Perspectives on the mechanistic underpinnings of choice biases

Konstantinos Tsetsos[1,2] and Yinan Cao[2]

[1] School of Psychological Science, University of Bristol, Bristol, United Kingdom
[2] Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Correspondence: k.tsetsos@bristol.ac.uk & ycaoneuro@gmail.com

Abstract

Early foundational work in the decision sciences carefully balanced empirical observations and theoretical explanations. Dating back to Daniel Bernoulli, a handful of behavioral regularities observed in theoretical lotteries ignited the refinement of normative theories and the development of new descriptive frameworks of valuation and choice. However, more recent tendencies in behavioral economics and psychology place empirical observations on a pedestal: modern behavioral science has identified more behavioral biases than it has explained. Coupled with replication and reliability crises in experimental psychology, this has resulted in an explanatory gap in the field, in-between the descriptive and predictive levels. Here, we aim to close this explanatory gap by asking how choice biases can emerge from certain decision computations. We demonstrate that biased and irrational choice behavior may arise from multiple, equally viable mechanisms, such as relative value coding and selective information sampling. We posit that this "multiple realizability" problem highlights a broader issue: inferring mechanisms of complex behavior solely from behavioral measures is an underdetermined exercise. We propose that using time-resolved neural recordings to track how attention serially parses complex information during multiattribute, multialternative decisions can resolve this "multiple realizability" issue and arbitrate between competing mechanistic explanations of choice biases.

1. Introduction

A large body of research within the cognitive and decision sciences has established that human decisions are often influenced by factors that should rationally be ignored (De Martino et al., 2006; Kahneman & Tversky, 1984; Summerfield & Tsetsos, 2015; Tversky & Kahneman, 1981). For example, we tend to stick with an energy plan because it is set as the default option (Baron & Ritov, 1994), we prefer volatile stocks when buying but dismiss them when selling (Shafir et al., 1993; Tsetsos, Chater, et al., 2012), or we choose salmon fillet over ribeye steak simply because we noticed that fish fingers are also available (Huber et al., 1982). These examples illustrate that choices are not solely determined by the properties of the available alternatives and the goals of the decision-maker but also by an array of normatively irrelevant factors including: the way the alternatives are presented, the framing of the choice, or the presence of dominated (inferior) alternatives in the choice-set (Usher et al., 2019).

The observed sensitivity of human decisions to normatively irrelevant factors (reflected in *choice biases*) has had profound impact on the behavioral sciences, resulting in the distinction between normative (how *should* we decide) and descriptive (how *do* we decide) theories of choice. One of the main goals of descriptive theories of choice is to specify how choice biases come about in the deciding brain. Achieving this goal can subsequently inform deeper multidisciplinary considerations, e.g., on why bias propensity varies among individuals (Aczel et al., 2015; Spektor et al., 2021) or across the lifespan (Parrish et al., 2024; Tentori et al., 2001); or on why biases have persisted despite evolutionary pressure for reward-maximizing choices (Moran & Tsetsos, 2018; Tsetsos et al., 2016; Webb et al., 2021). Furthermore, choice biases have taken center stage in applied behavioral science, representing predictable blind spots that can be harnessed in interventions to induce behavioral change (Thaler & Sunstein, 2009). Precisely understanding the mechanisms that mediate choice biases can aid the development of targeted approaches that could bolster the limited effectiveness of extant "nudging" interventions (Maier et al., 2022).

Despite the theoretical and applied importance of understanding the computational and neural mechanisms that lead to choice biases, existing theories of choice have not provided definitive insights. This is reflected in the multitude and disparity of frameworks proposed to explain choice biases, ranging from verbally formulated heuristics (Gigerenzer & Gaissmaier, 2011; Kahneman & Tversky, 1984; Shafir et al., 1993) and algebraic modifications of normative theories (Tversky & Kahneman, 1992; Tversky & Simonson, 1993) to Bayesian (Bhui & Xiang, 2021; Srivastava & Schrater, 2012) and dynamical models (Busemeyer et al., 2019). To date, these disparate explanations of choice biases, often casted at different levels of analyses (McClelland, 2009), have not been comprehensively related to underlying decision mechanisms. In this chapter, we aim to close this gap by describing how explanations of hallmark choice biases can be situated along the processing stages that occur during decision-making.

We begin by clarifying the notion of choice bias and proceed to show that hallmark choice biases could fall out from computations occurring at almost any stage—including relative and non-linear value coding (Louie & De Martino, 2014), selective information sampling (Usher et al., 2019), and non-linear accumulation dynamics (Cavanagh et al., 2020). Given this "multiple realizability" issue, we then address how distorting mechanisms along the processing pathway can be better identified. We highlight that during complex decisions, information is sampled partially and serially before reaching the decision formation level; and propose that tracing this information flow with non-invasive high-temporal-resolution neural recordings as decisions unfold can considerably constrain mechanistic inferences. We conclude that a central goal in the cognitive and neural sciences should be to understand the principles that orchestrate information sampling during decision-making.

2. Choice biases: innocuous and irrational

From a normative standpoint, an optimal agent should always (or more likely, in the presence of behavioral stochasticity (Loomes & Sugden, 1995)) choose the most desirable course of action in any given situation (Summerfield & Tsetsos, 2015). Thus, a prerequisite of optimal behavior is the ability to value choice alternatives by transforming their objective properties into "desirability" (or utility) scores based on a set of criteria that represent the goals and needs of the agent at any given moment (Juechems & Summerfield, 2019). A longstanding question in the decision sciences is whether human choices comply with this notion of optimality. This has been hard to assess since, in real-life decisions, decision-relevant criteria are inherently subjective and opaque. To circumvent this issue, decision theorists have resorted to two distinct approaches: i) studying choice behavior in simplified scenarios where decision-relevant criteria can be objectively defined; ii) specifying a set of rules (or axioms) that utility-maximizing agents must abide by. In the following, we define choice biases in relation to these two approaches.

2.1 Innocuous biases

To study human valuation and choice in a tractable way, researchers often rely on laboratory tasks with externally imposed objectives (Summerfield & Tsetsos, 2012). This approach has roots in the foundational years of probability theory, where theoretical lotteries were used to benchmark human behavior against statistical optimality (Stearns, 2000). In these cases, it can be assumed that the desirability of each lottery option is equivalent to the expected value of the corresponding payout distribution. Thus, statistically optimal agents should always choose the lottery with the largest expected value. However, even in these tractable situations, human choices depart from the statistical ideal. Consider a choice between two prospects *A* and *B*. Prospect *A* offers £40 for sure while prospect *B* offers £100 with 50% probability or £0 otherwise. Although prospect *B* has a higher expected value ($EV_A = £40 \text{ vs.} EV_B = £50$), most people in this scenario will choose prospect *A* (Kahneman & Tversky, 1979) (Fig. 1A). This example illustrates that human valuation is sensitive to the variances of the payout distributions, which is not relevant for maximizing expected rewards. More broadly, human economic choices disclose several idiosyncrasies, such as risk (Mata et al., 2018) and skewness preferences (Olschewski et al., 2024) or aversion to losses (Novemsky & Kahneman, 2005).

Idiosyncratic biases are not specific to theoretical lotteries but can be encountered in any choice task entailing a transparently defined payoff structure. In perceptual choice tasks, participants observe sensory evidence and are asked to make choices based on a criterion defined by the experimenter (e.g., choose the rectangle with the larger area). In these tasks, humans and other animals exhibit several suboptimal tendencies that prevent them from attaining optimal performance. For instance, across consecutive decisions, they exhibit choice history biases (e.g., repeating or avoiding a previous choice) (Braun et al., 2018; Urai et al., 2019); or, within a decision, they assign larger weight to information presented early in the trial (Tsetsos, Gao, et al., 2012). Similarly, in value-learning (bandit) tasks the objective of maximizing monetary reward is undermined by a list of suboptimal tendencies such as ignoring infrequent rewards or penalties (Hertwig et al., 2004), or overestimating the importance of extreme outcomes (Ludvig et al., 2014).

