

*Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Social Sciences 243*

Subjective value and preference

MATTIAS FORSGREN



ACTA UNIVERSITATIS
UPSALIENSIS
2026



UPPSALA
UNIVERSITET

Dissertation presented at Uppsala University to be publicly examined in Sal IX, Universitetshuset, Biskopsgatan 3, Friday, 12 June 2026 at 13:30 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Neil Stewart (Warwick Business School, The University of Warwick).

Abstract

Forsgren, M. 2026. Subjective value and preference. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 243. 121 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-2849-2.

Classical social science has often assumed "existing" preferences. These are informed by deeply held beliefs of subjective value and consequently relatively stable and logically coherent. In this view, people know what they like, and this is reflected in their preferences. Psychologists have challenged that account and instead suggested that many preferences are "constructed" in the moment. Such preferences are strongly influenced by the immediate context, making them unstable and potentially incoherent. In this view, people do not know what they like and need to figure it out on the spot. The purpose of this thesis is to critically examine the generality of the constructed preferences view by revisiting phenomena taken to support constructive theories. Study I tested the Decision by Sampling theory which postulates that subjective values of quantitative attributes (e.g. money) are constructed in the moment by ordinal comparisons between the target and a small sample of exemplars from long-term memory and recent experiences. If so, manipulating the distribution of recently experienced attribute values should shift people's subjective values, as revealed by pairwise choices. However, after addressing previous methodological limitations, we found evidence against that prediction. In Study II, we again found no effect of the same manipulation on pairwise choices, only on listwise choice. This effect was weaker for a group with substantial domain expertise. The effect direction is consistent with a previous explanation where the cognitive mechanism described by Decision by Sampling affects the interpretation of numerical information rather than subjective values per se. Study III investigated a criterion of logical coherence of preferences: "transitivity", which implies a mental ranking of options. Preferences for everyday objects like political parties and confectioneries, but also monetary gambles, were near-universally strongly stochastically transitive. That is, participants behaved as-if they knew what they liked but the subjective values were represented with noise. We only estimated a clearly non-zero prevalence of violations of transitivity for monetary gambles designed specifically to be difficult to choose between. To the extent that the scrutinised phenomena were taken to support constructed preferences, that may have been premature. Existing preferences may be more prevalent than commonly recognised.

Keywords: Subjective value, Preference, Decision by Sampling, Transitivity

Mattias Forsgren, Department of Psychology, Box 1225, Uppsala University, SE-75142 Uppsala, Sweden.

© Mattias Forsgren 2026

ISSN 1652-9030

ISBN 978-91-513-2849-2

URN urn:nbn:se:uu:diva-582031 (<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-582031>)

*Till mormor och morfar,
farmor och farfar,
mamma, pappa,
Sanna*

List of papers

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

- I Forsgren, M., Frimanson, L., & Juslin, P. (2025). A Preregistered Falsification Test of the Decision by Sampling Model and Rank-Order Effect. *Management Science*, *71*(10), 8218-8229.
<https://doi.org/10.1287/mnsc.2022.03611>
- II Forsgren, M., Frimanson, L., & Juslin, P. (2025). Preceding Options Affect Subsequent Listwise but Not Pairwise Choice, Even for Experts. *Journal of Behavioral Decision Making*, *38*(3), Article e70019.
<https://doi.org/10.1002/bdm.70019>
- III Forsgren, M., Karreskog Rehbinder, G., & Juslin, P. (2026). Preferences for Everyday Objects are Strongly Stochastically Transitive. *Manuscript*.

Reprints were made with permission from the publishers.

For Studies I and II, Mattias Forsgren contributed to the conceptualisation, methodology, and reviewing and editing of the original drafts. He held primary responsibility for writing the original drafts, coding, validation of code, formal analysis, data collection, data curation, visualisation, and project administration. For Study III, Mattias Forsgren held similar responsibilities except that Gustav Karreskog Rehbinder held primary responsibility for coding, validation of code, data collection and data curation. Mattias Forsgren assisted him in these areas but could never have done it without him. All co-authors made invaluable contributions to the projects.

Contents

1	Introduction	13
2	Background	16
2.1	Subjective value	16
2.1.1	Historical conceptions of subjective value	16
2.1.2	From subjective value to choice - psychophysical origins	18
2.1.3	Utility functions as descriptors of risk attitudes	19
2.1.4	Subjective value in contemporary psychological research	19
2.1.5	Decision by Sampling	22
2.2	Preference	24
2.2.1	Existing preferences	24
2.2.2	Constructed preferences	25
2.2.3	Constructivism as a spectrum	30
2.2.4	Transitivity of preferences	31
2.3	Summary of literature review and purpose of this thesis	36
3	Aims of the empirical section	38
4	Empirical work	40
4.1	Methods common to Study I and II	40
4.1.1	Design	41
4.1.2	Procedure	41
4.2	Study I	43
4.2.1	Background	43
4.2.2	Methods	44
4.2.3	Results	45
4.2.4	Discussion	45
4.3	Study II	49
4.3.1	Background	49
4.3.2	General methods	50
4.3.3	Experiment 1	50
4.3.4	Experiment 2	54
4.3.5	Experiment 3	56
4.3.6	Experiment 4	57
4.3.7	General discussion	60

4.4	Study III	60
4.4.1	Background	60
4.4.2	General methods	63
4.4.3	Experiment 1	67
4.4.4	Experiment 2	70
4.4.5	Experiment 3	71
4.4.6	Meta-analysis	78
4.4.7	General discussion	78
5	General Discussion	80
5.1	Summary and conclusions	80
5.2	Interpretation of results in Study I and II	81
5.3	Whither now for Decision by Sampling?	82
5.4	Conceptual preferences	84
5.5	SST and the evidence accumulation framework	85
5.6	Transitivity but not completeness of preference?	88
5.7	Great expectations: the promise of the mathematical psychologist's approach	89
5.8	Limitations	90
5.8.1	Psychological interpretation of expertise	90
5.8.2	Confounding of frequency, range, and mean	91
5.8.3	Limitations of average treatment effects	91
5.8.4	Two-alternative forced choice	92
5.8.5	Representative design	92
5.8.6	Model discriminability	93
5.9	Concluding remark	93
6	Acknowledgments	94

Abbreviations

EUT – Expected utility theory.

vmPFC – Ventromedial prefrontal cortex.

OFC – Orbitofrontal cortex.

DbS – Decision by Sampling.

IIA – Independence from irrelevant alternatives.

TAX – Transfer of Attention Exchange.

SST – Strong stochastic transitivity.

2AFC – Two-alternative forced choice.

AMT – Amazon Mechanical Turk.

GLMM – Generalised linear mixed model.

MMTP – Mixture model of transitive preferences.

MST – Moderate stochastic transitivity.

WST – Weak stochastic transitivity.

FMM – Finite mixture model.

UK – United Kingdom.

USA – United States of America.

EEG – Electroencephalographic/electroencephalogram.

The White Rabbit put on his spectacles. "Where shall I begin, please your Majesty?" he asked. "Begin at the beginning," the King said gravely, "and go on till you come to the end: then stop."

*– Lewis Carroll, Alice's Adventures in Wonderland,
Chapter XII: Alice's evidence*

1. Introduction

I give a series of four lectures to a group of exchange students at the start of every semester. While exploring Uppsala, my students might discover that a caffè latte from an unassuming Swedish coffee shop will set them back about 40 SEK - around €3.60. At this point, they will have some immediate impression of *subjective value*: €3.60 for a caffè latte is daylight robbery, or a bargain, or something in between. At the start of the first lecture, I ask each student to state whether they are a cat or a dog person - a matter of *preference*. Such subjective values and preferences may be revealed by the students acting to select a caffè latte, or indeed a cat, if one is offered as an option. That is, *subjective values* and *preferences* may support *choices*. The distinction between these three related concepts is the vantage point of the present thesis. But what cognitive process underlies that immediate impression of the subjective value of €3.60? And are our preferences for pets, hot drinks, and other everyday options logically coherent or self-contradictory?

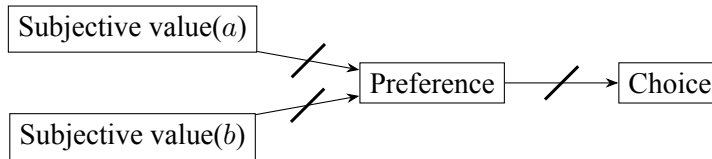
In this thesis, *subjective value* refers to a mental representation of the value of a single option, attribute, or set of attributes; we may speak of the subjective value of a caffè latte in isolation. Let us take *preference* to refer to covert, mental representations of which options are preferred over others. Unlike subjective value, a preference is a relation between at least two options. For example, a caffè latte is preferred to a cup of tea. In the sense I will use that word, we cannot speak of some preference for an option in isolation.¹ Having used the term preference for the covert relation, let us take *choice* to refer to the overt behaviour of choosing an option from some set. In this view, choices can only be said to “reveal” preferences (Richter, 1966) in so far as we have confidence in some (at least implicit) model linking overt choice behaviour to the covert cognitive representations that are the preferences. So, these three terms are distinct. People can make choices that do not reflect their preferences (Figure 1.1A). “Satisficing” (Artinger et al., 2022; Simon, 1955) is an example of a theory that suggests so: one identifies an option that is “good enough” without reference to the other options. People can also have preferences that do not reflect differences in subjective value (Figure 1.1B). “Multialternative decision field theory” (Roe et al., 2001) is an example, where aspects of the options are compared to each other to identify which is preferred but options are never valued in isolation. Alternatively, subjective values of options can support

¹The term preference, as used here, refers to which option is preferred over another. It does not say anything about how strongly the option is preferred. I will instead refer to that as *strength of preference*.

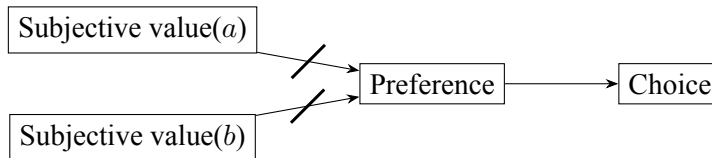
preferences over said options, and those preferences can then support overt choices (Figure 1.1C). The “Good based model” (Padoa-Schioppa, 2011) is an example, where a holistic valuation of each option is formed independently of the other options at hand.

Figure 1.1. Subjective value, preference and choice

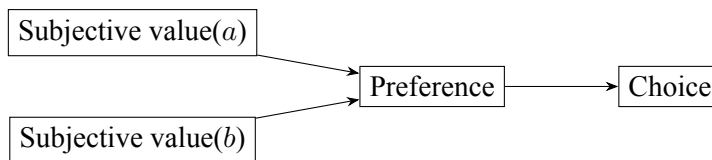
A



B



C



Note. Three possible relations between subjective value, preference and choice. Panel A: Choice is not based on preference or subjective value. Panel B: Choice is supported by preference, but preference is not based on subjective values of the options. Panel C: Choice is supported by preference which is supported by subjective values of options *a* and *b*.

Historically, research in social science has often assumed the latter version (Figure 1.1C). The subjective values exist, in some sense, prior to the choice scenario, ready to support a “clear intuition of preference” (Von Neumann & Morgenstern, 2004/1944, p. 17). Such preferences are typically presumed to be logically coherent. In addition, they are fairly stable over time and fairly insensitive to superficial aspects of the immediate environment that - by some normative standard - do not justify a change of mind. In contrast, a more recent research tradition within judgement and decision-making psychology has instead assumed that subjective values and preferences are constructed ad hoc (see Slovic, 1995). One view within this tradition is that there exist no deeply held subjective values to be revealed by choices; preferences are fickle things that are generated on the spot (Chater, 2018). Another view is that deeply held subjective values do exist but they are inaccessible to us because of lim-

itations to the information available and our limited capacity for self-control, attention, and computation (Sugden, 2009, p. 370). As a consequence of being constructed, subjective values and preferences are often conceived of as logically incoherent and unstable (McKenzie, 2018).

The purpose of this thesis is to critically examine the generality of the second research tradition by revisiting some claims relevant to the instability of subjective values and to the logical incoherence of preferences. Firstly, I discuss how we scrutinised a prominent theory - Decision by Sampling (Stewart et al., 2006) - of how subjective values are constructed (Study I) and whether the preference instability supposed by this theory is contingent on response format (Study II). Secondly, I discuss how we scrutinised the claim that preferences are incoherent - specifically, that they are “often” (American Psychological Association, 2025) intransitive (Study III). While these studies do not prove that the subjective values and preferences were *not* constructed, nor that subjective values and preferences are *never* constructed, they indicate that - at least in the fairly paradigmatic empirical examples considered here - such a process is not necessary to account for the results.

The thesis proceeds as follows. Beginning with subjective value, I briefly review its conceptual history within social science, its history in psychology (specifically, psychophysics), the idea of a subjective value function, and contemporary research in the field. I then introduce (i) the theory of Decision by Sampling (Stewart et al., 2006), which describes how subjective value is constructed, and (ii) the empirical paradigm that will be scrutinised. Moving to preferences, I review the historical debate on whether preferences are existing or constructed, and contemporary research connected with this debate. I then discuss a specific mathematical property of preferences - transitivity - that has been claimed to not hold, supposedly due to preferences being constructed. I proceed to review the empirical work conducted by my collaborators and I, and discuss its general implications for the debate between existing and constructed subjective values and preferences. I then conclude.

2. Background

2.1 Subjective value

Throughout this thesis, “subjective” should be taken to mean something that is experienced by a mind, while “objective” is something that is independent of any mind. Consequently, subjective value is that which is understood *with* reference to psychology. Objective value is that which is understood *without* reference to psychology. This usage is analogous to how a psychophysicist is interested in the mapping from (objective) physical quantities (light intensity in lumen) to (subjective) sensations (experienced light intensity).

2.1.1 Historical conceptions of subjective value

The necessity of a subjectivist account of value

The labour theory of value (A. Smith, 2025/1776a) suggests that the value of a commodity is determined by the labour required to produce it (Moscati, 2021). Note that this makes no reference to psychology. One can, in principle, determine something’s value by measuring a physical quantity like “working hours per unit”. Therefore, one cannot say that something has a labour theoretic value to me but a different labour theoretic value to you. In the present terminology: value is objective. There are, however, empirical problems with this account. Famously, it cannot explain why diamonds seem to be valued more than water in Paris while the opposite would be true for an (imagined) thirsty traveller traversing the Sahara (A. Smith, 2025/1776a). This is odd because water is essential to survival while diamonds fulfil a, less pressing, aesthetic and social need. So, value cannot be fully understood without making reference to a mind experiencing it: value is subjective.

When accepting the subjective nature of value, one could resolve the paradox of diamonds and water by assuming decreasing “marginal utility” (D. Bernoulli, 1954/1738): the increment in subjective value of some commodity was decreasing in the quantity of the commodity already held. When in the market for water, its subjective value depends on the current stock of water available to the individual. In Paris, water is abundant and the subjective value of any increment in its stock is trifling. In the Sahara, water is (supposedly) difficult to come by. The stock of water available to the imagined traveller is thus low and any increment to it would be of substantial subjective value. This resolved the paradox by assuming a particular functional form relating the quantity of a commodity to its subjective value.

In this thesis, I will use the term *subjective value function* for Bernoulli's idea of a functional relationship between a quantity of something and the subjective value of that quantity - for example, the different subjective values of different quantities of euros (€). This is analogous to how a psychophysical function relates a physical quantity on the objective scale (e.g. light intensity in lumen) to a perceived quantity on the psychological scale (e.g. perceived light intensity).

Five historical conceptions

Firstly, one view of subjective value has been that “there is no accounting for tastes”¹ (Montano, 1628). Already Hume (1987/1777, p. 229-230) suggested that this was seen as common sense. A second, alternative view has been that the subjective value of something reflects its “utility” - its usefulness in fulfilling some subjective need or desire of an individual (Bentham, 1789; Jevons, 1879, p. 52-53). In the first view, subjective values are valenced qualia (see Tye, 2021): reflections of what experiencing some stimulus feels like, and nothing more. In the second view, they are representational content (see Chalmers, 2010; Lycan, 2023): reflections of the needs-fulfilling properties of the stimulus. As a third view, economists have argued for the behaviourist (e.g. Watson, 1913) approach of blackboxing subjective value and focusing purely on observable behaviour (e.g. Fisher, 1892; Samuelson, 1938, see p. 11 and p. 61, respectively, in particular). This appears to have been a way of defending the social sciences as so-called special sciences (Fodor, 1974) that cannot be reduced to complex psychology (e.g. Schumpeter, 1984; M. Weber, 1975/1908).² A fourth view, also from economics (von Mises, 2009/1912, p. 46-47) envisioned subjective value as so deeply contextual that “there is no such thing as abstract value”. In that view, it only makes sense to speak of the subjective value of something relative to a specific alternative - which in the terminology used in this thesis would be a strength of preference rather than a subjective value. A fifth and final view, which has a long tradition in philosophy but also psychology, is “value pluralism” (Mason, 2023; Thomson, 1997). Here, there are different *kinds* of subjective value which cannot be compared in some “common currency” (Walasek & Brown, 2023), but may perhaps be prioritised relative to one another (Banfield, 1848; Maslow, 1943; Mill, 2009/1879).

¹While this is how we typically phrase the proverb in modern English, it is a slight corruption of the original Latin: “de gustibus non est disputandum; merito non valet actus”. In English: “About tastes is not to be debated; the act is not worth the merit.”

²Contrast this to methodological individualism (see Coleman, 1986; Schumpeter, 1909; Ylikoski, 2016) which argues that social behaviour can (and should) be understood through the behaviour of individuals, i.e. that social sciences indeed are reducible to complex psychology.

2.1.2 From subjective value to choice - psychophysical origins

“Weber’s law” (Fechner, 1860) describes the objective difference in intensity between two physical stimuli (e.g. two light intensities) required for it to be just-noticeable that the two stimuli are not the same (“just-noticeable difference”, see e.g. Grondin, 2010; Nash et al., 2016; Ross & Brodie, 1987). Fechner’s law states that our perception of a physical stimulus is proportional to the natural logarithm of its objective magnitude. So, Weber’s law introduced the notion of some limit on discriminative ability in perception while Fechner’s law introduced a concave psychophysical function (see Algom, 2021, for a review). Once articulated, this concave psychophysical function was taken as a justification for D. Bernoulli’s (1954/1738) previous assumption of a concave subjective value function (see M. Weber, 1975/1908, p. 27, who was critical of that assumption and argued that empirical subjective value functions should not be presumed to have any regularity too them).

Both Weber’s and Fechner’s law are deterministic in the sense that they always produce the same output for the same input.³ Thurstone (1927a, 1927b) was the first to explicitly model the process as stochastic in his Law of comparative judgement:

$$P(\phi_1 \succ \phi_2) = \frac{\phi_1 - \phi_2}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2r_{12}\sigma_1\sigma_2}} \quad (2.1)$$

where $P(\phi_1 \succ \phi_2)$ is the proportion of trials on which stimulus ϕ_1 is judged to be greater than stimulus ϕ_2 , σ_1^2 and σ_2^2 are the standard deviations of the perceptions of stimuli 1 and 2, respectively, and r_{12} is the correlation between σ_1^2 and σ_2^2 . That is, choice is deterministically guided by what stimulus is perceived to be more intense, but those perceptions are subject to normal noise (possibly with unequal variances). Because of the normality of the noise, the greater the objective difference in stimulus magnitude the less likely it is that the noise perturbs the perceptions enough for it to flip the choice. This idea was soon exported from psychophysics to the more cognitive domains of social attitudes (Thurstone & Chave, 1929) and aesthetic quality (Thurstone, 1954). Despite arguably being attributable to Thurstone (see Thurstone, 1931), these kinds of models have been called “Fechnerian” models (e.g. Blavatsky, 2018; Cavagnaro & Regenwetter, 2023; Davis-Stober et al., 2017).⁴

³But see also Dzharfarov and Colonius (2011) who argue that Fechner (1860) only assumed determinism to simplify his example - the “true” process was always thought to be probabilistic.

⁴Note how Equation 2.1 is compatible not only with Thurstone’s noisy representations, but with other sources of variability too. For example, Alós-Ferrer and Garagnani (2022a, 2022b) presented a gambling task where one option dominated the other on each trial. To identify which option that was, participants needed to perform some calculations (presumably in working memory). If those calculations were intuitive (Sundh et al., 2021) guesstimates, they would indeed yield data compatible with a Fechnerian model. Alternatively, the cognition proper could be deterministic and the noise purely attributable to the response stage.

2.1.3 Utility functions as descriptors of risk attitudes

The so-called Saint Petersburg paradox consists of the intuition that people would only be willing to pay a finite amount of money for the opportunity to play a game of dice which had an infinite expected value (N. Bernoulli, 1713). D. Bernoulli (1954/1738) introduced the idea of a concave subjective value function of outcomes to account for this: while a moderate amount of money had a tangibly greater subjective value than a small amount, adding a few zeros to an exceedingly large amount would not make much difference for the sum's subjective value. This resolved the paradox by assuming non-linear weighting of outcomes. However, it assumed linear weighting of probabilities. In the present terminology, the value of outcomes was subjective but the value of the probabilities of those outcomes was objective (see Section 2.1). This approach lives on in "utility" functions in economics, where the individual's "risk attitude" is subsumed in the curvature of what in our terminology is a single subjective value function of the outcome (see e.g. Schoemaker, 1982).

An alternative approach, more in line with the psychophysical tradition described above, is to claim that probabilities too are subjectively valued (e.g. Stewart et al., 2015). Whether a 90% probability of your side winning the next football derby is conceived of as a "good", "bad", or "alright" chance might differ, between and within people. This psychophysical approach will be the view taken for the rest of this thesis; when I use the term subjective value function, that might apply to any quantitative attribute and not just outcomes.

2.1.4 Subjective value in contemporary psychological research

The shape of the subjective value function

Research on the shapes of subjective value functions has often suggested and tested theoretical models (e.g. Birnbaum, 2008), with some notable, more atheoretical exceptions (e.g. Berlinghieri et al., 2023; Brandstätter & Brandstätter, 1996). Such models have often added complications to the simple subjective value function proposed by D. Bernoulli (1954/1738). Here, the perhaps most central proposition is cumulative prospect theory (Tversky & Kahneman, 1992). Building on Kahneman and Tversky (1979), it suggests a piecewise, S-shaped subjective value function with the junction point at "status quo" (no change in one's stock of subjective value). The curvature was supposed to be more severe below status quo than above it, which was described as "loss aversion". While there are empirical challenges to that model (e.g. Birnbaum, 2008; Birnbaum & Chavez, 1997; Peterson et al., 2021; Regenwetter et al., 2022) it has nevertheless become a prominent starting point for other research (e.g. Lopes & Oden, 1999; Zilker & Pachur, 2022; Zilker et al., 2020). Loomes and Sugden (1982) instead suggested that the subjective value of an outcome was determined by the difference between a pre-existing valuation of the outcome in isolation and the subjective value of the alternative outcomes on offer,

weighted by its probability. In other words, valuations of options are not driven only by what one has a chance of receiving but also by what one risks missing out on, why Loomes and Sugden (1982) named the model “regret theory”. This leads to a subjective value function with a momentarily shifting shape depending on what options are being considered. See Starmer (2000) for an overview.

Neural coding of subjective value

A prominent hypothesis in cognitive neuropsychology has been that subjective value is coded in the ventromedial prefrontal cortex (vmPFC) and orbitofrontal cortex (OFC, e.g. Barretto-Garcia et al., 2025; Chib et al., 2009; Clithero & Rangel, 2014; De Martino & Cortese, 2023; Levy & Glimcher, 2012; Padoa-Schioppa & Assad, 2006, 2008). In recent years, that hypothesis has become more nuanced.

Empirical studies have often used tasks where the participant must not only code the subjective value (function) of options, but also express that subjective value as, for example, a bid (e.g. Plassmann et al., 2007, 2010) in a relatively complex kind of auction (e.g. Becker et al., 1964). Thus, such tasks may confound the coding of the subjective values with the process by which participants formulate a bid. This critique is perhaps supported by evidence that judgements of the “pleasantness” of stimuli - which perhaps reflect the subjective values more directly - implicate vmPFC but not OFC (Lebreton et al., 2009). Similarly, Lin et al. (2011) found evidence of vmPFC encoding the subjective value of both social and monetary stimuli in a two-armed bandit task. Bartra et al. (2013) and Pessiglione and Delgado (2015) suggest a more complex view, where coding is highly context-dependent and positive and negative subjective value are coded partially separately (Ben-Zion & Levy, 2025; Knutson et al., 2007). These complications have led to a re-evaluation of the role of the OFC, which has been suggested to code a cognitive map (Tolman, 1948) of different states of the world (Knudsen & Wallis, 2022) while the subjective value of the states is coded in vmPFC (Moneta et al., 2024) in a “common currency” (Dang et al., 2024).

The neural codings have sometimes been taken to be absolute in the sense that they reflect the value of the object in isolation (e.g. Padoa-Schioppa, 2011). In such cases, they count as subjective values in the sense that term is used in this thesis. Others have suggested that they are inherently relative, reflecting how good an outcome is compared to the other options at hand (see e.g. Seymour & McClure, 2008). If so, in the terminology used here, they reflect strengths of preference rather than subjective value. The coding of such strengths of preference has also been claimed to implicate OFC (e.g. Hosokawa et al., 2007) rather than vmPFC, which might suggest that both are coded but in different areas of the brain.

All in all, this neuropsychological literature seems to presume a link of the form displayed in Figure 1.1C, where choices and preferences are grounded in subjective values.

Attitudes

The APA Dictionary of Psychology defines an attitude as “a relatively enduring and general evaluation of an object, person, group, issue, or concept on a dimension ranging from negative to positive.” (American Psychological Association, 2026). By that definition, an attitude is effectively a subjective value of some target. An influential account is that attitudes of this kind are retrieved automatically upon stimulus presentation (Fazio et al., 1986), as disclosed by priming effects (McNamara, 2005) for stimuli with congruent/incongruent valence (see Herring et al., 2013, for a meta-analysis). Such attitudes may have both affective and cognitive components (Fazio et al., 1986, p. 230).

Wicker (1969) provided an early, influential critique, arguing that attitudes predicted behaviour to only a very limited degree. In the present terminology, the link between subjective value and choice (Figure 1.1) was weak. This motivated work that refined the concept of attitude to overlap more distinctly with what I here would call a subjective value (see e.g. Fazio, 2007). It also motivated work that sought to unpack the link from subjective value to (often social) behaviour, where the Theory of Planned Behaviour (Ajzen, 1991) is a famous contribution. See Granados Samayoa and Fazio (2024, p. 679-682) for a historical overview.

Where does subjective value come from?

One stream of explanations is that subjective value is reflected in immediate affective reactions (Damasio, 1996; Slovic et al., 2007; Zajonc, 1980). Through experience, one learns to associate a particular stimulus with whatever affects that subsequently arise (cf. Kalman filter temporal difference learning, Gershman, 2015). These (valenced) affective reactions can then support preferences and choices (Damasio, 1994).

An alternative explanation comes from classical learning theory. Drive state theories (e.g. Dulberg et al., 2023; Hull, 1943) postulate the existence of one or several internal states (e.g. hunger, thirst, arousal) that are homeostatically regulated - when the current state diverges from the homeostatic state, making up part of that difference is perceived as rewarding. Correspondingly, being moved further from one’s homeostatic state is perceived as punishing. This tradition has been revived in “homeostatically regulated reinforcement learning” theory (Keramati & Gutkin, 2014; Yoshida et al., 2025), which suggests that drive reduction itself is not the carrier of perceived reward - rather, the predicted drive reduction caused by some outcome must be learnt associatively (Rescorla, 1988; Widrow & Hoff, 1960). The *predicted* reduction, then, is

what is rewarding.⁵ This idea has recently been widened beyond internal states to goals in general, the suggestion being that it thereby can constitute a general explanation of the origin of subjective value (Juechems & Summerfield, 2019). Perhaps because they arose from different research traditions, drive state theories tend to be phrased in different language than theories centred on a motivational valence system (see e.g. Burton & Balleine, 2022) that regulates a general “approach/avoidance motivation” (Elliot, 2013). Yet, the two perspectives seem quite similar to me. Importantly, both postulate a subjective value scale that goes from negative to positive rather than from 0 to positive (contrast this to the utility scales in Von Neumann & Morgenstern, 2004/1944, which are normalised to stretch from 0 to 1). That is, some unpleasant stimulus might not merely have a *low* subjective value but may instead be aversive and, consequently, actively avoided.

Note how both affective and drive reduction accounts seem to presume a representational conception of subjective value (see Section 2.1.1). Affects and drive reduction indicate the individual’s internal state, and subjective values derived from them indicate the target’s usefulness in regulating that internal state.

2.1.5 Decision by Sampling

Although the shapes of subjective value functions have been a topic of discussion for some time (see Section 2.1.4 above), their cognitive origins have received surprisingly little attention. While there exists a literature on how subjective value comes to be (see the preceding paragraph), it does not directly address the subjective value of different quantities of the same object, which is what a subjective value *function* refers to in this thesis. Even if it turns out that I make decisions by weighting losses in a different way than gains (Tversky & Kahneman, 1992), why does a loss of €5 sting this or that much harder than a loss of €3? Why do some people cognise a gain of €2 as pocket money while others find it to be a fortune?

The Decision by Sampling (DbS, Stewart et al., 2006) theory provides an elegant cognitive explanation of how subjective value functions get their shapes, why that shape differs, and where the cognised subjective values come from, deriving them from frequencies of attribute values in the individual’s natural environment. It suggests that subjective value u is the ranking of a target stimulus x in a small sample of exemplars:

$$u(x) = \frac{\sum_{i=1}^n \mathbf{I}(x > y_i)}{\text{card}(Y)}, y_i \in Y \quad (2.2)$$

⁵While typically not explicated in associative theories, reinforcement learning does not preclude that there also exist at least some predispositions (see e.g. Birch, 1999) that set the starting point of, and thus influence, the learning trajectory (cf. Denrell, 2005).

where \mathbf{I} is an indicator function that is equal to 1 if the condition $x > y_i$ is true. Y is populated both from long-term memory and recent experiences. n is not specified but assumed to be “small”. The cognitive process is sequential: one exemplar is sampled and compared to the target as per the indicator function. A tally is initialised and the result is stored. Another exemplar is sampled and compared. A tally of the results is updated. This continues until the process is terminated (see Stewart & Simpson, 2008, for a suggested stopping rule). This temporal aspect is not described in Equation 2.2. The proportion of Y sampled from recent experience rather than long-term memory is, to the best of my knowledge, not explained. It is, however, claimed to be large enough for recent experiences to affect choices (see e.g. Stewart et al., 2015; Ungemach et al., 2011).

In this way, DbS explains cognised subjective value as given by the (valenced) rarity of some quantity in one’s environment. How does this relate to the historical accounts of value that we reviewed in Section 2.1.1? Perhaps it makes sense to think of DbS as a heuristic (Gigerenzer & Gaissmaier, 2011) for representational subjective value. This interpretation seems consistent with Vlaev et al. (2011) who list it as being a theory “without value computation” (see also Walasek & Brown, 2023, who seem to take a similar view).

Note an important theoretical difference between DbS and other exemplar models that often assume that the judgement or decision relies on some measure of “similarity” (e.g. Hodgetts & Hahn, 2025; Juslin et al., 2003; Nosofsky, 1984). DbS only computes whether the target is greater than the exemplar or not. *How much* greater it is (or is not) never enters the cognitive process (G. D. A. Brown & Walasek, 2023).

