin, M. (Eds.). (2011). *Advances in behavioral economics*. Princeton university press.

Cardano, G. (1663). Opus novum de proportionibus. In *idem, Opera omnia* (C. Sponi, Ed.) (10 vols.) (Vol. IV, pp. 463–603). Leiden.

Carlin, P. S. (1990). Is the Allais paradox robust to a seemingly trivial change of frame?. *Economics Letters, 34*(3), 241-244.

Carlson, B. W. (1990). Anchoring and adjustment in judgments under risk. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 665.

Cavazzuti, M. (2013). Robust design analysis. *Optimization Methods*, pp. 131-143). Springer, Berlin, Heidelberg.

Cheung, J. H., Burns, D. K., Sinclair, R. R., & Sliter, M. (2017). Amazon Mechanical Turk in organizational psychology: An evaluation and practical recommendations. *Journal of Business and Psychology, 32*(4), 347-361.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT press.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making, 4,* pp. 20-33.

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1346), 1413-1420.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. OUP USA.

Descartes, R. (1664). *L'Homme: Treatise on Man*.
Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin, 130*(6), 959.
Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51*(4), 380.
Edwards, W. (1961). Behavioral decision theory. *Annual Review of Psychology, 12*(1), 473-498.
Edwards, W. (1962). Subjective probabilities inferred from decisions. *Psychological Review, 69*(2), 109.
Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgement and choice. *Annual Review of Psychology*, *32*(1), 53-88.
Epley, N., & Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter-provided anchors. *Psychological Science, 12*(5), 391-396.
Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science, 17*(4), 311-318.
Erev, I., Ert, E., Plonsky, O., Cohen, D., & Cohen, O. (2017). From anomalies to forecasts: Toward a descriptive model of decisions under risk, under ambiguity, and from experience. *Psychological Review, 124*(4), 369.
Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology.*, *59*, 255-278.
Eysenck, M. W., & Keane, M. T. (1990). *Cognitive psychology: A student's handbook*, Hove. East Sussex: Lawrence Erlbaum Associates Publishers.
Fan, Y., Budescu, D. V., & Diecidue, E. (2019). Decisions with compound lotteries. *Decision*, *6*(2), 109.
Farashahi, S., Donahue, C. H., Hayden, B. Y., Lee, D., & Soltani, A. (2019). Flexible combination of reward information across primates. *Nature Human Behaviour, 3*(11), 1215-1224.
Fellner, W. (1961). Distortion of subjective probabilities as a reaction to uncertainty. *The Quarterly Journal of Economics*, pp. 670-689.
Fishburn, P. C. (1978). A probabilistic expected utility theory of risky binary choices. *International Economic Review*, pp. 633-646.
Fox, C. R., Erner, C., & Walters, D. J. (2015). Decision under risk: From the field to the laboratory and back. *The Wiley Blackwell handbook of judgment and decision making, 2,* 41-88.
Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances, 3*(10), e1701381.
Galanter, E., & Pliner, P. (1974). Cross-modality matching of money against other continua. In *Sensation and measurement*, pp. 65-76., Springer, Dordrecht.
Ganzach, Y. (1996). Preference reversals in equal-probability gambles: a case for anchoring and adjustment. *Journal of Behavioral Decision Making,* 9(2), 95-109.
Garcia-Retamero, R., Sobkow, A., Petrova, D., Garrido, D., & Traczyk, J. (2019). Numeracy and Risk Literacy: What Have We Learned so Far?. *The Spanish Journal of Psychology*, p. 22.
Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review, 103*(4), 650.
Gigerenzer, G., & Selten, R. (2001). Rethinking rationality. *Bounded rationality: The adaptive toolbox*, 1, 12.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1*(1), 107-143.
Gigerenzer, G. E., Hertwig, R. E., & Pachur, T. E. (2011). *Heuristics: The foundations of adaptive behavior*. Oxford University Press.
Gigerenzer, G., Swijtink, Z., Porter, T., & Daston, L. (1990). *The empire of chance: How probability changed science and everyday life*. Cambridge University Press.
Glimcher, P., & Fehr, E. (2014). *Neuroeconomics: Decisions and the brain.*
Glöckner, A., & Betsch, T. (2008). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1055.
Glöckner, A. (2007). Does intuition beat fast and frugal heuristics? A systematic empirical analysis. *Intuition in judgment and decision making*, pp. 309-325.
Gosling, S. D., & Mason, W. (2015). Internet research in psychology. *Annual Review of Psychology*, p. 66.
Hammond, K. R., & Stewart, T. R. (Eds.). (2001). *The essential Brunswik: Beginnings, explications, applications*. Oxford University Press.
Hammond, K. R. (1955). Probabilistic functioning and the clinical method. *Psychological Review, 62*(4), 255.
Hammond, K. R. (1999). Mats Björkman and Swedish studies of judgment and decision making. In *Judgment and decision making: Neo-Brunswikian and process-tracing approaches*, 305-320.
Hammond, K. R., Hursch, C. J., & Todd, F. J. (1964). Analyzing the components of clinical inference. *Psychological Review, 71*(6), 438.
Hansson, P., Juslin, P., & Winman, A. (2008). The role of short-term memory capacity and task experience for overconfidence in judgment under uncertainty. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1027.
Hauser, D. J., & Schwarz, N. (2016). Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods, 48*(1), 400-407.
Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. Perspectives on *Psychological Science, 12*(6), 973-986.
Hertwig, R., Pleskac, T. J., & Pachur, T. (2019). *Taming uncertainty*. Mit Press.
Hey, J. D., Morone, A., & Schmidt, U. (2009). Noise and bias in eliciting preferences. *Journal of Risk and Uncertainty*, *39*(3), 213-235.
Hoffman, P. J. (1960). The paramorphic representation of clinical judgment. Psychological bulletin, 57(2), 116.
Hoffmann, J. A., von Helversen, B., & Rieskamp, J. (2014). Pillars of judgment: How memory abilities affect performance in rule-based and exemplar-based judgments. *Journal of Experimental Psychology: General, 143*(6), 2242.
Hoffmann, J. A., von Helversen, B., & Rieskamp, J. (2016). Similar task features shape judgment and categorization processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(8), 1193.
Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive psychology, 24*(1), 1-55.
Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: a review and theoretical analysis. *Psychological Bulletin, 125*(5), 576.
Huygens, C. (1657). *De ratiociniis in ludo aleae. (On reckoning at Games of Chance).* T. Woodward, London.