The presence of this type of choice biases in tasks with clearly defined objectives underlines that "desirability" is an inescapably subjective notion. That is, even when the task dictates maximizing economic value, people will not limit themselves to just that. Instead, in addition to accruing economic value, people seem to be gaining extra utility by avoiding risky prospects and losses, by repeating their previous choices, or by avoiding committing errors in perceptual choices at the expense of dwelling for too long on a given decision (Bogacz et al., 2010). Although these tendencies lead to biased choices and curtail reward accrual, they may satisfy other latent, non-economic metrics. Thus, biased choice behavior can be "rationalized" as maximizing a stable, albeit idiosyncratic, utility function. Because the kind of choice biases we reviewed here do not rule out utility-maximizing behavior under a more liberal definition of utility, we deem them *innocuous*[1].

[1] The point that choice biases reflect idiosyncratic preferences for non-normative aspects is rather theoretical, aiming to emphasise that the presence of certain biases does not falsify an expanded notion

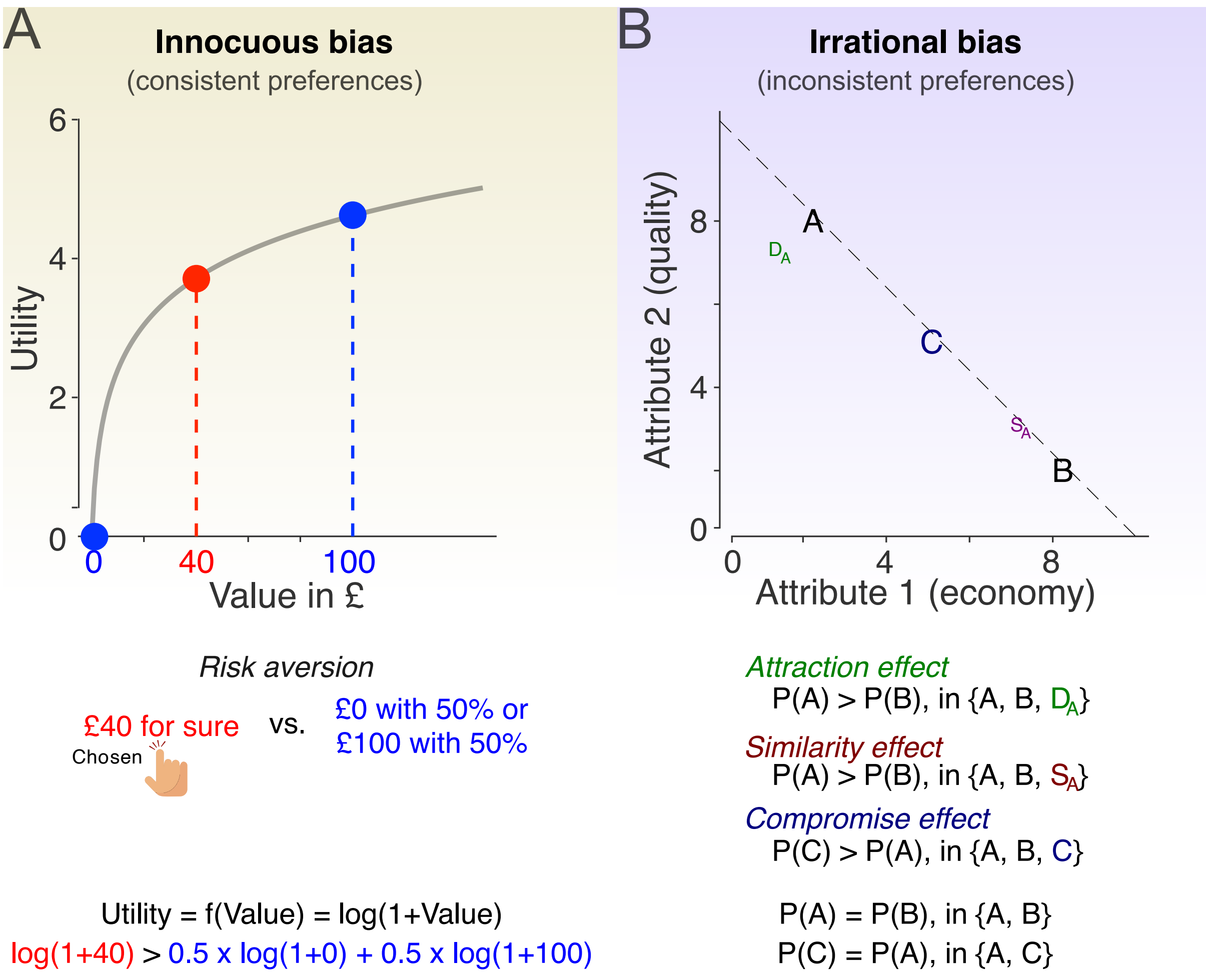

Figure 1. Illustrative examples of innocuous and irrational choice biases. (A) A stable concave (diminishing returns) utility function predicts a risk-aversion bias in a choice between two hypothetical gambles with explicitly described rewards and probabilities. (B) Contextual preference reversal in multiattribute choice. Alternatives vary across two attributes. For illustration purposes, we assume that any two alternatives positioned on the negative diagonal are equally preferred in the respective binary choices. The attraction effect is a choice bias for a target alternative *A* over a competitor *B* occurring in ternary choices after an inferior decoy (D_A) is introduced near the target. The similarity effect is a choice bias for the target *A* after a non-dominated decoy (S_A) is introduced near the competitor *B*. The compromise effect is a choice bias for the all-average alternative *C* over *A* or *B*, occurring in ternary trials featuring another extreme alternative *B*.

2.2 Irrational biases

As described above, prominent choice biases can be absorbed into a putative utility function that the agent's choices could be maximizing. Is the premise of utility-maximization even falsifiable? In their influential work on rational choice theory, von Neumann and Morgenstern demonstrated that the premise of utility-maximization is falsifiable: it can hold only if preferences (as revealed by overt choices) satisfy certain rationality axioms (Savage, 1972; von Neumann & Morgenstern, 2007). A stable utility function can be defined only when preferences satisfy these axioms. Accordingly, if preferences violate any of these axioms, then the utility-maximizing narrative falls apart. In the below we adopt a bird's-eye-view and describe the core intuition underlying the axioms of rational choice theory. More detailed and

---

of optimal (utility-maximizing) behavior. Practically, certain choice biases likely stem from processing bottlenecks in biological brains rather than from explicit preferences for non-normative aspects.

formal expositions of these axioms can be found elsewhere (Regenwetter & Davis-Stober, 2012; Rieskamp et al., 2006).

Rational choice theory axioms are not concerned with the specific preferences of agents but with the internal consistency (or rationality) of those preferences (Allingham, 2002). Thus, they do not prescribe how choice alternatives should be mapped onto utilities; instead, they just ensure that this mapping does not change due to irrelevant factors. One such irrelevant factor is the choice framing (Kahneman & Tversky, 1984). Do preferences change when a choice is framed as “select the best” versus when it is framed as “reject the worst”? According to rational choice theory, in both frames, one should consider their needs and wants and assign utility scores to alternatives. In the “selection” frame the alternative with the highest score should be chosen, and in the logically equivalent “rejection” frame, the alternative with the lowest utility score should be eliminated. Thus, when two alternatives are available, selecting *A* coheres with rejecting *B*, and vice versa. However, it has been shown that when alternative *A* is mediocre (e.g., a not-so-expensive and dull holiday destination) and *B* is more extreme (e.g., an expensive but exciting destination) people tend to *both* select and reject the extreme alternative (Shafir, 1993; Tsetsos, Chater, et al., 2012). This behavioral pattern discloses inconsistent preferences and cannot be reconciled under the maximization of a stable utility function that rational choice theory anticipates.