Empirical evidence

An important observational argument for DbS is the idea that previously suggested shapes of subjective value functions (e.g. piecewise, S-shaped subjective value function, Tversky & Kahneman, 1992) can be explained in terms of how the objective quantity (e.g. money gained and money lost) is distributed in the individual’s ecology (Stewart, 2009). Stewart et al. (2006, p. 4-17) presented data on credits and debits to British bank accounts. Most credits and debits were small, becoming increasingly rare as the amount credited/debited increased. However, the distribution of debits trailed off earlier than the distribution of credits. The cumulative distribution of credits and debits thus mirrored the S-shape of Tversky and Kahneman (1992), including the steeper slope below zero (“loss aversion”). They reported corresponding results for the (dis)value of (i) waiting time for, and (ii) probability of, a reward. The data taken to indicate the ecological distributions of those quantities were (i) the number of hits on www.google.com when searching for different time delays (e.g. “1 day”, “1 year”) and (ii) a corpus of British English (but see Stewart et al., 2006, p. 13-19, for some complexities in how assumed numeric probabilities were calculated from the corpus data). If these distributions too

were representative of what was held in people’s memories, then it was claimed that DbS would predict so-called hyperbolic discounting (Ainslie & Haslam, 1992) and overweighting of small probabilities and underweighting of large ones (but see Hertwig et al., 2004, who claim the opposite empirical pattern).

The perhaps most important source of empirical support for DbS is studies that expose participants to different distributions of some attribute (or record such exposure in observational studies, e.g. Study 1A in Ungemach et al., 2011). This has often been done by presenting participants with options where the attribute values of the options (e.g. potential rewards) are manipulated between groups (see Stewart, 2009). One group might be exposed to predominantly (i) low values while the other group is exposed to predominantly (ii) high values. (i) should increase the subjective value of an intermediate target because the target will win if put through an ordinal comparison against the new, low exemplars. Conversely, (ii) should decrease the value of the intermediate target because if those exemplars are sampled the target will lose the ordinal comparison. Sometimes the dependent variable has been judgements or choices (e.g. certainty equivalent or chosen option, Alempaki et al., 2019). Other times, the dependent variable has been a (maximum likelihood) value of some free parameter in a theoretical model (e.g. Stewart et al., 2015). We employed this empirical paradigm in Study I and II, with choices and judgements as dependent variables.

2.2 Preference

2.2.1 Existing preferences

Foundational work in philosophy (A. Smith, 2025/1776a, p. 364), economics (Arrow, 1950), and sociology (M. Weber, 1947, p. 158-159) has taken preferences to reflect fairly stable, deeply held beliefs of subjective value (i.e. “axiological statuses”, Schroeder, 2024).⁶ I borrow some terminology from MacLean (2006) and call such preferences “existing preferences”. Existing preferences presume the link displayed in Figure 1.1C. The assumption of existing preferences has allowed social scientists to treat preferences as exogenous - as something taken as a given and put beyond current investigations. By assuming that people know what they prefer and do their best to achieve their preferred outcome, one can explain social phenomena in terms of the structure of the strategic, social interaction in which such existing preferences are expressed (e.g. Akerlof, 1978; Schelling, 1971; V. L. Smith, 1962; Weibull,

⁶Note how at least Max Weber seemed to take a representational view (see Section 2.1.1) of subjective value, defining economic behaviour as that “concerned with the satisfaction of a desire for *Nutzleistungen*”. The translators translate “*Nutzleistungen*” to “utilities” but a more literal translation would be “use-services”; the needs-fulfilling properties of the target. I am grateful to Leonie Polzer for a helpful discussion.

2004). Explanations can be “purely” societal (cf. Forsgren, 2025), without reference to psychology.

These existing preferences were later taken to obey certain axioms, the most famous being those of Von Neumann and Morgenstern (2004/1944): continuity, completeness, mixture independence, and transitivity. Together, the axioms guaranteed the existence of a differentiable, well-behaved utility function (see Section 2.1.3) and constituted the “expected utility theory” (EUT). A famous alternative axiom is “Independence of irrelevant alternatives” (IIA) which, in short, states that adding new options to the menu should not affect the preferences for the original options.⁷

Right from the start (e.g. Allais, 1953; Allais & Hagen, 1979; Tversky, 1975), critique of existing preferences was bound up with critique of the empirical viability of the axioms: criticising one often involved some critique of the other, with differences in where the emphasis lay. One stream of critique leaned more into showing that preferences did not seem to be supported by fairly stable subjective values of outcomes; preferences could be affected by experimental manipulation of supposedly irrelevant aspects of the immediate environment. For example, studies showing that some people’s choices could occasionally (see e.g. Huber et al., 2014) be swayed by whether identical outcomes were framed as gains or losses (Kahneman & Tversky, 1979; Tversky & Kahneman, 1981), or whether a “decoy” option was added or removed from the menu (e.g. Castillo, 2020; Huber et al., 1982).⁸ Other critiques mainly concerned the axioms (e.g. common consequence & common ratio effects, Birnbaum & Chavez, 1997; Blavatsky et al., 2022, 2023), rather than the existing preferences per se.

Together, and despite objections by Morgenstern (1979), these critiques helped build an impression that the notion of “existing preferences” was questionable. The “deeply held beliefs” seemed quite superficial.

2.2.2 Constructed preferences

The evidence that people do not always behave as-if they expressed existing preferences - because their apparent preferences were malleable and/or because

⁷IIA was originally introduced as a condition that an acceptable aggregation of preferences (for example, vote shares in a general election) should fulfil, not as applying to the preferences of any one individual (Arrow, 1951, Chapter III, section 3). On the individual level, Arrow (1951) merely assumed completeness and transitivity (these axioms are discussed in Section 2.2.4 below). Yet, IIA seems to have come to be interpreted as an axiom of individual preferences (Debreu, 1960).

⁸Allais (1953) also offered a normative criticism: to him, a given prospect existed in the context of the other prospects entailed by the same option, but also in the context of choices made in the past. Thinking of preferences as divorced from such context made little sense because there was “no such thing as abstract value” (cf. Section 2.1.1 above). Thus, behaviour *should* be guided as per Figure 1.1B. See Mongin (2019) for a historical overview.

they violated axioms - motivated a new stream of research in psychological science. Rather than being based on deeply held beliefs, decisions were said to rely on preferences that were “constructed” (Slovic, 1995) ad hoc and online. Let me briefly review some of the main lines of inquiry.

Heuristics

A prolific tradition under the constructed preferences umbrella is the idea that people rely on “heuristics” (e.g. Todd & Gigerenzer, 2000; Tversky & Kahneman, 1974): simplifying decision rules that implement a form of bounded rationality (Simon, 1990) by ignoring some information and avoiding complex computations (Gigerenzer, 2008).

One example is the “default heuristic” (Gigerenzer & Gaissmaier, 2011) where the decision rule states that whatever is the pre-selected option (if one exists) is chosen. If preferences were based on deeply held subjective values, this superficial aspect of the interface should not sway people’s choices. The idea of a default heuristic has seen substantial applied use to “nudge” (Thaler & Sunstein, 2008) citizens into behaviour that aligns with governmental policy goals (see e.g. Cronqvist et al., 2018; Szaszi et al., 2017; Tannenbaum et al., 2017), although there are questions about whether effects of defaults are empirically robust (e.g. Altmann et al., 2019; Bronchetti et al., 2013; Forsgren, Mandl, & Karreskog Reh binder, 2025; Jachimowicz et al., 2019).

Another example is “noncompensatory” heuristics. These construct preferences such that they form lexicographic (semi)orders (see Fishburn, 1974) by checking which option is superior in a given aspect, one at a time, until a choice is made (e.g. Birnbaum & Gutierrez, 2007; Brandstätter et al., 2006). Being a lot better in one aspect can thereby not compensate for being a little worse in another, hence the moniker. Tversky (1972) used the example of a choice between (i) a trip to Paris, (ii) a trip to Paris and a small cash bonus, and (iii) a trip to Rome and a small cash bonus. Because (i) and (ii) are equalised in all respects but the cash bonus, the addition of the cash bonus would result in a preference favouring (ii). Conversely, (ii) and (iii) are equalised in the cash bonus but set apart by the destination of the trip, why that would determine the preference. While noncompensatory heuristics share family resemblance, they differ in exactly how they suppose that the aspects are checked and evaluated. One of the more radical examples is the “take the best” heuristic (Gigerenzer & Goldstein, 1996) which bases a choice purely on the “cue” with the highest “validity” (Hirsch et al., 1964): the single piece of information that, out of those available, is the superior predictor of some outcome. For example, a cue might be whether some business is an established company or not and an outcome might be return on shares in that company (Bröder, 2000; Newell & Shanks, 2003). A less radical version along the same lines is the “priority heuristic” (Brandstätter et al., 2006) which compares the attributes of options one by one until it finds a sufficiently decisive tie breaker.

Apart from these two examples, there exists a large library of suggested heuristics (see e.g. Gigerenzer & Todd, 1999; Tversky & Kahneman, 1974). Some appear motivated by empirical phenomena that seem to contradict existing preferences (e.g. default heuristic). Other studies work from a *presumption* that preferences are constructed, and then present theories of how that could work (e.g. noncompensatory heuristics). While there is a long-running debate on their normative status (see e.g. Gigerenzer, 1996, 2025; Kahneman & Tversky, 1996), heuristics thus share the feature that they do not ground choices in subjective values (Figure 1.1C). Instead, they draw on more proximal features of the options to form preferences (Figure 1.1B) or determine choices directly (Figure 1.1A). For example, the default heuristic determines choices rather than preferences because the default option is chosen without comparing options. An individual employing the default heuristic would not stop to ask “irrespective of the other options, how good is option *a* in and of itself?”

Walasek and Brown (2023) have made a philosophical argument for constructing preferences through heuristics, presuming a value pluralist conception of subjective value (see Section 2.1.1). They use the example of a professor who might be a so-and-so good teacher and a so-and-so good researcher, claiming that “there is no higher-level, utility-like, abstract common currency such as ‘academic quality’.” (Walasek & Brown, 2023, p. 1013). In this view, there cannot exist a unitary subjective value of each option, why preferences cannot be based on comparing them (Figure 1.1C). Instead, people must resort to “making choices without trade-offs” (Brandstätter et al., 2006) through use of noncompensatory heuristics. Such processes go through different attributes of options - which correspond to different kinds of value - to find one that can work as a tie-breaker. I return to this argument in the General discussion.

Sequential evaluation

When choosing between *a* and *b*, adding an option *c* that is similar in its aspects to *a* might increase the probability of someone choosing *b* (“the similarity effect”, Tversky, 1972). Alternatively, when *c* is dominated by *a* in all aspects but not by *b*, this might increase the probability of choosing *a* (“the attraction/asymmetric dominance effect”, Huber et al., 1982). Lastly, when the aspects of *c* are designed such that *b* constitutes a middle ground between *a* and *b*, this might increase the probability of *b* being chosen (“the compromise effect”, Simonson, 1989).⁹ See also Noguchi and Stewart (2014). The similarity, attraction, and compromise effects show that preferences are unstable; they are affected by the similarities and differences between the presented op-

⁹These effects are often called “context effects” but this invites us to understand “context” in a quite narrow sense. In this thesis, I will generally prefer the phrase “effects of the option set”, or variations thereof. “Context” will typically be reserved for the original psychological meaning of that term: the (massive) set of stimuli presented concurrently with some focal stimulus of interest.

tions and not only the characteristics of the individual options. This should, the argument goes, not happen for existing preferences based on comparisons between deeply held subjective values (Figure 1.1C).

While the similarity, attraction, and compromise effects are far from being universal (see e.g. Figure 3 in Cataldo & Cohen, 2018) they still seem to have become something of a litmus test (see e.g. Noguchi & Stewart, 2018) for theories in the “sequential evaluation” tradition. In such theories, preference construction (Figure 1.1B) is thought to happen by sequential evaluation and comparison of different aspects of options, where the result of each comparison is accumulated until a decision threshold is reached (e.g. Roe et al., 2001; Trueblood et al., 2014; E. U. Weber & Johnson, 2006). The “adaptive decision-maker” programme (Payne et al., 1992) posits a number of more fundamental processes of sequential evaluation and comparison that can then be combined to flexibly form strategies for constructing preferences (cf. “architectural” approaches, Ritter et al., 2018).

Sampling theories

The similarity, attraction, and compromise effects have been interpreted as evidence of option attributes being compared sequentially. However, they have also been interpreted as a kind of priming effect (A. S. Brown, 1979) where the presented options cue memory retrieval (e.g. Bhatia, 2013). It has also been noted that, just like perceptual decisions (e.g. which stimulus is the largest), preferential choice is subject to a speed/accuracy trade-off: the longer one contemplates, the higher the probability that one selects the option with the highest subjective value (see e.g. Gluth et al., 2015; Shadlen & Shohamy, 2016) - where subjective values are inferred from a separate judgement task (e.g. Becker et al., 1964).

These phenomena have been taken as support for the proposition that some kind of information is sampled from (long-term) memory and integrated to construct a preference (Shadlen & Shohamy, 2016). The sampled information might consist of potential subjective values (Gonzalez & Aggarwal, 2023; Gonzalez & Dutt, 2011; Zylberberg et al., 2024)¹⁰ but also arguments in favour of or against making a certain choice (Shafir, 1993; E. U. Weber & Johnson, 2011).

Some envision the sampled information to potentially have episodic qualities (Shadlen & Shohamy, 2016; E. U. Weber & Johnson, 2006). Such episodic sampling accounts would seem to me to complement the attitudinal and affective views on subjective value discussed in Section 2.1.4: there is evidence that episodic simulation of events that we *have not* experienced is generated by the same system as that which we use to remember events that we *have*

¹⁰One recent model with a similar structure is the Autocorrelated Bayesian Sampler (Zhu et al., 2024). However, to the best of my knowledge it has only been applied to questions of fact rather than preference. This orientation is perhaps indicated by the authors referring to the information sampled as “hypotheses”.

experienced (e.g. Schacter et al., 2008; Szpunar et al., 2014). If so, it stands to reason that sampling (simulating) episodic pros and cons would evoke similar attitudinal and affective reactions as thinking back on them. The subjective values of those attitudes and affects could then be integrated to construct preferences.

In sum, sampling theories can propose either that subjective values are unstable, leading to unstable preferences further down the line, or that preferences themselves are constructed and therefore inherently unstable.

Risk preferences

In the traditions reviewed above, some important empirical phenomena concerned adding or removing options from the presented set. When options are risky (in the sense of their outcomes being explicitly probabilistic, see Knight, 1957/1921), one can also imagine adding, removing, or manipulating so-called “branches”. A branch is in essence a potential outcome at a potential probability, where the option is defined by all its potential outcomes and their probabilities. As discussed in Section 2.1.3, one theoretical approach has been to subsume the subjective (dis)value of taking risks (sensu Knight, 1957/1921) into the shape of the subjective value function. Alternatively, some models weight subjective values based on their relative rank or distance within the set of available outcomes. (e.g. Transfer of Attention Exchange, “TAX”; salience theory, Birnbaum & Navarrete, 1998; Bordalo et al., 2022). Here, subjective values pre-exist but apparent risk preferences are constructed through the weighting, which is supposed to reflect putative attentional processes. These kinds of models often draw support from experiments where manipulating the branches of options affect choices (e.g. Birnbaum, 2008; Birnbaum & Chavez, 1997), even when the expected values of options are identical (“Mao pairs”, Dertwinkel-Kalt & Köster, 2019).

Others have taken more of an existing preference approach by investigating risk taking as a trait-like property of the individual (e.g. Frey et al., 2017; Millroth et al., 2020; E. U. Weber et al., 2002). Here, the motivation is perhaps not some specific empirical phenomenon per se but rather the open question of to what extent there exists stable individual differences.

Lastly, there exists a third way that perhaps does not neatly slot into the existing/constructed preferences spectrum: a focus on how the probabilities themselves are weighted depending on whether and how they are learnt from experience (e.g. Biele et al., 2009; Hertwig & Erev, 2009; Hertwig et al., 2004). Such learnt weights may be beliefs or perhaps associations (Forsgren et al., 2023). However, they are not subjective values, nor do they imply constructed preferences, in the senses those terms are used here.

The most radically constructivist take on risk preferences does perhaps come from the literature concerning (e.g. Stewart & Simpson, 2008) or adjacent to (e.g. Vlaev et al., 2007b) DbS. There it is argued that apparent risk preferences are a by-product of the subjective values of outcomes, probabilities, and other

attributes being constructed in the moment. This view will be interrogated in Study I & II.

Implications of constructed preferences

It has been claimed that because preferences are constructed, choices may not be taken to reveal subjective value (Figure 1.1A & 1.1B). To give an applied example, Slovic (1995) cites a discussion concerning informed consent from an unpublished manuscript by Douglas MacLean. It goes something like this. Consider a patient who needs a, possibly painful and/or risky, treatment. The physician wants “what is best for the patient” - they want the patient to have the best subjective experience possible, all things considered. What should the physician’s role be in the decision on whether to go through with the treatment or not? Under what we here call the existing preferences view, the whole business is fairly straightforward. The physician acts as an “expert advisor” (Slovic, 1995, p. 370) to the patient. Their job is to make sure that the patient understands the lay of the land. As long as the patient does so well, they can be trusted to employ their existing preferences to choose whatever is best for them. If there are no existing preferences, and the preferences are being constructed under the influence of the physician, the patient cannot sensibly hold primary responsibility for making the decision even when well-informed.¹¹ This is an early example of the “choice architect” view. Because preferences are constructed, and predictably so, the locus of agency is not within the individual making the choice. The distal cause (Krieger, 2008) of the choice is actions taken by whoever designed the choice environment. If so, treating preferences as given would be a serious limitation of work in the tradition of Arrow (1950), A. Smith (2025/1776b), and M. Weber (1947).

2.2.3 Constructivism as a spectrum

Theories of preference may come in different shades of constructivism. For example, while noncompensatory heuristics postulate that preferences for options are constructed on the spot, they still seem to require existing preference for the attributes themselves. To check if an option is “superior” in an attribute, one must have some kind of preference over the different states each attribute can take. In the example from Tversky (1972), one must first know whether Paris is preferred to Rome and whether a small cash bonus is preferred to no cash bonus before one can apply a noncompensatory heuristic. In sequential evaluation models, every single evaluation performed could be seen as a preference for one state of an attribute over another state of that attribute. Where

¹¹In a similar vein, Fredrickson and Kahneman (1993) claimed that participants’ retrospective evaluations of episodes (stated subjective values) did not match the sum of their stated affective state during the episode (putatively correct subjective values). Such “duration neglect” could motivate relegating decision-making to the physician.

these more fundamental preferences come from is not explained in the theories. While one could presumably suggest a theory where the fundamental preferences too are constructed, it seems like one at some point will have to draw on some notion of what is better and what is worse that pre-exists the constructive process. Conversely, one could claim that there strictly speaking cannot exist “deeply held” subjective values, and thus no truly existing preferences, because all memory is constructed (e.g. Loftus, 1997; Schacter & Addis, 2007).¹²

While there certainly are intricacies to this terminology (see S. M. Smith & Krajbich, 2021, for an alternative), the core issue is perhaps stability and malleability. Preferences are “highly” constructed when details of the immediate environment are “highly” important in determining the end result, as evidenced by effects of manipulations of, for example, the option set. Conversely, preferences are “highly” existing when they are fairly stable, as-if they were sitting in the mind just waiting to be expressed. Thinking versus knowing, if you will. A given theory will sit on a spectrum between those end points.

With this in mind, one position could be that there exists a mix of the end points: some preferences are (by and large) existing while others are (by and large) constructed. This possibility has been acknowledged explicitly (Bettman et al., 1998; Fischhoff, 2006, p. 188 & p. 654, respectively), but those authors still seem to view highly constructed preferences as constituting the greater component of the mix (Bettman et al., 1998, 2008; Fischhoff, 2006, p. 199-207, p. 171-172, & p. 655, respectively) - or at least treat it as the component of greater scientific interest.

2.2.4 Transitivity of preferences

One could in principle imagine both (highly) existing and (highly) constructed preferences that violate any axiom. Yet, violations of transitivity of preferences have often been presented as support for theories of preference construction (e.g. Birnbaum & Gutierrez, 2007; Brandstätter et al., 2006). Perhaps this is so because it is difficult to see how preferences based on deeply held beliefs of subjective value could violate an axiom that, at some level, seems so sensible (see below). In what follows, I explicate the meaning of transitivity and describe a prominent descriptive model of transitive preferences (Strong stochastic transitivity, SST). That model is consistent with the path displayed in Figure 1.1C but with noisy subjective values in the spirit of Thurstone’s idea

¹²If memory is constructive, then thinking of subjective values as “stored” in memory (e.g. Atkinson & Shiffrin, 1977/1968) is misleading. Moscovitch (2007) and Schacter (2007) emphasised a clear distinction between the engram - the neural substrate storing information - and memory. The latter is constructed from the (existing) engram. “It follows, therefore, that *memory does not exist until it is recovered.*” (Moscovitch, 2007, p. 17, emphasis in original). If so, all subjective values are generated in the moment from a lasting engram.

(Equation 2.1). Then, I reflect on the normative status of the transitivity axiom. Lastly, I provide an overview of the current state of the empirical debate.

Study III will scrutinise whether intransitivity of preferences really is as prevalent as has been claimed (e.g. American Psychological Association, 2025). It will do so by estimating the prevalence of some descriptive models, including SST.

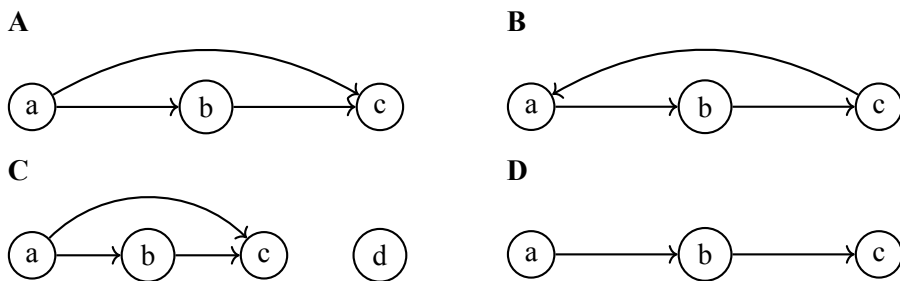
Definition and implications

Transitivity means that some relation is “transferred” or “inherited” by successive elements. Ancestry is transitive because if Adam is an ancestor of Abraham, and Abraham is an ancestor of Moses, then Adam is an ancestor of Moses. Dominance in Rock, paper, scissors is not transitive (*intransitive*) because Rock beats Scissors, and Scissors beats Paper, but this does not imply that Rock beats Paper. We can write aRb and bRc for some arbitrary binary relation on $a, b, c \in \mathcal{O}$. We say that the relation is transitive when the following condition holds: $aRb, bRc \Rightarrow aRc$. For the two examples above, the relation R was “is an ancestor of” and “beats”, respectively. For preferences, R will be something that expresses the preference relation: “preferred to”, “not preferred to”, or “indifferent to”. Figure 2.1 illustrates transitive and intransitive orderings.

On its own, transitivity tells us something about the relations between some subset of options but we need additional properties to fully characterise the entire set of preferences. If we add the property of *completeness* - that the relation R exists for all elements in \mathcal{O} - then \mathcal{O} is weakly ordered; there exists a ranking, with ties allowed, of all options. In other words, transitivity and completeness together ensure that all options are comparable - the individual either prefers one to the other or is indifferent between them, but we never have a case of “apples and oranges” where they simply cannot say. Again see Figure 2.1. To exclude the possibility of indifference - that two options are ranked identically - we need a property called *irreflexivity*, which together with transitivity happens to imply *asymmetry*,¹³ which basically means that one option is always strictly preferred over the other. In empirical studies, one usually rules out the possibility of indifference by design: participants are forced to select one option or the other, and thus cannot express indifference even if that were to be their view. If R is transitive, complete, and irreflexive, we have a strict total order: a strict ranking of all options.

¹³To see why, the following proof might be helpful to some readers: consider some arbitrary $a, b \in \mathcal{O}$. Let us assume that some binary relation R is transitive and irreflexive. Let us further assume that R is not asymmetric: aRb but also bRa . By transitivity, $aRb, bRa \Rightarrow aRa$. This contradicts irreflexivity: $a\mathcal{R}a$. Thus, by contradiction, we have proven that for $\forall a, b \in \mathcal{O}$, transitivity and irreflexivity imply asymmetry. This concludes the proof.

Figure 2.1. Illustration of transitive and intransitive orderings



Note. Panel A: transitive and complete ordering over three elements a , b , and c . aRb , bRc , and aRc , so the directed graph is acyclic. Panel B: intransitive and complete ordering over the same elements. Here, aRb , bRc , but cRa . Thus, we have a cycle in the directed graph. Panel C: transitive and incomplete ordering. aRb , bRc , and aRc , but d is incomparable to all other elements. A degenerate example for three elements would be when aRb and aRc but b is incomparable to c . That ordering is transitive because the implication $aRb, bRc \Rightarrow aRc$ cannot be evaluated when one of the two antecedent conditions is missing. Panel D: intransitive and incomplete ordering. $aRb, bRc \not\Rightarrow aRc$ because a and c are incomparable. In all panels, R is irreflexive because there are no self-links.

Strong stochastic transitivity of preferences

As cognitive psychologists, we are faced with the challenge of how to infer whether transitivity holds for someone's covert preferences based on their overt choices. If an individual deterministically makes transitive choices (and we have reason to believe that they are indeed choosing the options they earnestly prefer), it seems safe to conclude that this is the systematic expression of transitive preferences. However, variation is ubiquitous in human behaviour (Sundh et al., 2025). To make inference to the preferences, we must model the source of this variation (Cavagnaro & Davis-Stober, 2014; Regenwetter & Davis-Stober, 2008). A famous such model is strong stochastic transitivity (SST, Davidson & Marschak, 1959):

Definition 1. An individual's preferences are SST if, for all $a, b, c \in \mathcal{O}$,
 $P(a \succ b) \geq 0.5, P(b \succ c) \geq 0.5 \Rightarrow P(a \succ c) \geq \max\{P(a \succ b), P(b \succ c)\}$

That is, options may be ordered from worst to best, the probability of choosing an option ranked higher in the ordering must exceed 50%, and this probability must increase the further apart in the ordering two options are. Importantly, this is a descriptive model - on its own, it has no cognitive interpretation. Let us consider some interesting cognitive models that predict SST.

Firstly, we have models where the subjective valuations are noise-free but choice is noisy. One example is the strong utility model (Cavagnaro & Re-

genwetter, 2023). Another is Luce’s choice rule (Luce, 1959). These models assume that a noiseless scalar subjective value is attributed to each option. The subjective values are then compared to form a preference, but the expression of this preference is noisy: the greater the difference between the subjective values, the less noise.

Secondly, we have the random utility model (Block & Marschak, 1960) where the subjective values U are random variables.¹⁴ For preferences to be SST, we must restrict the distributions of the subjective values to be (i) symmetric and (ii) homogenous.¹⁵ The probability of some individual preferring option a to option b is simply the probability that the subjective value of the former exceeds the latter at a given moment. This probability increases with the mean difference - the greater the difference in subjective value between the options, the lower the discriminative difficulty. If we make the more restrictive assumption that all subjective values are normally distributed with equal variance, then we have a version of the Thurstonian model (Equation 2.1, Thurstone, 1927a).

In sum, cognitive models that predict SST share the family resemblance that they suppose the existence of one unitary subjective value per option. This is inconsistent with many theories of preference construction (see e.g. Dertwinkel-Kalt & Köster, 2019; Rieskamp et al., 2006; Roe et al., 2001; Trueblood et al., 2014; Tversky, 1972; Tversky & Edward Russo, 1969; Vlaev et al., 2011). Evidence that SST holds thus challenges that literature, or at least its generality. Because it is highly parsimonious (see Study III), the principle of Ockham’s razor suggests that SST should be preferred over more flexible models unless that extra flexibility is needed (see e.g. Akaike, 1974; Hansen & Yu, 2001; Pitt et al., 2002).

Normative status of transitivity of preferences

SST suggests that behaviour is stochastic. Why would systematic violations of transitivity (i.e. intransitivity) be any worse than if similar deviations arose because of noise? Let me make three points.

Firstly, intransitive preferences open one up to being exploited. A shrewd swindler might exploit the “cyclical” of an intransitive preference ordering (Figure 2.1B) to “money pump” a victim (Danan, 2010; Davidson et al., 1955), leading them down a series of choices that ultimately leave them worse off than where they started. Note that “money pumping” does not require agency: if cyclical choices involve any cost in terms of fitness (e.g. extra time spent on cyclical foraging) then one is effectively being money pumped by nature.¹⁶

¹⁴Formally, $P(a \succ b) = P(U_a > U_b)$ where $u_i \in U$ are subjective values.

¹⁵To preview, in the Empirical work we will also test a model that relaxes these restrictions. See Polania et al. (2018) for a neuroscientific suggestion, based on efficient coding (Barlow, 1961; Laughlin, 1981), of how such a model might work.

¹⁶Arkes et al. (2016) argue that there is no empirical evidence for the existence of money pumps so violations of transitivity are not costly, but that way of looking at it seems to me to be re-

Secondly, there is something intuitively whimsical about intransitive preferences. It seems, at least to me, that the victim of the money pump scheme above has been bamboozled; they should have known better!

Lastly, intransitive preferences are aimless. From Skinner (1958) to Maslow (1943), behaviour is presumed to be goal-directed. This is so because animals have evolved to be fitness maximising (cf. Axelrod & Hamilton, 1981; Dawkins, 1976). Transitivity of preferences ensures that there exists a ladder to climb, a way to subjectively better oneself. With intransitive preferences, we are like Sisyphus: consumed by an eternal cycle of futile labour. It has been claimed that we must imagine Sisyphus happy (Camus, 1955) but that is the kind of thing you say when you have never rolled a massive boulder up a hill for all eternity. All in all, there seems to be a good case to be made for transitivity of preferences being regarded as normatively desirable.

Testing for transitivity of preferences

A famous study claimed to evidence intransitive preferences was Tversky (1969) where participants had been screened for their susceptibility to making intransitive choices. In Experiment 2 in that study, the screening session was also used to tailor subsequent stimuli so as to make more choices violate transitivity. More recently, Butler and Pogrebnina (2018) have used a related approach to claim that preferences are intransitive under tailored stimuli (but see Birnbaum, 2020, for a statistical critique and reanalysis, and Butler, 2020, for a response). The experiments in Tversky (1969) were “not addressed to the question of whether human preferences are, in general, transitive; but rather to the question of whether reliable intransitivities can be produced, and under what conditions.” Despite this, some authors have promoted such existence proofs to “well-established evidence ... that cyclical choice is a robust and reasonably general phenomenon.” (Starmer, 2000). That notion appears to be fairly widespread (e.g. American Psychological Association, 2025; Bikhchandani & Segal, 2021; Brandstätter et al., 2006; Fishburn, 1991; Grether & Plott, 1979; Kalenscher, 2010; Lee et al., 2009; Rieskamp et al., 2006; Schoemaker, 1982; Veit, 2024; Walasek & Brown, 2023).

The evidence supporting this was most seriously called into question by Regenwetter et al. (2010, 2011). Fundamentally, their critique rested on how previous work had insufficiently addressed the complications inherent in evaluating a deterministic, algebraic property of the covert cognition based on stochastic choice data. They imported statistical advances by Klugkist and Hoijtink (2007) into psychology that allowed the bridging of this gap. In short, the method consists of setting up a number of inequality conditions that define the subspace of the sample space compatible with some model of transitive preferences. Definition 1 provides the inequality conditions for SST.

versing the potential causality. Rather than intransitivity of preferences being costless because of an inexistence of money pumps, maybe there are no money pumps because transitivity of preferences means that they do not work.

The state of the debate is thus that the transitivity of preferences is an open empirical question. The small number of studies that have used appropriate methods have found some participants to have intransitive preferences but these have usually relied on monetary gambles (Birnbaum, 2020, 2023a, 2023b; Birnbaum & Schmidt, 2008; Cavagnaro & Davis-Stober, 2014; Cavagnaro et al., 2024; Regenwetter et al., 2011).¹⁷ When participants instead gambled for uncertain quantities of everyday items (e.g. “9/50 probability of winning about 15 sandwiches”, see the “noncash” condition in Regenwetter et al., 2011) instead of money, there was no indication of intransitive preferences. See also Cavagnaro and Davis-Stober (2014) for a Bayesian reanalysis.