James, W. (1890). *The principles of psychology*. London: Macmillan.
JASP team. (2019). JASP (version 0.10. 2)[computer software]. See https://jasp-stats.org.
Jervis, R. (1992). Political implications of loss aversion. *Political Psychology*, 187-204.
Joag, S. G., Mowen, J. C., & Gentry, J. W. (1990). Risk perception in a simulated industrial purchasing task: The Effects of single versus multi-play decisions. *Journal of Behavioral Decision Making, 3*(2), 91-108.
Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of endowment: a query theory of value construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 461.
Juslin, P., Karlsson, L., & Olsson, H. (2008). Information integration in multiple cue judgment: A division of labor hypothesis. *Cognition, 106*(1), 259-298.
Juslin, P., Nilsson, H., Winman, A., & Lindskog, M. (2011). Reducing cognitive biases in probabilistic reasoning by the use of logarithm formats. *Cognition, 120*(2), 248-267.
Juslin, P., Nilsson, H., & Winman, A. (2009). Probability theory, not the very guide of life. *Psychological Review, 116*(4), 856.
Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment*, *49*, 81.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-291.
Kahneman, D. (2011). Thinking, fast and slow. Macmillan.
Kant, I. (1781). *Critique of pure reason*. New York: St. Martins.
Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin, 134*(3), 404.
Karlsson, L., Juslin, P., & Olsson, H. (2007). Adaptive changes between cue abstraction and exemplar memory in a multiple-cue judgment task with continuous cues. *Psychonomic Bulletin & Review, 14*(6), 1140-1146.
Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge university press.
Keynes, J. M. (1921/2013). *A treatise on probability*. Courier Corporation.
Krische, S. D. (2019). Investment Experience, Financial Literacy, and Investment-Related Judgments. *Contemporary Accounting Research*, *36*(3), 1634-1668.
Lembcke, T. B., Engelbrecht, N., Brendel, A. B., & Kolbe, L. (2019). To nudge or not to nudge: ethical considerations of digital nudging based on its behavioral economics roots. *ECIS Proceedings*, pp. 1-17).
Levin, J. A. (1981). Estimation techniques for arithmetic: Everyday math and mathematics instruction. *Educational Studies in Mathematics, 12*(4), 421-434.
Levy, J. S. (1996). Loss aversion, framing, and bargaining: The implications of prospect theory for international conflict. *International Political Science Review, 17*(2), 179-195.
Li, Y., Vanni-Mercier, G., Isnard, J., Mauguière, F., & Dreher, J. C. (2016). The neural dynamics of reward value and risk coding in the human orbitofrontal cortex. *Brain, 139*(4), 1295-1309.
Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology, 89*(1), 46.
Lichtenstein, S., & Slovic, P. (Eds.). (2006). *The construction of preference*. Cambridge University Press.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*(2), 267.
Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, *92*(368), 805-824.
Lopes, L. L., & Ekberg, P. H. S. (1980). Test of an ordering hypothesis in risky decision making. *Acta Psychologica, 45*(1-3), 161-167.
Lopes, L. L. (1976). Model-based decision and inference in stud poker. *Journal of Experimental Psychology: General, 105*(3), 217.
Lopes, L. L. (1982). *Toward a procedural theory of judgment*. Wisconsin Human Information Processing Program Madison.
Lopes, L. L. (1985). Averaging rules and adjustment processes in Bayesian inference. *Bulletin of the Psychonomic Society, 23*(6), 509-512.
Lopes, L. L. (1995). Algebra and process in the modeling of risky choice. *Psychology of Learning and Motivation*, *32*, 177-220.
Lovallo, D., & Sibony, O. (2010). The case for behavioral strategy. *McKinsey Quarterly, 2*(1), 30-43.
Luce, R. D. (2000). *Utility of gains and losses: Measurement-theoretical and experimental approaches*. Psychology Press.
Lusardi, A., Mitchell, (2009). *Financial literacy and financial sophistication among older Americans*. National Bureau of Economic Research.
Luhrmann, M., Serra-Garcia, M., & Winter, J. (2018). The impact of financial education on adolescents' intertemporal choices. American Economic Journal: *Economic Policy, 10*(3), 309-32.
Mallpress, D. E., Fawcett, T. W., Houston, A. I., & McNamara, J. M. (2015). Risk attitudes in a changing environment: An evolutionary model of the fourfold pattern of risk preferences. *Psychological Review, 122(*2), 364.
Markowitz, H. (1952). The utility of wealth. *Journal of Political Economy, 60*(2), 151-158.
Marr, D. (1982). *Vision: The philosophy and the approach*. Freeman.
Marshall, A. (1890). *Principles of Economics*. London, Mcmillan.
Mata, R., Frey, R., Richter, D., Schupp, J., & Hertwig, R. (2018). Risk preference: A view from psychology. *Journal of Economic Perspectives, 32*(2), 155-72.
McCrink, K., & Spelke, E. S. (2016). Non-symbolic division in childhood. *Journal of Experimental Child Psychology*, *142*, 66-82.
McDermott, R., Fowler, J. H., & Smirnov, O. (2008). On the evolutionary origin of prospect theory preferences. *The Journal of Politics, 70*(2), 335-350.
Mellers, B. A., & Chang, S. J. (1994). Representations of risk judgments. *Organizational Behavior and Human Decision Processes, 57*, 167-167.
Mellers, B. A., Chang, S. J., Birnbaum, M. H., & Ordonez, L. D. (1992). Preferences, prices, and ratings in risky decision making. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 347.
Mellers, B., Weiss, R., & Birnbaum, M. (1992). Violations of dominance in pricing judgments. *Journal of Risk and Incertainty, 5*(1), 73-90.
Melnikoff, D. E., & Bargh, J. A. (2018). The mythical number two. *Trends in Cognitive Sciences*, *22*(4), 280-293.
Mendell, H. (2017). Aristotle and mathematics. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), available online at: https://plato.stanford.edu/archives/fall2019/entries/aristotle-mathematics/