A second factor that provokes inconsistent choice patterns is the composition of the choice-set. Following the rational choice schema, the utility assigned to an alternative should solely be a function of its inherent properties and the goals and needs of the decision-maker. The utility of an alternative should thus be independent of the properties of other alternatives that are available for choice (i.e., the *independence-from-irrelevant-alternatives* axiom). A logical consequence of this schema is that if *A* is preferred over *B* when only these two alternatives are offered, then *A* should still be preferred over *B* when a third alternative *C* is available for choice. However, it has been shown time and again—even in non-primate species including amoebae and bees (Latty & Beekman, 2011; Tan et al., 2015)—that preferences change as a function of the choice-set composition (Evangelidis et al., 2024). To illustrate, a preference for an all-inclusive holiday to Berlin (*A*) over an all-inclusive holiday to Rome (*B*) can reverse when a holiday to Rome where you must pay for breakfast (*C*) is introduced in the choice-set. In this example, the presence of the inferior alternative (*C*) boosts the desirability of its more similar alternative (*B*). This so-called *attraction effect* (Huber et al., 1982) (Fig. 1B) and similar phenomena where preferences change as alternatives are added or removed from the choice-set, are collectively referred to as *contextual preference reversals* (Tsetsos et al., 2010). Unlike innocuous biases, the framing and choice-set preference reversals outlined here are puzzling as they cannot be “rationalized” under a utility-maximizing narrative. Therefore, we deem these choice biases *irrational*. An open question, which we will explore in the next sections, is whether innocuous and irrational biases have distinct or common mechanistic underpinnings.

2.3 Limitations in existing accounts of choice biases

Within the judgment and decision-making literature, mainstay accounts of biased behavior consist of ad-hoc formulations that effectively re-describe human behavior without providing deeper explanations. For example, the “take-the-best” heuristic posits that decisions are settled exclusively based on the most important cue or attribute (Gigerenzer & Goldstein, 1996). Accordingly, it is assumed that people actively *use* this heuristic rule when making multiattribute decisions. In this case, the explanation (take-the-best heuristic) and the explanandum (choice bias in favor of the alternative that is better in the most important attribute) are almost indiscernible. This issue is not exclusive to the heuristic and biases framework: in experience-based decisions (Hertwig & Erev, 2009) certain behavioral regularities (e.g., recency or ignoring rare events) are ascribed to homologue processing biases (i.e., underweighting early and rare events); or in algebraic models like in Prospect Theory (Kahneman & Tversky, 1979), the shape of psychoeconomic functions and the position of the reference point fit the patterns of human behavior but have no *a priori* motivation. Thus, extant influential accounts of choice biases stay too close to the observed behavioral effects, offering little explanatory depth.

Here, we do not claim that ad-hoc accounts should be dispensed with as they provide a useful and abstract way to summarize how human behavior deviates from the normative expectations. It is even conceivable that certain ad-hoc formulations, like explicit loss-aversion, reflect hardwired asymmetries in the way the brain processes information (Tom et al., 2007). Instead, we posit that this level of theorizing cannot readily provide an encompassing answer to the question: how do choice biases occur in the deciding brain? In the next section we sketch an alternative and deeper level of theorizing choice biases.

3. Mechanistic explanations of choice biases

Here we explore the idea that choice biases *emerge* from dynamic decision computations rather than stemming from a rigid set of ad-hoc rules. This approach can have multiple advantages. First, it can afford precise quantitative predictions, which can be valuable in predicting novel biases or in anticipating how people would respond in different contexts. Second, it can offer a natural interface between behavioral and neural levels (Gold & Shadlen, 2007), enabling the understanding of altered decision-making during ageing or in neuropsychiatric disorders. Third, with certain decision computations serving adaptive functions in biological brains (Summerfield & Tsetsos, 2020), linking choice biases to these computations can help reconcile the normative-descriptive gap. Finally, a mechanistic framework can, in principle, offer unifying and parsimonious explanations. Multiple choice biases can arise from variations in a single or a minimal set of mechanisms, thereby reducing the dimensionality of the tangled ontology of behavioral biases (Hallsworth, 2023).

However, to a large extent, mechanistic inferences are under-constrained by empirical data (Pirrone & Tsetsos, 2023), especially in relatively complex behavioral domains (e.g., multiattribute choices where choice biases abound). As a result, mechanistic models can become overparametrized and arbitrary (Anderson, 2013), often ending up as ad-hoc as heuristic formulations. Indeed, several influential multiattribute models can be criticized for being overly flexible, combining algebraic ad-hoc and dynamic (and biologically grounded) mechanisms to explain preference reversals (Roe et al., 2001; Trueblood et al., 2014; Usher & McClelland, 2004). Even though we acknowledge the merit and influence of these more complex models (for a comprehensive review see Busemeyer et al., 2019), we here consider choice biases within the minimal evidence-accumulation framework developed for simple perceptual decisions. We then expand this framework with a small set of additional mechanisms intended to help agents navigate the rich information involved in multiattribute, multialternative decisions.

3.1 Information processing during decision-making

In a very generic description of the decision process, choices arise from first assigning utilities to available alternatives (valuation) and then selecting the alternative that has the highest utility (comparison) (Platt & Plassmann, 2014; Vlaev et al., 2011). In most choice theories in psychology and economics, valuation and comparison are formulated in a stylized fashion using algebraic operations and functions (such as weighted summation, Keeney & Raiffa, 1993; or the *softmax* function). In contrast to these static and stylized formulations, representations and computations in biological brains are shaped by noisy and dynamical processes (Miller et al., 2024). What are the dynamical processes underpinning decision-making?

Akin to biologists using the Drosophila as their model organism for studying more complex organisms, psychologists and neuroscientists have used simple sensorimotor decisions to approximate generalizable decision processes (Shadlen & Kiani, 2013). Behavioral and neural data (Platt & Glimcher, 1999; Ratcliff & McKoon, 2008; Yang & Shadlen, 2007) indicate that during simple decisions—such as categorizing an ambiguous image as a face or a house or determining the dominant direction of motion in a random dot kinematogram (Heekeren et al., 2008)—noisy information is sampled and accumulated over time into a growing confidence signal, until a criterial degree of confidence (or *boundary*) is reached. Following the law of large numbers, accumulation over time alleviates the detrimental influence of noise and improves decision accuracy. The boundary on confidence controls how long accumulation lasts for, effectively determining how the observer trades off the speed and accuracy of the decision (Bogacz et al., 2010).

Within this *accumulation-to-bound* framework, valuation and comparison are dynamic and temporally multiplexed processes. Valuation arises from *representing*[2], *sampling*[3] and *accumulating* noisy evidence (held in different *accumulators*) in favor of the available choice alternatives (Fig. 2A). Choices are made once the activity of one of the accumulators exceeds the decision boundary, thereby implementing the comparison operation (Lo & Wang, 2006). The specifics implementation of the comparison process can vary across models, particularly in the form and degree of competition among the accumulators (Teodorescu & Usher, 2013). Having outlined the fundamental mechanisms implicated in simple decisions, we next ask how these mechanisms can lead to choice biases.

[2] In any decision-task, external information needs to be mapped onto decision-relevant information. For example, if the task is to determine the rectangle that has the largest width, then information about the height or color of the rectangles is not relevant and needs to be discarded. Neural representations are subject to non-linear transduction, typically captured by a concave (e.g., logarithmic) transformation of the objective information (see Weber-Fechner law).

[3] Decision tasks can entail either dynamic or static information. Dynamic information is naturally chunked into monetary samples, which are subject to serial accumulation. In static tasks, such as a face/house discrimination based on an ambiguous photograph, the assumption is that internal sampling turns takes noisy snapshots of the external information, which are then serially accumulated over time.