2.3 Summary of literature review and purpose of this thesis

Social science has historically presumed that people perceive options to have fairly stable subjective values, leading to fairly stable and coherent preferences that obey axioms like transitivity. Behaviour can then sensibly be thought of as individuals expressing such “existing” preferences. This has been criticised, by psychologists and others, who argue that preferences are “constructed” in the moment. Empirical evidence taken to support this view primarily comes from experiments showing that manipulations of the immediate environment - such as supposedly irrelevant characteristics of options or the inclusion/exclusion of options from the menu of alternatives - can affect apparent preferences. However, even theories in this constructive tradition might sometimes require some conception of preferences that pre-exist the choice situation.

Decision by Sampling is a theory that in different ways stands out from previous work. Instead of proposing a shape of the subjective value function, it postulates a cognitive process that brings the shape about. Instead of attributing the origins of subjective value to associative learning of affects or drive-state reduction, it attributes it to the rarity of a target quantity in the (historical) environment of the organism. Specifically, it suggests that subjective value of quantitative attributes (e.g. money) is constructed momentarily by sampling a small sample of exemplars from memory. For each exemplar, the algorithm makes an ordinal comparison of whether the target is greater than the exemplar or not. The subjective value is simply the proportion of such comparisons that favour the target. A central experimental paradigm for studying Decision by Sampling has been to manipulate the (quantitative) attribute values of recently experienced options. Having recently experienced predominantly lower attribute values should increase the subjective value of subsequent intermedi-

¹⁷But see Cha et al. (2013) who criticise the methods of Birnbaum (2020, 2023a, 2023b) and Birnbaum and Schmidt (2008).

ate attribute values, and vice versa. This is because the intermediate target will win ordinal comparisons against the lower values.

While theories may sit somewhere in-between “pure” existence and “pure” construction of preference, there does seem to exist a view that the lion’s share of our preferences are largely constructed (e.g. Bettman et al., 2008). A clear example of this is the idea that preferences are “often” intransitive (American Psychological Association, 2025), which is typically attributed to preference construction. An alternative account would be that preferences are strongly stochastically transitive - that it is as-if they were supported by comparing stable (existing) subjective values that are corrupted by noise. However, inferring a covert property of preferences (such as transitivity) based on overt choices is fraught with analytical challenges which have only been addressed quite recently (Klugkist & Hooijink, 2007). Consequently, there only exists a small number of recent studies that investigate transitivity of preferences with appropriate methods. These do find that some participants have intransitive preferences, but have almost exclusively used so-called monetary gambles as stimuli.

While it is hard to deny that people *sometimes* construct preferences, the question is how much prominence such explanations warrant. The purpose of this thesis is therefore to critically examine the generality of the constructed preferences account by revisiting certain previous research that has been interpreted as support for constructive processes. I will scrutinise the central experimental paradigm used to investigate Decision by Sampling. Does the effect of recently experienced attribute values obtain when we address some methodological limitations (Study I)? More generally, does such an effect obtain across response modes and is it robust to relevant expertise (Study II)? If not, perhaps it is not necessary to postulate that the subjective values and preferences are constructed momentarily. I will also scrutinise whether preferences are often or rarely intransitive, and whether this differs between (unfamiliar) monetary gambles and (more familiar) everyday objects (Study III). To what extent is strong stochastic transitivity - which is consistent with an existing preferences account - a sufficient description?

3. Aims of the empirical section

The purpose of this thesis - to critically examine the generality of the constructed preferences account - was addressed through two aims.

The first aim of the Empirical work was to revisit certain empirical phenomena - previous demonstrations of which had been criticised on account of methodological limitations - that had been taken as support for theories of constructed preferences.

Study I sought to investigate a theory of the subjective value function - Decision by Sampling - and the phenomenon that had been its principal source of empirical support: that manipulating the distribution of attribute values that one observed had a causal effect on subsequent choices. It did so using a two-alternative forced choice paradigm.

Study II studied the same empirical phenomenon but in two response formats: two-alternative forced choices (2AFCs) and a vertically arranged list of several options. The latter format was taken from Bird and Harris (2018) who had presented it as an “investment mean variance measure” inspired by Vlaev et al. (2007b); a tool for studying risk preferences. Bird and Harris (2018, p. 606) interpreted the effect of preceding options in that study as “consistent with theories of relative judgement, including Adaptation Level Theory (ALT; Helson, 1947), Range Frequency Theory (RFT; Parducci, 1965, 1995), and Decision by Sampling (DbS; Stewart, Chater, & Brown, 2006)”, thus connecting with the literature on subjective value functions.

Study III moved from investigating the subjective value function to studying preference. Specifically, it sought to investigate the transitivity of preferences and to what extent (in)transitivity may be understood as a result of novel, artificial stimuli. It further investigated whether preferences for pairwise choices were compatible with a descriptive model - Strong stochastic transitivity - that could be interpreted cognitively as choice being supported by preference derived from a greater-than comparison between noisy subjective values (Figure 1.1C).

The second aim was to investigate if the empirical phenomena - presuming they obtained - would be attenuated by what we, perhaps for lack of a better word, originally thought of as “expertise” or “experience”. In Section 2.2.3, I introduced the idea that it could be that some preferences are existing while others must be constructed. We speculated that this mix may ultimately be determined by knowledge. With lived experience, people will gradually build knowledge of representational (Section 2.1.1) subjective values. These can then support relatively stable and coherent preferences. When people lack such well-known subjective values, they must construct subjective

values and/or preferences on the spot.¹ Whether someone exhibits stable, coherent (presumably existent) preferences or unstable, incoherent (presumably constructed) preferences would then differ based on a person \times domain interaction, as revealed by a difference between “expert” and “non-expert” demographics when stimuli cued the experts’ domain of expertise.² To preview, the empirical results would prevent us from fully addressing this aim.

¹While the perhaps most famous definition of “expertise” (Ericsson & Lehmann, 1996) is performance-centred - an expert is someone who consistently outperforms their cohort - note that the sense in which I am using the term also has some precedent (e.g. Chase & Simon, 1973; de Groot, 1978/1965).

²Thus, rather than being a psychological construct, expertise of this kind is a mere demographic label. It is shorthand for “having long-term memory knowledge of relevant subjective values that are cued by the task at hand”.

4. Empirical work

4.1 Methods common to Study I and II

Study I and II are the result of a series of experiments that my collaborators and I ran with a view towards scrutinising the Decision by Sampling theory and an empirical phenomenon that had motivated it: an effect of recently experienced attribute values on subsequent choice. As mentioned, DbS supposes that the attribute values one has recently come across will populate the distribution of exemplars that one's mind samples from (Y in Equation 2.2 above). When estimating the subjective value of options according to the DbS mechanism, the subjective value of a target attribute value should therefore be increased by experiencing lower exemplars and decreased by experiencing higher exemplars.¹ The subjective values will then support preferences, which will then support choices, leading to an observable effect. Originally, we expected to find that effect and were looking forward to a series of experiments that would investigate its generality. We were suspicious of whether someone who could fall back on substantial experience would display the fickle subjective values, preferences, and choices suggested by DbS. This plan quickly took a turn: while we were gearing up to start collecting data, we begun to realise that the effect of interest (and thus, DbS) was perhaps more controversial than we had originally thought. An exchange of views between Walasek et al. (2021) and André and de Langhe (2021a, 2021b) regarding Walasek and Stewart (2015) had captured our attention. While the authors did not fully agree, it seemed clear to us that there were methodological limitations to address. So, we changed our plans. It took two experiments (run almost in parallel) before we could come up with an experimental design that we felt was a sufficiently serious, and fair, test of DbS. This experiment became Study I. We then went back and (re)analysed the data from the first three experiments and added a fourth one. This became Study II. Study I allowed us to say, with some confidence, that an effect was not there even though DbS said it should be. Study II allowed us to say that we often seemed able to produce the effect when several options were presented in a list, replicating the results of Bird and Harris (2018). Here is the set-up we used in the experiments.

¹Of course, for this to work the subsequent options needed to be defined by the same attributes as the preceding options. This was always the case in our experiments and will not be discussed further here.

4.1.1 Design

All designs involved manipulating the set of attribute values that characterised the training options while holding the attribute values of the test options constant (“measurement invariance”, Meredith, 1993). See Table 4.1 for option domains and demographic groups in each experiment in the series. One group would be exposed to predominantly lower training attribute values, one to predominantly higher attribute values. That is, the participants’ recent experiences were manipulated. This was the independent variable of principal interest. We also, sometimes, manipulated the option domain between groups: half of the conditions would choose between options presented as representing monetary gambles. The other half would choose between numerically identical options presented as representing prospective jobs on the crowdworking website where we recruited participants. We thought of the former as more abstract and the latter as more contextualised, providing richer cues for long-term memory retrieval. Two experiments only used a single contextualised option domain. Lastly, we recruited “expert” and “non-expert” participants. The former were members of a demographic group that we *a priori* had reason to believe would have more expertise in the contextualised option domain. The latter were recruited from a group that we had reason to believe would have lower expertise in the contextualised option domain.

It is difficult to randomise expertise while holding other characteristics of participants’ backgrounds constant.² Consequently, different demographic groups differ in many ways, expertise being only one of them. Any of these non-manipulated characteristics could produce a difference in choice and judgement behaviour. When we then seek to interpret the difference as being due to differences in expertise, we might be attributing it to the wrong construct. However, it is perhaps more difficult to imagine characteristics that do not form part of our conception of expertise, yet happen to affect behaviour only in the domain of expertise.³ Any interpretation of the “main effect” of demographic group should thus be tentative, but any interaction effect involving demographic group and option domain should perhaps be a bit less tentative.

4.1.2 Procedure

The procedure was similar across all experiments in Study I and II, see Table 4.2. Participants read the instructions, responded to an attention check that tested whether they had read a second set of mock instructions, performed a

²An experimental psychologist might argue that this is in fact possible through repeated training in the tradition of Ebbinghaus (1913). I do not dispute this, but suspect that such expertise may be different in kind to that obtained from continual navigation of a natural environment. If we are after expertise in this latter sense, randomisation might be impractical.

³Perhaps it makes sense to think of this as a kind of “placebo analysis” (see Athey & Imbens, 2017; Eggers et al., 2024) - the interpretation of the effect as stemming from expertise is strengthened by confirming whether it *does not* appear when that interpretation suggests it should not.

Table 4.1. *Option domains and demographic groups for each experiment in the series*

Exp.	Appeared in Study	Option domain	Demographic
1	II	Stock funds	General population & bankers
2	II	Monetary gambles & AMT jobs	Low & High AMT crowdworking experience
3	I & II	Monetary gambles & AMT jobs	Low & High AMT crowdworking experience
4	II	Inflation adjustment to municipality budget	General population & municipality accountants

Note. Exp. stands for Experiment, in chronological order. AMT stands for Amazon Mechanical Turk, an online crowdworking platform.

two-alternative forced choice (2AFC) task for a number of trials, performed the test task(s), and responded to some demographic questions, a manipulation check question and some sincerity checks. There were two kinds of test tasks: (i) a small number of 2AFC trials (similar to the training phase trials; pairwise format) and (ii) a choice of a single option from a vertically arranged list (listwise format). The test phase consisted of (i) pairwise format test trials followed by (ii) a listwise format test trial, except for in Experiment 4 in Study II which only included a (ii) listwise format test trial. After each test choice, participants stated their strength of preference for the chosen option on a half-range scale (N. Weber & Brewer, 2003). We later realised that the half-range scale data from the listwise choices was a bit difficult to interpret: we had asked for a strength of preference for the chosen option vis-à-vis *all* other options. For that reason, we never analysed those data.

Table 4.2. *General procedure*

Phase 1	Phase 2	Phase 3	Phase 4	Phase 5
Instructions	Attention check	Manipulated training trials	Common test trials (choice & preference strength)	Demographics, manipulation and sincerity checks

Note. Procedure of the experiments in Study I and II.

4.2 Study I

4.2.1 Background

Study I aimed to carefully scrutinise a key prediction of DbS: that recently experienced attribute values influence subjective values and, consequently, preferences and choices. Specifically, being exposed to attribute values lower than some target attribute value should increase the apparent subjective value, as revealed by choice, of that target. We addressed three limitations of previous work that had become a topic of conversation in the debate. Firstly, previous research had fit theoretical models to data and interpreted differences in parameter values between conditions (e.g. Rakow et al., 2020; Walasek & Stewart, 2015, 2019). This is different to traditional statistical modelling where we design the model based on knowledge about the data generating process obtained from knowing the experiment design. The problem is that if the theory does not match the “true” data generating process, the parameter values may be biased (Nassar & Gold, 2013).

Secondly, the same previous work had violated measurement invariance (Meredith, 1993) by testing participants on different stimuli depending on condition. Thirdly, reanalyses that were aware of the two former limitations (Alem-paki et al., 2019; André & de Langhe, 2021a; Walasek et al., 2021) had picked out the trials that did not violate measurement invariance and not found the predicted effect. Since those trials were not originally intended to be analysed separately, they appeared at different points throughout the experiments. This effectively meant that testing was interleaved. Thus, even if the full set of attribute values was designed for DbS to predict an effect on subsequent choice, a given subset of attribute values that preceded a given trial might still not have the right distribution for an effect to be predicted. One could of course go back and check, but we felt that it was more straightforward to instead collect new data where we could guarantee the distribution on the preceding trials.

We also wanted to investigate the generality of any effect of recently experienced attribute values. To the extent that subjective values were indeed constructed in the manner postulated by DbS, we argued that this might be a process the mind defaults to when there are no existing subjective values that can support preferences and choices. When the options are contextualised, participants with substantial experience of that context might predominantly rely on their existing subjective values, thereby displaying an attenuated effect of recently experienced attribute values.

We implemented a “triple lock” consisting of three analyses that should reveal the predicted effect if it existed: (i) choices, (ii) strength of preference judgements reported on a half-range scale (N. Weber & Brewer, 2003), and (iii) model-free subjective value functions calculated as the cumulative preference strength judgement favouring one class of option across trials. Ensuring that all three analyses could be straightforwardly interpreted as aligning/not aligning with DbS required careful design of the stimuli. Together, (i)-(iii)

provided a good chance for an effect to be observed if it existed. For example, if there would not turn out to be an effect on choices, a proponent of DbS could argue that maybe there did exist an effect but it was too small to flip any decisions. If so, an effect should still emerge in the strength of preference judgements which were more fine-grained and thus more sensitive. Alternatively, if all three analyses spoke against the predicted effect, we could be fairly confident that it truly was not there.

4.2.2 Methods

The experiment was a 2x2x2 between-groups factorial design along the lines described in Section 4.1. The factors were Training (low/high values), Option domain (monetary gambles/AMT jobs), and Experience (low/high experience of completing jobs on Amazon Mechanical Turk, abbreviated as “AMT”). To obtain one group with higher and one with lower experience of the contextualised option domain, we posted one data collection that was only visible to AMT accounts that had been used to perform either fewer than 50 (low experience) or more than 5 000 (high experience) AMT jobs. We excluded participants who during Phase 5 (see Table 4.2) stated an experience in performing AMT jobs that did not roughly match that of their account information (presumably indicating that they were using someone else’s account or owned several accounts). This resulted in 300 valid low experience participants and 428 valid high experience participants for a total of 728 valid participants, out of 1020 originally recruited.

Training and test stimuli were as displayed in Figure 4.1. There were 60 training trials and 6 test trials. Apart from addressing the previously mentioned methodological limitations through the experiment design, these stimuli were designed to also address the following possible loop holes. Firstly, they held range of training values (as good as) constant. This is important because unlike in range-frequency theory (e.g. Parducci, 1965), it is only the number of exemplars above or below the target that should matter. If an effect is produced by manipulating the range, then that effect cannot be explained by DbS. Secondly, we held training probabilities constant and tried to minimise the impact that they, according to DbS, would have on choice (see note to Figure 4.1). Thirdly, we used simulations to verify that DbS predicted an effect given our training and test stimuli and that this effect was not dependent on which one of three reasonable “preference mechanisms” is assumed. DbS suggests that the results of the ordinal comparisons are integrated additively and sequentially (Stewart, 2009; Stewart & Simpson, 2008). In closed form, this can be written as Equation 2.2 being evaluated for each attribute and the results added together. However, one could imagine a version where each attribute is valued separately and the valuations then multiplied. One could also imagine that the attributes are valued separately and the average of those valuations is

the valuation of the entire option. We designed the stimuli such that DbS predicts an effect regardless of which of these mechanisms is assumed. Lastly, we implemented the triple lock mentioned earlier.

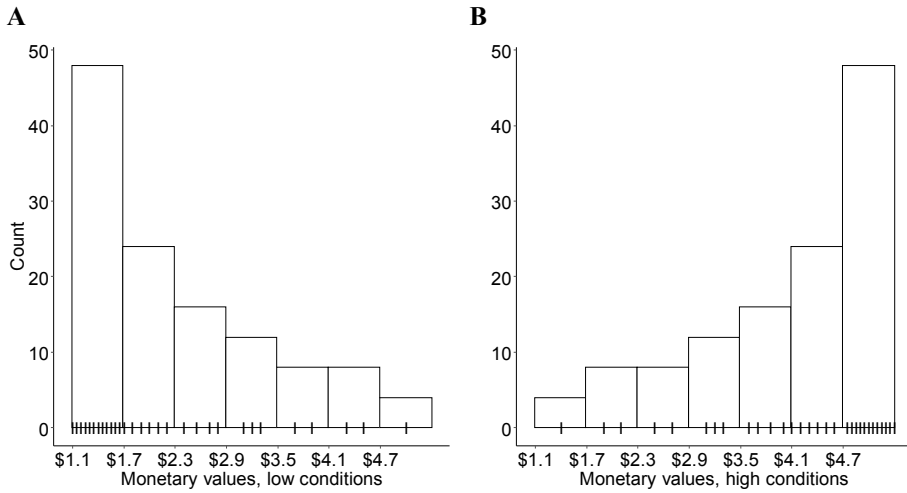
4.2.3 Results

Bayesian three-way ANOVAs (JASP Team, 2023) indicated strong evidence in favour of the point null hypothesis of no difference between the training conditions, both for the proportion of risky options chosen ($BF_{10} = 0.082$) and for strength of preference for the risky option ($BF_{10} = 0.079$). For the choices, there was also moderate evidence against all interactions between Training and Option domain and/or Experience ($BF_{10} \leq 0.270$). For the judged preference strengths, there was inconclusive evidence against an interaction between Training and Option domain ($BF_{10} = 0.956$), but otherwise moderate evidence against interactions with Training ($BF_{10} \leq 0.363$). There was also strong evidence against a difference between the two Experience groups for both dependent variables ($BF_{10} \leq 0.099$). The only effects in these data were that participants were more likely to choose and had a stronger preference for the risky options when they were described as prospective AMT jobs ($BF_{10} > 280000$), rather than monetary gambles. However, this apparent main effect of option domain was mainly driven by its interaction with experience ($BF_{10} > 2400$): the difference between option domains was limited for high experience participants and substantial only for low experience participants.

These (absent) effects are most clearly displayed in the averaged model-free subjective value functions (Figure 4.2. For details on how the functions were computed, see p. 8224-8225 in Study I). The only clear separation is between the AMT jobs and monetary gambles for the low experience participants. Let me lean into an additional result that is not captured by the Bayesian ANOVAs: the coloured lines in Figure 4.2 are individual level subjective value functions. These do not coalesce nicely around the means; indeed, they do not seem to coalesce much at all. Such visual impressions will be affected by design choices made when constructing the figure (see e.g. Franconeri et al., 2021; Hullman & Diakopoulos, 2011). Even so, it seems clear that there exists substantial heterogeneity.

4.2.4 Discussion

The main take home message of Study I is that the effect that DbS predicts does not seem to exist in this 2AFC task. We can claim this with some confidence because we addressed the previously mentioned limitations of previous work, designed stimuli that should be sufficient for the effect to emerge, verified this sufficiency using simulations, and provided three tests that should allow us to spot any appreciable tendency towards an effect. This does not mean



	Safe option	Risky option
Test trial 1	\$1.1 × 100%	\$1.7 × 90%
Test trial 2	\$1.7 × 100%	\$2.3 × 90%
Test trial 3	\$2.3 × 100%	\$2.9 × 90%
Test trial 4	\$2.9 × 100%	\$3.5 × 90%
Test trial 5	\$3.5 × 100%	\$4.1 × 90%
Test trial 6	\$4.1 × 100%	\$4.7 × 90%

Figure 4.1. Training values and test options used in Study I

Note. The sets of monetary values used in the Training conditions. Panel A: Monetary values used in the Low training conditions. Ticks indicate individual values used. Panel B: Monetary values used in the High training conditions. Each set of monetary values was combined with the same set of probabilities to produce options. The set of probabilities consisted of the following values: 60% and 61% appeared six times each. Each value from 62% to 89% in 1% increments appeared eight times each. 90% and 100% each appeared twice. This held probabilities constant between training conditions and increased the strength of the training values manipulation: remember, 90% and 100% are the highest values in the set of probabilities. 100% would only win two more ordinal comparisons if compared to all attribute values in the training set. Although there, according to DbS, would be additional exemplars in long-term memory that would make 100% of higher subjective value than 90%, any exemplar (but two) sampled from recent experiences would dilute that difference. The (manipulated) monetary values should therefore more often serve as the deciding factor. Panel C: Test trial option pairs. Each test trial consisted of one option involving a certain gain (“safe option”) and one option involving a probabilistic gain (“risky option”).

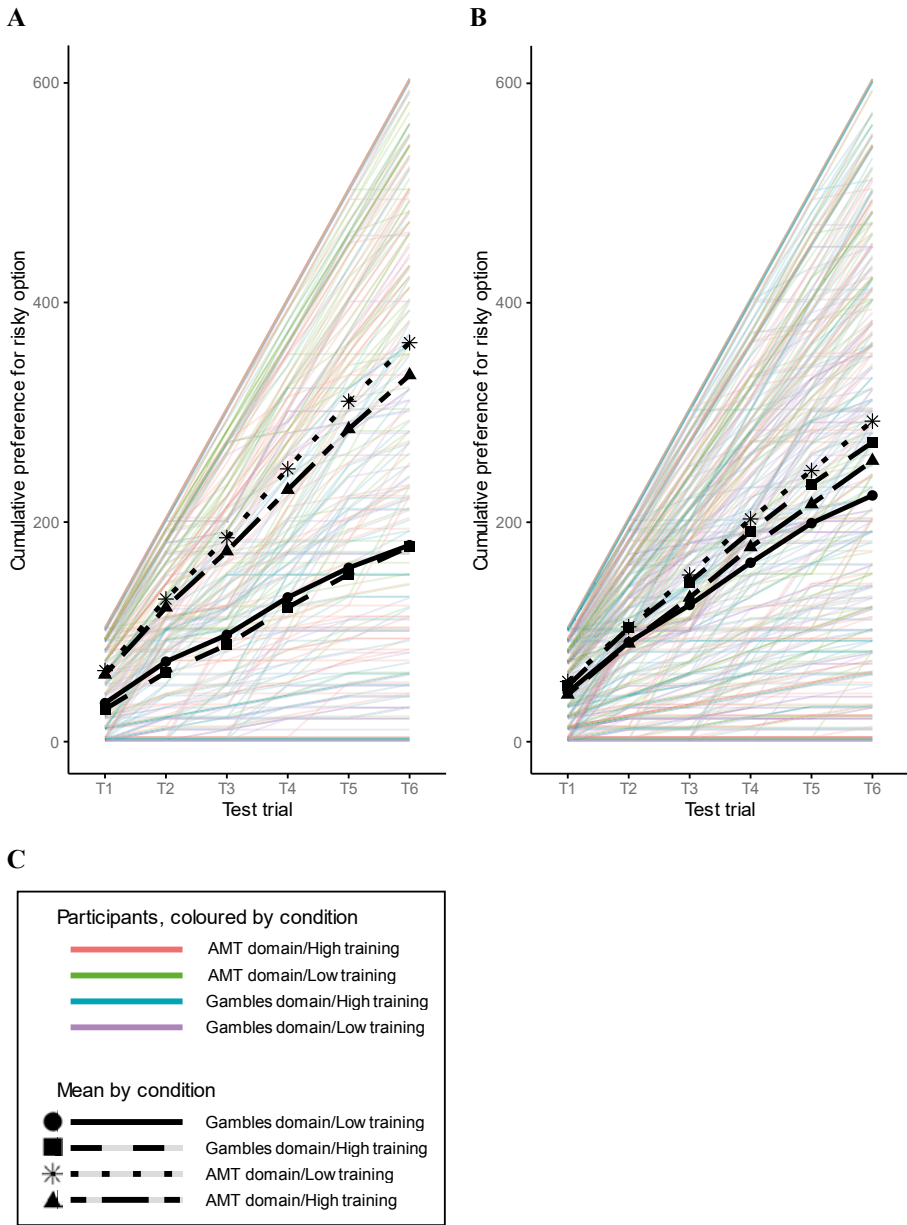


Figure 4.2. Model-free subjective value functions from Study I

Note. Panel A: Model-free subjective value functions for individual and averaged low experience participants by condition. Panel B: Likewise but for high experience participants. Panel C: Legend common to Panel A and B.

that preceding options can never affect subsequent choice; just that it did not do it when DbS said it should. Because we did not find an effect of recently experienced attribute values to begin with, we were unable to investigate our hypothesis that relevant experience would attenuate it.

The only substantial effect was that participants who had low experience of performing AMT jobs tended to choose differently depending on whether numerically identical options were contextualised as either monetary gambles or prospective AMT jobs. For participants with high experience of AMT jobs, the same effect was tangibly smaller. One speculative explanation would be that the participants' beliefs about the generative process was the cause (cf. Forsgren, Juslin, & van den Berg, 2025). We presented participants with frequentist probabilities of receiving a reward that we claimed were based on historical data on real AMT jobs. However, an unconditional frequentist probability may be different to the probability when conditioning on some individual. There is no logical contradiction in saying that the historical frequency of a randomly drawn AMT worker being rewarded is 90%, whereas I have a 95% probability of being rewarded if I were to perform the same job. The data generating process might differ for different people! A logical contradiction - if one exists - can only be demonstrated on group level, such as when more than 50% of car drivers say that they are more skilled than the median driver (Svenson, 1981). For aleatory devices like monetary gambles, however, this potential difference between the group and the individual is not the case. A (well-functioning) roulette wheel has got the same probability of turning up red regardless of who spins it. If participants recognised that difference, and a tendency to overestimate one's personal ability to sway the odds was more prevalent amongst low experience participants than high experience participants, this could explain both the main effect and the interaction. This does not seem unreasonable - one reason why we chose AMT jobs as the contextualised option domain was that the believed capriciousness of AMT job posters (e.g. McInnis et al., 2016) meant that we could credibly present the outcomes as probabilistic in nature. It might take extensive experience with AMT jobs to realise that there indeed exists some stochasticity that is beyond one's control.

The high heterogeneity in reported preference strengths (Figure 4.2) suggests that responses are not primarily determined by the experimental manipulations, but there is also no consensus between participants on the (apparent) relative subjective value of the options. Some participants always prefer the safe option with utmost clarity, some always prefer the risky option with utmost clarity, and most are in between. DbS, and many other theories, could in principle explain such results by reference to idiosyncrasies in participants' learning histories. Yet, the heterogeneity in judgement and choice, rather than the mean differences, should maybe receive particular attention in future work (cf. Sundh et al., 2025).

4.3 Study II

4.3.1 Background

In Study I, we had stuck to the two-alternative forced choice (2AFC) format that had been employed in some of the perhaps most famous previous studies on DbS (e.g. Stewart et al., 2015; Ungemach et al., 2011). As mentioned in Section 3, Bird and Harris (2018) had also used a different format: an “investment mean variance measure” (see Appendix 3, Vlaev et al., 2009). See also Vlaev et al. (2007a, 2007b). This was a list of hypothetical pension funds that varied in their proportions of (more volatile) stocks and (less volatile) bonds, thereby intended to disclose risk preferences. In their Experiments 1 and 2, Bird and Harris (2018) hypothesised, and confirmed, that being exposed to higher percentages of stocks in preceding options would cause an increase in the average percentage of stocks in the option chosen from a subsequent list. They interpreted this as consistent with DbS, Adaptation Level Theory (Helson, 1947), and Range Frequency Theory (Parducci, 1965, 1995). More generally, it supported “the proposition that prospects in the immediate choice environment are evaluated relative to recently experienced samples.” In Study II, we analysed the full series of experiments we had performed, including data on choice in the “listwise” format from Bird and Harris (2018).

Study II provided tentative answers to the following questions. Firstly, to what extent was the lack of effect of preceding options on two-alternative forced choices (2AFCs) we observed in Study I a general phenomenon, rather than being tightly connected to the precautions we took to target DbS specifically? Secondly, did it hold also in the listwise format? Thirdly, while we at the time felt that separating participants based on how many AMT jobs they had performed was fairly clever, it could be argued to reflect a fairly mild difference in expertise. Maybe we could observe the attenuation we had expected if the level of expertise was just high enough? Lastly, as is often the case in psychology, we had been and would be working with choice and judgement data. Such data violate assumptions of regular ANOVA because the sample space is discrete and bounded, rather than continuous and unbounded. There exists a vast statistical literature that investigates when those violations will matter (e.g. Knief & Forstmeier, 2021; Lumley et al., 2002; Schmidt & Finan, 2018; Yang et al., 2019). While such theoretical work is certainly important for establishing consensus about best-practices, we saw Study II as an opportunity to avoid certain violations altogether by fitting Bayesian Generalised Linear Mixed Models (GLMMs). These respect the binary and ordinal nature of the dependent variables that we would be working with, making the most of our data and showing whether the results in Study I were driven by our choice of statistical model.

4.3.2 General methods

We distinguished between “pairwise” and “listwise” choices. The former were 2AFCs, the latter were vertically arranged lists of options presented in random order. See Figure 4.3. While the pairwise format had been used in previous studies on Decision by Sampling (e.g. Stewart & Simpson, 2008; Stewart et al., 2015; Ungemach et al., 2011), we had (at the time) only seen the listwise format in Bird and Harris (2018). The experiments progressed as per Table 4.2. Phase 4 consisted of common pairwise test trials followed by a listwise test trial. To facilitate communication, we designated a “target attribute value” which was manipulated in the training phase and according to which we ranked the options from lowest to highest. In Experiment 1 and 4, manipulating the target attribute value also involved manipulating the other attribute value(s) that defined the options. In Experiment 2 and 3, this was not the case. There, we held the other attribute value (probability of receiving the target attribute value) as good as and exactly constant, respectively, between conditions.

Bayesian generalised linear mixed models common to Experiment 1 to 4

We fit Bayesian logistic regression models to the pairwise choice data and Bayesian ordinal logistic regressions to the listwise choice data. For the logistic regression models, we included a participant-level random effect to account for the fact that we have several pairwise test trials (repeated measures). We also included item-level fixed effects. These accounted for the fact that each test trial item is different to the others, a source of variation that should be modelled.⁴ We used recommended weakly informative priors (Vehtari, 2025). We used sum coding rather than dummy coding. This allows us to interpret the results as main and interaction effects, just like in an ANOVA. In addition to these analyses, we performed (non-parametric) Mann-Whitney tests. These generally aligned with the GLMMs and will not be further discussed here.

4.3.3 Experiment 1

Experiment 1 was an initial attempt at conceptual replication of Bird and Harris (2018) within the same domain of expertise (investment decisions). Would we be able to observe an effect of preceding options on subsequent choice in the pairwise and listwise format, respectively? If so, would it be attenuated when experts made decisions within their domain of expertise?