Meertens, R. M., & Lion, R. (2008). Measuring an individual's tendency to take risks: The risk propensity scale. *Journal of Applied Social Psychology, 38*(6), 1506-1520.
Merton, R. K. (1948). The self-fulfilling prophecy. *The Antioch Review, 8*(2), 193-210.
Mill, J. S. (1861). *Utilitarianism*. Hackett Publishin Co. Indianapolis.
Millroth, P., & Juslin, P. (2015). Prospect evaluation as a function of numeracy and probability denominator. *Cognition, 138*, 1-9.
Millroth, P., Juslin, P., Winman, A., Nilsson, H., & Lindskog, M. Preference or ability: Exploring the relations between risk preference, personality, and cognitive abilities. *Journal of Behavioral Decision Making*.
Minsky, M. (1960). Steps toward artificial intelligence. Republished in 2006. *AI Magazine, 27*(4), 109.
Minsky, M. (2006). *The emotion machine: commonsense thinking, artificial intelligence, and the future of the human mind*. Simon & Schuster.
Mischel, W. (2005). Alternative futures for our science. *APS Observer, 18*(3).
Mischel, W. (2008). The toothbrush problem. *APS Observer, 21*(11).
Mishra, S., Barclay, P., & Sparks, A. (2017). The relative state model: Integrating need-based and ability-based pathways to risk-taking. *Personality and Social Psychology Review, 21*(2), 176-198.
Montgomery, H., & Adelbratt, T. (1982). Gambling decisions and information about expected value. *Organizational Behavior and Human Performance*, *29*(1), 39-57.
Morey, R. D., Rouder, J. N., & Jamil, T. (2015). BayesFactor: Computation of Bayes factors for common designs. R package version 0.9. 12-2.
Mudrik, L., Faivre, N., & Koch, C. (2014). Information integration without awareness. *Trends in Cognitive Sciences, 18*(9), 488-496.
Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs,
O'Doherty, J. P., Cockburn, J., & Pauli, W. M. (2017). Learning, reward, and decision making. *Annual Review of Psychology, 68*, 73-100.
O'Neill, M., & Schultz, W. (2010). Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron, 68*(4), 789-800.
Oppenheimer, D. M., & Kelso, E. (2015). Information processing as a paradigm for decision making. *Annual Review of Psychology, 66*, 277-294.
Pachur, T., & Olsson, H. (2012). Type of learning task impacts performance and strategy selection in decision making. *Cognitive Psychology, 65*(2), 207-240.
Pachur, T., Suter, R. S., & Hertwig, R. (2017). How the twain can meet: Prospect theory and models of heuristics in risky choice. *Cognitive Psychology, 93*, 44-73.
Parducci, A. (1965). Category judgment: a range-frequency model. *Psychological Review, 72*(6), 407.
Pascal, B. (1654/1991). Treatise on the arithmetical triangle. In *Great Books of the Western*.
Pavlov, I. P. (1927). *Conditioned reflexes.* London: Oxford.
Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 534.
Payne, J. W., Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge university press.
Platzer, C., & Bröder, A. (2013). When the rule is ruled out: Exemplars and rules in decisions from memory. *Journal of Behavioral Decision Making, 26*(5), 429-441.
Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization, 3*(4), 323-343.