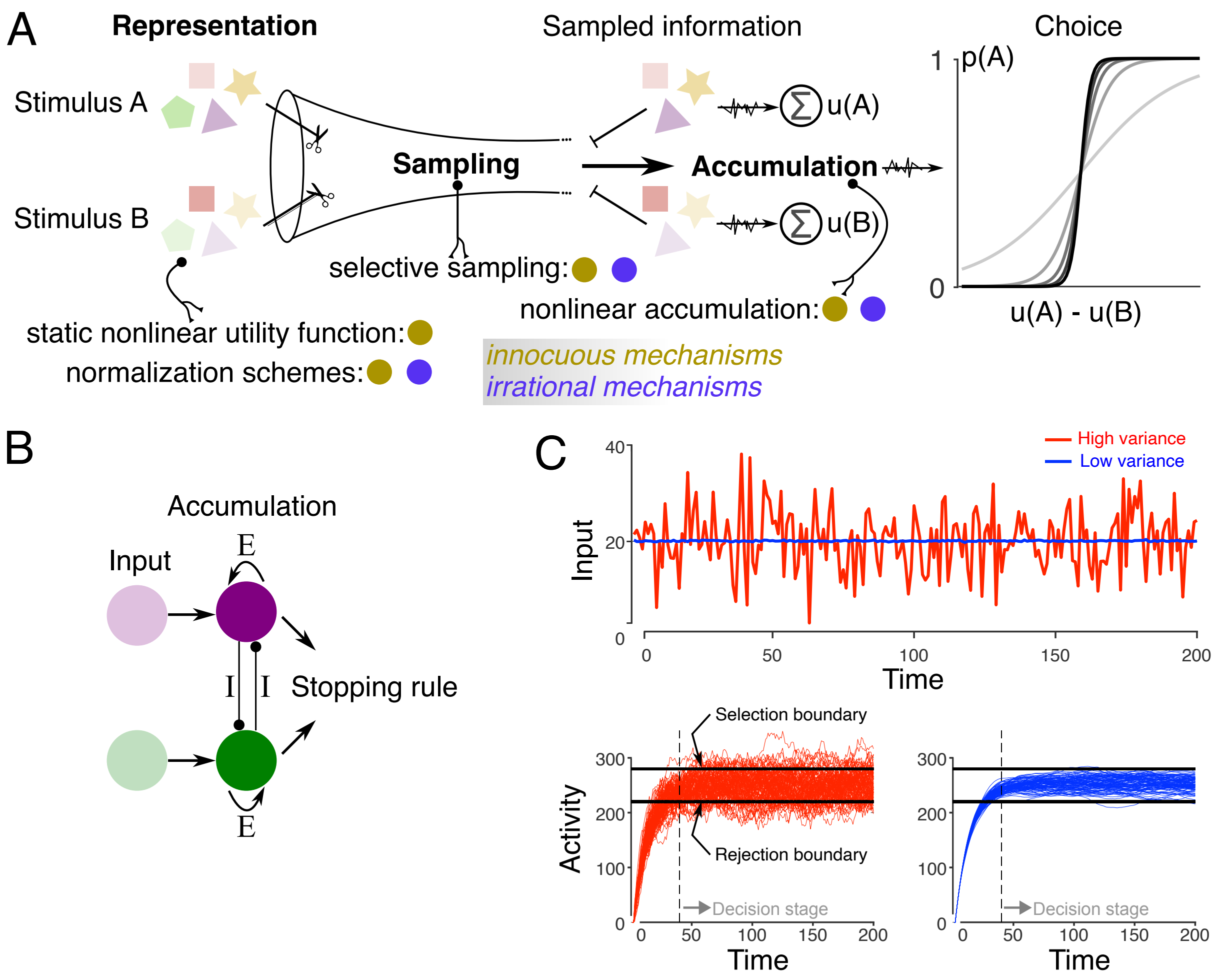

Figure 2. Choice biases within a generalized mechanistic framework. (A) Schematic of the processing hierarchy in decision-making: from stimulus representation to sampling, evidence accumulation, and ultimately, choice. (B) The leaky competing accumulator (LCA) model architecture. “I” denotes lateral-inhibitory connections and “E” self-excitatory connections. With self-excitation being < 1, information is subject to dissipation (leakage). The balance between leakage and lateral inhibition strength controls the profile of temporal weighting (see main text). (C) Risk preferences within a bounded accumulation framework. In the example scenario a high-variance (red) and a low-variance (blue) alternative compete for choice. Following an initial pre-decisional period (left relative to the dashed vertical line), evidence is accumulated in two accumulators that are weakly coupled with inhibition. Under the selection (rejection) framing, once an accumulator breaches the upper (lower) boundary a selection choice is made in its favor. The high-variance accumulator has large deflections and thus exceeds both boundaries more often than the low-variance accumulator. This predicts risk-seeking under selection and risk-aversion under rejection.

3.2 Choice biases within the standard mechanistic framework

Can the standard mechanistic framework outlined above produce innocuous and irrational choice biases? At first glance, this seems like a tall order for accumulation-to-bound models because they essentially stem from the framework of optimal sequential hypothesis testing (Bogacz et al., 2006; Wald, 2004; Wald & Wolfowitz, 1948). However, as described above, accumulation-to-bound models can vary in their implementational details. Due to such variations, some accumulation-to-bound models can deviate from statistically optimal computations.

One characteristic example of suboptimal computations is the non-uniform temporal weighting of evidence emerging from the accumulation dynamics in competing accumulator models. In the leaky competing accumulator model (Usher & McClelland, 2001), the activity of a given accumulator increases with the corresponding incoming evidence and decreases due to self-dissipating activity (or leakage) and lateral inhibition coming from competing accumulators (Fig. 2B). When the leakage and inhibition are balanced, the model mimics the diffusion model showing equal sensitivity to early and late evidence (Bogacz et al., 2007). When the inhibition is stronger than the leakage, the model operates in an impulsive regime, exhibiting larger sensitivity to early evidence (*primacy*) through strong winner-take-all dynamics. Conversely, when the inhibition is weaker than the leakage, the model becomes forgetful or "leaky", being more sensitive to late evidence (*recency*) (Tsetsos, Gao, et al., 2012). Similar temporal weighting profiles fall out from variations in the excitation/inhibition ratio in a biophysical cortical circuit model of evidence accumulation (Lam et al., 2022).

Non-uniform temporal weighting can lead to innocuous choice biases when decision-relevant information is processed in a fixed order. Consider the choice between an affordable but dull holiday destination (*A*) and an expensive and exciting destination (*B*). If the price information is conveyed first, agents with a primacy weighting profile will consistently choose *A*. Interestingly, non-uniform temporal weighting can lead to irrational choice biases if the framing of the task or the choice-set composition alters the order in which information is processed. For example, an agent with a primacy profile will disclose a preference reversal if they first process positive information (*B* is exciting) in the "select the best" framing and negative information (*B* is expensive) in the "reject the worst" framing. Thus, deviations from statistically optimal computations can open the door to both innocuous and irrational choice biases, depending on certain assumptions about the order in which information is considered. Irrespective of assumptions about the order of information processing, how can the standard mechanistic framework produce well-established choice biases?

We consider the innocuous preferences that humans have towards less or more variable alternatives. Empirical findings suggest that people are risk-averse in description-based lotteries but risk (variance)-seeking when value information is experienced sequentially (Tsetsos, Chater, et al., 2012). For presentation purposes, we describe both choice biases using a hypothetical choice between two holiday destinations. A preference for the mediocre holiday destination *A* over the more extreme one *B* (risk-aversion) can be explained by a concave transduction function that maps objective values onto internal subjective counterparts (see also footnote 2 and Fig. 1A). This is the classical explanation of risk-aversion adopted in expected-utility theory and prospect theory (Kahneman & Tversky, 1979). Beyond this rather rigid representational distortion, in Figure 2C we show how the opposite risk-seeking bias (Tsetsos, Chater, et al., 2012) naturally emerges within a minimal accumulation-to-bound framework involving two independent (or weakly competing) accumulators racing towards a boundary. Due to its more variable input, the accumulator corresponding to the more extreme alternative *B* shows larger deflections. These deflections translate into an increased likelihood of crossing the decision boundary. Similarly, under a rejection frame—and assuming that elimination happens when an accumulator breaches a lower boundary—*B* will be more frequently eliminated resulting in an irrational (framing) reversal of risk preferences (Shafir, 1993).