Methods

The experiment was a 2x2 between-groups factorial design with the factors Training (low/high percentages of stocks) and Expertise (general population/banker). We recruited 87 experts - mainly bankers - and 73 participants from a general

⁴While I will not present any formal robustness checks, these fixed effects did not matter for the conclusions.

A

Which fund would you prefer to invest in?

Fund 1 Stocks: 35% Preferred stocks: 65% Min: 188 000 Average: 296 500 Max: 450 500	Fund 2 Stocks: 0% Preferred stocks: 100% Min: 230 000 Average: 230 000 Max: 230 000
---	---

B

Out of the following funds, please select the one you would prefer to invest in.

- 80% stocks, 20% pref., min: 134 000, average: 382 000, max: 734 000
- 5% stocks, 95% pref., min: 224 000, average: 239 500, max: 261 500
- 95% stocks, 5% pref., min: 116 000, average: 410 500, max: 828 500
- 65% stocks, 35% pref., min: 152 000, average: 353 500, max: 639 500
- 35% stocks, 65% pref., min: 188 000, average: 196 500, max: 450 500
- 20% stocks, 80% pref., min: 260 000, average: 268 000, max: 356 000

Figure 4.3. Example of pairwise and listwise choices

Note. Panel A: Example of pairwise task from Experiment 1. Panel B: Example of listwise task from Experiment 1. Note the typos: the average return of the fifth option from above should be 296 500 rather than 196 500. The minimum return of the final option from above should be 206 000 rather than 260 000. While these typos did not change how the options ranked in terms of spread between the minimum and maximum return, and while the proportions of stocks and preferred stocks were correct, this is an inelegance of Study II. There were no typos in the other experiments.

population. The options were different hypothetical funds consisting of stocks and preferred stocks, in proportions from 0% to 100% in 5% increments. Thus, no option dominated another but participants faced a trade-off between higher expected earnings and higher spread in earnings. We excluded the 50/50 option (half each of stocks and preferred stocks). This yielded 20 different options. Following Bird and Harris (2018), who in turn followed Vlaev et al. (2009), the options were defined by their proportion stocks, proportion preferred stocks, minimum return, mean return, and maximum return. See Figure 4.3. To generate the options, we first recorded the (i) minimum, (ii) mean, and (iii) maximum return of all stocks and (iv) average return of all preferred stocks traded on the Nasdaq OMX Stockholm exchange. Each option was then characterised by a proportion of (risky) stocks v. (risk-free) preferred stocks. The options' stated minimum, mean, and maximum returns were calculated as weighted averages between (i) and (iv), (ii) and (iv), and (iii) and (iv), respectively, depending on the proportion of stocks. The options thereby consisted of different trade-offs between potential return and variability in return. We told participants that there was a probability greater than 99% that the return would fall in between the minimum and maximum and that the options were based on the Nasdaq OMX Stockholm exchange. We split this range of options into a low and a high training phase option set, see Figure 4.3 for the distributions and test stimuli.

Results

The distribution of training options one had experienced did not appear to have any effect on subsequent choice in the pairwise format (Odds ratio: 0.97. 95% credible interval: 0.69, 13.4). The experts tended to choose options with a higher percentage of stocks (OR: 1.43. CI: 1.11, 1.86.). There was no evidence of an interaction (OR: 1.12. CI: 0.81, 1.55). In the listwise format, preceding options did seem to affect subsequent choice in the same direction as in Bird and Harris (2018). Being exposed to high proportions of stocks in the training phase seemed to cause participants to choose an option with a higher proportion of stocks from the list (OR: 1.60. CI: 1.19, 2.16). There was no clear evidence of the experts tending to select options with more stocks (OR: 1.29. CI: 0.97, 1.71). Interestingly, there was no particular evidence of the attenuation that we had expected (OR: 1.11. CI: 0.84, 1.48).

Discussion

We succeeded in replicating Bird and Harris (2018): preceding options did seem to affect listwise choice such that participants previously exposed to options with higher expected returns tended to select an option with a higher expected return. However, we did not uncover a similar effect on pairwise choices. We did not find any particular evidence for the effect, when it emerged, being attenuated for experts making choices within their domain of expertise.

Table 4.3. *Stimuli for Experiment 1*

Option	Low training values	High training values	Listwise test option	Pairwise test option
0% stocks, 100% pref.	8			
5% stocks, 95% pref.	8		Yes	
10% stocks, 90% pref.	8			
15% stocks, 85% pref.	8			
20% stocks, 80% pref.	4		Yes	
25% stocks, 75% pref.	4			Item 1
30% stocks, 70% pref.	4			Item 2
35% stocks, 65% pref.	2		Yes	Item 1
40% stocks, 60% pref.	2			Item 2
45% stocks, 55% pref.	2			
55% stocks, 45% pref.		2		
60% stocks, 40% pref.		2		Item 3
65% stocks, 35% pref.		2	Yes	Item 4 & Item 5
70% stocks, 30% pref.		4		Item 3
75% stocks, 25% pref.		4		Item 4
80% stocks, 20% pref.		4	Yes	Item 5
85% stocks, 15% pref.		8		
90% stocks, 10% pref.		8		
95% stocks, 5% pref.		8	Yes	
100% stocks, 0% pref.		8		

Note. Number of times each option appeared in the training phase and whether the option appeared in the listwise and pairwise test trials, by condition. “Preferred stocks” is abbreviated “pref.” for legibility. Item number indicates which options were paired to form a pairwise choice item.

4.3.4 Experiment 2

Experiment 2 concerned the same research questions as Experiment 1. We introduced certain methodological improvements that allowed a clearer interpretation of any results in terms of attribute values being weighted relative to their rank amongst recently experienced attribute values. Specifically, we limited ourselves to two different attributes, holding one of them constant while manipulating the frequency but not the range of values of the other attribute. To achieve this, we returned to an option domain which had been used in previous studies on DbS (e.g. Stewart & Simpson, 2008; Stewart et al., 2015; Ungemach et al., 2011): monetary gambles. To investigate domain expertise, we also presented numerically identical monetary gambles recontextualised as prospective tasks on the crowdworking website on which we would recruit participants (Amazon Mechanical Turk; AMT).

Methods

The experiment was a 2x2x2 between-groups factorial design with the factors Training (Low/High values), Option domain (monetary gambles/AMT jobs), and Experience (low/high experience of completing jobs on AMT). We recruited 805 participants on AMT through two simultaneous data collections, one only visible to AMT users whose accounts had completed more than 5 000 jobs on the website and one visible to accounts that had completed fewer than 50 jobs. After the experiment, we asked participants how many jobs they had performed on AMT during their lifetime to check that the participant was truly experienced and had not, for example, borrowed someone else's account or registered a second one. If they self-reported having completed fewer than 1 000 AMT jobs/more than 100, we excluded them from the high/low experience condition, respectively. This yielded a final result of 352 high experience participants and 309 low experience participants, for a total of 661 valid participants.

In Experiment 1, we had manipulated the frequency of high and low attribute values in the training sets, but also the range. We realised that this confounded the interpretation of any effect of training. Remember, as far as DbS is concerned, rank rather than range is all that matters. Therefore, in Experiment 2, we made sure all participants were exposed to all options in the training set at least once (Table 4.4A). We also held probability *near* constant (Table 4.4B). The pairwise test stimuli were between (reward/probability of reward): \$2.35/97.5% or \$3.40/80%; \$5.45/97.5% or \$7.50/80%; \$6.45/97.5% or \$8.50, 80%. After each test item, participants stated their strength of preference for the chosen option on a half-range scale (N. Weber & Brewer, 2003).

Results

In the pairwise format, we found no clear evidence of effects of preceding options on subsequent 2AFCs (OR: 0.97. CI: 0.69, 1.34). However, there

Table 4.4. *Stimuli for Experiment 2***A**

Reward value	Low training values	High training values	Listwise test option
\$0.30	30	1	\$0.30 · 100%
\$1.35	13	1	\$1.35 · 97.5%
\$2.35	9	1	\$2.35 · 95%
\$3.40	1	1	\$3.40 · 92.5%
\$4.40	1	1	\$4.40 · 90%
\$5.45	1	1	\$5.45 · 87.5%
\$6.45	1	9	\$6.45 · 85%
\$7.50	1	13	\$7.50 · 82.5%
\$8.50	1	30	\$8.50 · 80%

B

Probability	Low training values	High training values
80%	6	7
82.5%	6	7
85%	6	7
87.5%	6	7
90%	6	6
92.5%	7	6
95%	7	6
97.5%	7	6
100%	7	6

Note. Panel A: Number of times each reward value appeared in the training phase by condition and the options that appeared in the listwise test phase. Panel B: The number of times each probability appeared in the training phase by condition.

appeared to be a floor effect: more than half of participants always selected the 2AFC option with lower potential reward (and thus higher probability of receiving the reward). We therefore focused on the strengths of preference for the chosen pairwise options that participants reported after each choice. These data did not suffer from any floor effect. There appeared to be an effect of training on strength of preference, but only for the low experience group when choosing between monetary gambles: in this subgroup, those who have been exposed to high training values subsequently tended to report a higher preference strength for the option with a higher potential reward (see Appendix S4 in Study II).

In the listwise format, however, we did conceptually replicate the finding of Experiment 1 and Bird and Harris (2018): there was evidence that participants who had been exposed to high values in the training options tended to select an option with a higher reward (OR: 1.26. CI: 1.09, 1.44). There was no clear evidence of any other effects in this format.

Discussion

In the pairwise format, we again did not find the expected general effect of preceding options. While the 2AFC data suffered from an unfortunate floor effect, where participants tended to select the option with the highest probability of a payoff, the judged preference strengths for the 2AFCs revealed an interaction. There appeared to be an effect of training values for one of four subgroups (low experience participants for whom options were presented as monetary gambles) on preference strengths. One could speculate why this may be, but ultimately I have no sound theoretical argument on which to anchor such speculations. On the other hand, it is perhaps too convenient to just chalk up any unexpected finding as a spurious result of multiple testing. Experiment 3 would arbitrate on this matter by providing a replication with certain methodological improvements.

In the listwise format in Experiment 2, we again found an effect of preceding options: participants were more likely to select an option with a higher reward from the list if they had been exposed to options with higher rewards. Contrary to our expectations, there was again no evidence of an attenuating effect of expertise.

4.3.5 Experiment 3

Experiment 3 investigated the same research questions and hypotheses as Experiment 2. It used the same stimulus domains and participant pool, but shaped up the limitations reported above. Fundamentally, it was an improved version of Experiment 2 that had the potential to provide more clear-cut evidence in favour of or against the DbS theory.

Methods

Methods were as in Section 4.1 and Study I. The listwise test options were \$1.10/90%, \$1.70/85%, \$2.30/80%, \$2.90/75%, \$3.50/70%, \$4.10/65%, and \$4.70/60%.

Results

The Bayesian logistic regression on the pairwise choices confirmed the results of the Bayesian ANOVAs reported in Study I: there was no main effect of training (OR: 0.99. CI: 0.86, 1.13). The only effects for which there was clear evidence were those reported in Study I and apparent in Figure 4.2: a main effect of option domain (OR: 2.26, CI: 1.96, 2.62) and its interaction with experience (OR: 0.65. CI: 0.57, 0.75). The judged preference strengths mirrored the results for the choices - the three-way interaction in Experiment 2 did not replicate in Experiment 3.

For the listwise choices, we found no clear evidence of any effects apart from the interaction between experience and option domain (OR: 0.74. CI: 0.65, 0.84), mirroring the same interaction in the pairwise format. That is, we could not replicate the effect from Bird and Harris (2018) this time around.

Discussion

Firstly, we could confirm the stated results of Study I using a different, arguably superior, analysis: there did not appear to be any effect of manipulating the distribution of recently experienced attribute values on subsequent 2AFCs or preference strengths. Secondly, we did not replicate the finding from Experiment 2 that low experience participants' judged preference strengths were affected by training. This supports the notion that it was either a spurious or not very robust finding, being sensitive to minor methodological changes. Thirdly, the effect of training on subsequent choice in the listwise format that we had observed in Experiment 1 and 2 was absent. There again seems to be two possible explanations of this: sampling error or differences in design between the otherwise highly similar Experiments 2 and 3. Lastly, because we did not observe any effect of training we could not tell if the attenuating effect of expertise that we had hypothesised was there.

4.3.6 Experiment 4

Through connections, we had gained the opportunity to collect data with Kommunalekonomernas förening - the Swedish Association for Municipality Accountants. We ran two simultaneous data collections where participants were randomised to either Experiment 4 or a different branch with two shorter experiments that fall outside the topic of this thesis.

We had so far uncovered two supposed insights. Firstly, it appeared possible to produce an effect of training in the listwise format, at least sometimes.

Secondly, when it appeared, this effect seemed largely robust across different demographic groups (that differed in expertise but of course also in other ways). Experiment 4 was a return to the conceptual replication of Bird and Harris (2018) in Experiment 1, but with an option domain tailored to relate to the municipality accountants' expertise. After discussions with our two contact persons at the association, we concluded that budgeting was something all or next to all municipality accountants would find themselves doing. In such budgeting, an important step is to select an appropriate level of inflation adjustment - in essence, a percentage increase in some part or all of the budget to account for rising costs. If the inflation adjustment ends up being lower than the eventual inflation, we have a net decrease of the real amount budgeted. So, participants had to express a kind of vicarious preference - choosing on behalf of their municipality rather than themselves - for different levels of inflation adjustment.

Our participants would have substantial - often decades-long - experience of making such budgets. They would also have substantial knowledge about the workings of their municipality and the specific conditions it faces. All of this could be used to support their choice. Participants recruited from the general population - or at least from a population with no particular experience of working within a municipality - would instead have very little to fall back on to support their choice.

Methods

The experiment was a 2x2 between-groups factorial design with the factors Training (Low/High values) and Expertise (General population/Municipality accountant). We only included a listwise format test item and no pairwise format test items. We recruited 395 municipality accountants within the data collection, 189 of which were randomised to Experiment 4 (rather than to the other branch intended for a different study). These participants had a median of 4 years of university or college education, with an interquartile range of 0. 150 of them worked in a municipality as a municipality accountant, 9 worked as a municipality accountant in a region (a "secondary municipality" primarily responsible for running health services), 22 worked in neither, 7 worked in "other" and 1 person worked both in a municipality and a region. They had been with their current employer for 5 years in the median (Q1: 3, Q3: 11). After the experiment, we asked the participants to what extent they had "used knowledge and experience from their everyday working life" when completing the task. There are well-known problems with comparing responses to Likert scale questions of this kind between groups (see e.g. Birnbaum, 1999; Chituc et al., 2026). Yet, one might take some comfort in that the median response was 5 on a seven-point Likert scale (Q1: 4, Q3: 6). All in all, our municipality accountant participants appeared to have substantial expertise.

We also recruited 201 general population participants from Prolific, an online crowdsourcing platform designed specifically for recruiting research par-

participants for online studies (Palan & Schitter, 2018; Peer et al., 2017).⁵ Participants had to be located in Sweden and speak fluent Swedish to be eligible. None of them currently worked for a municipality as a municipality accountant. 1 had done so previously for 1 year. Another had worked as a municipality accountant both for a municipality and a region for 2 years each. The median response to the question of whether they had used working life knowledge and experience when completing the task was 2 (Q1: 1, Q3: 4). It seemed like this sample had hardly any experience of municipality accountancy.

The stimuli were different levels of inflation adjustment, defined by the adjustment in percent and the budget cuts in percent that the adjustment would entail. In practice, this meant that we implied a forecasted inflation level for the participants to assume. This was because inflation forecasting is performed centrally by Sveriges Kommuner och Regioner - the Swedish Association of Local Authorities and Regions. The actual cost increases faced by a specific municipality might of course differ from the forecast and a savvy municipality accountant might be able to predict whether or when they will. Such intricacies were glossed over here. Training and test stimuli are displayed in Table 4.5.

Table 4.5. *Stimuli for Experiment 4*

Option	Low training values	High training values	Listwise test option
1.8% adjustment, 4.0% cuts	5		Yes
2.2% adjustment, 3.6% cuts	5		Yes
2.6% adjustment, 3.2% cuts	5		Yes
3.0% adjustment, 2.8% cuts	5		Yes
3.4% adjustment, 2.4% cuts	5		Yes
3.8% adjustment, 2.0% cuts	5	5	Yes
4.2% adjustment, 1.6% cuts		5	Yes
4.6% adjustment, 1.2% cuts		5	Yes
5.0% adjustment, 0.8% cuts		5	Yes
5.4% adjustment, 0.4% cuts		5	Yes
5.8% adjustment, 0% cuts		5	Yes

Note. Number of times each option appeared in the training phase by condition and whether it appeared in the test phase.

Results

We found the strongest main effect of preceding options on subsequent choice so far (OR: 2.02. CI: 1.66, 2.44), such that having been exposed to options with a high adjustment made participants select an option with a higher adjustment. For the first time, we also identified the kind of attenuation that we

⁵We originally intended to recruit 200 participants but one seemed to have forgotten to collect their reward, leading to their slot being reopened and filled by a new participant.

had hypothesised - the effect was smaller for municipality accountants than for the Prolific sample (OR: 0.83. CI: 0.70, 0.99). There was also evidence that municipality accountants tended to select a higher adjustment regardless of condition (OR: 1.53. CI: 1.28, 1.82).

Discussion

In Experiments 1 and 2, we did not find that expertise attenuated the effect of preceding options on subsequent listwise choice, In Experiment 3, we found no evidence of any such effect and consequently could not identify any potential attenuation. Experiment 4 arguably provided the strongest “dose” of expertise yet, using an option domain that we have reason to believe was quite alien to the Prolific sample and well-known to the municipality accountants. When doing so, we did finally observe an attenuating effect of expertise - at least in this setting (cf. Cartwright, 2013). It is worth noting, however, that this attenuation is far from complete. The effect is smaller, but still quite clear.

4.3.7 General discussion

These results seem to suggest that the effect of preceding choices investigated in Bird and Harris (2018), Stewart et al. (2015), and Ungemach et al. (2011) and elsewhere is dependent on, or at least facilitated by, presenting options in a listwise format. When the effect emerges, it goes “from high to high”: being exposed to preceding options that primarily have higher values on some target attribute causes an increased probability of selecting an option from the list that has got similarly high values of the same attribute. This replicates the finding from Bird and Harris (2018). There seems to exist circumstances under which expertise can lead to an attenuation of the effect, although we should keep in mind that expertise was not manipulated and remains confounded with other demographic differences.

It is important to recognise the heterogeneity in behaviour. While we can typically see effects in the listwise format when aggregating over hundreds of participants, this translates to an increase in the odds rather than a universal separation. (This appears to mirror Bird & Harris, 2018, see e.g. their Figure 3). The greatest main effect we find is an odds ratio of 2.02. Thus, while listwise choices are sometimes *affected* by the training phase there are clearly other important factors influencing choice too.

4.4 Study III

4.4.1 Background

Studies I and II were ultimately concerned with subjective value. In Study III, we shifted to studying preference.

A received view within psychology is that preferences are “often” intransitive (e.g. American Psychological Association, 2025; Bikhchandani & Segal, 2021; Brandstätter et al., 2006; Edwards, 1954; Fishburn, 1991; Kalenscher, 2010; Lee et al., 2009; Rieskamp et al., 2006; Veit, 2024; Walasek & Brown, 2023). Some economists subscribe to it too (e.g. Grether & Plott, 1979; Schoemaker, 1982; Starmer, 2000). We provided evidence against this received view in Study III.

As has famously been shown by Regenwetter et al. (2010, 2011), the literature on transitivity of preferences suffers from serious analytical and methodological limitations. While their critique is wide-ranging, the core of it is perhaps that one must use appropriate modelling methods to make inferences about properties of (covert) preferences based on (overt) choices.

Because of these limitations, there is only a small number of studies that can be marshalled in support of (in)transitivity of preferences. Those studies have almost⁶ exclusively used monetary gambles as stimuli (see Birnbaum, 2020; Cavagnaro & Davis-Stober, 2014; Regenwetter et al., 2011). Here are some reasons why such results may not generalise to the kinds of stimuli we face in our everyday lives:

1. **Risky decision making is the exception, not the rule.** Few, if any, decision making situations outside casinos, computer or board games, and laboratories involve options defined by exhaustive descriptions of the outcomes and probabilities of said outcomes. Usually, one or both of these are uncertain *sensu* Knight (1957/1921).
2. **Monetary gambles are novel options.** Because they hardly, if at all, exist outside of the laboratory, monetary gambles will be new to many participants. Those who have experience of participating in experiments might have come across *some* monetary gambles before, but they are unlikely to be familiar with the exact gambles used in a particular experiment.
3. **Tailored stimuli.** Previous studies have often tailored their stimuli in order to provide existence proofs of some theory of preference construction (e.g. Müller-Trede et al., 2015; Tversky, 1969).
4. **Confounding preferences with mathematical ability.** There is evidence that choice behaviour for monetary gambles and in related tasks correlate with numeracy and other measures of cognitive ability (Mill-

⁶The exceptions that I am aware of are the following. Firstly, the “noncash gambles” in Regenwetter et al. (2011), where participants faced lotteries with outcomes like “about 7 paperback books”. This condition revealed evidence against intransitive preferences for every participant. Secondly, Müller-Trede et al. (2015) who used hypothetical sound systems defined by three attributes. However, that study might have used too few stimuli to allow discrimination between the transitive models they implement, see e.g. p. 183-184 in Zwilling et al. (2019). In addition, Experiment 2 in Tversky (1969) used hypothetical job applicants defined by their ranking on three attributes, which participants were told “may not be perfectly reliable”. I am not aware of any reanalysis of those data.

roth et al., 2019, 2020). This raises the question of to what extent such behaviour reflects the difficulty in comprehending what the options entail and identifying what option one prefers (cf. Oprea, 2024).

5. **Results for composite options might not generalise to unitary options.** Tversky (1972) and Luce (1959) distinguished between options that are evaluated attribute-by-attribute (composite) and those evaluated as wholes (unitary). We should perhaps not presume that the same cognitive mechanisms support preferences for both kinds.

Items 1, 2, and 3 suggest that (tailored) monetary gambles might be a kind of stimulus for which participants will not have had the opportunity to develop existing preferences. Item 3 suggests that experimenters might have selected stimuli that participants are particularly unlikely to have existing preferences for. Item 4 suggests that they might struggle with expressing any existing preferences, if they do exist. Item 5 suggests that monetary gambles might interrogate a different data generating process than unitary stimuli would (cf. Forsgren, 2025). Together, these items might yield an experimental situation where participants are particularly unlikely to rely on any existing preferences they may have and instead have to construct preferences to solve the task.

We speculated that such existing preferences could be supported by conceptual knowledge of the subjective values of different options. (I return to this idea, which was crystallised and fleshed out gradually over the course of the project, in the General discussion, see Section 5.4). This speculation led to the suspicion that preferences would tend to be strongly stochastically transitivity (SST, see Definition 1). That descriptive model can be interpreted cognitively as a “random utility model”: a straight choice of whichever option has the highest subjective value, where retrieval of existing subjective values is corrupted by symmetric, homogenous noise (Figure 1.1C).

In Study III, we addressed the limitations of previous work by (i) appropriately comparing models of the link between preferences and choices and (ii) comparing monetary gambles to representative samples of everyday unitary stimuli. We mainly obtained such stimuli from Yougov⁷ polls which report what percentage of a representative sample of the Great British public that have heard of some object. If more than half of the sample had heard of an object, we took that to mean that it was something that the majority of the population would come across. Such a representative design (see Brunswik, 1956; Dhimi et al., 2004) allows us to say something about how “often” we might expect people to have transitive preferences for everyday objects.

⁷Yougov (www.yougov.co.uk) is a major UK polling company.

4.4.2 General methods

Materials

Participants made 2AFCs between options and stated their strength of preference on a half-range scale (N. Weber & Brewer, 2003). Within each category, all pairwise combinations of five different options were presented three times in random order and at random positions (left or right). Thus, there were 30 trials of 10 items per category. The order of the categories was randomised at the start of the experiment and then maintained throughout. Since we analysed each category separately, this procedure ensured that there were always filler items between items from the same category.

In Experiments 1 and 2, participants also rated their familiarity with (Pliner & Hobden, 1992) and liking of (Wichchukit & O'Mahony, 2014) a sample of items from the master sets. The sample always included the five options from that category that the participants had been choosing between in the 2AFCs.

When one of the categories was monetary gambles (Experiments 1 and 3), the experiment was incentivised: one of the chosen gambles was played for real at the end of the experiment. Any winnings were kept as a bonus. Participants were recruited using Prolific, an online platform.

Modelling methods

We leveraged modelling methods described in Cavagnaro and Davis-Stober (2014), Regenwetter et al. (2010, 2011), and Zwilling et al. (2019). Because these methods are not widely disseminated and the literature describing them highly technical, I now provide an overview.

In the present modelling approach, models are implemented as unions of convex polytopes that live inside the sample space. For data on choice proportions from 2AFCs, the sample space is a unit hypercube with as many dimensions as there are pairwise combinations of stimuli, thus describing all possible choice probabilities. Each “corner” of the cube represents one (transitive or intransitive) ordering of the options. With five options, the sample space thus has 10 dimensions and $2^{10} = 1024$ “corners”. Note that a transitive ordering is effectively a permutation of the options. Thus, $5! = 120$ of the “corners” represent transitive orderings. The remaining 904 “corners” represent intransitive orderings. We already here get some intuition for how transitive preferences is a much more restrictive hypothesis than intransitive preferences when there are five different options. We can define a convex polytope by providing all its vertices (V-form) or defining all the facets (H-form). We have actually done this already: Definition 1 in Section 2.2.4 provides the inequality conditions that define the facets of the convex polytope of SST. All models were implemented using such H-form definitions. The defined convex polytope describes what regions of the sample space - what possible data - are consistent with the SST model. For transitive models, these regions will always extend out from the transitive corners. The models we implemented were agnostic as to which

particular transitive ordering a participant would have, so they extend out from each of the 120 transitive corners. If one's hypothesis stipulated that participants would follow some particular transitive order (e.g. "always choose the option with the highest expected value"), one could build this into the model by restricting which transitive corners it would extend from.

Note that there are no free parameters when we define models in this way. Also, we can get a precise measure of model parsimony: the proportion of the sample space that is occupied by a model's convex polytope. The smaller that proportion is, the fewer possible datasets exist that would be compatible with the model. SST is the most parsimonious model of those tested here, see Study III for details.

The Bayes factor for any pair of models is bounded by their relative parsimony (i.e. relative volume). We can calculate the Bayes factor of any model compared to an "encompassing" baseline model, so called because it encompasses the entire sample space. This is a catch-all model that allows both transitive and intransitive preferences. If all other models imply transitive preferences, intransitive preferences will be attributed to the encompassing model while transitive preferences will be attributed to a transitive model because of their superior parsimony.

These Bayes factors consider each individual in isolation. One can, of course, aggregate them by multiplication but that is a very crude mode of analysis. A more sophisticated method of aggregation, used by Cavagnaro and Davis-Stober (2014), is to fit a finite mixture model (FMM) using the likelihoods obtained as part of computing the Bayes factors. Here, we essentially ask "Assuming that all participants conform to one of the implemented models, what proportion of the population the participants were sampled from is estimated to conform to each model?" That is, any person is taken to conform to exactly one of the models but there is uncertainty as to which model that is. This approach leverages tendencies in the data across individuals but also translates more neatly to generalising statements about prevalence in the population. Unlike Cavagnaro and Davis-Stober (2014), we model the full experimental design by including fixed and random effects. This analysis is essentially a regular multinomial logistic regression but where the dependent variable is an unobserved class membership, which the data support with varying clarity, rather than an observed classification. We used the fixed effect prior $N(0,1)$ in an attempt to mimic the uniform prior in Cavagnaro and Davis-Stober (2014) and tested null hypotheses of no difference between contrasts using the Savage-Dickey density ratio.

We had to make two methodological developments for these methods to be sufficient. Firstly, Cavagnaro and Davis-Stober (2014) used a very simple algorithm (performed in an Excel document) for approximating the posterior distribution of the FMM. We discovered that the approximation did not perform well in model recovery analyses. We therefore implemented the FMM using PyStan (Riddell et al., 2021) which supplies an effective sampling algorithm.

This allowed us to approximate the posterior more accurately. Secondly, the rejection sampling algorithm used by Cavagnaro and Davis-Stober (2014) to compute the likelihoods did not provide high enough numerical precision. This is because some of the models we tested are extremely parsimonious: they occupy a very small proportion of the sample space. Under rejection sampling, one has to draw many tens of millions of samples to get an accurate estimate of their volumes. We therefore followed a suggestion in Zwilling et al. (2019) and invented a new direct simulation algorithm that improves efficiency (see below).

In addition to SST, we also tested three other prominent models of transitive preferences: Weak stochastic transitivity (WST, e.g. Tversky, 1969), Moderate stochastic transitivity (MST, e.g. Tversky, 1972), and the Mixture model of transitive preferences (MMTP, e.g. Becker & DeGroot, 1974). Within the context of a short laboratory experiment, we suspected that preferences would tend to align with SST. In the long run, however, I believe some kind of Mixture model of transitive preferences seems more reasonable, not least because it allows changes of mind.

Any model comparison exercise is relative - the results depend not only on the model in question but on what other models are included. Given this, we cannot understand the results of a model comparison exercise without understanding all of the models that were considered. We discussed SST already in Section 2.2.4 but let us go through the other models.

Models

The *mixture model of transitive preferences* (MMTP, Becker & DeGroot, 1974) allows all choice probabilities that one can obtain through different mixtures of transitive preference orderings.

Definition 2. *An individual's preferences conform to MMTP if, for all $a, b, c \in \mathcal{O}$, $P(a \succ b) + P(b \succ c) - P(a \succ c) \leq 1$*

This can have at least two cognitive interpretations. Firstly, it can reflect changes of mind between different transitive preference orderings. I may first believe that $a \succ b \succ c$, but later conclude that $b \succ a \succ c$. At no point are my preferences intransitive, but my choices may violate any single transitive ordering. Secondly, it can represent a random utility model (see Section 2.2.4) but without constraints on the distributions of the subjective values - the noise distributions are allowed to be heteroskedastic, irregular, correlated, and even degenerate. Because of this, the random perturbations of a single preference order can produce choice probabilities that mimic changes of mind. We may note in passing that DbS conforms to MMTP.

Moderate stochastic transitivity (MST, Tversky, 1972) assumes a single transitive preference ordering but relaxes SST's assumption that the relation-

ship between the distance in subjective value and choice probability is the same for all options:

Definition 3. *An individual's preferences are MST if, for all $a, b, c \in \mathcal{O}$, $P(a \succ b) \geq 0.5, P(b \succ c) \geq 0.5 \Rightarrow P(a \succ c) \geq \min\{P(a \succ b), P(b \succ c)\}$*

This is predicted by cognitive models where the probability of making a choice is determined by the difference in subjective value between options - just like in SST - but where some option pairs are more difficult to compare⁸ than others, which moderates the choice probability (“moderate utility model”, Cavagnaro & Regenwetter, 2023; He & Natenzon, 2024). Moderate stochastic transitivity is also predicted by a noncompensatory heuristic called “elimination by aspects” (Tversky, 1972).

Weak stochastic transitivity (WST, Fishburn, 1973) again postulates a single transitive preference order but places only minimal restrictions on the noise in choices.

Definition 4. *An individual's preferences are WST if, for all $a, b, c \in \mathcal{O}$, $P(a \succ b) \geq 0.5, P(b \succ c) \geq 0.5 \Rightarrow P(a \succ c) \geq 0.5$*

This can be realised by a participant exhibiting lapses at a rate that is allowed to differ between each option pair but never exceeds 50% (e.g. Regenwetter et al., 2011). Alternatively, one can again imagine a stochastic response selection process (as in a strong and moderate utility model) but where the stochasticity differs between option pairs (“weak utility model”, Luce & Suppes, 1965; Navarick & Fantino, 1975).