Ramsey, F. P. (1931). Theories. The foundations of mathematics, 212-236.
Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2011). Transitivity of preferences. *Psychological Review, 118*(1), 42.
Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and individual Differences, 7*(1), 1-75.
Reyna, V. F., & Brainerd, C. J. (2011). Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. *Developmental Review, 31*(2-3), 180-206.
Reyna, V. F. (2008). A theory of medical decision making and health: fuzzy trace theory. *Medical Decision Making, 28*(6), 850-865.
Rouder, J. N., & Morey, R. D. (2012). Default Bayes factors for model selection in regression. *Multivariate Behavioral Research, 47*(6), 877-903.
Savage, L. J. (1954). *The foundations of statistics*. Courier Corporation.
Schley, D. R., & Peters, E. (2014). Assessing "economic value" symbolic-number mappings predict risky and riskless valuations. *Psychological Science, 25*(3), 753-761.
Schlottmann, A., & Anderson, N. H. (1994). Children's judgments of expected value. *Developmental Psychology, 30*(1), 56.
Schlottmann, A. (2001). Children's probability intuitions: Understanding the expected value of complex gambles. *Child Development, 72*(1), 103-122.
Schmeidler, D. (1989). Subjective probability and expected utility without additivity. Econometrica: *Journal of the Econometric Society*, 571-587.
Schoemaker, P. J. (1979). The role of statistical knowledge in gambling decisions: Moment vs risk dimension approaches. *Organizational Behavior and Human Performance*, *24*(1), 1-17.
Schoemaker, P. J. (1982). The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature*, 529-563.
Schurz, G., & Hertwig, R. (2019). Cognitive success: A consequentialist account of rationality in cognition. *Topics in Cognitive Science, 11*(1), 7-36.
Schwartz, A. (2009). *Information integration theory*. In M. W. Kattan (Ed.), Encyclopedia of medical decision making (pp. 622-622). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412971980.n18
Shanteau, J. C. (1970). An additive model for sequential decision making. *Journal of Experimental Psychology, 85*(2), 181.
Shanteau, J. (1974). Component processes in risky decision making. *Journal of Experimental Psychology, 103*(4), 680.
Shanteau, J. (1975). An information integration analysis of risky decision making. *Human Judgement and Decision Processes*, 109-137.
Shanteau, J. (1980). *The concept of weight in judgment and decision making: A review and some unifying proposal*s. Colorado University at Boulder Center for Research on Judgment and Policy.
Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62*(4), 273-296.
Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69*(1), 99-118.
Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*(2), 129.
Simon, H. A. (1976). From substantive to procedural rationality. *25 years of economic theory* (pp. 65-86). Springer, Boston, MA.