A preference for the more variable alternative (*B*) can also arise within a biophysical cortical model of evidence accumulation due to convex evidence transduction and non-linear accumulation dynamics (Cavanagh et al., 2020). Although it is not obvious how the biophysical model could produce a framing reversal of this bias, this model can produce a choice-set reversal. In choices among three alternatives that vary in their decision values, increasing the value of the worst alternative improves the relative discrimination accuracy between the two highest-value alternatives (Chau et al., 2014) (but see Cao and Tsetsos (2022) for an alternative interpretation of this effect). This *positive distractor* effect violates the independence-from-irrelevant-alternatives principle. The biophysical model captures the positive distractor effect because increasing the value of the worst alternative raises the level of the pooled inhibition. This, in turn, adjusts the accumulation dynamics to a regime closer to optimal, approximating the diffusion model that achieves the highest discrimination accuracy (Bogacz et al., 2007).

3.3 Choice biases within an extended mechanistic framework

We showed that some innocuous and irrational choice biases can emerge within the standard accumulation-to-bound framework. However, this framework cannot readily explain multiattribute choice-set reversals (Fig. 1B). Notably, the standard mechanistic framework was built around simple choice tasks, where the amount of decision-relevant information typically falls below the processing capacity of the cognitive system (Donner et al., 2009; Gold & Shadlen, 2007). However, choice biases, especially irrational ones, occur in more complex decision domains involving multiple (more than two) alternatives that often vary in more than one attribute (Busemeyer et al., 2019). In these cases, parsing information in parallel becomes impossible and the cognitive system needs to find ways to efficiently navigate the increased complexity given its processing bottlenecks. Below, we discuss two classes of mechanisms that can help the brain efficiently process large amounts of information. As a byproduct, these mechanisms enable dynamic and context-dependent valuation, and choice biases ensue.

Relative coding

The first class of mechanisms, collectively referred to as *relative coding* (Summerfield & Tsetsos, 2020), entails dynamic and non-linear distortions impacting the representation of decision-relevant information. More specifically, in contrast to static and context-independent transduction non-linearities (see footnote 2), relative coding schemes adjust representations as a function of the temporal (i.e., the recent history of stimulation) or the immediate (i.e., the choice-set) context. The exact form of these adjustments is motivated by the theory of efficient neural coding widely evidenced in sensory systems (Simoncelli, 2003). This theory states that neurons minimize redundancy by increasing their representational resolution for the most frequently occurring stimuli (Barlow, 1961). Below we provide an overview of prominent relative coding schemes with regards to choice biases.

In the divisive normalization model, the “raw” utility of each alternative is divided by the sum of the raw utilities of all alternatives in the choice set (Louie et al., 2013). Therefore, the model predicts that increasing the utility of the worst alternative reduces the discrimination accuracy between the two high-utility alternatives, a *negative distractor* effect at odds with the independence-from-irrelevant-alternatives axiom (but see Gluth et al. (2020) for a failure to replicate this effect). In the range normalization model (Rustichini et al., 2017), raw utilities are divided by the range of raw utilities $(\max - \min)$ encountered in the choice-set, that way producing a positive distractor effect (Fig. 3). The range normalization principle together with auxiliary assumptions can also explain contextual preference reversals in multiattribute choice, including the attraction, the similarity, and the compromise effects (Soltani et al., 2012) (Fig. 1B). These multiattribute effects can also be captured by a recurrent version of the divisive normalization model, where the attribute value of an alternative is divided by itself plus the mean choice-set values on that attribute (Dumbalska et al., 2020). Practically, relative coding schemes implement context-dependent utility functions thereby generating irrational choice biases. One downside of relative coding schemes is that they afford compressive representations. Thus, these models cannot readily capture the risk-seeking bias obtained in decisions from experience (Tsetsos, Chater, et al., 2012).

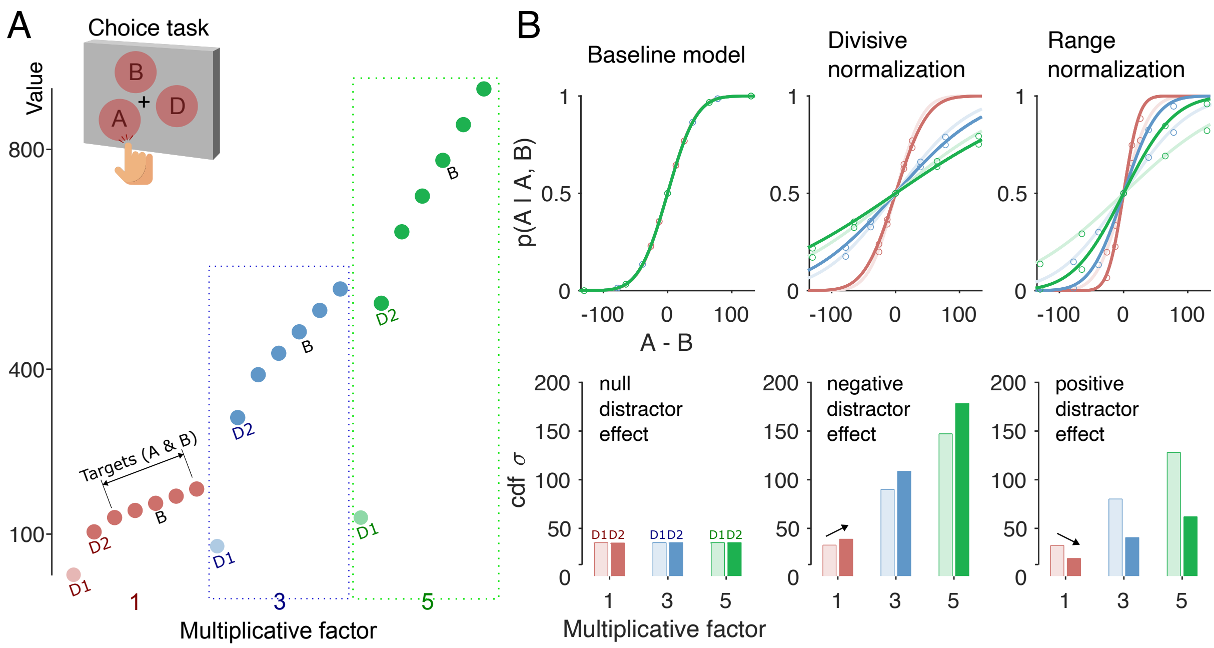

Figure 3. Distractor effects under relative coding schemes. (A) In a ternary choice task, observers choose one alternative among three candidates (*A*, *B*, and a distractor *D*) on each trial. The distractor's value is always lower than the values of targets *A* and *B*, so it should not influence the choice between *A* and *B*, as predicted by the baseline model in panel B. In a thought experiment, observers perform this task in two additional contexts where the option values are multiplicatively scaled up by 3 or 5. This manipulation creates an ideal situation for testing prominent normalization theories. The distractor effect is defined as the change in sensitivity (slope measure of a Gaussian cumulative-density-function fit to choice probabilities) to the value difference between *A* and *B* (fixed), modulated by changes in the distractor value. (B) Divisive normalization assumes the option value is transformed into the mean firing rate $\mu_i = KV_i/(\sigma_h + \Sigma wV_j)$, where V_i is the raw value of the option under consideration, $K > 0$, $\sigma_h > 0$, and $w > 0$ represent gain, semi-saturation, and weight, respectively (Louie et al., 2013). When $w = 0$, the model reduces to the baseline version, meaning the value coding is independent of other options in the choice-set. Range normalization assumes $\mu_i = KV_i/(\sigma_h + w(\max(V_j) - \min(V_j)))$, meaning the denominator involves the range of the values rather than the summation. Both normalization models predict an increase in the distractor effect as the multiplicative factor increases, but they predict the distractor effect in opposite directions. These very specific predictions can be contrasted with the predictions of other non-normalization models of distractor effects (e.g., decision-by-sampling). Code for reproducing this figure can be found at: https://github.com/YinanCao/bookchapter

Selective sampling

The second class of mechanisms we review operate downstream the early distortions that relative coding schemes induce. These mechanisms govern selective information sampling by determining, at each moment, which aspects of the available information should be passed on for accumulation. Selective information sampling is a pragmatic solution to the challenge posed by the rich information involved in multialternative and multiattribute decisions. Indeed, it is commonly observed that during complex decisions, attention serially traverses across attributes and alternatives (Fiedler & Glöckner, 2012; Russo & Dosher, 1983), driving the online construction of preferences (Slovic, 1995). Recent work has incorporated attentional fluctuations into the accumulation-to-bound framework by positing that the gain of processing increases for attended items (Krajbich et al., 2010). However, the principles that orchestrate these attentional fluctuations, and the reasons why sampling can end up being partial and biased, remain unknown. Below we review proposals that infer principles of information sampling through explaining contextual preference reversals.