Lastly, the *encompassing model* (Cavagnaro & Davis-Stober, 2014) is a catch-all model that allows any preferences.

Definition 5. *An individual's preferences conform to the encompassing model if, for all $a, b, c \in \mathcal{O}$, $0 \leq P(a \succ b), P(b \succ c), P(a \succ c) \leq 1$*

Because it is the only model that allows intransitive preferences, we can interpret evidence in favour of it as evidence for intransitive preferences - if preferences align with any of the transitive models, those will “win” on account of their greater parsimony.

Direct simulation algorithm

The direct simulation algorithm builds on the insight that we can represent the parameters (i.e. choice probabilities) in a matrix. Every sample is generated as described in Figure 4.4. After a sample is generated, the results are inputted

⁸See Blavatsky (2007), Blavatsky and Pogrebna (2009), and Buschena and Zilberman (2000) for different interpretations of what would determine this difficulty.

into the likelihood function, which is a product of binomials due to assuming independent observations (hence the importance of including filler stimuli in the task). The results are stored. This procedure is repeated, in our case, 1 000 000 times with 10 000 burn-in samples that are discarded. We can get some informal sense of the gain in efficiency here by comparing it to the tens of millions of samples that were not sufficient under the rejection sampling algorithm. The mean of the results is our approximation of the posterior marginal likelihood of the data given the model in question. This is what we used to fit FMMs.

The choice proportion matrix

Here is how we can get some preliminary, descriptive insight into which models are likely to fit well. First, we relabel the options according to each participant's best-fitting transitive preference ordering. If you have chosen $a \succ b$ on $\frac{2}{3}$ occasions, $a \succ c$ on $\frac{3}{3}$, and $b \succ c$ on $\frac{3}{3}$, then your best-fitting ordering is $a \succ b \succ c$. That is, the best-fitting ordering is the one with the least Hamming distance (Hamming, 1950, p. 154-155) to the observed ordering. Then, we calculate the average choice proportion for each pair of options (a v. b , b v. c , and a v. c) across participants. We now have what we need to construct empirical tables similar to those displayed in Figure 4.4. (If we had enough data per participant, we could do this on the individual level rather than on group level). Let us call such matrices “choice proportion matrices”. If the proportions displayed are increasing as we read from left to right and from bottom to top, this is consistent with a “large” proportion of participants following SST or MST. Under other models, the pattern might be highly irregular. While new to the present literature, similar tables have been used elsewhere (e.g. “confusion matrices”, Clarke, 1957).

4.4.3 Experiment 1

Experiment 1 was an initial attempt at comparing the typical stimulus category of monetary gambles with two categories of everyday objects: confectioneries and political parties. While stimuli were from the United Kingdom (UK), we recruited participants from both the UK and the United States of America (USA) to obtain groups that differed in their familiarity with stimuli. However, this difference was smaller than expected, and would not turn out to matter much when it came to transitivity of preferences, why we only recruited from the UK in subsequent experiments.

Methods

We recruited 600 participants, half from the UK and half from the USA. Participants needed to be of the respective nationality and be currently located in that country, according to Prolific, to be eligible. There were three categories: monetary gambles (taken from Experiment 1 in Tversky, 1969, and

Figure 4.4. Illustration of the direct simulation algorithm

A

	A	B	C	D	E
A	0.5	0.6	0.7	0.8	0.9
B	-	0.5	0.6	0.7	0.8
C	-	-	0.5	0.6	0.7
D	-	-	-	0.5	0.6
E	-	-	-	-	0.5

B

	A	B	C	D	E
A	0.5	0.54	0.7	0.8	0.9
B	-	0.5	0.6	0.7	0.8
C	-	-	0.5	0.6	0.7
D	-	-	-	0.5	0.6
E	-	-	-	-	0.5

C

	A	B	C	D	E
A	0.5	0.54	0.62	0.89	0.97
B	-	0.5	0.61	0.70	0.88
C	-	-	0.5	0.59	0.79
D	-	-	-	0.5	0.69
E	-	-	-	-	0.5

D

	D	B	A	E	C
D	0.5	0.54	0.62	0.89	0.97
B	-	0.5	0.61	0.70	0.88
A	-	-	0.5	0.59	0.79
E	-	-	-	0.5	0.69
C	-	-	-	-	0.5

Note. Panel A: the algorithm initialises by drawing a set of choice probabilities that respect the inequality conditions that define the model in question. In this example, the model is SST. Panel B: one off-diagonal cell is selected (indicated by bold font). A new value is randomised such that the matrix as a whole still respects the conditions for the model. Panel C: the same procedure is repeated for each cell, one at a time. This constitutes Gibbs sampling. Panel D: rows/columns are scrambled. The resultant matrix is our first direct simulation sample.

with outcomes in British pounds instead of US dollars, the currency participants were compensated in), Great British confectioneries, and the political parties of the UK House of Commons (elected lower house). The confectioneries were scraped from a Yougov poll. All participants chose between options from all categories. The experiment was thus a 2 x 3 mixed factorial design.

Results

The choice proportion matrices (Table 4.6) give us a hint of what is to be revealed by the modelling: for the everyday objects, the proportion of choices violating the best-fitting transitive preference ordering decreases as the distance between the options, according to that ordering, increases.

Table 4.6. *Proportion of choices in line with the best-fitting transitive ordering, by condition*

	Option B	C	D	E
Gambles				
A	0.898	0.915	0.916	0.927
B	-	0.914	0.927	0.928
C	-	-	0.905	0.904
D	-	-	-	0.923
Confectioneries				
A	0.951	0.978	0.988	0.996
B	-	0.935	0.979	0.991
C	-	-	0.936	0.978
D	-	-	-	0.963
Political parties				
A	0.942	0.968	0.983	0.989
B	-	0.912	0.946	0.977
C	-	-	0.915	0.964
D	-	-	-	0.942

Note. Choice proportion matrices for data from Experiment 1. Group level choice proportions are consistently as predicted by Strong and Moderate stochastic transitivity: they increase as we read the rows from left to right and decrease as we read the columns from top to bottom. That is, choice probabilities increase with distance between options.

The FMM (Table 4.7) indicates, for both UK and USA participants, that preferences for these everyday objects are transitive. Specifically, they typically seem to be SST. The prevalence of the other transitive models is estimated, according to the credible intervals, to range from about zero to one and a half percent. For monetary gambles, intransitive preferences are rare. The credible intervals indicate them to be non-zero, perhaps in the order of a few

percent. This apparent difference in the prevalence of intransitive preferences between confectioneries & political parties and monetary gambles is confirmed by evidence against the point null hypothesis of no difference ($BF_{10} \approx 357$). There was, however, evidence against a difference in the prevalence of intransitive preferences both between UK and USA participants ($BF_{10} = 0.138$) and for the interaction between which country participants were recruited from and the category being an everyday object category ($BF_{10} = 0.224$).

Table 4.7. Results of finite mixture model for Experiment 1

Stimulus type	SST	Transitive models		MMTP	Intransitive model
		MST	WST		Encompassing
UK					
Gambles	0.907 (0.817, 0.963)	0.026 (0.003, 0.100)	0.022 (0.003, 0.072)	0.007 (0.001, 0.019)	0.037 (0.015, 0.067)
Confectioneries	0.990 (0.971, 0.998)	0.003 (0, 0.016)	0.002 (0, 0.008)	0.001 (0, 0.003)	0.004 (0.001, 0.010)
Political parties	0.990 (0.971, 0.998)	0.003 (0, 0.017)	0.002 (0, 0.008)	0.001 (0, 0.003)	0.004 (0.001, 0.011)
US					
Gambles	0.946 (0.860, 0.987)	0.017 (0.001, 0.088)	0.014 (0.001, 0.058)	0.003 (0, 0.011)	0.020 (0.005, 0.044)
Confectioneries	0.996 (0.984, 1.000)	0.001 (0, 0.011)	0.001 (0, 0.004)	0 (0, 0.002)	0.001 (0, 0.005)
Political parties	0.996 (0.985, 1.000)	0.001 (0, 0.010)	0.001 (0, 0.004)	0 (0, 0.002)	0.001 (0, 0.005)

Note. Mean (95% credible interval) proportion of population estimated to be associated with each model. For the monetary gambles, a few percent are estimated to have intransitive preferences, but for the everyday objects the mean proportion is near zero. Strong stochastic transitivity has the highest prevalence.

Discussion

So far, there is no evidence that preferences are “often” intransitive (American Psychological Association, 2025). Rather, they seem to be “rarely” intransitive for monetary gambles and “very rarely to virtually never” intransitive for the two categories of everyday objects that we have tested so far. However, one could argue that maybe we had happened to pick two object categories for which people were particularly likely to have transitive preferences. Experiment 2 addressed this concern.

4.4.4 Experiment 2

In Experiment 2, we expanded the representative design to include a greater and more haphazard selection of everyday object categories. If we could replicate the findings of Experiment 1 in these new sets of stimuli, one would be

more confident in making the general claim that preferences for everyday objects indeed are transitive.

Methods

We recruited 400 participants from the UK. The categories were: the movies from the IMDB Top 100 list, magazines, cars, Great British cities as holiday destinations, charitable donations to charity organisations, dinners (mainly Great British cuisine), establishments (cafés, restaurants, fast food outlets), and fruits. All but the movies were scraped from Yougov polls. Participants were randomised to four out of the eight categories.

Results

The results for the eight new categories of everyday objects replicate those for the object categories from Experiment 1. The choice proportion matrices (Table 4.8) consistently indicate the pattern expected under SST and MST: as the distance between the options grows in the (best-fitting) preference ordering, the probability of selecting the higher-ranked option increases. As for the FMM (Table 4.9), the prevalence of intransitive preferences is again consistently estimated to be close to zero while the prevalence of SST is high. There is consistently evidence in favour of the null for all contrasts between all categories (all $BF_{S_{10}} \leq 0.06$). That is, it seems like preferences are typically SST regardless of the category of everyday objects.

Discussion

Again we find evidence that SST is the most prevalent model and intransitive preferences are very rare to virtually inexistent.

4.4.5 Experiment 3

Here are two reasonable objections to the conclusions of Experiment 1 and 2. Firstly, one could speculate that the representative samples of everyday objects produced fairly easy choices. Down the shops, the shopkeeper has probably curated the selection to only include the items that are most popular with consumers. Our putatively representative samples would not be representative of that context. Secondly, and conversely, the gambles from Tversky (1969) have been tailored to be very tricky for participants. What if the results for monetary gambles were mainly driven by this greater discriminative difficulty? In other words, maybe the tailoring, rather than the other reasons I mentioned in Section 4.4.1, makes for the observed difference between monetary gambles and objects. Alternatively, it could be that occasional intransitivity is a persistent special case of the cognitive process that supports preferences for monetary gambles, but not that which supports preferences for everyday objects, regardless of whether discriminative difficulty is high or moderate. These objections were addressed in Experiment 3.

Table 4.8. *Proportion of choices in line with the best-fitting transitive ordering, by condition*

Option	B	C	D	E	Option	B	C	D	E
Movies					Charitable donations				
A	0.948	0.967	0.98	0.993	A	0.955	0.976	0.986	0.99
B	-	0.92	0.959	0.989	B	-	0.915	0.953	0.997
C	-	-	0.915	0.958	C	-	-	0.921	0.979
D	-	-	-	0.923	D	-	-	-	0.963
Magazines					Dinners				
A	0.956	0.972	0.985	0.992	A	0.966	0.983	0.99	1.0
B	-	0.924	0.972	0.982	B	-	0.954	0.988	0.997
C	-	-	0.91	0.97	C	-	-	0.947	0.985
D	-	-	-	0.934	D	-	-	-	0.976
Cars					Establishments				
A	0.939	0.965	0.982	0.993	A	0.941	0.971	0.988	0.997
B	-	0.946	0.962	0.985	B	-	0.915	0.963	0.993
C	-	-	0.947	0.975	C	-	-	0.932	0.98
D	-	-	-	0.955	D	-	-	-	0.949
Holiday					Fruits				
A	0.955	0.99	0.993	1.0	A	0.961	0.98	0.998	1.0
B	-	0.936	0.967	0.984	B	-	0.951	0.983	0.998
C	-	-	0.926	0.976	C	-	-	0.956	0.99
D	-	-	-	0.936	D	-	-	-	0.966

Note. Choice proportion matrices for data from Experiment 2. Again, group level choice proportions are consistently as predicted by Strong and Moderate stochastic transitivity: they increase as we read the rows from left to right and decrease as we read the columns from top to bottom.

Table 4.9. Results of finite mixture model for Experiment 2

Stimulus type	SST	Transitive models		MMTP	Intransitive model
		MST	WST		Encompassing
Movies	0.994 (0.977, 0.999)	0.002 (0, 0.013)	0.001 (0, 0.006)	0.001 (0, 0.004)	0.002 (0, 0.007)
Magazines	0.994 (0.977, 0.999)	0.002 (0, 0.014)	0.001 (0, 0.006)	0.001 (0, 0.004)	0.002 (0, 0.007)
Cars	0.992 (0.979, 0.998)	0.003 (0, 0.010)	0.001 (0, 0.005)	0.001 (0, 0.004)	0.003 (0.001, 0.008)
Holiday destinations	0.994 (0.977, 0.999)	0.002 (0, 0.014)	0.001 (0, 0.006)	0.001 (0, 0.004)	0.002 (0, 0.007)
Charitable donations	0.994 (0.977, 0.999)	0.002 (0, 0.013)	0.001 (0, 0.006)	0.001 (0, 0.004)	0.002 (0, 0.007)
Dinners	0.994 (0.976, 0.999)	0.002 (0, 0.014)	0.001 (0, 0.007)	0.001 (0, 0.004)	0.002 (0, 0.007)
Establishments	0.994 (0.977, 0.999)	0.002 (0, 0.013)	0.001 (0, 0.006)	0.001 (0, 0.004)	0.002 (0, 0.007)
Fruits	0.994 (0.977, 0.999)	0.002 (0, 0.013)	0.001 (0, 0.006)	0.001 (0, 0.004)	0.002 (0, 0.007)

Note. Mean (95% credible interval) proportion of population estimated to have preferences consistent with each model. A large majority is consistently estimated to have SST preferences. Intransitive preferences are estimated to be rare.

Methods

We recruited another 400 participants from the UK. The categories were monetary gambles and four everyday object categories. To determine which, we calculated the mean liking of every object and compared the difference between the first and fifth most liked option within each category. By only using the most attractive options, we figured, we should increase discriminative difficulty. The category with the least difference was cars, but four out of the five most liked cars were Audis. The major problem with this was that the pictures we had scraped from the Yougov poll were taken in angles such that it was difficult to tell that it was not four pictures of the same car. We felt that this might confuse participants, so we excluded that category. We instead included the second through fifth categories with the least difference in liking within their top five most liked options: charitable donations, confectioneries, magazines, and holiday destinations.

Each category came in a “curated” and “representative” or “generated” version. For the everyday objects, “representative” meant a random sample of five options from the representative option sets in Experiment 1 and 2. “Curated” meant only the five options with the highest mean liking rating. For the monetary gambles, “generated” meant a set of 100 gambles generated to have a range in expected value double that of the gambles in Tversky (1969)

(i.e. £0.74 rather than £0.37) and probabilities in the range $[0.2, 0.8]$ rather than $[\frac{7}{24}, \frac{11}{24}]$. “Curated” meant the gambles from Tversky (1969) used in Experiment 1.

In addition to the modelling, we also calculated the average number of deviations from the best-fitting transitive ordering (the “Hamming distance”, Hamming, 1950). We calculated credible intervals around these using a Poisson distribution with a gamma-distributed parameter. The prior was $\text{gamma}(2, 1)$ but note that the prior is massively dominated by the data. I also report a Poisson regression that was not preregistered.

Results

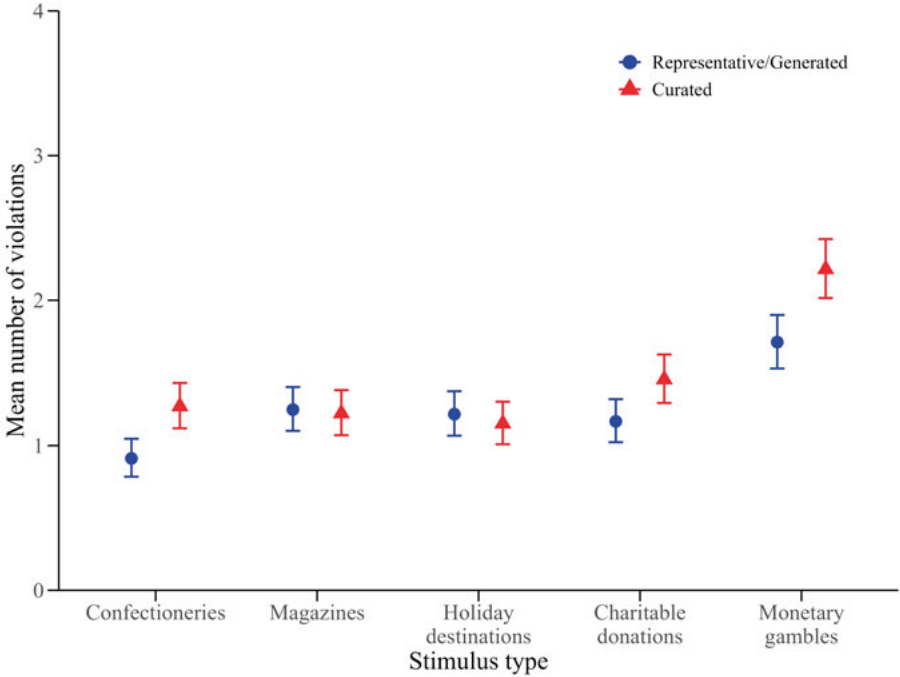
As indicated by the credible intervals in Figure 4.5, curating the most liked options only clearly succeeded in increasing discriminative difficulty for one out of four categories. For one, there was some but not clear evidence and for the remaining two there was no evidence of any difference. The generated monetary gambles did seem to incur fewer violations than those curated by Tversky (1969), however. This is corroborated by a Bayesian Poisson regression where we model the experimental design and include random slopes: we find evidence against a main effect of everyday objects v. monetary gambles ($\text{BF}_{S_{10}} \approx 0$), and anecdotal evidence of a main effect of curating the options ($\text{BF}_{S_{10}} = 2.7$). Instead, we find evidence in favour of the interaction between monetary gambles and curation ($\text{BF}_{S_{10}} = 8.07$).

In the choice proportion matrices (Table 4.10), the everyday objects generally display the pattern expected under SST and MST with the exception of two cells for the representative charitable donations. As for the curated monetary gambles, we again do not observe the pattern of choice probabilities increasing with distance between options, as in Experiment 1. When we generate the gambles, and apparently decrease discriminative difficulty (Figure 4.5), that pattern does obtain. The FMM (Table 4.11) reveals that across all everyday object categories, SST is the most prevalent model and intransitive preferences are estimated to be very rare to inexistent. Again we find that SST is the most prevalent model for monetary gambles too. For the gambles that had been curated by Tversky (1969), there is a tendency towards a lower prevalence of SST preferences and a higher, non-zero prevalence of intransitive preferences. However, Bayes factors consistently indicate evidence favouring the null hypothesis of no difference in the prevalence of the different models between everyday objects & monetary gambles ($\text{BF}_{S_{10}} \leq 0.325$), representative/generated & curated options ($\text{BF}_{S_{10}} \leq 0.268$) and their interaction ($\text{BF}_{S_{10}} \leq 0.204$).

Discussion

Whether we curate only the most liked everyday objects or not only had occasional effects on the apparent discriminative difficulty. It also does not seem to matter for the prevalence of SST or intransitivity. So, the first reasonable

Figure 4.5. Mean number of violations of the closest transitive preference ordering, by category and condition



Note. Whiskers indicate 95% credible intervals.

Table 4.10. Proportion of choices in line with the best-fitting transitive ordering, by condition

Representative/Generated					Curated				
Option	B	C	D	E	Option	B	C	D	E
Confectioneries					Confectioneries				
A	0.944	0.982	0.993	0.993	A	0.961	0.97	0.98	0.993
B	-	0.944	0.977	0.988	B	-	0.926	0.96	0.963
C	-	-	0.934	0.975	C	-	-	0.934	0.951
D	-	-	-	0.967	D	-	-	-	0.939
Magazines					Magazines				
A	0.947	0.979	0.982	0.99	A	0.941	0.964	0.986	0.991
B	-	0.941	0.95	0.982	B	-	0.925	0.965	0.988
C	-	-	0.91	0.96	C	-	-	0.917	0.965
D	-	-	-	0.942	D	-	-	-	0.951
Holiday destinations					Holiday destinations				
A	0.951	0.963	0.986	0.993	A	0.943	0.97	0.993	0.997
B	-	0.914	0.956	0.988	B	-	0.923	0.957	0.989
C	-	-	0.915	0.971	C	-	-	0.926	0.967
D	-	-	-	0.958	D	-	-	-	0.952
Charitable donations					Charitable donations				
A	0.929	0.967	0.975	0.988	A	0.919	0.956	0.978	0.983
B	-	0.936	0.977	0.992	B	-	0.907	0.951	0.978
C	-	-	0.917	0.985	C	-	-	0.919	0.968
D	-	-	-	0.946	D	-	-	-	0.955
Gambles					Gambles				
A	0.925	0.947	0.964	0.978	A	0.887	0.925	0.93	0.946
B	-	0.913	0.939	0.969	B	-	0.936	0.931	0.954
C	-	-	0.925	0.937	C	-	-	0.894	0.949
D	-	-	-	0.932	D	-	-	-	0.908

Note. Choice proportion matrices for data from Experiment 3. SST and MST predict that the choice proportions should increase as we read the rows from left to right and as we read the columns from bottom to top. For objects, this is the case except for two cells in Representative charitable donations. For monetary gambles, choice probabilities consistently increase with distance between options when the discriminative difficulty is relatively lower (Generated) but not when it is high (Curated).

Table 4.11. Results of finite mixture model for Experiment 3

Stimulus type	SST	Transitive models		MMTP	Intransitive model
		MST	WST		Encompassing
<u>Representative/Generated</u>					
Confectioneries	0.997 (0.984, 1.000)	0.001 (0, 0.010)	0 (0, 0.003)	0 (0, 0.002)	0.001 (0, 0.005)
Magazines	0.997 (0.983, 1.000)	0.001 (0, 0.010)	0 (0, 0.003)	0 (0, 0.002)	0.001 (0, 0.004)
Holiday destinations	0.997 (0.986, 1.000)	0.001 (0, 0.009)	0 (0, 0.003)	0 (0, 0.002)	0.001 (0, 0.004)
Charitable donations	0.997 (0.985, 1.000)	0.001 (0, 0.010)	0 (0, 0.004)	0 (0, 0.002)	0.001 (0, 0.005)
Monetary gambles	0.991 (0.972, 0.999)	0.003 (0, 0.017)	0.001 (0, 0.007)	0.001 (0, 0.004)	0.003 (0, 0.011)
<u>Curated</u>					
Confectioneries	0.990 (0.965, 0.999)	0.004 (0, 0.021)	0.002 (0, 0.008)	0.001 (0, 0.006)	0.004 (0.001, 0.011)
Magazines	0.991 (0.969, 0.999)	0.003 (0, 0.019)	0.001 (0, 0.008)	0.001 (0, 0.006)	0.004 (0.001, 0.011)
Holiday destinations	0.991 (0.969, 0.999)	0.003 (0, 0.019)	0.001 (0, 0.007)	0.001 (0, 0.005)	0.003 (0, 0.010)
Charitable donations	0.990 (0.968, 0.999)	0.003 (0, 0.020)	0.002 (0, 0.008)	0.001 (0, 0.006)	0.004 (0.001, 0.012)
Monetary gambles	0.972 (0.934, 0.992)	0.008 (0.001, 0.033)	0.005 (0.001, 0.015)	0.003 (0, 0.011)	0.011 (0.003, 0.025)

Note. Mean (95% credible interval) proportion of population estimated to have preferences consistent with each model. SST is the most prevalent model, just as in Experiment 1 and 2. Intransitive preferences are estimated to be rare, with higher estimates for the monetary gambles.

objection does not seem to apply. One could indeed devise more severe ways of curating stimuli⁹ but that will eventually leave us with an experiment that cannot be called a representative design.

For the monetary gambles, generating them with greater differences in expected value did seem to decrease discriminative difficulty. This translated to a tendency towards a decrease in the prevalence of intransitive preferences, but this was not supported by the test of the point null hypothesis.

4.4.6 Meta-analysis

We performed a meta-analysis where we pooled data from all experiments. The results are displayed in Table 4.12. For representative samples of everyday objects, about 100% of the population from which we have sampled participants is estimated to have SST preferences. Importantly, arriving at this exact figure is contingent on modelling choices, such as which parameter priors to use, and numerical approximation. It is therefore perhaps prudent to interpret the estimate as “nearly everyone” rather than literally 100%. Curating only the most liked of those representative sets of objects does not appear to increase the prevalence of SST ($BF_{10} < 0.109$). There is similarly no clear difference in the prevalence of SST between the generated monetary gambles and the representative ($BF_{10} = 0.487$) and curated ($BF_{10} = 0.138$) everyday object categories. Only for the monetary gambles from Tversky (1969), that have been designed to produce difficult choices, do we find clear evidence of a lower prevalence of SST preferences compared to the representative/curated everyday objects (both $BF_{s_{10}} > 1000$) and generated monetary gambles ($BF_{10} = 6.524$). Mirroring this, we estimate that a few percent have intransitive preferences for the curated monetary gambles. This is more than for the representative/curated everyday objects (both $BF_{s_{10}} > 1000$) and the generated monetary gambles ($BF_{10} = 6.003$). For those stimuli, intransitive preferences are estimated to be virtually inexistent, with prevalences close to zero percent.

4.4.7 General discussion

We find no evidence that preferences are “often” intransitive (American Psychological Association, 2025). Instead, intransitive preferences appear to be possibly inexistent for everyday objects - at least in pairwise choices. To the extent that intransitive preferences exist for monetary gambles, this tentatively

⁹I have been playing around with an idea for a procedure that would maximise discriminative difficulty. One would photograph a serving of, for example, meatballs. One would then rotate the plate, say, sixty degrees and snap another photo. One would proceed like this to photograph from five angles. The resulting pictures would be the stimuli in a 2AFC task like in the experiments reported here. It is an interesting question whether preferences for such literally indiscriminable objects would be SST too.

Table 4.12. *Meta-analysis of data from Experiment 1, 2 and 3*

Stimulus type	SST	Transitive models		MMTP	Intransitive model
		MST	WST		Encompassing
<u>Everyday objects</u>					
Representative	1.000 (0.998, 1.000)	0 (0, 0.001)	0 (0, 0.001)	0 (0, 0)	0 (0, 0.001)
Curated	0.996 (0.988, 0.999)	0.001 (0, 0.007)	0.001 (0, 0.003)	0 (0, 0.001)	0.002 (0, 0.005)
<u>Gambles</u>					
Generated	0.983 (0.948, 0.997)	0.005 (0, 0.030)	0.003 (0, 0.014)	0.001 (0, 0.005)	0.007 (0.001, 0.019)
Curated	0.933 (0.874, 0.968)	0.019 (0.002, 0.067)	0.013 (0.002, 0.037)	0.004 (0.001, 0.012)	0.031 (0.016, 0.048)

Note. Mean (95% credible interval) proportion of population estimated to be associated with each model. Representative everyday objects refers to the data from all object categories in Experiment 1 and 2 and the Representative objects conditions from Experiment 3. Curated everyday objects refers to data from all Curated object conditions in Experiment 3. Generated gambles refers to data from the Generated gambles condition in Experiment 3. Curated gambles are the gambles from Tversky (1969), used in Experiment 1 and the Curated gambles condition in Experiment 3.

appears to be driven by stimuli that have been tailored to be difficult to discriminate between. When discriminative difficulty is lower, the prevalence of intransitive preferences for monetary gambles is closer to that for everyday objects.

SST is a prevalent model of preferences for monetary gambles and a largely sufficient model for preferences for objects. Let us recapitulate what this means: we can model choices between objects at least “as if” they were supported by preferences that were supported by the comparison of each object’s subjective value (see Figure 1.1C). While modelling alone cannot tell us if these subjective values are existing, it suggests that it is unnecessary to postulate that preferences are constructed based on comparisons between aspects of options to explain these data.

5. General Discussion

5.1 Summary and conclusions

In social science and psychological science, there are two grand narratives of preference. The first is that of “existing preferences”: preferences reflect deeply held beliefs of subjective value and are, as a consequence, logically coherent and mathematically well-behaved. The second is “constructed preferences”: preferences are fickle and malleable, and perhaps say more about the context in which the choice was made than the person making it. While theories can exist on a spectrum between these end points, research in psychology has tended to emphasise preference construction. The purpose of this thesis was to examine the generality of this constructed preference account. This was done through two aims.

The first aim was to revisit some empirical phenomena that had been taken to support theories of constructed preference, but which have more recently come under criticism due to methodological limitations. Here, we found that the previously reported phenomena often did not obtain. Manipulating the distributions of recently experienced attribute values did not affect subsequent two-alternative forced choices (Study II) even under more tightly controlled circumstances tailored to test Decision by Sampling (DbS, Study I). We only found an effect in the listwise format, such that being exposed to higher values of some attribute on average led to an increase in the probability of selecting an option with similarly high values (Study II). When we investigated the claim that preferences were “often” (American Psychological Association, 2025) intransitive, we instead found the prevalence of intransitive preferences to be near-zero. This was particularly clear for various categories of everyday objects, but even for monetary gambles that had been curated to yield difficult choices did the prevalence of intransitive preferences only amount to a few percent. Instead, preferences were strongly stochastically transitive, which is consistent with a “random utility model”, where choices are made through comparison of subjective values of options that are retrieved with noise but stable in expectation.

The second aim was to investigate if relevant expertise in the option domain would attenuate the phenomena. We largely failed in addressing this aim because we did not find the phenomena to begin with (Study I & III). However, when we did so there was some evidence that sufficiently substantial expertise may decrease the effect size (Study II). One way to think about this is perhaps as a “floor effect”, combined with a sometimes small difference in expertise, concealing any attenuating effect.

All in all, the results indicate that it is not necessary to postulate malleable subjective values and constructed preferences to account for much of the present empirical data. To the extent the scrutinised phenomena have provided support for the claim that preferences are typically or often constructed, that support appears to have been premature. Consequently, existing preferences may play a larger role than commonly recognised.

5.2 Interpretation of results in Study I and II

DbS is fundamentally a theory of the subjective value function of attributes of options. To say anything about choices, it requires an additional mechanism that describes how those subjective values support preferences which then support choices (as hinted already by Stewart & Simpson, 2008, p. 262). I will call this a “preference mechanism”.

I am aware of two preference mechanisms that have been proposed in previous work. Firstly, Stewart and Simpson (2008) introduced a sequential accumulation mechanism. Two accumulators race against each other, counting each option’s successes in ordinal comparisons between its attribute values and exemplars. The option whose accumulator first reaches a certain threshold is chosen. We can write a closed form version of the model:

$$u(a) = \sum_{k=1}^2 \sum_{y_{ki} \in Y_k} \mathbb{I}(x_k > y_{ki} + \varepsilon_k) \quad (5.1)$$

where $a \in \mathcal{O}$ indicates the option, k indicates its attributes, x_k is again the stated value of attribute k (e.g. €5), Y_k denotes the sets of exemplars from memory, and ε_k are attribute-specific just-noticeable differences. When applied to the valuation of a single attribute in isolation, this amounts to the numerator in Equation 2.2 in Section 2.1.5. Secondly, Noguchi and Stewart (2018) have suggested a more sophisticated preference mechanism. It puts some structure on the sequential search, makes the ordinal comparisons probabilistic rather than sharp, and implements the decision threshold on the difference between the accumulators rather than the accumulated sums (see Noguchi & Stewart, 2018, p. 514-515). I call both of these “additive” preference mechanisms because the subjective values are added together.