Simon, H. A. (1991). Bounded rationality and organizational learning. *Organization Science, 2*(1), 125-134.
Simon, H. A. (2000). Bounded rationality in social science: Today and tomorrow. *Mind & Society, 1*(1), 25-39.
Simonsohn, U. (2014). Citing Prospect Theory. Available at http://datacolada.org/15.
Sjöberg, L. (1968). Studies of the rated favorableness of offers to gamble. *Scandinavian Journal of Psychology, 9*(1), 257-273.
Skinner, B. F. (1938). The Behavior of Organisms: An Experimental Analysis. New York: Appleton-Century
Slovic, P., & Lichtenstein, S. (1968). Relative importance of probabilities and payoffs in risk taking. *Journal of Experimental Åsychology, 78*(3p2), 1.
Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance, 6*(6), 649-744.
Sobkow, A., Fulawka, K., Tomczak, P., Zjawiony, P., & Traczyk, J. (2019). Does mental number line training work? The effects of cognitive training on real-life mathematics, numeracy, and decision making. *Journal of Experimental Psychology: Applied*.
Sokolowska, J., & Pohorille, A. (2000). Models of risk and choice: challenge or danger. *Acta Psychologica, 104*(3), 339-369.
Stewart, N. (2011). Information integration in risky choice: Identification and stability. *Frontiers in Psychology, 2*, 301.
Stewart, N. (2009). Decision by sampling: The role of the decision environment in risky choice. *The Quarterly Journal of Experimental Psychology, 62*(6), 1041-1062.
Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. Cognitive psychology, 53(1), 1-26.
Stewart, N., Reimers, S., & Harris, A. J. (2015). On the origin of utility, weighting, and discounting functions: How they get their shapes and how to change their shapes. *Management Science, 61*(3), 687-705.
Sundh, J., & Juslin, P. (2018). Compound risk judgment in tasks with both idiosyncratic and systematic risk: The "Robust Beauty" of additive probability integration. *Cognition, 171*, 25-41.
Sundh, J., Millroth, P., Collsiöö, A., & Juslin, P. (2020). Precise/Not Precise (PNP): A Brunswikian model that relies on the judgment error distributions to identify the cognitive processes. *Manuscript in revision.*
Svenson, O. (1983). Decision rules and information processing in decision making. *Human decision making*, 131-162.
Taleb, N. N. (2018). *Skin in the game: Hidden asymmetries in daily life*. Random House Trade Paperbacks.
Thorngate, W. (1980). Efficient decision heuristics. *Behavioral Science, 25*(3), 219-225.
Tolman, E. C., & Brunswik, E. (1935). The organism and the causal texture of the environment. *Psychological Review, 42*(1), 43.
Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55*(4), 189.
Traczyk, J., & Fulawka, K. (2016). Numeracy moderates the influence of task-irrelevant affect on probability weighting. *Cognition, 151*, 37-41.
Traczyk, J., Sobkow, A., Fulawka, K., Kus, J., Petrova, D., & Garcia-Retamero, R. (2018). Numerate decision makers don't use more effortful strategies unless it pays:

A process tracing investigation of skilled and adaptive strategy selection in risky decision making. *Judgment and Decision Making, 13*, 372–381.
Tversky, A. (1967). Utility theory and additivity analysis of risky choices. *Journal of Experimental Psychology, 75*(1), 27.
Tversky, A. (1969). Intransitivity of preferences. *Psychological Review, 76*(1), 31.
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. *Science, 185*(4157), 1124-1131.
Tversky, A., & Kahneman, D. (1975). *Judgment under uncertainty: Heuristics and biases* (pp. 141-162). Springer Netherlands.
Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297-323.
Tversky, A., & Thaler, R. H. (1990). Anomalies: preference reversals. *Journal of Economic Perspectives, 4*(2), 201-211.
Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity.* Cambridge university press.
Weber, E. U., Blais, A. R., & Betz, N. E. (2002). A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making, 15*(4), 263-290.
Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual Review of Psychology*, *60*, 53-85.
Viegas, R. G., Oliveira, A. M., Garriga-Trillo, A., & Grieco, A. (2012). A functional model for the integration of gains and losses under risk: Implications for the measurement of subjective value. *Psicológica, 33*(3), 711-733.
Wiggins, N., & Hoffman, P. J. (1968). Three models of clinical judgment. *Journal of Abnormal Psychology, 73*(1), 70.
Vlaev, I., Chater, N., Stewart, N., & Brown, G. D. (2011). Does the brain calculate value?. *Trends in Cognitive Sciences, 15*(11), 546-554.
von Helversen, B., & Rieskamp, J. (2008). The mapping model: A cognitive theory of quantitative estimation. *Journal of Experimental Psychology: General, 137*(1), 73.
von Helversen, B., & Rieskamp, J. (2009). Models of quantitative estimations: Rule-based and exemplar-based processes compared. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(4), 867.
Neumann, J. V., & Morgenstern, O. (1944/2007). *Theory of games*. J. Wiley.
Wagenmakers, E. J. (2007a). A practical solution to the pervasive problems of p-values. *Psychonomic Bulletin & Review*, *14*(5), 779-804.
Wagenmakers, E.J. (2007b). Corrigendum for "A Practical Solution to the Pervasive Problems of p Values", Psychonomic Bulletin & Review, 14, 779–804. Available at: http://www.ejwagenmakers.com/2007/CorrigendumPvalues.pdf
Wundt, W. M. (1907). *Outlines of psychology*. London: Swan Sonnenschein
Wundt, W. (1904). *Principles of physiological psychology*. London: Swan Sonnenschein.
Yarosh, H. L., Hyatt, C. J., Meda, S. A., Jiantonio-Kelly, R., Potenza, M. N., Assaf, M., & Pearlson, G. D. (2014). Relationships between reward sensitivity, risk-taking and family history of alcoholism during an interactive competitive fMRI task. *PloS one, 9*(2).

Zhang, D. C., Highhouse, S., & Nye, C. D. (2019). Development and validation of the General Risk Propensity Scale (GRiPS). *Journal of Behavioral Decision Making*, *32*(2), 152-167.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 178

Editor: The Dean of the Faculty of Social Sciences

A doctoral dissertation from the Faculty of Social Sciences, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences".)

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2020

Distribution: publications.uu.se
urn:nbn:se:uu:diva-407865