Various models of multiattribute choice assume that only one attribute can be processed at a time, with attention stochastically fluctuating across attributes over time (Roe et al., 2001; Turner et al., 2018; Tversky, 1972). More recent models additionally assume that within each attended attribute, only a subset of the available alternatives is considered at a given instance (Wollschläger & Diederich, 2012). In the decision-by-sampling multiattribute model (Noguchi & Stewart, 2018), two alternatives are compared within each sampled attribute, with more similar alternatives forming comparison pairs more frequently. The binary outcome of the comparison updates the "counting" accumulator of the winner. Due to these principles and some auxiliary assumptions, the model can explain the attraction, similarity, and compromise effects (Fig. 1B). A recent adaptation of the decision-by-sampling framework can also explain a positive distractor effect in single-attribute decisions (Tohidi-Moghaddam & Tsetsos, 2024). Another sampling model, the selective integration model, assumes that within a focused attribute, attentional selection prioritizes the processing (i.e., assigns a larger gain) of high-valued alternatives at the expense of low-valued alternatives, a principle that leads to the attraction and other contextual preference reversal effects (Tsetsos, 2012; Tsetsos, Chater, et al., 2012) including violations of transitivity (Tsetsos et al., 2016). The selective integration model also predicts risk-seeking that reverses under a rejection frame where observers prioritize the processing of lower values (Usher et al., 2019). Taken together, in the models described above, selective sampling leads to innocuous and irrational choice biases by prioritizing the processing of certain choice aspects at the expense of others.

4. Identifying the mechanisms underpinning choice biases

The previous section highlights that, even within a limited search space of mechanisms, there are multiple and equally viable[4] explanations for choice biases situated at all stages of the processing pathway (Table 1). This is a stark reminder that the cognitive and neural mechanisms of more complex behaviors are grossly underdetermined by empirical data (Pirrone & Tsetsos, 2023). How can the mechanisms underlying the various choice biases be securely identified?

[4] Here, arbitrating competing mechanisms based on auxiliary criteria, like biological plausibility or normative justification, seems fruitless. For instance, given efficient codes in the brain, two competing schemes, divisive and range normalization, seem equally biologically plausible. Similarly, while relative coding schemes can maximize information transfer while keeping metabolic costs bounded, the selective integration can maximize a different metric, namely decision accuracy in the presence of late noise (Tsetsos, K., Moran, R., Moreland, J., Chater, N., Usher, M., & Summerfield, C. (2016). Economic irrationality is optimal during noisy decision making. *Proc Natl Acad Sci U S A*, *113*(11), 3102-3107. https://doi.org/10.1073/pnas.1519157113).

Table 1. Summary of the standalone predictions of various mechanisms in relation to innocuous and irrational choice biases. “R” stands for representation, “S” for sampling, and “A” for accumulation. “?” is inserted when certain effects have not been adequately explored in relation to certain choice biases.

Mechanism		**Choice biases**					
		Risk-seeking (under selection)	Risk-aversion (under selection)	Framing risk-reversal	Negative distractor	Positive distractor	Multiattribute preference reversals
Stable non-linear function (R)	Convex	Yes	No	No	No	No	No
	Concave	No	Yes	No	No	No	No
Relative value coding (R)	Divisive normalization	No	Yes	?	Yes	?	No
	Range normalization	No	Yes	?	No	Yes	Yes
	Recurrent divisive normalization	No	Yes	?	?	?	Yes
Selective sampling (S)	Selective integration	Yes	No	Yes	?	?	Yes
	Decision-by-sampling	?	?	?	No	Yes	Yes
Non-linear accumulation (A)	Bounded accumulator	Yes	No	Yes	No	No	No
	Biophysical cortical circuit	Yes	No	?	No	Yes	No

Identifying the mechanisms that underlie a cognitive process requires the ability to experimentally manipulate or observe the process input while simultaneously observing the output at the behavioral and neural levels. As we alluded to in previous sections, this approach is exemplified in the study of sensorimotor decisions, where the use of well-controlled psychophysical stimuli have uncovered the neural underpinnings of perceptual decision-making (Gold & Shadlen, 2007). However, in multiattribute choice problems—which entail information across different attributes and require comparisons between multiple alternatives—the input to the decision process is not under experimental control. Instead, due to the rich information exceeding the parallel processing capacities of the cognitive system, attention can voluntarily fluctuate over time in numerous different ways (i.e., examining one alternative on all attributes at a given time, dividing attention between two alternatives, focusing on all alternatives in one attribute etc.). Observing these fluctuations is necessary to uncover the flow of the input that drives the decision-making process. In turn, knowing the dynamical input of the decision process can radically constrain mechanistic inferences and answer how choice biases come about. Simply put, understanding *how* the brain processes information requires knowing *what* information it processes.

Mainstay approaches have attempted to empirically characterize information sampling using eye-tracking techniques. However, interpreting eye movements and fixations is not straightforward in the context of decision-making tasks. Eye movements can also correspond to operations unrelated to the decision-making process (i.e., merely reading the information), which cannot be dissociated from decision-related oculomotor activity (i.e., accumulating the fixated information). Instead, it is often asserted that only fixations in the middle part of the deliberation are decision-related, while early fixations correspond to the “scanning” stage and late fixations to the “validation” stage (Bettman & Kakkar, 1977). Importantly, even if decision-related fixations could be dissociated from unrelated operations, knowing where people fixate does not say much about what they think. Does looking at the price of holiday destination *A* guarantees that the price of holiday *B* is not covertly considered at the same time in a comparative manner?