Bird and Harris (2018, p. 616) suggested a third, “interpretative” mechanism to explain their results. Participants have, as per E. U. Weber et al. (2002), an existing preference for risk. To Bird and Harris (2018), this might be an existing preference for, for example, the 30th percentile on some “riskiness” scale. An option’s position on this riskiness scale - how risky is €5 × 70%, else €0? - is in turn constructed as per some “theory of relative judgement”

(e.g. DbS). Participants then select the option whose constructed riskiness level most closely matches their existing riskiness preference.¹

At least in Study I, there “should” be an effect in the pairwise format given an additive preference mechanism. When we in Study II find an effect in the listwise format, it goes “from low to low”. This is consistent with the interpretative mechanism from Bird and Harris (2018) but does not seem consistent with the preference mechanisms of Stewart and Simpson (2008) and Noguchi and Stewart (2018). Under that interpretation, the effects in the listwise format - when they obtain - are consistent with preceding options affecting the interpretation of the significance of numerical stimuli, rather than an effect on the subjective values of options per se.

Zooming out, it seems clear that what preference or interpretative mechanism we assume matters for our interpretation of the results. When we say that some choice data are consistent or inconsistent with DbS, we are actually saying that they are consistent or inconsistent with the DbS core mechanism *combined* with some preference (or interpretative) mechanism(s). Choice data will treat these as a package.

5.3 Whither now for Decision by Sampling?

DbS was originally described as a process theory (Stewart et al., 2006), thus operating on the algorithmic level (see Marr, 2010/1982). In a single sentence: subjective value is constructed by performing sequential ordinal comparisons between the target attribute value and a small number of exemplars in working memory. However, it also comes with an interesting computational-level proposal. Again in a sentence: the subjective value of something is determined by its rarity in the (historical) ecology of the individual. This thesis has indicated that there probably is something not quite right with the algorithmic level description, but does not indicate what that something is. The computational level description, however, remains unaddressed.

Here are four ways in which a proponent of DbS could proceed. Firstly, they may overturn the results of Studies I and II. This can be done by failed replications (i.e. future studies finding an effect of preceding options where these did not) or by identifying methodological limitations. Secondly, they may propose a new preference mechanism that allows one to rationalise the results. Thirdly, they may modify the algorithmic level theory while maintaining the computational level theory. Lastly, they may abandon the theory on both levels of description. There is nothing presented in this thesis that would recommend the fourth option over the former three.

¹In the General discussion of Study II, we speculated in a satisficing process that would be consistent with the results without making reference to neither subjective value nor preference (Figure 1.1A). At <https://doi.org/10.17605/OSF.IO/6SWBV>, I provide code implementing such a model and a brief description.

DbS has been a pioneer effort in the tradition of constructed preferences. If one nevertheless were to take one's business elsewhere, then the following might be some preliminary suggestions based on other traditions reviewed in this thesis.

The historical conception that “there is no such thing as abstract value” (Section 2.1.1) provides a slightly provocative alternative: there is no cognitive origin of subjective value functions. Now, as a researcher, one can always present participants with different magnitudes of stimuli and observe their response: if we give participants a task, they will always do *something*. By mapping the stimulus-response relationship, one can measure a subjective value function in a purely descriptive sense. However, that does not mean that the stimulus-response relationship is generated by a dedicated cognitive mechanism. Instead, descriptive subjective value functions may be the result of a menu of cognitive mechanisms, recruited somewhat erratically and bearing no particular family resemblance. If so, descriptive individual-level subjective value functions would presumably be highly idiosyncratic, and perhaps context-specific, to an even greater extent than DbS already suggests.

Alternatively, homeostatically regulated reinforcement learning (Keramati & Gutkin, 2014; Yoshida et al., 2025, see Section 2.1.4) might provide a foundation for “existing” subjective value functions. Suppose presentation of a given quantity of, say, money tends to predict some drive state reduction. For example, handling 40 SEK might predict whatever drive state reduction is afforded by a caffè latte. That contingency might then be learnt associatively. If the associative learning system in question has some capacity for interpolating between training stimuli (as in e.g. optimal linear associative memory models, Lind, 2024), learning the associative values of a relatively small number of training stimuli might be sufficient for learning a subjective value function. As usual in associative learning (see e.g. Rosas et al., 2013), one might expect such a function to possibly differ between contexts but be relatively stable within contexts once full learning is achieved.

These suggestions may serve as some tentative inspiration for how one might think about subjective value functions beyond DbS. At present, there is only limited *atheoretical* work that tries to establish the descriptive properties of subjective value functions from the bottom up (e.g. Brandstätter & Brandstätter, 1996). Future work might want to do so within individuals. To what degree are the descriptive functions stable over time, context-specific, and sensitive to preceding options? When are they concave, convex, linear, or neither?² Establishing certain facts about their empirical character might

²One might claim that Parducci (1995) has done this already. However, that work seems to rest on the use of Likert-type rating scales. As a psychophysical method, these come with limitations (see e.g. Birnbaum, 1999; Chituc et al., 2026). Indeed, linguists interpret similar experiments as indicating participants' interpretations of the scales rather than the represented magnitude (see e.g. Weicker & Schulz, 2020).

allow us to triage the suggestions above. This could be done using standard psychophysical methods like magnitude estimation and magnitude production.

5.4 Conceptual preferences

In Study III we speculated in the existence of “conceptual preferences”. The idea was that preferences could be supported by conceptual knowledge. In a nutshell: upon stimuli presentation, the individual will automatically categorise (see e.g. Hahn & Chater, 1997; Hodgetts & Hahn, 2025) each option as being a so-and-so. This causes retrieval of conceptual knowledge of that category (e.g. Collins & Loftus, 1975; Quillian, 1966). The proposition, then, is that subjective value is part of that conceptual knowledge.³ Thereby, one may form a preference by comparing a memory of the subjective value of option *a* to a memory of the subjective value of option *b* (Figure 1.1C). Assuming that these memories will be corrupted by (symmetric and homogenous) noise upon retrieval, we have a random utility model and thus SST preferences. Although the results in Study III turned out to be more decisive than we could have expected, an evolving version of this idea guided our expectation of SST preferences for the objects: categorisation would work well for familiar everyday objects like political parties and confectioneries, but less well for “artificial” options like monetary gambles. On account of their novelty, the latter would not elicit strong activation of any category or, if anything, be categorised as belonging to the same category.

The results, however, indicated that preferences were typically SST even for monetary gambles. The difference that did exist was at least to a substantial degree accounted for by discriminative difficulty. To be clear: this goes against what I was expecting before the fact. Yet, remember that the descriptive models we implemented (like SST) are predicted by multiple cognitive processes. SST is consistent with retrieval of noisy subjective values from memory, but also with “intuitive” (Sundh et al., 2021) expected value calculations. Ultimately, the methodology employed in Study III is useful for investigating the properties of the preferences. This allows us to narrow down the set of viable cognitive theories, but it is insufficient for pinning down the exact cognitive process. This is not, I claim, a limitation of the modelling per se but of its reliance on choice data. Choices are the end points of a cognitive process; when the processes we want to discriminate between share end points, such data are not diagnostic (cf. “multiple realisability”, Shapiro, 2000). Future work might want to include electroencephalographic (EEG) data or other methods more suitable for process tracing. By triangulating several methods, converging evidence might point to a more unequivocal interpretation.

³As such, subjective values would primarily constitute semantic, rather than episodic, memory (Tulving, 1972). Episodic memory would, in this view, primarily be important for preference construction, see Section 2.2.2.

5.5 SST and the evidence accumulation framework

Study III estimated SST preferences to be near-universal, at least for ten different categories of everyday objects. This begs the question: should we expect this near-universality to hold near-universally or did we just manage to set up an experimental context which favoured SST?

To answer that question, it might be useful to align SST with some theory of the decision-making process. Under a random utility interpretation, SST is perhaps most readily interpretable in terms of memory retrieval or intuitive (Sundh et al., 2021) processes, as opposed to explicit reasoning. One well-established framework for this is sequential evidence accumulation (Ratcliff, 1978) which will be used as an example here.⁴ In sequential evidence accumulation, “evidence” is assumed to be sampled from a normal distribution. Figure 5.1 provides an illustrative example. Note that assuming a normal distribution implies a Thurstonian random utility model (Section 2.2.4) and thus SST preferences. This puts an interesting spin on the high prevalence of SST in Study III: the results suggest that an already well-established model of decision-making is widely applicable. Framed in that way, the results are perhaps not so surprising after all.

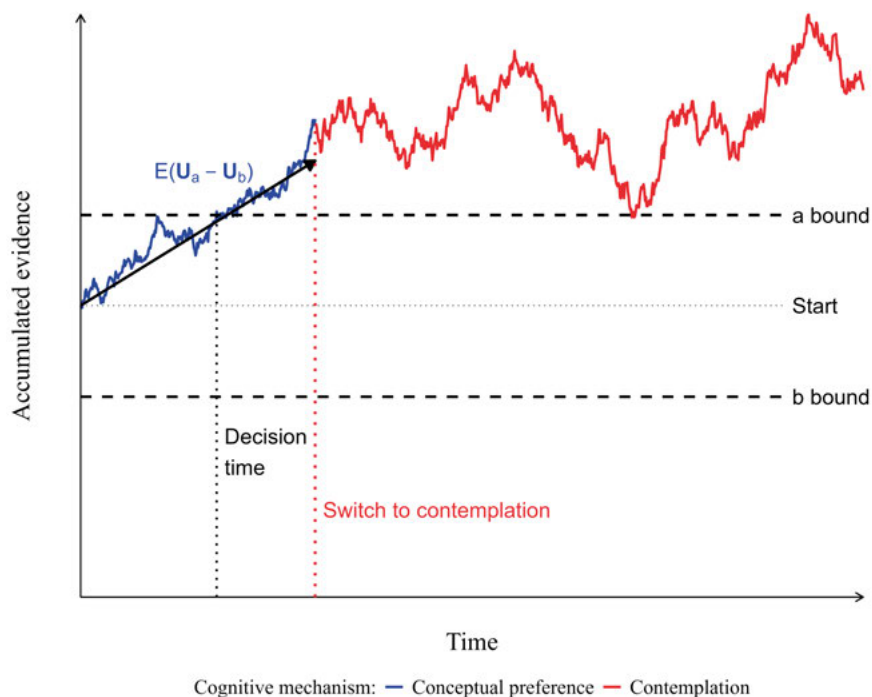
In what follows, I leverage the evidence accumulation framework to illustrate how one may think about the generalisability of the near-universal prevalence of SST suggested by Study III.

Value pluralism does not contradict SST

In Study III, choices were always made between options from within the same category. For each category, we stated or at least implied a shared goal for the options: dishes to have for dinner, fruits to have as an afternoon snack, et cetera. A value pluralist (Section 2.1.1) might say that we stated what *kind* of value that was relevant on a given trial. We can also imagine a situation where a choice is made between two options from different categories and where different kinds of value are relevant: “Would you prefer to have a cup of tea [to drink] or a ticket to the cinema [to watch a movie]?”. As mentioned in Section 2.2.2, Walasek and Brown (2023) have recently argued that in such situations, subjective values cannot support choices. Instead, people must rely on noncompensatory heuristics - which will typically not predict SST preferences (e.g. Tversky, 1972). So, a critic might argue - based on value pluralism - that Study III provided conditions where preferences supported by subjective values (and thus SST preferences) were particularly likely to be sufficient but noncompensatory heuristics (and thus non-SST preferences) were particularly unlikely to be called for. While that is ultimately an empirical question, I will now claim that value pluralism is in principle compatible with SST preferences instantiated in a sequential evidence accumulation model.

⁴See Alós-Ferrer et al. (2021) and Berlinghieri et al. (2023) for more sophisticated attempts at bridging the random utility and drift-diffusion frameworks.

Figure 5.1. Example of evidence accumulation process under present theory



Note. Hypothetical example of conceptual preference within evidence accumulation framework. A participant chooses between options a and b . As per the Thurstonian random utility model, they observe options a and b and integrate their conceptual subjective values $u_a - u_b$. This causes an initial drift towards the corresponding decision bound by that much. In this example, it exceeds the bound and a choice is made. Note how the actual evidence (solid blue line) from $u_a - u_b$ deviates from the “idealised” mean random utility $\mathbb{E}(U_a - U_b)$ (solid black arrow) because of random perturbations upon retrieval. For illustrative purposes, we imagine that evidence accumulation continues for a while after the choice. Having accumulated evidence from the conceptual subjective values, the individual may switch to instead seeking support for their choice from contemplation (red). This will appear as them starting to accumulate evidence from a different distribution, with a different drift rate and standard deviation. Even if the conceptual preference is insufficient to surpass the threshold and thus support choice, it will still “move the needle”, bringing the individual closer to choosing the conceptually preferred option. Subscripts of u are simplified for legibility.

When options relate to different goals, one simply has to assume that the individual may prioritise between those goals (cf. Maslow, 1943): slaking thirst is so-and-so more important than being entertained. This can then enter as a bias to the starting point (“Start” in Figure 5.1), making the eventual decision jointly determined by the priority of the affordance and the subjective value itself. So, even if we concede the value pluralist position that there are different *kinds* of subjective value, this is in itself not enough to reject the idea that those subjective values may support preferences under conflicting goals - the well-established sequential evidence accumulation framework already incorporates a mechanism for handling it. How prioritisation between goals might work lies outside the scope of this thesis, but there exists a large literature addressing the question (e.g. Anselme & Robinson, 2019; Bindra, 1968; K. Smith, 1984; Toates, 1986).⁵

Indecision as a limiting condition

A sequential evidence accumulation process describes a state of indecision that is ultimately resolved when the accumulated evidence crosses a decision bound. If the current difference between two subjective values is not great enough for the evidence to exceed the decision bound, the strength of preference is insufficient to support a choice. If we accept this, any existing preference ordering may potentially be a semiorder (Luce, 1956) rather than a strict total order (see Section 2.2.4). In a semiorder, only objects that differ by more than a certain “distance” are comparable. Objects that differ by less than that distance are incomparable (see Figure 5.2).

If two options are incomparable - be that because of a small difference in existing subjective value or wide decision bounds - some kind of tiebreaker must be imposed for a choice to be made. I suggest that this is a limiting condition for a random utility interpretation of SST and the point where theories that construct preferences might come in (Section 2.2.2). While some constructive processes may generate SST preferences, others will not (e.g. Tversky, 1972), why SST will not hold near-universally.

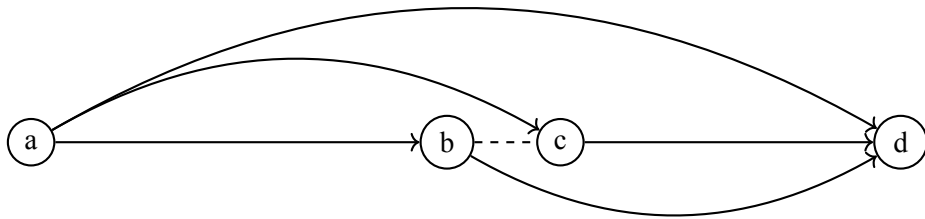
We tried to address the question of large differences in subjective value in Experiment 3 in Study III. We did not try to manipulate the decision bounds. If these are indeed set optimally (Bogacz et al., 2006), one could presumably manipulate their width by manipulating the costs and benefits of judicious choice. I speculate that the time spent per trial might be the most important cost in experiments like those reported in this thesis, at least when collecting data on the internet.

In sum, I suggest a two-stage process. Upon presentation of some focal stimuli that constitute options, existing (potentially conceptual) subjective values are accumulated to form a preference (Figure 1.1C). If this fails to support

⁵Note that these authors might speak of motivation to behaviours rather than goals. For present purposes, that is much the same thing.

a choice, further deliberation ensues to construct a preference (e.g. Figure 1.1B). See Figure 5.1 for an illustration. Compare this to Shadlen and Shohamy (2016): “One speculative possibility is that these different circuits support two different kinds of decisions: those that involve retrieval of a specific, well-learned value association and those that involve integration of learned associations to support new decisions, inferences, and prospection.”

Figure 5.2. Illustration of semiorder



Note. Semiorder over four options a , b , c , and d . Options b and c are close together and are therefore incomparable. The distances between options a and b , a and c , a and d , b and d , and c and d are larger why those pairs are comparable.

5.6 Transitivity but not completeness of preference?

Study III considered the transitivity axiom and found it to hold, with very limited exception, in a two-alternative forced choice task. To the extent these results hold up to future scrutiny, it would suggest that transitivity is a realistic axiom of preference. An axiom that has received a lot less interest is the completeness axiom (Section 2.2.4).⁶ As seen in the previous section, if preferences are determined by sequential evidence accumulation, we would expect existing preference semiorders. Semiorders violate completeness.

This leads to the following conjecture: when we have existing preferences, those preferences are transitive. When we face options for which we do not have existing preferences, or when the strength of preference is insufficient to support choice, we must construct preferences. In such cases, there does not necessarily exist a nice, clean axiomatisation of the preference set. Maybe that is the limiting condition for when many “anomalies” (e.g. Tversky & Thaler, 1990) arise.

⁶In health economics, Ryan and Miguel (2002) have claimed to provide an empirical test but appear to have confused completeness with indifference and strict preference. In addition, they do not address “Luce’s challenge” (see Regenwetter et al., 2011). Others have criticised it from a normative standpoint (see Anand, 1987; Aumann, 1962), but see Aumann (1964) for a corrigendum.

In two-alternative *forced* choice paradigms, participants must select one option and thus cannot express incomplete preferences. On the other hand, allowing participants to explicitly state indecision (e.g. including a “Can’t choose” option) is in and of itself probably not a solution. That a participant selected “Can’t choose” indicates that the accumulated evidence crossed the decision bound for that option. It does not imply the conjectured shift in the cognitive process. Completeness is, just like transitivity, an axiom of covert preferences. To test the transitivity axiom based on overt choice data, we needed models of stochastic transitivity. To test the completeness axiom, we may similarly need models of “stochastic completeness”. I am at present not aware of any such models. Devising them, and methods for testing them, may be an important challenge for future work.

Sher et al. (2025) have recently argued that incomplete preferences may explain certain empirical results such as the endowment effect. Their perspective differs from mine in at least three ways. Firstly, they seem to equate absence of preference with genuine indiscriminability between options. I have instead viewed it as arising from a temporary state of indecision, where options are discriminable but not clearly enough so to yield a preference. Secondly, and less surely, I interpret Sher et al. (2025) to take value pluralism to be their normative foundation for incomplete preferences. Here, I have argued from a representationalist perspective on subjective value (see Section 2.1.1). Thirdly, in the absence of preference, Sher et al. (2025) envision people to adopt choice policies that implement “rational arbitrariness”. Such policies do not, in the present terminology, seek to construct subjective values or preferences. Instead, they determine choices directly (Figure 1.1A) in such a way that, if the policy is applied consistently, a sequence of choices will avoid certain “incoherence traps” (cf. money pumps, Davidson et al., 1955). I have instead suggested preference construction as a secondary step for resolving indecision. Perhaps “rationally arbitrary” decision rules would be a third, final step after preference construction has failed to yield a resolution.

In sum, my perspective and that of Sher et al. (2025) suggest two different cognitive origins and interpretations of incomplete preferences.

5.7 Great expectations: the promise of the mathematical psychologist’s approach

It has been claimed that “Science is essentially a conversation in which people respond to what others have most recently said, or to the ideas that are currently dominant.” (Kahneman, 2010). (See also Rakow, 2010). Interesting research, others say (Davis, 1971), is that which challenges current scientific thought. Taken together, the advice would seem to be to contradict whatever is *en vogue*.

However, “responding to what others have most recently said” without reading “old books” (Baron, 2010) risks reducing the scientific project to a tele-

phone game. The phrase “toothbrush problem” (Mischel, 2005, 2008) refers to the alleged unwillingness of psychological scientists to use other people’s models. This is perhaps one reason why psychological science is not cumulative to the extent we want it to be (Mischel, 2009).

Study III relied heavily on the work of Regenwetter et al. (2010, 2011) and other mathematical psychologists (e.g. Davidson & Marschak, 1959; Luce, 1956). That literature is characterised by a dogged interest in “roots and beginnings” (Tolkien, 1954). Through careful analysis, mathematical psychologists build a detailed understanding of the implications of some (family of) model(s). This allows “strong” (Platt, 1964) empirical tests of the models. Most importantly, the cycle of analysis and testing is repeated with reference to the same, relatively small, set of models. Mathematical psychologists share their toothbrushes! While mathematical psychologists take a mathematical approach, this might not always be necessary. Simulations require substantially less mathematical proficiency but may often yield insights into how and when a model works (cf. Gelman & Shalizi, 2012). More generally, one might perhaps meditate on whether cognitive psychology needs a division into theoretical and empirical cognitive psychology, like in physics and economics. Or rather, maybe it needs the specialisation that such a division is testament to.

5.8 Limitations

5.8.1 Psychological interpretation of expertise

In this thesis, and Studies I and II in particular, “expert” and “high experience” have been used as demographic labels. They reflected the general idea that some people would have more relevant knowledge in long-term memory that could be brought to bear on a particular choice problem. However, the psychological meaning of that idea is not as well-defined as one would want. In addition, such knowledge is confounded with unobserved variables. While it seems safe to conclude that persons who had worked as municipality accountants for many years knew more about municipality accountancy than those who had never done so (Study II), we are ultimately not able to rule out that some other factor caused the difference. On the flip side, inducing expertise through experimental manipulation would perhaps not be feasible. A more successful strategy might be to combine several approaches: (i) recruiting a non-expert control group that more closely matches the demographic characteristics of the expert group, (ii) adjusting statistically for differences in observable demographic characteristics, and (iii) credibly ensuring that any attenuation is specific to tasks cueing the relevant domain of expertise. In Studies I and II, we employed only (iii), and only in half of the experiments. While (i)-(iii) would not rule out confounds from unobservables, they might make the claim of an effect of expertise more credible.

5.8.2 Confounding of frequency, range, and mean

All training sets in Study II manipulated the frequencies of different attribute values between groups. However, only Experiment 2 and 3 held the range of values (at least as good as) constant. All experiments confounded frequency with differences in mean. Thus, while the effect in the listwise format - when it obtained - may have been driven by frequency, we cannot say with certainty that this was the sole “active ingredient”. Note that the theories of “relative judgement” that Bird and Harris (2018) cited make different predictions. ALT (Helson, 1947) predicts an effect of the mean of recently experienced attribute values, RFT (Parducci, 1965, 1995) predicts an effect of frequency and minimum/maximum, and DbS (Stewart et al., 2006) predicts an effect of frequency only. When combined with the methodological developments argued for below, manipulating frequency, range, and mean one at a time might be informative.

5.8.3 Limitations of average treatment effects

Already Meehl (1967) cautioned against testing theories through a mere directional effect. Even if our “theory” would be nothing more than a correlation coefficient drawn randomly from a uniform distribution, he argued, the probability of it being empirically vindicated goes towards 0.5 as power and precision increase. That is, merely inspecting whether the probability of choosing options of type A increased or not is not a severe test (Mayo, 1991) of *any* theory. Thus, while Studies I and II can provide evidence against a theory, the main analyses we reported would not provide strong evidence in favour of any theory.

How can we do better? Firstly, one could investigate effects on individual trials and not only on average across trials, viewing any theory as the conjunction of its predictions across trials (Davis-Stober & Regenwetter, 2019).⁷ This could be done using a carefully specified GLMM (cf. Tan & Jaeger, 2025) but also through “stimulus plots” (Simonsohn et al., 2025) where individual data-points are plotted for each trial (cf. Figure 4 in Study I). Such visualisation can provide an undemanding, presumably intuitive alternative to statistical modelling.⁸

A second improvement would be to investigate heterogeneity across participants. The universality of a theory across individuals may, I think, be suggested by the universality of the postulated cognitive mechanism. For example, automatic processes may be highly universal while controlled processes

⁷One could also perform placebo analyses (checking whether any effect is absent when it should be absent, Athey & Imbens, 2017) on training set trials.

⁸Inference from visualisations has historically been emphasised as the remit of so-called “exploratory” analysis (Tukey, 1993) but the distinction between that and “confirmatory” analysis is perhaps more porous than is sometimes recognised (see e.g. Hullman & Gelman, 2021).

may be more liable to cultural moderation and thus be displayed with more exceptions. While single-N designs (Fischer et al., 2023) allow individual-level results, within-groups designs can still provide a distribution of effect sizes and directions.⁹

5.8.4 Two-alternative forced choice

Study III investigated straight choices between two options. Under such circumstances, preferences appear to predominantly be strongly stochastically transitive. However, the predominance of SST identified in Study III should not be naïvely generalised beyond the 2AFC format. As complications are added to the task, we may expect, for example, early attentional and perceptual processes to play a greater role in determining overt behaviour, why SST may be violated as a descriptive model even if the preferences themselves are still generated as per a random utility model.¹⁰ How generally SST might be applied will probably not be answered by any single research article, but be gradually revealed by a comprehensive research programme. It should be noted, however, that the evidence claimed to speak *against* transitive preferences has also come from 2AFC tasks (e.g. Ranyard, 1977; Tversky, 1969).

5.8.5 Representative design

Study III claimed to implement a representative design. This begs the question: representative of what? Study III covered ten different categories of everyday objects, including abstract entities (political parties), places (e.g. holiday destinations), actions (charitable donations), durables (e.g. cars), and consumables (e.g. fruits). While it is perhaps difficult to argue that those kinds of stimuli do not abound in our natural environments, one could argue that they are distinct from the options we face in other important decisions in life - for example, which job candidate to select for a vacant position.¹¹ While I personally suspect that “most” decisions we make concern everyday objects of the kind studied here, this speculation is probably difficult or impossible to verify empirically. Ultimately, the ten categories of stimuli are - at best - representative

⁹If only a between-groups design is possible, it turns out that the test statistics computed as part of a Kolmogorov-Smirnov test can be interpreted as the upper and lower bound on the proportion of participants who exhibit some effect (Fan & Park, 2009; Simonsohn, 2024).

¹⁰In addition to complications to the task, one might think of more complex options than those employed here - for example, bundles of many objects or gambles with many branches. Unless the bundle lends itself to categorisation (e.g. “a selection of fruits”), complex options might induce inexhaustive search which adds stochasticity (cf. “garden of forking paths”, Gelman & Loken, 2013). If this “search noise” is not well-behaved enough to preserve symmetry and homogeneity, random utility-based preferences for complex options might conform to MMTP rather than SST. This is an open empirical question and should be investigated.

¹¹I am grateful to Henry Montgomery for making this point.

of a section of human experience. That section is probably wider than the section covered by monetary gambles, but a totalising representative design remains elusive.

5.8.6 Model discriminability

Study III used three repetitions of every combination of five options. This number of retests is insufficient for good individual-level model discriminability between the transitive models, but we were able to achieve decent population-level discriminability through finite mixture modelling. Yet, model discriminability between the transitive models remains imperfect due to the models being nested. It is not clear (to me) that this is fixable. In particular, strong stochastic transitivity and moderate stochastic transitivity are extremely similar models. This will continue to cause issues even with extreme numbers of repetitions and participants. One will probably always be able to imagine that some proportion of participants have *barely* moderately stochastically transitive preferences, that violate SST only by a hair. While we can say with some confidence that preferences were not intransitive, we will probably always remain less confident in the exact distribution of the transitive models.

5.9 Concluding remark

Baruch Fischhoff (Fischhoff & Slovic, 2025) recently released a (presumably excellent) book called “Decisions - Studying and supporting people facing hard choices”. The cover features a corvid - probably a raven - stood in front of a pile of jumbled red, blue, green, light green, orange, and greige cables. I think it is supposed to look exasperated, but you never know with corvids! It is this archetype of a pensive, challenged (although usually flightless) decision maker that I have sought to nuance in this thesis. Christopher Summerfield has argued that researchers of artificial intelligence have overemphasised the importance of *thinking* over *knowing* for a system’s capacity for intelligent behaviour (Scarfe, 2023-present). As in AI research, so in psychology: sometimes people need to think about what they like, but sometimes we just know. I tentatively submit, based on the results presented here, that we currently have not got that balance right.

6. Acknowledgments

I wish to make the following acknowledgements:

I acknowledge **mormor and morfar, farmor and farfar, mamma, pappa, Sanna**. Thank you, for everything.

I acknowledge **my friends**. I am immensely fond of you all. I don't know half of you half as well as I should like; and I like less than half of you half as well as you deserve.

I acknowledge my supervisors, **Peter Juslin** and **Lars Frimanson**. Thank you for your patience, trust, cheerfulness, kindness, and support. I know better than to take it for granted.

I acknowledge **Ronald van den Berg**, the unmoved mover of all my work. I always pay my debts.

I acknowledge **my colleagues** who have made the department a home to me.

I acknowledge **you** who reads books till you come to the end, then stops.