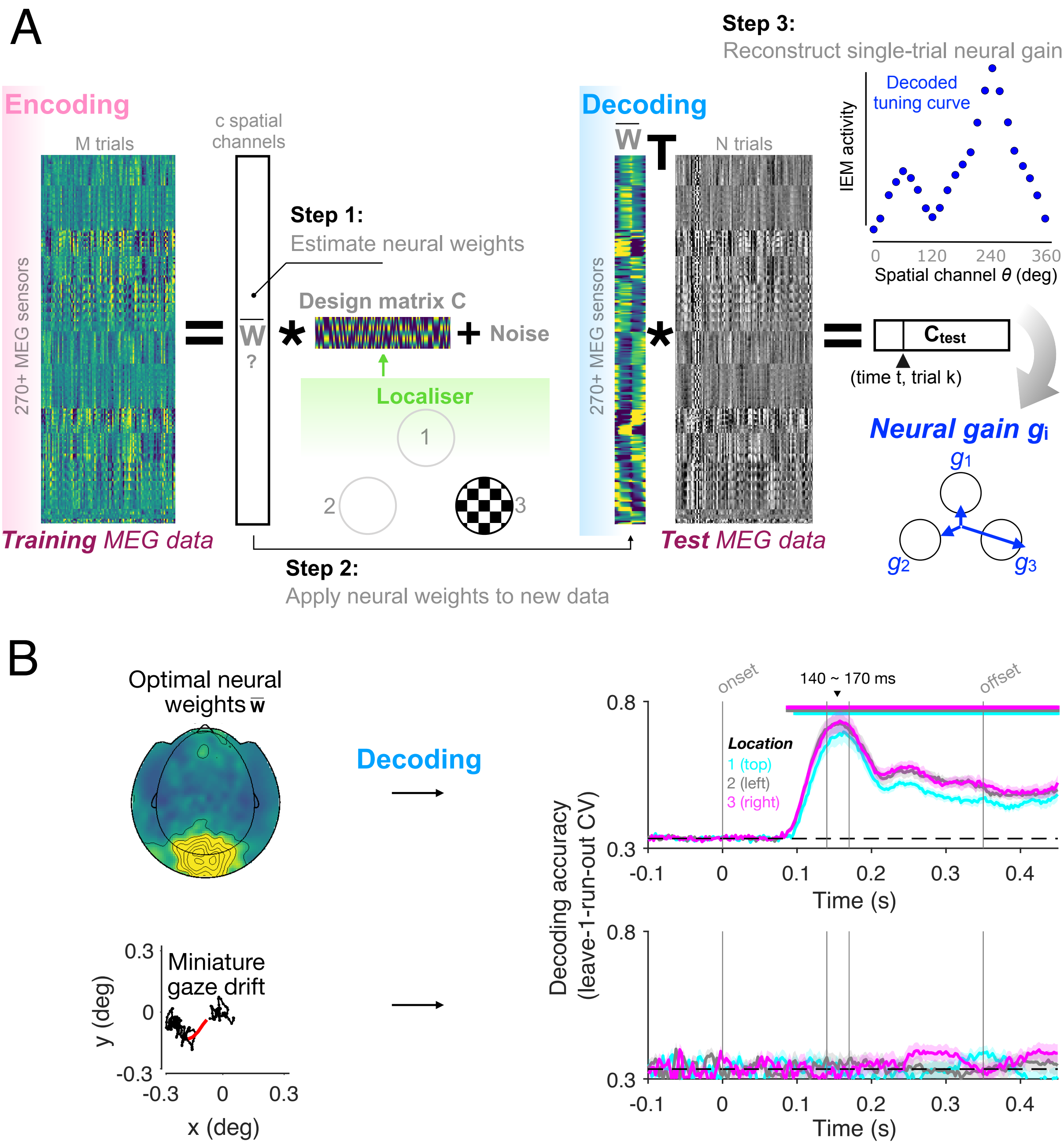

Figure 4. Decoding the locus of spatial attention using M/EEG. (A) In the localizer task, a checkerboard patch appears at one location on each trial while participants fixate centrally and press a button as soon as the fixation cue changes color (a rare event: ~10% of trials). The encoding-decoding model (IEM) involves three main steps. First, optimal neural weights are estimated using a linear model that maps multi-channel M/EEG activities onto hypothetical spatial channels (rectified sinusoids). The M/EEG

amplitude measured at each sensor during the localizer task (“Training data”) is modelled with spatial channels, each selectively tuned to a different angular position. Next, the encoding model is inverted to estimate the channel responses from the pattern of M/EEG signals across the scalp in another task (“Test data”), such as multi-alternative decision-making or a classic bandit task in reinforcement learning. Finally, neural gains projected spatially across the stimulus space are reconstructed for each single trial. (B) Decoding performance revealed by training IEM on neural data vs. concurrent eye-gaze data in the spatial localizer task. Upper panel: Topography shown as PCA loadings that maximally differentiate between the three spatial stimulus locations, indicating where the neural activation patterns across the scalp best diversify across these locations. Cross-validation shows that the IEM trained on neural signals exhibits excellent and highly reliable decoding performance, peaking around 140 to 170 ms after stimulus onset, and is consistent across individual locations. In stark contrast, because the localizer task prohibits eye movement, IEM trained on eye positions contains no decodable information about stimulus locations (lower panel). This demonstrates that the method reliably captures covert neural processes and neural gain modulation in the visual-parietal regions.

Thus, eye-tracking techniques are limited to overt attention, while the computational role of the tracked measures is unclear. What has been missing is a way to track both the locus of attention (overt and covert) and at the same time the state of the accumulators as complex decisions take shape. Recent work from our lab (Siems et al., 2023) and others (Mostert et al., 2018) has offered new possibilities for continuously tracking the locus of covert spatial attention using non-invasive neural recording techniques that have high temporal resolution (magneto/electro-encephalography (M/EEG)). This relies on a dedicated functional localizer task and an encoding model that is inverted to estimate the locus of attentional allocation on a *single trial* basis from the pattern of M/EEG signals across the scalp in another task, such as multialternative decision-making or a classic bandit task in reinforcement learning (Fig. 4A). The key technical advantage of this approach is the clear dissociation of covert processes from oculomotor “contaminations”, with the spatial location decoding of attention being robust across multiple locations and unaffected by eye gaze shifts, no matter how minuscule they are (Fig. 4B). Combining the tracking of covert attention with well-established M/EEG signals that track the state of decision accumulators (e.g., the beta-band lateralization in parietal and pre-motor cortices) (Donner et al., 2009; O'connell et al., 2012) can reveal regularities and biases in information sampling (Siems et al., 2023), decisively constraining mechanistic inferences about the general decisions processes as well as those that generate innocuous and irrational choice biases.

5. Conclusions

Research in behavioral economics and psychology has identified several choice biases that still lack conclusive mechanistic explanations. In this chapter, we explored how choice biases can be mapped onto the neural and computational mechanisms underlying decision-making. Using as our starting point the standard accumulation-to-bound framework developed for simple decisions, we described how choice biases can emerge from decision computations throughout the processing pathway: from representing, to sampling, to accumulating decision-relevant information. While non-linear accumulation dynamics within the standard accumulation-to-bound framework can explain some choice biases, a more complete explanation of irrational preference reversals requires invoking relative coding at the representation level or selective information sampling. Given that relative coding and selective sampling are both descriptively adequate, biologically plausible, and normatively motivated, distinguishing between these two types of mechanisms seems impossible. This resonates with the broader issue that identifying the mechanisms underlying complex decision-making is underdetermined, given that the effective input to the decision process is opaque and intractable with conventional process tracing techniques. We outlined how using non-invasive time-resolved neural recordings can track attentional fluctuations during decision-making, that way measuring the effective decision input feeding to downstream decision computations.

We argue that observing how information is being sampled during complex decisions can unlock the mechanistic understanding of puzzling behavioral regularities. As we described in this chapter, non-linear accumulation dynamics can lead to non-uniform temporal weighting of information (Tsetsos, Gao, et al., 2012). Across these lines, more recent findings suggest that evidence accumulation is not simply a feed-forward process, with the state of downstream accumulators biasing the way incoming evidence is weighted (Talluri et al., 2018). Thus, with information being unequally weighted over time, the order in which information is being parsed can be the major determinant of choice. If choice-set or framing manipulations do not systematically alter the order of information sampling, then relative coding computations would appear necessary to explain irrational choice biases. However, if framing and choice-set manipulations systematically alter the order of information processing, then puzzling choice biases could simply result from a combination of non-linear accumulation dynamics and specific patterns of information sampling (presumably consistent with extant sampling proposals, such as selective integration and decision-by-sampling). In this latter case, better understanding the principles and mechanisms that orchestrate information sampling during decision-making would become a critical new goal in the cognitive and decision sciences.

Acknowledgments

This work was supported by the EU Horizon 2020 Research and Innovation Program (ERC starting grant no. 802905) to K.T.

References

Aczel, B., Bago, B., Szollosi, A., Foldes, A., & Lukacs, B. (2015). Measuring individual differences in decision biases: Methodological considerations. *Frontiers in psychology*, *6*, 163396.

Allingham, M. (2002). *Choice theory: A very short introduction*. OUP Oxford.

Anderson, J. R. (2013). *The adaptive character of thought*. Psychology Press.

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. *Sensory communication*, *1*(01), 217-233.