*Här är mitt öde. Blås signal i hornet:
"Junker Roland kom till Det mörka tornet."
– Robert Browning, Junker Roland kom till det mörka tornet,
övers. av John-Henri Holmberg*

References

- Ainslie, G., & Haslam, N. (1992). Hyperbolic discounting. In G. Loewenstein & J. Elster (Eds.), *Choice over time* (pp. 57–92). Russell Sage Foundation.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-t](https://doi.org/10.1016/0749-5978(91)90020-t)
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. <https://doi.org/10.1109/tac.1974.1100705>
- Akerlof, G. A. (1978). The market for “lemons”: Quality uncertainty and the market mechanism. In *Uncertainty in economics* (pp. 235–251). Elsevier. <https://doi.org/10.1016/b978-0-12-214850-7.50022-x>
- Alempaki, D., Canic, E., Mullett, T. L., Skylark, W. J., Starmer, C., Stewart, N., & Tufano, F. (2019). Reexamining how utility and weighting functions get their shapes: A quasi-adversarial collaboration providing a new interpretation. *Management Science*, 65(10), 4841–4862. <https://doi.org/10.1287/mnsc.2018.3170>
- Algom, D. (2021). The weber–fechner law: A misnomer that persists but that should go away. *Psychological Review*, 128(4), 757–765. <https://doi.org/10.1037/rev0000278>
- Allais, M. (1953). Le comportement de l’homme rationnel devant le risque: Critique des postulats et axiomes de l’ecole americaine. *Econometrica*, 21(4), 503. <https://doi.org/10.2307/1907921>
- Allais, M., & Hagen, O. (1979). *Expected utility hypotheses and the allais paradox : Contemporary discussions of the decisions under uncertainty with allais’ rejoinder* (1st ed. 1979.). Springer Netherlands.
- Alós-Ferrer, C., Fehr, E., & Netzer, N. (2021). Time will tell: Recovering preferences when choices are noisy. *Journal of Political Economy*, 129(6), 1828–1877. <https://doi.org/10.1086/713732>
- Alós-Ferrer, C., & Garagnani, M. (2022a). Strength of preference and decisions under risk. *Journal of Risk and Uncertainty*, 64(3), 309–329. <https://doi.org/10.1007/s11166-022-09381-0>
- Alós-Ferrer, C., & Garagnani, M. (2022b). The gradual nature of economic errors. *Journal of Economic Behavior & Organization*, 200, 55–66. <https://doi.org/10.1016/j.jebo.2022.05.015>

- Altmann, S., Falk, A., Heidhues, P., Jayaraman, R., & Teirlinck, M. (2019). Defaults and donations: Evidence from a field experiment. *The Review of Economics and Statistics*, *101*(5), 808–826. https://doi.org/10.1162/rest_a_00774
- American Psychological Association. (2025). Intransitivity. In *APA Dictionary of Psychology - dictionary.apa.org*. <https://dictionary.apa.org/intransitivity>
 Accessed: 2025-03-21. Last updated: 2023-11-15.
- American Psychological Association. (2026). *Attitude*. APA Dictionary of Psychology. Retrieved March 6, 2026, from <https://dictionary.apa.org/attitude>
- Anand, P. (1987). Are the preference axioms really rational? *Theory and Decision*, *23*(2), 189–214. <https://doi.org/10.1007/bf00126305>
- André, Q., & de Langhe, B. (2021a). How (not) to test theory with data: Illustrations from walasek, mullett, and stewart (2020). *Journal of Experimental Psychology: General*, *150*(12), 2671–2674. <https://doi.org/10.1037/xge0001053>
- André, Q., & de Langhe, B. (2021b). No evidence for loss aversion disappearance and reversal in walasek and stewart (2015). *Journal of Experimental Psychology: General*, *150*(12), 2659–2665. <https://doi.org/10.1037/xge0001052>
- Anselme, P., & Robinson, M. J. F. (2019, February). Incentive motivation. In *The cambridge handbook of motivation and learning* (pp. 163–182). Cambridge University Press.
- Arkes, H. R., Gigerenzer, G., & Hertwig, R. (2016). How bad is incoherence? *Decision*, *3*(1), 20–39. <https://doi.org/10.1037/dec0000043>
- Arrow, K. J. (1950). A difficulty in the concept of social welfare. *Journal of Political Economy*, *58*(4), 328–346. <https://doi.org/10.1086/256963>
- Arrow, K. J. (1951). *Social choice and individual values*. Wiley.
- Artinger, F. M., Gigerenzer, G., & Jacobs, P. (2022). Satisficing: Integrating two traditions. *Journal of Economic Literature*, *60*(2), 598–635. <https://doi.org/10.1257/jel.20201396>
- Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, *31*(2), 3–32. <https://doi.org/10.1257/jep.31.2.3>
- Atkinson, R., & Shiffrin, R. (1977/1968). Human memory: A proposed system and its control processes. In *Human memory* (pp. 7–113). Elsevier. <https://doi.org/10.1016/b978-0-12-121050-2.50006-5>
- Aumann, R. J. (1962). Utility theory without the completeness axiom. *Econometrica*, *30*(3), 445. <https://doi.org/10.2307/1909888>
- Aumann, R. J. (1964). Utility theory without the completeness axiom: A correction. *Econometrica*, *32*(1/2), 210. <https://doi.org/10.2307/1913746>
- Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390–1396. <https://doi.org/10.1126/science.7466396>

- Banfield, T. (1848). *The organization of industry: Explained in a course of lectures*. Longman, Brown, Green, Longmans.
- Barlow, H. B. (1961). Possible principles underlying the transformations of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). MIT press.
- Baron, J. (2010). Comment: Past work is one source of inspiration among many [Published as a comment within Rakow (2010)]. *Judgment and Decision Making*, 5(6), 465. <https://doi.org/10.1017/s1930297500001303>
- Barretto-Garcia, M., Tu, J. C., & Padoa-Schioppa, C. (2025). Neuronal activity in orbitofrontal cortex during trinary choices under risk. <https://doi.org/10.1101/2025.09.19.677423>
- Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of bold fmri experiments examining neural correlates of subjective value. *NeuroImage*, 76, 412–427. <https://doi.org/10.1016/j.neuroimage.2013.02.063>
- Becker, G. M., & DeGroot, M. H. (1974). Stochastic models of choice behavior (1963). In *Economic information, decision, and prediction* (pp. 254–279). Springer Netherlands. https://doi.org/10.1007/978-94-010-9276-0_11
- Becker, G. M., Degroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3), 226–232. <https://doi.org/10.1002/bs.3830090304>
- Bentham, J. (1789). *An introduction to the principles of morals and legislation. printed in the year 1780, and now first published. by jeremy bentham, of lincoln's inn, esquire*. T. Payne; son.
- Ben-Zion, Z., & Levy, I. (2025). Representation of anticipated rewards and punishments in the human brain. *Annual Review of Psychology*, 76(1), 197–226. <https://doi.org/10.1146/annurev-psych-022324-042614>
- Berlinghieri, R., Krajbich, I., Maccheroni, F., Marinacci, M., & Pirazzini, M. (2023). Measuring utility with diffusion models. *Science Advances*, 9(34). <https://doi.org/10.1126/sciadv.adf1665>
- Bernoulli, D. (1954/1738). Exposition of a new theory on the measurement of risk. *Econometrica*, 22(1), 23. <https://doi.org/10.2307/1909829>
- Bernoulli, N. (1713). Letter to pierre rémond de montmort (9 september 1713). In P. R. de Montmort (Ed.), *Essay d'analyse sur les jeux de hazard* (2nd ed.). Quillau.
- Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research*, 25(3), 187–217. <https://doi.org/10.1086/209535>
- Bettman, J. R., Luce, M. F., & Payne, J. W. (2008). Preference construction and preference stability: Putting the pillow to rest. *Journal of Consumer Psychology*, 18(3), 170–174. <https://doi.org/10.1016/j.jcps.2008.04.003>

- Bhatia, S. (2013). Associations and the accumulation of preference. *Psychological Review*, 120(3), 522–543. <https://doi.org/10.1037/a0032457>
- Biele, G., Erev, I., & Ert, E. (2009). Learning, risk attitude and hot stoves in restless bandit problems. *Journal of Mathematical Psychology*, 53(3), 155–167. <https://doi.org/10.1016/j.jmp.2008.05.006>
- Bikhchandani, S., & Segal, U. (2021). Intransitivity in the small and in the large. *Journal of Risk and Uncertainty*, 63(3), 257–273. <https://doi.org/10.1007/s11166-021-09367-4>
- Bindra, D. (1968). Neuropsychological interpretation of the effects of drive and incentive-motivation on general activity and instrumental behavior. *Psychol. Rev.*, 75(1), 1–22.
- Birch, L. L. (1999). Development of food preferences. *Annual Review of Nutrition*, 19(1), 41–62. <https://doi.org/10.1146/annurev.nutr.19.1.41>
- Bird, S., & Harris, A. J. (2018). Robust, domain-specific effects of prior context in risk preferences for pension choice. *Journal of Behavioral Decision Making*, 31(5), 605–618. <https://doi.org/10.1002/bdm.2077>
- Birnbaum, M. H. (1999). How to show that $9 > 221$: Collect judgments in a between-subjects design. *Psychological Methods*, 4(3), 243–249. <https://doi.org/10.1037/1082-989x.4.3.243>
- Birnbaum, M. H. (2008). New paradoxes of risky decision making. *Psychological Review*, 115(2), 463–501. <https://doi.org/10.1037/0033-295x.115.2.463>
- Birnbaum, M. H. (2020). Reanalysis of Butler and Pogrebna (2018) using true and error model. *Judgment and Decision Making*, 15(6), 1044–1051. <https://doi.org/10.1017/s1930297500008238>
- Birnbaum, M. H. (2023a). Dominance and transitivity. *Judgment and Decision Making*, 18. <https://doi.org/10.1017/jdm.2023.43>
- Birnbaum, M. H. (2023b). Testing transitivity of preference in individuals. *Decision*, 10(2), 153–180. <https://doi.org/10.1037/dec0000185>
- Birnbaum, M. H., & Chavez, A. (1997). Tests of theories of decision making: Violations of branch independence and distribution independence. *Organizational Behavior and Human Decision Processes*, 71(2), 161–194. <https://doi.org/10.1006/obhd.1997.2721>
- Birnbaum, M. H., & Gutierrez, R. J. (2007). Testing for intransitivity of preferences predicted by a lexicographic semi-order. *Organizational Behavior and Human Decision Processes*, 104(1), 96–112. <https://doi.org/10.1016/j.obhdp.2007.02.001>
- Birnbaum, M. H., & Navarrete, J. B. (1998). Testing descriptive utility theories: Violations of stochastic dominance and cumulative independence. *Journal of Risk and Uncertainty*, 17(1), 49–79. <https://doi.org/10.1023/a:1007739200913>
- Birnbaum, M. H., & Schmidt, U. (2008). An experimental investigation of violations of transitivity in choice under uncertainty. *Journal of Risk*

- and *Uncertainty*, 37(1), 77–91. <https://doi.org/10.1007/s11166-008-9043-z>
- Blavatskyy, P. R. (2007). Stochastic expected utility theory. *Journal of Risk and Uncertainty*, 34(3), 259–286. <https://doi.org/10.1007/s11166-007-9009-6>
- Blavatskyy, P. R. (2018). Fechner’s strong utility model for choice among $n > 2$ alternatives: Risky lotteries, savage acts, and intertemporal payoffs. *Journal of Mathematical Economics*, 79, 75–82. <https://doi.org/10.1016/j.jmateco.2018.08.005>
- Blavatskyy, P. R., Ortman, A., & Panchenko, V. (2022). On the experimental robustness of the allais paradox. *American Economic Journal: Microeconomics*, 14(1), 143–163. <https://doi.org/10.1257/mic.20190153>
- Blavatskyy, P. R., Panchenko, V., & Ortman, A. (2023). How common is the common-ratio effect? *Experimental Economics*, 26(2), 253–272. <https://doi.org/10.1007/s10683-022-09761-y>
- Blavatskyy, P. R., & Pogrebna, G. (2009). Models of stochastic choice and decision theories: Why both are important for analyzing decisions. *Journal of Applied Econometrics*, 25(6), 963–986. <https://doi.org/10.1002/jae.1116>
- Block, H. D., & Marschak, J. (1960). Random orderings and stochastic theories of responses. In I. Olkin, S. Ghurye, H. Hoefding, W. Madow, & H. Mann (Eds.), *Contributions to probability and statistics* (pp. 97–132). Stanford University Press.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4), 700–765. <https://doi.org/10.1037/0033-295x.113.4.700>
- Bordalo, P., Gennaioli, N., & Shleifer, A. (2022). Salience. *Annual Review of Economics*, 14(1), 521–544. <https://doi.org/10.1146/annurev-economics-051520-011616>
- Brandstätter, E., & Brandstätter, H. (1996). What’s money worth? determinants of the subjective value of money. *Journal of Economic Psychology*, 17(4), 443–464. [https://doi.org/10.1016/0167-4870\(96\)00019-0](https://doi.org/10.1016/0167-4870(96)00019-0)
- Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, 113(2), 409–432. <https://doi.org/10.1037/0033-295x.113.2.409>
- Bröder, A. (2000). Assessing the empirical validity of the “take-the-best” heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1332–1346. <https://doi.org/10.1037/0278-7393.26.5.1332>
- Bronchetti, E. T., Dee, T. S., Huffman, D. B., & Magenheimer, E. (2013). When a nudge isn’t enough: Defaults and saving among low-income tax filers. *National Tax Journal*, 66(3), 609–634. <https://doi.org/10.17310/ntj.2013.3.04>

- Brown, A. S. (1979). Priming effects in semantic memory retrieval processes. *Journal of Experimental Psychology: Human Learning and Memory*, 5(2), 65–77. <https://doi.org/10.1037/0278-7393.5.2.65>
- Brown, G. D. A., & Walasek, L. (2023, June). Homo ordinalus and sampling models: The past, present, and future of decision by sampling. In *Sampling in judgment and decision making* (pp. 35–65). Cambridge University Press. <https://doi.org/10.1017/9781009002042.003>
- Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2. ed., rev. and enl.). University of California Press.
- Burton, T. J., & Balleine, B. W. (2022). The positive valence system, adaptive behaviour and the origins of reward (A. L. Milton, Ed.). *Emerging Topics in Life Sciences*, 6(5), 501–513. <https://doi.org/10.1042/etls20220007>
- Buschena, D., & Zilberman, D. (2000). Generalized expected utility, heteroscedastic error, and path dependence in risky choice. *Journal of Risk and Uncertainty*, 20(1), 67–88. <https://doi.org/10.1023/a:1007814719863>
- Butler, D. J. (2020). Intransitive preferences or choice errors? A reply to Birnbaum. *Judgment and Decision Making*, 15(6), 1052–1053. <https://doi.org/10.1017/s193029750000824x>
- Butler, D. J., & Pogrebna, G. (2018). Predictably intransitive preferences. *Judgment and Decision Making*, 13(3), 217–236. <https://doi.org/10.1017/s193029750000766x>
- Camus, A. (1955, December). *The myth of sisyphus*. Hamish Hamilton.
- Carroll, L. (1865). Alice’s adventures in wonderland. Macmillan.
- Cartwright, N. (2013). Knowing what we are talking about: Why evidence doesn’t always travel. *Evidence and Policy*, 9(1), 97–112. <https://doi.org/10.1332/174426413x662581>
- Castillo, G. (2020). The attraction effect and its explanations. *Games and Economic Behavior*, 119, 123–147. <https://doi.org/10.1016/j.geb.2019.10.012>
- Cataldo, A. M., & Cohen, A. L. (2018). The comparison process as an account of variation in the attraction, compromise, and similarity effects. *Psychonomic Bulletin & Review*, 26(3), 934–942. <https://doi.org/10.3758/s13423-018-1531-9>
- Cavagnaro, D. R., & Davis-Stober, C. P. (2014). Transitive in our preferences, but transitive in different ways: An analysis of choice variability. *Decision*, 1(2), 102–122. <https://doi.org/10.1037/dec0000011>
- Cavagnaro, D. R., & Regenwetter, M. (2023). Probabilistic choice induced by strength of preference. *Computational Brain & Behavior*, 6(4), 569–600. <https://doi.org/10.1007/s42113-023-00176-3>
- Cavagnaro, D. R., Yang, X., & Regenwetter, M. (2024). Choose for others as you would choose for yourself? a layered analysis of probabilistic preferential choice across social distances. *Journal of Economic Psychology*, 104, 102754. <https://doi.org/10.1016/j.joep.2024.102754>

- Cha, Y.-s., Choi, M., Guo, Y., Regenwetter, M., & Zwilling, C. (2013). Reply: Birnbaum's (2012) statistical tests of independence have unknown type-I error rates and do not replicate within participant. *Judgment and Decision Making*, 8(1), 55–73. <https://doi.org/10.1017/s1930297500004502>
- Chalmers, D. J. (2010). The representational character of experience. In *The character of consciousness*. Oxford University Press.
- Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. *Visual Information Processing*, 215–281. <https://doi.org/10.1016/b978-0-12-170150-5.50011-1>
- Chater, N. (2018, August). *The mind is flat: The remarkable shallowness of the improvising brain*. Yale University Press.
- Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *The Journal of Neuroscience*, 29(39), 12315–12320. <https://doi.org/10.1523/jneurosci.2575-09.2009>
- Chituc, V., Crockett, M., & Scholl, B. J. (2026). How to show that a cruel prank is worse than a war crime: Shifting scales and missing benchmarks in the study of moral judgment. *Cognition*, 266, 106315. <https://doi.org/10.1016/j.cognition.2025.106315>
- Clarke, F. R. (1957). Constant-ratio rule for confusion matrices in speech communication. *The Journal of the Acoustical Society of America*, 29(6), 715–720. <https://doi.org/10.1121/1.1909023>
- Clithero, J. A., & Rangel, A. (2014). Informatic parcellation of the network involved in the computation of subjective value. *Social Cognitive and Affective Neuroscience*, 9(9), 1289–1302. <https://doi.org/10.1093/scan/nst106>
- Coleman, J. S. (1986). Social theory, social research, and a theory of action. *American Journal of Sociology*, 91(6), 1309–1335. <https://doi.org/10.1086/228423>
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407–428. <https://doi.org/10.1037/0033-295x.82.6.407>
- Cronqvist, H., Thaler, R. H., & Yu, F. (2018). When nudges are forever: Inertia in the Swedish premium pension plan. *AEA Papers and Proceedings*, 108, 153–158. <https://doi.org/10.1257/pandp.20181096>
- Damasio, A. R. (1994). Descartes' error and the future of human life. *Scientific American*, 271(4), 144–144. Retrieved March 5, 2026, from <http://www.jstor.org/stable/24942888>
- 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. <https://doi.org/10.1098/rstb.1996.0125>

- Danan, E. (2010). Randomization vs. selection: How to choose in the absence of preference? *Management Science*, 56(3), 503–518. <https://doi.org/10.1287/mnsc.1090.1116>
- Dang, S., Antono, J. E., Kagan, I., & Pooresmaeili, A. (2024). Modality-specific and modality-general representations of subjective value in frontal cortex. *Communications Biology*, 7(1). <https://doi.org/10.1038/s42003-024-07253-8>
- Davidson, D., & Marschak, J. (1959). Experimental tests of a stochastic decision theory. In C. W. Churchman & P. Ratoosh (Eds.), *Measurement: Definitions and theories* (pp. 233–269). Wiley.
- Davidson, D., McKinsey, J. C. C., & Suppes, P. (1955). Outlines of a formal theory of value, i. *Philosophy of Science*, 22(2), 140–160. <https://doi.org/10.1086/287412>
- Davis, M. S. (1971). That’s interesting!: Towards a phenomenology of sociology and a sociology of phenomenology. *Philosophy of the Social Sciences*, 1(2), 309–344. <https://doi.org/10.1177/004839317100100211>
- Davis-Stober, C. P., Brown, N., Park, S., & Regenwetter, M. (2017). Recasting a biologically motivated computational model within a fechnerian and random utility framework. *Journal of Mathematical Psychology*, 77, 156–164. <https://doi.org/10.1016/j.jmp.2016.10.009>
- Davis-Stober, C. P., & Regenwetter, M. (2019). The ‘paradox’ of converging evidence. *Psychological Review*, 126(6), 865–879. <https://doi.org/10.1037/rev0000156>
- Dawkins, R. (1976, October). *The selfish gene*. Oxford University Press.
- de Groot, A. D. (1978/1965). Thought and choice in chess. Mouton Publishers.
- De Martino, B., & Cortese, A. (2023). Goals, usefulness and abstraction in value-based choice. *Trends in Cognitive Sciences*, 27(1), 65–80. <https://doi.org/10.1016/j.tics.2022.11.001>
- Debreu, G. (1960). *The American Economic Review*, 50(1), 186–188. Retrieved September 9, 2025, from <http://www.jstor.org/stable/1813477>
- Denrell, J. (2005). Why most people disapprove of me: Experience sampling in impression formation. *Psychological Review*, 112(4), 951–978. <https://doi.org/10.1037/0033-295x.112.4.951>
- Dertwinkel-Kalt, M., & Köster, M. (2019). Saliency and skewness preferences. *Journal of the European Economic Association*, 18(5), 2057–2107. <https://doi.org/10.1093/jeaa/jvz035>
- Dhmi, M. K., Hertwig, R., & Hoffrage, U. (2004). The Role of Representative Design in an Ecological Approach to Cognition. *Psychological Bulletin*, 130(6), 959–988. <https://doi.org/10.1037/0033-2909.130.6.959>
- Dickens, C. (1861). *Great Expectations by Charles Dickens: In three volumes*. Chapman; Hall.

- Dulberg, Z., Dubey, R., Berwian, I. M., & Cohen, J. D. (2023). Having multiple selves helps learning agents explore and adapt in complex changing worlds. *Proceedings of the National Academy of Sciences*, *120*(28). <https://doi.org/10.1073/pnas.2221180120>
- Dzhafarov, E. N., & Colonius, H. (2011). The fechnerian idea. *The American Journal of Psychology*, *124*(2), 127–140. <https://doi.org/10.5406/amerjpsyc.124.2.0127>
- Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology*. (H. A. Ruger & C. E. Bussenius, Eds.). Teachers College Press. <https://doi.org/10.1037/10011-000>
- Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, *51*(4), 380–417. <https://doi.org/10.1037/h0053870>
- Eggers, A. C., Tuñón, G., & Dafoe, A. (2024). Placebo tests for causal inference. *Am. J. Pol. Sci.*, *68*(3), 1106–1121.
- Elliot, A. J. (2013). Approach and avoidance motivation. In A. J. Elliot (Ed.), *Handbook of approach and avoidance motivation* (pp. 3–11). Psychology Press.
- Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*(1), 273–305. <https://doi.org/10.1146/annurev.psych.47.1.273>
- Fan, Y., & Park, S. S. (2009). Sharp bounds on the distribution of treatment effects and their statistical inference. *Econometric Theory*, *26*(3), 931–951. <https://doi.org/10.1017/s0266466609990168>
- Fazio, R. H. (2007). Attitudes as object–evaluation associations of varying strength. *Social Cognition*, *25*(5), 603–637. <https://doi.org/10.1521/soco.2007.25.5.603>
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*(2), 229–238. <https://doi.org/10.1037/0022-3514.50.2.229>
- Fechner, G. T. (1860). *Elemente der psychophysik. t. 2*. Breitkopf & Hartel.
- Fischer, H., Nilsson, M. E., & Ebner, N. C. (2023). Why the single-n design should be the default in affective neuroscience. *Affective Science*, *5*(1), 62–66. <https://doi.org/10.1007/s42761-023-00182-5>
- Fischhoff, B. (2006). Constructing preferences from labile values. In *The construction of preference* (pp. 653–667). Cambridge University Press.
- Fischhoff, B., & Slovic, P. (2025, October). *Decisions*. MIT Press.
- Fishburn, P. C. (1973). Binary choice probabilities: On the varieties of stochastic transitivity. *Journal of Mathematical Psychology*, *10*(4), 327–352. [https://doi.org/10.1016/0022-2496\(73\)90021-7](https://doi.org/10.1016/0022-2496(73)90021-7)
- Fishburn, P. C. (1974). Exceptional paper—lexicographic orders, utilities and decision rules: A survey. *Management Science*, *20*(11), 1442–1471. <https://doi.org/10.1287/mnsc.20.11.1442>

- Fishburn, P. C. (1991). Nontransitive preferences in decision theory. *Journal of Risk and Uncertainty*, 4(2), 113–134. <https://doi.org/10.1007/bf00056121>
- Fisher, I. (1892). *Mathematical investigations in the theory of value and prices* [Doctoral dissertation, Yale University].
- Fodor, J. A. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese*, 28(2), 97–115. <https://doi.org/10.1007/bf00485230>
- Forsgren, M. (2025). *Against analogy in behavioural laboratory experiments: A cognitive psychologists' take on parallelism* [Preprint]. <https://doi.org/10.2139/ssrn.5203596>
- Forsgren, M., Juslin, P., & van den Berg, R. (2023). Further perceptions of probability: In defence of associative models. *Psychological Review*, 130(5), 1383–1400. <https://doi.org/10.1037/rev0000410>
- Forsgren, M., Juslin, P., & van den Berg, R. (2025). Further perceptions of probability: Accurate, stepwise updating is contingent on prior information about the task and the response mode. *Psychonomic Bulletin & Review*, 32(3), 1284–1296. <https://doi.org/10.3758/s13423-024-02604-2>
- Forsgren, M., Mandl, B., & Karreskog Rehbinder, G. (2025). Probabilistic functionalism as a limiting condition for robustness. *Scientific Reports*, 15(1). <https://doi.org/10.1038/s41598-025-32580-z>
- Franconeri, S. L., Padilla, L. M., Shah, P., Zacks, J. M., & Hullman, J. (2021). The science of visual data communication: What works. *Psychological Science in the Public Interest*, 22(3), 110–161. <https://doi.org/10.1177/15291006211051956>
- Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology*, 65(1), 45–55. <https://doi.org/10.1037/0022-3514.65.1.45>
- 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). <https://doi.org/10.1126/sciadv.1701381>
- Gelman, A., & Loken, E. (2013, November). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time. https://stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf
- Gelman, A., & Shalizi, C. R. (2012). Philosophy and the practice of bayesian statistics. *British Journal of Mathematical and Statistical Psychology*, 66(1), 8–38. <https://doi.org/10.1111/j.2044-8317.2011.02037.x>
- Gershman, S. J. (2015). A unifying probabilistic view of associative learning (J. Diedrichsen, Ed.). *PLOS Computational Biology*, 11(11), e1004567. <https://doi.org/10.1371/journal.pcbi.1004567>

- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review*, *103*(3), 592–596.
- Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science*, *3*(1), 20–29. <https://doi.org/10.1111/j.1745-6916.2008.00058.x>
- Gigerenzer, G. (2025). The rationality wars: A personal reflection. *Behavioural Public Policy*, *9*(3), 495–515. <https://doi.org/10.1017/bpp.2024.51>
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*(1), 451–482. <https://doi.org/10.1146/annurev-psych-120709-145346>
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*(4), 650–669. <https://doi.org/10.1037/0033-295x.103.4.650>
- Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. Oxford University Press.
- Gluth, S., Sommer, T., Rieskamp, J., & Büchel, C. (2015). Effective connectivity between hippocampus and ventromedial prefrontal cortex controls preferential choices from memory. *Neuron*, *86*(4), 1078–1090. <https://doi.org/10.1016/j.neuron.2015.04.023>
- Gonzalez, C., & Aggarwal, P. (2023, June). Sequential decisions from sampling: Inductive generation of stopping decisions using instance-based learning theory. In *Sampling in judgment and decision making* (pp. 289–310). Cambridge University Press. <https://doi.org/10.1017/9781009002042.017>
- Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review*, *118*(4), 523–551. <https://doi.org/10.1037/a0024558>
- Granados Samayoa, J. A., & Fazio, R. H. (2024, August). Attitudes and social cognition: The synergistic interface. In *The Oxford handbook of social cognition, second edition* (pp. 671–693). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780197763414.013.24>
- Grether, D. M., & Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *The American Economic Review*, *69*(4), 623–638. Retrieved October 22, 2024, from <http://www.jstor.org/stable/1808708>
- Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics*, *72*(3), 561–582. <https://doi.org/10.3758/app.72.3.561>
- Hahn, U., & Chater, N. (1997, October). Concepts and similarity. In *Knowledge, concepts, and categories* (pp. 43–92). The MIT Press. <https://doi.org/10.7551/mitpress/4071.003.0006>

- Hamming, R. W. (1950). Error detecting and error correcting codes. *Bell System Technical Journal*, 29(2), 147–160. <https://doi.org/10.1002/j.1538-7305.1950.tb00463.x>
- Hansen, M. H., & Yu, B. (2001). Model selection and the principle of minimum description length. *Journal of the American Statistical Association*, 96(454), 746–774. <https://doi.org/10.1198/016214501753168398>
- He, J., & Natenzon, P. (2024). Moderate utility. *American Economic Review: Insights*, 6(2), 176–195. <https://doi.org/10.1257/aeri.20230085>
- Helson, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *The American Journal of Psychology*, 60(1), 1. <https://doi.org/10.2307/1417326>
- Herring, D. R., White, K. R., Jabeen, L. N., Hinojos, M., Terrazas, G., Reyes, S. M., Taylor, J. H., & Crites, S. L. (2013). On the automatic activation of attitudes: A quarter century of evaluative priming research. *Psychological Bulletin*, 139(5), 1062–1089. <https://doi.org/10.1037/a0031309>
- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, 15(8), 534–539. <https://doi.org/10.1111/j.0956-7976.2004.00715.x>
- Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, 13(12), 517–523. <https://doi.org/10.1016/j.tics.2009.09.004>
- Hodgetts, C. J., & Hahn, U. (2025). The limited place in cognitive space. *Topics in Cognitive Science*, 17(3), 662–680. <https://doi.org/10.1111/tops.70001>
- Hosokawa, T., Kato, K., Inoue, M., & Mikami, A. (2007). Neurons in the macaque orbitofrontal cortex code relative preference of both rewarding and aversive outcomes. *Neuroscience Research*, 57(3), 434–445. <https://doi.org/10.1016/j.neures.2006.12.003>
- Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, 9(1), 90. <https://doi.org/10.1086/208899>
- Huber, J., Payne, J. W., & Puto, C. P. (2014). Let’s be honest about the attraction effect. *Journal of Marketing Research*, 51(4), 520–525. <https://doi.org/10.1509/jmr.14.0208>
- Hull, C. L. (1943). *Principles of behavior : An introduction to behavior theory*. Appleton-Century-Crofts.
- Hullman, J., & Diakopoulos, N. (2011). Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12), 2231–2240. <https://doi.org/10.1109/tvcg.2011.255>

- Hullman, J., & Gelman, A. (2021). Designing for interactive exploratory data analysis requires theories of graphical inference. *Harvard Data Science Review*. <https://doi.org/10.1162/99608f92.3ab8a587>
- Hume, D. (1987/1777). Essays, moral, political, and literary. In E. F. Miller (Ed.), *Essays, moral, political, and literary* (Rev. ed.). Liberty Classics.
- Hursch, C. J., Hammond, K. R., & Hursch, J. L. (1964). Some methodological considerations in multiple-cue probability studies. *Psychological Review*, *71*(1), 42–60. <https://doi.org/10.1037/h0041729>
- Jachimowicz, J. M., Duncan, S., Weber, E. U., & Johnson, E. J. (2019). When and why defaults influence decisions: A meta-analysis of default effects. *Behavioural Public Policy*, *3*(02), 159–186. <https://doi.org/10.1017/bpp.2018.43>
- JASP Team. (2023). JASP (Version 0.17.3)[Computer software]. <https://jasp-stats.org/>
- Jevons, W. S. (1879). *The theory of political economy*. (2. ed. / Revised and enlarged). Macmillan; Co.
- Juechems, K., & Summerfield, C. (2019). Where does value come from? *Trends in Cognitive Sciences*, *23*(10), 836–850. <https://doi.org/10.1016/j.tics.2019.07.012>
- Juslin, P., Olsson, H., & Olsson, A.-C. (2003). Exemplar effects in categorization and multiple-cue judgment. *Journal of Experimental Psychology: General*, *132*(1), 133–156. <https://doi.org/10.1037/0096-3445.132.1.133>
- Kahneman, D. (2010). Comment: This is not how science is done [Published as a comment within Rakow (2010)]. *Judgment and Decision Making*, *5*(6), 465. <https://doi.org/10.1017/s1930297500001303>
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263. <https://doi.org/10.2307/1914185>
- Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, *103*(3), 582–591. <https://doi.org/10.1037/0033-295x.103.3.582>
- Kalenscher, T. (2010). Neural signatures of intransitive preferences. *Frontiers in Human Neuroscience*. <https://doi.org/10.3389/fnhum.2010.00049>
- Keramati, M., & Gutkin, B. (2014). Homeostatic reinforcement learning for integrating reward collection and physiological stability. *eLife*, *3*. <https://doi.org/10.7554/elife.04811>
- Klugkist, I., & Hoijtink, H. (2007). The bayes factor for inequality and about equality constrained models. *Computational Statistics & Data Analysis*, *51*(12), 6367–6379. <https://doi.org/10.1016/j.csda.2007.01.024>
- Knief, U., & Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods*, *53*(6), 2576–2590. <https://doi.org/10.3758/s13428-021-01587-5>

- Knight, F. H. (1957/1921). *Risk, uncertainty and profit*. (1st ed. reprinted with an additional chapter, 8th impression). London School of Economics & Political Science.
- Knudsen, E. B., & Wallis, J. D. (2022). Taking stock of value in the orbitofrontal cortex. *Nature Reviews Neuroscience*, 23(7), 428–438. <https://doi.org/10.1038/s41583-022-00589-2>
- Knutson, B., Rick, S., Wimmer, G. E., Prelec, D., & Loewenstein, G. (2007). Neural predictors of purchases. *Neuron*, 53(1), 147–156. <https://doi.org/10.1016/j.neuron.2006.11.010>
- Krieger, N. (2008). Proximal, distal, and the politics of causation: What's level got to do with it? *American Journal of Public Health*, 98(2), 221–230. <https://doi.org/10.2105/ajph.2007.111278>
- Laughlin, S. (1981). A simple coding procedure enhances a neuron's information capacity. *Zeitschrift für Naturforschung C*, 36(9–10), 910–912. <https://doi.org/10.1515/znc-1981-9-1040>
- Lebreton, M., Jorge, S., Michel, V., Thirion, B., & Pessiglione, M. (2009). An automatic valuation system in the human brain: Evidence from functional neuroimaging. *Neuron*, 64(3), 431–439. <https://doi.org/10.1016/j.neuron.2009.09.040>
- Lee, L., Amir, O., & Ariely, D. (2009). In search of homo economicus: Cognitive noise and the role of emotion in preference consistency. *Journal of Consumer Research*, 36(2), 173–187. <https://doi.org/10.1086/597160>
- Levy, D. J., & Glimcher, P. W. (2012). The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology*, 22(6), 1027–1038. <https://doi.org/10.1016/j.conb.2012.06.001>
- Lin, A., Adolphs, R., & Rangel, A. (2011). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience*, 7(3), 274–281. <https://doi.org/10.1093/scan/nsr006>
- Lind, M. (2024). Ensembles of optimal linear associative memory banks. <https://doi.org/10.31234/osf.io/hp7wu>
- Loftus, E. F. (1997). Creating false memories. *Scientific American*, 277(3), 70–75. Retrieved December 27, 2025, from <http://www.jstor.org/stable/24995913>
- Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, 92(368), 805. <https://doi.org/10.2307/2232669>
- Lopes, L. L., & Oden, G. C. (1999). The role of aspiration level in risky choice: A comparison of cumulative prospect theory and sp/a theory. *Journal of Mathematical Psychology*, 43(2), 286–313. <https://doi.org/10.1006/jmps.1999.1259>
- Luce, R. D. (1956). Semiorders and a theory of utility discrimination. *Econometrica*, 24(2), 178. <https://doi.org/10.2307/1905751>
- Luce, R. D. (1959). *Individual choice behavior : A theoretical analysis*. J. Wiley.

- Luce, R. D., & Suppes, P. (1965). Handbook of mathematical psychology. vol. 3, chapters 15-21. In *Handbook of mathematical psychology. vol. 3, chapters 15-21*. Wiley.
- Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health, 23*(1), 151–169. <https://doi.org/10.1146/annurev.publhealth.23.100901.140546>
- Lycan, W. (2023). Representational Theories of Consciousness. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (Winter 2023). Metaphysics Research Lab, Stanford University.
- MacLean, D. (2006, August). Informed consent and the construction of values. In *The construction of preference* (pp. 668–681). Cambridge University Press. <https://doi.org/10.1017/cbo9780511618031.037>
- Marr, D. C. (2010/1982, July). The philosophy and the approach. In *Vision* (pp. 8–38). The MIT Press. <https://doi.org/10.7551/mitpress/9780262514620.003.0002>
- Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*(4), 370–396. <https://doi.org/10.1037/h0054346>
- Mason, E. (2023). Value Pluralism. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (Summer 2023). Metaphysics Research Lab, Stanford University.
- Mayo, D. G. (1991). Novel evidence and severe tests. *Philosophy of Science, 58*(4), 523–552. <https://doi.org/10.1086/289639>
- McInnis, B., Cosley, D., Nam, C., & Leshed, G. (2016). Taking a hit: Designing around rejection, mistrust, risk, and workers’ experiences in amazon mechanical turk. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2271–2282. <https://doi.org/10.1145/2858036.2858539>
- McKenzie, C. M. R. (2018). Constructed preferences, rationality, and choice architecture. *Review of Behavioral Economics, 5*(3–4), 337–370. <https://doi.org/10.1561/105.00000091>
- McNamara, T. P. (2005). *Semantic priming : Perspectives from memory and word recognition* (1st ed.). Psychology Press.
- Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science, 34*(2), 103–115. <https://doi.org/10.1086/288135>
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525–543. <https://doi.org/10.1007/bf02294825>
- Mill, J. (2009/1879). *Utilitarianism*. Floating Press.
- 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, 33*(4), 477–491. <https://doi.org/10.1002/bdm.2171>

- 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. <https://doi.org/10.1017/s1930297500006161>
- Mischel, W. (2005). Alternative futures for our science. *APS Observer*, 18.
- Mischel, W. (2008). The toothbrush problem. *APS Observer*, 21.
- Mischel, W. (2009). Becoming a cumulative science. *APS Observer*, 22, 18.
- Moneta, N., Grossman, S., & Schuck, N. W. (2024). Representational spaces in orbitofrontal and ventromedial prefrontal cortex: Task states, values, and beyond. *Trends in Neurosciences*, 47(12), 1055–1069. <https://doi.org/10.1016/j.tins.2024.10.005>
- Mongin, P. (2019). The allais paradox: What it became, what it really was, what it now suggests to us. *Economics and Philosophy*, 35(3), 423–459. <https://doi.org/10.1017/s0266267118000469>
- Montano, H. (1628). *Repetitio legis imperialem de prohibita feudi alienat per fridericum*. Roncaliolus. <https://books.google.se/books?id=EH91AAAaAAJ>
- Morgenstern, O. (1979). Some reflections on utility. In *Expected utility hypotheses and the allais paradox* (pp. 175–183). Springer Netherlands. https://doi.org/10.1007/978-94-015-7629-1_6
- Moscatti, I. (2021, November). The routledge handbook of the philosophy of economics. In C. Heilmann & J. Reiss (Eds.). Routledge.
- Moscovitch, M. (2007). Part 1. memory. chapter 3: Why the engram is elusive. In *Science of memory* (pp. 17–22). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195310443.003.0002>
- Müller-Trede, J., Sher, S., & McKenzie, C. R. M. (2015). Transitivity in context: A rational analysis of intransitive choice and context-sensitive preference. *Decision*, 2(4), 280–305. <https://doi.org/10.1037/dec0000037>
- Nash, C. J., Cole, D. J., & Bigler, R. S. (2016). A review of human sensory dynamics for application to models of driver steering and speed control. *Biological Cybernetics*, 110(2–3), 91–116. <https://doi.org/10.1007/s00422-016-0682-x>
- Nassar, M. R., & Gold, J. I. (2013). A healthy fear of the unknown: Perspectives on the interpretation of parameter fits from computational models in neuroscience (O. Sporns, Ed.). *PLoS Computational Biology*, 9(4), e1003015. <https://doi.org/10.1371/journal.pcbi.1003015>
- Navarick, D. J., & Fantino, E. (1975). Stochastic transitivity and the unidimensional control of choice. *Learning and Motivation*, 6(2), 179–201. [https://doi.org/10.1016/0023-9690\(75\)90021-1](https://doi.org/10.1016/0023-9690(75)90021-1)
- Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? factors influencing “one-reason” decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(1), 53–65. <https://doi.org/10.1037/0278-7393.29.1.53>

- Noguchi, T., & Stewart, N. (2014). In the attraction, compromise, and similarity effects, alternatives are repeatedly compared in pairs on single dimensions. *Cognition*, *132*(1), 44–56. <https://doi.org/10.1016/j.cognition.2014.03.006>
- Noguchi, T., & Stewart, N. (2018). Multialternative decision by sampling: A model of decision making constrained by process data. *Psychological Review*, *125*(4), 512–544. <https://doi.org/10.1037/rev0000102>
- Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 104–114. <https://doi.org/10.1037/0278-7393.10.1.104>
- Oprea, R. (2024). Decisions under risk are decisions under complexity. *American Economic Review*, *114*(12), 3789–3811. <https://doi.org/10.1257/aer.20221227>
- Padoa-Schioppa, C. (2011). Neurobiology of economic choice: A good-based model. *Annual Review of Neuroscience*, *34*(1), 333–359. <https://doi.org/10.1146/annurev-neuro-061010-113648>
- Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223–226. <https://doi.org/10.1038/nature04676>
- Padoa-Schioppa, C., & Assad, J. A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature Neuroscience*, *11*(1), 95–102. <https://doi.org/10.1038/nn2020>
- Palan, S., & Schitter, C. (2018). Prolific.ac—a subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. <https://doi.org/10.1016/j.jbef.2017.12.004>
- Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review*, *72*(6), 407–418. <https://doi.org/10.1037/h0022602>
- Parducci, A. (1995). *Happiness, pleasure, and judgment: The contextual theory and its applications*. Lawrence Erlbaum Associates.
- Payne, J. W., Bettman, J. R., Coupey, E., & Johnson, E. J. (1992). A constructive process view of decision making: Multiple strategies in judgment and choice. *Acta Psychologica*, *80*(1–3), 107–141. [https://doi.org/10.1016/0001-6918\(92\)90043-d](https://doi.org/10.1016/0001-6918(92)90043-d)
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. <https://doi.org/10.1016/j.jesp.2017.01.006>
- Pessiglione, M., & Delgado, M. R. (2015). The good, the bad and the brain: Neural correlates of appetitive and aversive values underlying decision making. *Current Opinion in Behavioral Sciences*, *5*, 78–84. <https://doi.org/10.1016/j.cobeha.2015.08.006>
- Peterson, J. C., Bourgin, D. D., Agrawal, M., Reichman, D., & Griffiths, T. L. (2021). Using large-scale experiments and machine learning to dis-

- cover theories of human decision-making. *Science*, 372(6547), 1209–1214. <https://doi.org/10.1126/science.abe2629>
- Pitt, M. A., Myung, I. J., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review*, 109(3), 472–491. <https://doi.org/10.1037/0033-295x.109.3.472>
- Plassmann, H., O’Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *The Journal of Neuroscience*, 27(37), 9984–9988. <https://doi.org/10.1523/jneurosci.2131-07.2007>
- Plassmann, H., O’Doherty, J. P., & Rangel, A. (2010). Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *The Journal of Neuroscience*, 30(32), 10799–10808. <https://doi.org/10.1523/jneurosci.0788-10.2010>
- Platt, J. R. (1964). Strong inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science*, 146(3642), 347–353. <https://doi.org/10.1126/science.146.3642.347>
- Pliner, P., & Hobden, K. (1992). Development of a scale to measure the trait of food neophobia in humans. *Appetite*, 19(2), 105–120. [https://doi.org/10.1016/0195-6663\(92\)90014-w](https://doi.org/10.1016/0195-6663(92)90014-w)
- Polanía, R., Woodford, M., & Ruff, C. C. (2018). Efficient coding of subjective value. *Nature Neuroscience*, 22(1), 134–142. <https://doi.org/10.1038/s41593-018-0292-0>
- Quillian, M. (1966). *Semantic memory* [Unpublished Ph.d. dissertation]. Carnegie Institute of Technology.
- Rakow, T. (2010). Risk, uncertainty and prophet: The psychological insights of frank h. knight. *Judgment and Decision Making*, 5(6), 458–466. <https://doi.org/10.1017/s1930297500001303>
- Rakow, T., Cheung, N. Y., & Restelli, C. (2020). Losing my loss aversion: The effects of current and past environment on the relative sensitivity to losses and gains. *Psychonomic Bulletin & Review*, 27(6), 1333–1340. <https://doi.org/10.3758/s13423-020-01775-y>
- Ranyard, R. (1977). Risky decisions which violate transitivity and double cancellation. *Acta Psychologica*, 41(6), 449–459. [https://doi.org/10.1016/0001-6918\(77\)90003-8](https://doi.org/10.1016/0001-6918(77)90003-8)
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108. <https://doi.org/10.1037/0033-295x.85.2.59>
- Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2010). Testing Transitivity of Preferences on Two-Alternative Forced Choice Data. *Frontiers in Psychology*, 1. <https://doi.org/10.3389/fpsyg.2010.00148>
- Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2011). Transitivity of preferences. *Psychological Review*, 118(1), 42–56. <https://doi.org/10.1037/a0021150>
- Regenwetter, M., & Davis-Stober, C. P. (2008). There are many models of transitive preference: A tutorial review and current perspective. In *De-*

- cision modeling and behavior in complex and uncertain environments* (pp. 99–124). Springer New York. https://doi.org/10.1007/978-0-387-77131-1_5
- Regenwetter, M., Robinson, M. M., & Wang, C. (2022). Four internal inconsistencies in tversky and kahneman’s (1992) cumulative prospect theory article: A case study in ambiguous theoretical scope and ambiguous parsimony. *Advances in Methods and Practices in Psychological Science*, 5(1). <https://doi.org/10.1177/25152459221074653>
- Rescorla, R. A. (1988). Pavlovian conditioning: It’s not what you think it is. *American Psychologist*, 43(3), 151–160. <https://doi.org/10.1037/0003-066x.43.3.151>
- Richter, M. K. (1966). Revealed preference theory. *Econometrica*, 34(3), 635. <https://doi.org/10.2307/1909773>
- Riddell, A., Hartikainen, A., & Carter, M. (2021, March). Pystan (3.0.0).
- Rieskamp, J., Busemeyer, J. R., & Mellers, B. A. (2006). Extending the bounds of rationality: Evidence and theories of preferential choice. *Journal of Economic Literature*, 44(3), 631–661. <https://doi.org/10.1257/jel.44.3.631>
- Ritter, F. E., Tehranchi, F., & Oury, J. D. (2018). Act□: A cognitive architecture for modeling cognition. *WIREs Cognitive Science*, 10(3). <https://doi.org/10.1002/wcs.1488>
- Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionst model of decision making. *Psychological Review*, 108, 370–392. <https://doi.org/10.1037/0033-295X.108.2.370>
- Rosas, J. M., Todd, T. P., & Bouton, M. E. (2013). Context change and associative learning. *WIREs Cognitive Science*, 4(3), 237–244. <https://doi.org/10.1002/wcs.1225>
- Ross, H. E., & Brodie, E. E. (1987). Weber fractions for weight and mass as a function of stimulus intensity. *The Quarterly Journal of Experimental Psychology Section A*, 39(1), 77–88. <https://doi.org/10.1080/02724988743000042>
- Ryan, M., & Miguel, F. S. (2002). Revisiting the axiom of completeness in health care. *Health Economics*, 12(4), 295–307. <https://doi.org/10.1002/hec.730>
- Samuelson, P. A. (1938). A note on the pure theory of consumer’s behaviour. *Economica*, 5(17), 61–71. <https://doi.org/10.2307/2548836>
- Scarfe, T. ((2023-present, February). Prof. chris summerfield: Natural general intelligence (no. 104) [audio podcast episode] [Audio podcast episode]. <https://www.podcasts.nu/episodes/machine-learning-street-talk/104-prof-chris-summerfield-natural-general-intelligence-special-edition-8zonEGbuHU>

- Schacter, D. L. (2007). Part 1. memory. chapter 4: Delineating the core. In *Science of memory: Concepts* (pp. 23–28). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195310443.003.0002>
- Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481), 773–786. <https://doi.org/10.1098/rstb.2007.2087>
- Schacter, D. L., Addis, D. R., & Buckner, R. L. (2008). Episodic simulation of future events: Concepts, data, and applications. *Annals of the New York Academy of Sciences*, 1124(1), 39–60. <https://doi.org/10.1196/annals.1440.001>
- Schelling, T. C. (1971). Dynamic models of segregation†. *The Journal of Mathematical Sociology*, 1(2), 143–186. <https://doi.org/10.1080/0022250x.1971.9989794>
- Schmidt, A. F., & Finan, C. (2018). Linear regression and the normality assumption. *Journal of Clinical Epidemiology*, 98, 146–151. <https://doi.org/10.1016/j.jclinepi.2017.12.006>
- Schoemaker, P. J. H. (1982). The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature*, 20(2), 529–563. Retrieved March 21, 2025, from <http://www.jstor.org/stable/2724488>
- Schroeder, M. (2024). Value Theory. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (Winter 2024). Metaphysics Research Lab, Stanford University.
- Schumpeter, J. (1909). On the concept of social value. *The Quarterly Journal of Economics*, 23(2), 213. <https://doi.org/10.2307/1882798>
- Schumpeter, J. (1984). The meaning of rationality in the social sciences. *Zeitschrift für die gesamte Staatswissenschaft / Journal of Institutional and Theoretical Economics*, 140(4), 577–593. Retrieved August 31, 2025, from <http://www.jstor.org/stable/40750743>
- Seymour, B., & McClure, S. M. (2008). Anchors, scales and the relative coding of value in the brain. *Current Opinion in Neurobiology*, 18(2), 173–178. <https://doi.org/10.1016/j.conb.2008.07.010>
- Shadlen, M. N., & Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron*, 90(5), 927–939. <https://doi.org/10.1016/j.neuron.2016.04.036>
- Shafir, E. (1993). Reason-based choice. *Cognition*, 49(1–2), 11–36. [https://doi.org/10.1016/0010-0277\(93\)90034-s](https://doi.org/10.1016/0010-0277(93)90034-s)
- Shapiro, L. A. (2000). Multiple realizations. *The Journal of Philosophy*, 97(12), 635. <https://doi.org/10.2307/2678460>
- Sher, S., Müller-Trede, J., & McKenzie, C. R. M. (2025). Choices without preferences: Principles of rational arbitrariness. *Psychological Review*, 132(6), 1375–1395. <https://doi.org/10.1037/rev0000529>

- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99. <https://doi.org/10.2307/1884852>
- Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, 41(1), 1–20. <https://doi.org/10.1146/annurev.ps.41.020190.000245>
- Simonsohn, U. (2024, September). *Off-label smirnov: How many subjects show an effect in between-subjects experiments?* [Blog post]. <https://datacolada.org/120>
- Simonsohn, U., Montealegre, A., & Evangelidis, I. (2025). Stimulus sampling reimaged: Designing experiments with mix-and-match, analyzing results with stimulus plots. *Journal of Personality and Social Psychology*, 129(1), 71–90. <https://doi.org/10.1037/pspa0000449>
- Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, 16(2), 158. <https://doi.org/10.1086/209205>
- Skinner, B. F. (1958). Reinforcement today. *American Psychologist*, 13(3), 94–99. <https://doi.org/10.1037/h0049039>
- Slovic, P. (1995). The construction of preference. *American Psychologist*, 50(5), 364–371. <https://doi.org/10.1037/0003-066x.50.5.364>
- Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research*, 177(3), 1333–1352. <https://doi.org/10.1016/j.ejor.2005.04.006>
- Smith, A. (2025/1776a, March). An inquiry into the nature and causes of the wealth of nations. In M. Garnier (Ed.). Anson Street Press.
- Smith, A. (2025/1776b, March). An inquiry into the nature and causes of the wealth of nations. In M. Garnier (Ed.). Anson Street Press.
- Smith, K. (1984). "drive": In defense of a concept. *Behaviorism*, 12(1), 71–114. Retrieved March 4, 2024, from <http://www.jstor.org/stable/27759036>
- Smith, S. M., & Krajbich, I. (2021). Mental representations distinguish value-based decisions from perceptual decisions. *Psychonomic Bulletin & Review*, 28(4), 1413–1422. <https://doi.org/10.3758/s13423-021-01911-2>
- Smith, V. L. (1962). An experimental study of competitive market behavior. *Journal of Political Economy*, 70(2), 111–137. <https://doi.org/10.1086/258609>
- Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, 38(2), 332–382. <https://doi.org/10.1257/jel.38.2.332>
- Stewart, N. (2009). Eps prize lecture: Decision by sampling: The role of the decision environment in risky choice. *Quarterly Journal of Experimental Psychology*, 62(6), 1041–1062. <https://doi.org/10.1080/17470210902747112>

- Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology*, 53(1), 1–26. <https://doi.org/10.1016/j.cogpsych.2005.10.003>
- Stewart, N., Reimers, S., & Harris, A. J. L. (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. <https://doi.org/10.1287/mnsc.2013.1853>
- Stewart, N., & Simpson, K. (2008, March). A decision-by-sampling account of decision under risk. In *The probabilistic mind*: (pp. 261–276). Oxford University Press/Oxford. <https://doi.org/10.1093/acprof:oso/9780199216093.003.0012>
- Sugden, R. (2009). On nudging: A review of nudge: Improving decisions about health, wealth and happiness by richard h. thaler and cass r. sunstein. *International Journal of the Economics of Business*, 16(3), 365–373. <https://doi.org/10.1080/13571510903227064>
- Sundh, J., Collsiö, A., Millroth, P., & Juslin, P. (2021). Precise/not precise (pnp): A brunswikian model that uses judgment error distributions to identify cognitive processes. *Psychonomic Bulletin & Review*, 28, 351–373. <https://doi.org/10.3758/s13423-020-01805-9>
- Sundh, J., Millroth, P., Collsiö, A., & Juslin, P. (2025). Enriching psychological research by exploring the source and nature of noise. *Perspectives on Psychological Science*, 20(3), 540–554. <https://doi.org/10.1177/17456916241235889>
- Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, 47(2), 143–148. [https://doi.org/10.1016/0001-6918\(81\)90005-6](https://doi.org/10.1016/0001-6918(81)90005-6)
- Szaszi, B., Palinkas, A., Palfi, B., Szollosi, A., & Aczel, B. (2017). A systematic scoping review of the choice architecture movement: Toward understanding when and why nudges work. *Journal of Behavioral Decision Making*, 31(3), 355–366. <https://doi.org/10.1002/bdm.2035>
- Szpunar, K. K., Spreng, R. N., & Schacter, D. L. (2014). A taxonomy of prospection: Introducing an organizational framework for future-oriented cognition. *Proceedings of the National Academy of Sciences*, 111(52), 18414–18421. <https://doi.org/10.1073/pnas.1417144111>
- Tan, M., & Jaeger, T. F. (2025). Learning to understand an unfamiliar talker: Testing distributional learning as a model of rapid adaptive speech perception. *Cognition*, 265, 106195. <https://doi.org/10.1016/j.cognition.2025.106195>
- Tannenbaum, D., Fox, C. R., & Rogers, T. (2017). On the misplaced politics of behavioural policy interventions. *Nature Human Behaviour*, 1(7). <https://doi.org/10.1038/s41562-017-0130>
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

- Thomson, J. J. (1997). The right and the good. *The Journal of Philosophy*, 94(6), 273. <https://doi.org/10.2307/2564542>
- Thurstone, L. L. (1927a). A law of comparative judgment. *Psychological Review*, 34(4), 273–286. <https://doi.org/10.1037/h0070288>
- Thurstone, L. L. (1927b). Psychophysical analysis. *The American Journal of Psychology*, 38(3), 368. <https://doi.org/10.2307/1415006>
- Thurstone, L. L. (1931). The indifference function. *The Journal of Social Psychology*, 2(2), 139–167. <https://doi.org/10.1080/00224545.1931.9918964>
- Thurstone, L. L. (1954). The measurement of values. *Psychological Review*, 61(1), 47–58. <https://doi.org/10.1037/h0060035>
- Thurstone, L. L., & Chave, E. J. (1929). Theory of attitude measurement. In *The measurement of attitude: A psychophysical method and some experiments with a scale for measuring attitude toward the church*. (pp. 1–21). The University of Chicago Press. <https://doi.org/10.1037/11574-001>
- Toates, F. M. (1986, February). *Problems in the behavioural sciences: Motivational systems series number 4*. Cambridge University Press.
- Todd, P. M., & Gigerenzer, G. (2000). Précis of simple heuristics that make us smart. *Behavioral and Brain Sciences*, 23(5), 727–741. <https://doi.org/10.1017/s0140525x00003447>
- Tolkien, J. R. R. (1954, July 29). *The fellowship of the ring* [First edition, first impression]. George Allen & Unwin.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4), 189–208. <https://doi.org/10.1037/h0061626>
- Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological Review*, 121, 179–205. <https://doi.org/10.1037/a0036137>
- Tukey, J. W. (1993). *Exploratory data analysis: Past, present and future* (Technical Report No. 302). Princeton University. Princeton, NJ.
- Tulving, E. (1972). Episodic and Semantic Memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory* (pp. 381–403). Academic Press.
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76(1), 31–48. <https://doi.org/10.1037/h0026750>
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, 79(4), 281–299. <https://doi.org/10.1037/h0032955>
- Tversky, A. (1975). A critique of expected utility theory: Descriptive and normative considerations. *Erkenntnis* (1975-), 9(2), 163–173. Retrieved September 2, 2025, from <http://www.jstor.org/stable/20010465>
- Tversky, A., & Edward Russo, J. (1969). Substitutability and similarity in binary choices. *Journal of Mathematical Psychology*, 6(1), 1–12. [https://doi.org/10.1016/0022-2496\(69\)90027-3](https://doi.org/10.1016/0022-2496(69)90027-3)

- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458. <https://doi.org/10.1126/science.7455683>
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. <https://doi.org/10.1007/bf00122574>
- Tversky, A., & Thaler, R. H. (1990). Anomalies: Preference reversals. *Journal of Economic Perspectives*, *4*(2), 201–211. <https://doi.org/10.1257/jep.4.2.201>
- Tye, M. (2021). Qualia. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2021). Metaphysics Research Lab, Stanford University.
- Ungemach, C., Stewart, N., & Reimers, S. (2011). How incidental values from the environment affect decisions about money, risk, and delay. *Psychological Science*, *22*(2), 253–260. <https://doi.org/10.1177/0956797610396225>
- Vehtari, A. (2025, April). Prior choice recommendations [Edited page on the stan-dev/stan GitHub wiki. Accessed: 2025-10-30.]. <https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations>
- Veit, W. (2024). What to make of intransitive preferences? <https://philsci-arch.ive.pitt.edu/23047/>
- Vlaev, I., Chater, N., & Stewart, N. (2007a). Relativistic financial decisions: Context effects on retirement saving and investment risk preferences. *Judgment and Decision Making*, *2*(5), 292–311. <https://doi.org/10.1017/s1930297500000619>
- Vlaev, I., Chater, N., & Stewart, N. (2007b). Financial prospect relativity: Context effects in financial decision-making under risk. *Journal of Behavioral Decision Making*, *20*(3), 273–304. <https://doi.org/10.1002/bdm.555>
- Vlaev, I., Chater, N., & Stewart, N. (2009). Dimensionality of risk perception: Factors affecting consumer understanding and evaluation of financial risk. *Journal of Behavioral Finance*, *10*(3), 158–181. <https://doi.org/10.1080/15427560903167720>
- Vlaev, I., Chater, N., Stewart, N., & Brown, G. D. (2011). Does the brain calculate value? *Trends in Cognitive Sciences*, *15*(11), 546–554. <https://doi.org/10.1016/j.tics.2011.09.008>
- von Mises, L. (2009/1912). *The theory of money and credit* (J. E. Batson, Trans.). Ludwig von Mises Institute. <https://cdn.mises.org/Theory%20of%20Money%20and%20Credit.pdf>
- Von Neumann, J., & Morgenstern, O. (2004/1944). *Theory of games and economic behavior* (Sixtieth-Anniversary ed.). Princeton University Press.

- Walasek, L., & Brown, G. D. A. (2023). Incomparability and incommensurability in choice: No common currency of value? *Perspectives on Psychological Science*, *19*(6), 1011–1030. <https://doi.org/10.1177/17456916231192828>
- Walasek, L., Mullett, T. L., & Stewart, N. (2021). Acceptance of mixed gambles is sensitive to the range of gains and losses experienced, and estimates of lambda (λ) are not a reliable measure of loss aversion: Reply to andré and de langhe (2021). *Journal of Experimental Psychology: General*, *150*(12), 2666–2670. <https://doi.org/10.1037/xge0001054>
- Walasek, L., & Stewart, N. (2015). How to make loss aversion disappear and reverse: Tests of the decision by sampling origin of loss aversion. *Journal of Experimental Psychology: General*, *144*(1), 7–11. <https://doi.org/10.1037/xge0000039>
- Walasek, L., & Stewart, N. (2019). Context-dependent sensitivity to losses: Range and skew manipulations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(6), 957–968. <https://doi.org/10.1037/xlm0000629>
- Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, *20*, 158–177. <https://doi.org/10.1037/h0074428>
- 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. <https://doi.org/10.1002/bdm.414>
- Weber, E. U., & Johnson, E. J. (2006, August). Constructing preferences from memory. In *The construction of preference* (pp. 397–410). Cambridge University Press. <https://doi.org/10.1017/cbo9780511618031.022>
- Weber, E. U., & Johnson, E. J. (2011). Query theory: Knowing what we want by arguing with ourselves. *Behavioral and Brain Sciences*, *34*(2), 91–92. <https://doi.org/10.1017/s0140525x10002797>
- Weber, M. (1947). The theory of social and economic organization [Chapter II, Section 1: “The Concept of Economic Action”] (A. M. Henderson & T. Parsons, Trans.). In T. Parsons (Ed.). Oxford University Press.
- Weber, M. (1975/1908). Marginal utility theory and “the fundamental law of psychophysics” (L. Schneider, Trans.). *Social Science Quarterly*, *56*(1), 21–36. Retrieved August 1, 2025, from <http://www.jstor.org/stable/42859467>
- Weber, N., & Brewer, N. (2003). The effect of judgment type and confidence scale on confidence-accuracy calibration in face recognition. *Journal of Applied Psychology*, *88*(3), 490–499. <https://doi.org/10.1037/0021-9010.88.3.490>
- Weibull, J. W. (2004). Testing game theory. In *Advances in understanding strategic behaviour* (pp. 85–104). Palgrave Macmillan UK. https://doi.org/10.1057/9780230523371_6

- Weicker, M., & Schulz, P. (2020). Not all gradable adjectives are vague – experimental evidence from adults and children. *Proceedings of Sinn und Bedeutung*, Vol 24 No 2 (2020): Proceedings of Sinn und Bedeutung 24. <https://doi.org/10.18148/SUB/2020.V24I2.905>
- Wichchukit, S., & O'Mahony, M. (2014). The 9-point hedonic scale and hedonic ranking in food science: Some reappraisals and alternatives. *Journal of the Science of Food and Agriculture*, 95(11), 2167–2178. <https://doi.org/10.1002/jsfa.6993>
- Wicker, A. W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, 25(4), 41–78. <https://doi.org/10.1111/j.1540-4560.1969.tb00619.x>
- Widrow, B., & Hoff, M. E. (1960). *Adaptive switching circuits* (tech. rep.). Stanford Electronics Laboratories.
- Yang, K., Tu, J., & Chen, T. (2019). Homoscedasticity: An overlooked critical assumption for linear regression. *General Psychiatry*, 32(5), e100148. <https://doi.org/10.1136/gpsych-2019-100148>
- Ylikoski, P. (2016). Methodological individualism [eBook]. In L. McIntyre & A. Rosenberg (Eds.), *The routledge companion to philosophy of social science* (1st, p. 12). Routledge.
- Yoshida, N., Sprekeler, H., & Gutkin, B. (2025). Linking homeostasis to reinforcement learning: Internal state control of motivated behavior. *Current Opinion in Behavioral Sciences*, 66, 101611. <https://doi.org/10.1016/j.cobeha.2025.101611>
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35(2), 151–175. <https://doi.org/10.1037/0003-066x.35.2.151>
- Zhu, J.-Q., Sundh, J., Spicer, J., Chater, N., & Sanborn, A. N. (2024). The autocorrelated bayesian sampler: A rational process for probability judgments, estimates, confidence intervals, choices, confidence judgments, and response times. *Psychological Review*, 131(2), 456–493. <https://doi.org/10.1037/rev0000427>
- Zilker, V., Hertwig, R., & Pachur, T. (2020). Age differences in risk attitude are shaped by option complexity. *Journal of Experimental Psychology: General*, 149(9), 1644–1683. <https://doi.org/10.1037/xge0000741>
- Zilker, V., & Pachur, T. (2022). Nonlinear probability weighting can reflect attentional biases in sequential sampling. *Psychological Review*, 129(5), 949–975. <https://doi.org/10.1037/rev0000304>
- Zwilling, C. E., Cavagnaro, D. R., Regenwetter, M., Lim, S. H., Fields, B., & Zhang, Y. (2019). QTest 2.1: Quantitative testing of theories of binary choice using bayesian inference. *Journal of Mathematical Psychology*, 91, 176–194. <https://doi.org/10.1016/j.jmp.2019.05.002>
- Zylberberg, A., Bakkour, A., Shohamy, D., & Shadlen, M. N. (2024). Value construction through sequential sampling explains serial dependencies in decision making. *eLife*, 13. <https://doi.org/10.7554/elife.96997.3>

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 243

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”.)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-582031



ACTA UNIVERSITATIS
UPSALIENSIS
2026