Baron, J., & Ritov, I. (1994). Reference Points and Omission Bias. *Organizational Behavior and Human Decision Processes*, *59*(3), 475-498. https://doi.org/DOI 10.1006/obhd.1994.1070

Bettman, J. R., & Kakkar, P. (1977). Effects of information presentation format on consumer information acquisition strategies. *Journal of Consumer Research*, *3*(4), 233-240.

Bhui, R., & Xiang, Y. (2021). A rational account of the repulsion effect.

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. *Psychol Rev*, *113*(4), 700-765. https://doi.org/10.1037/0033-295X.113.4.700

Bogacz, R., Hu, P. T., Holmes, P. J., & Cohen, J. D. (2010). Do humans produce the speed–accuracy trade-off that maximizes reward rate? *Quarterly journal of experimental psychology*, *63*(5), 863-891.

Bogacz, R., Usher, M., Zhang, J., & McClelland, J. L. (2007). Extending a biologically inspired model of choice: multi-alternatives, nonlinearity and value-based multidimensional choice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1485), 1655-1670.

Braun, A., Urai, A. E., & Donner, T. H. (2018). Adaptive history biases result from confidence-weighted accumulation of past choices. *Journal of Neuroscience*, *38*(10), 2418-2429.

Busemeyer, J. R., Gluth, S., Rieskamp, J., & Turner, B. M. (2019). Cognitive and Neural Bases of Multi-Attribute, Multi-Alternative, Value-based Decisions. *Trends Cogn Sci*, *23*(3), 251-263. https://doi.org/10.1016/j.tics.2018.12.003

Cao, Y., & Tsetsos, K. (2022). Clarifying the role of an unavailable distractor in human multiattribute choice. *Elife*, *11*. https://doi.org/10.7554/eLife.83316

Cavanagh, S. E., Lam, N. H., Murray, J. D., Hunt, L. T., & Kennerley, S. W. (2020). A circuit mechanism for decision-making biases and NMDA receptor hypofunction. *Elife*, *9*. https://doi.org/10.7554/eLife.53664

Chau, B. K., Kolling, N., Hunt, L. T., Walton, M. E., & Rushworth, M. F. (2014). A neural mechanism underlying failure of optimal choice with multiple alternatives. *Nat Neurosci*, *17*(3), 463-470. https://doi.org/10.1038/nn.3649

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, *313*(5787), 684-687. https://doi.org/10.1126/science.1128356

Donner, T. H., Siegel, M., Fries, P., & Engel, A. K. (2009). Buildup of choice-predictive activity in human motor cortex during perceptual decision making. *Curr Biol*, *19*(18), 1581-1585. https://doi.org/10.1016/j.cub.2009.07.066

Dumbalska, T., Li, V., Tsetsos, K., & Summerfield, C. (2020). A map of decoy influence in human multialternative choice. *Proc Natl Acad Sci U S A*, *117*(40), 25169-25178. https://doi.org/10.1073/pnas.2005058117

Evangelidis, I., Bhatia, S., Levav, J., & Simonson, I. (2024). 50 Years of Context Effects: Merging the Behavioral and Quantitative Perspectives. *Journal of Consumer Research*, *51*(1), 19-28.

Fiedler, S., & Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in psychology*, *3*, 335.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, *62*, 451-482.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, *103*(4), 650.

Gluth, S., Kern, N., Kortmann, M., & Vitali, C. L. (2020). Value-based attention but not divisive normalization influences decisions with multiple alternatives. *Nature Human Behaviour*, *4*(6), 634-645.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535-574. https://doi.org/10.1146/annurev.neuro.29.051605.113038

Hallsworth, M. (2023). A manifesto for applying behavioural science. *Nat Hum Behav*, *7*(3), 310-322. https://doi.org/10.1038/s41562-023-01555-3

Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience*, *9*(6), 467-479.

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.

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

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), 90-98. https://doi.org/Doi 10.1086/208899

Juechems, K., & Summerfield, C. (2019). Where Does Value Come From? *Trends Cogn Sci*, *23*(10), 836-850. https://doi.org/10.1016/j.tics.2019.07.012

Kahneman, D., & Tversky, A. (1979). Prospect Theory - Analysis of Decision under Risk. *Econometrica*, *47*(2), 263-291. https://doi.org/Doi 10.2307/1914185

Kahneman, D., & Tversky, A. (1984). Choices, Values, and Frames. *American Psychologist*, *39*(4), 341-350. https://doi.org/Doi 10.1037/0003-066x.39.4.341

Keeney, R. L., & Raiffa, H. (1993). *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge university press.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nat Neurosci*, *13*(10), 1292-1298. https://doi.org/10.1038/nn.2635

Lam, N. H., Borduqui, T., Hallak, J., Roque, A., Anticevic, A., Krystal, J. H., Wang, X.-J., & Murray, J. D. (2022). Effects of altered excitation-inhibition balance on decision making in a cortical circuit model. *Journal of Neuroscience*, *42*(6), 1035-1053.

Latty, T., & Beekman, M. (2011). Irrational decision-making in an amoeboid organism: transitivity and context-dependent preferences. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1703), 307-312.

Lo, C.-C., & Wang, X.-J. (2006). Cortico–basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nature neuroscience*, *9*(7), 956-963.

Loomes, G., & Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review*, *39*(3-4), 641-648.
Louie, K., & De Martino, B. (2014). The neurobiology of context-dependent valuation and choice. In *Neuroeconomics* (pp. 455-476). Elsevier.
Louie, K., Khaw, M. W., & Glimcher, P. W. (2013). Normalization is a general neural mechanism for context-dependent decision making. *Proc Natl Acad Sci U S A*, *110*(15), 6139-6144. https://doi.org/10.1073/pnas.1217854110
Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2014). Extreme outcomes sway risky decisions from experience. *Journal of Behavioral Decision Making*, *27*(2), 146-156.
Maier, M., Bartos, F., Stanley, T. D., Shanks, D. R., Harris, A. J. L., & Wagenmakers, E. J. (2022). No evidence for nudging after adjusting for publication bias. *Proc Natl Acad Sci U S A*, *119*(31), e2200300119. https://doi.org/10.1073/pnas.2200300119
Mata, R., Frey, R., Richter, D., Schupp, J., & Hertwig, R. (2018). Risk preference: A view from psychology. *Journal of Economic Perspectives*, *32*(2), 155-172.
McClelland, J. L. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science*, *1*(1), 11-38.
Miller, E. K., Brincat, S. L., & Roy, J. E. (2024). Cognition is an emergent property. *Current Opinion in Behavioral Sciences*, *57*, 101388.
Moran, R., & Tsetsos, K. (2018). The standard Bayesian model is normatively invalid for biological brains. *Behav Brain Sci*, *41*, e237. https://doi.org/10.1017/S0140525X18001449
Mostert, P., Albers, A. M., Brinkman, L., Todorova, L., Kok, P., & de Lange, F. P. (2018). Eye Movement-Related Confounds in Neural Decoding of Visual Working Memory Representations. *eNeuro*, *5*(4). https://doi.org/10.1523/ENEURO.0401-17.2018
Noguchi, T., & Stewart, N. (2018). Multialternative decision by sampling: A model of decision making constrained by process data. *Psychological review*, *125*(4), 512.
Novemsky, N., & Kahneman, D. (2005). The boundaries of loss aversion. *Journal of Marketing Research*, *42*(2), 119-128.
O'connell, R. G., Dockree, P. M., & Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nature neuroscience*, *15*(12), 1729-1735.
Olschewski, S., Spektor, M. S., & Le Mens, G. (2024). Frequent winners explain apparent skewness preferences in experience-based decisions. *Proceedings of the National Academy of Sciences*, *121*(12), e2317751121.
Parrish, A. E., Dawes, J., & Thompson, H. L. (2024). Exploring the Impact of Decoys on Decision-Making by Young Children. *Journal of Behavioral Decision Making*, *37*(3), e2385.
Pirrone, A., & Tsetsos, K. (2023). Toward an Atlas of Canonical Cognitive Mechanisms. *Cogn Sci*, *47*(2), e13243. https://doi.org/10.1111/cogs.13243
Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, *400*(6741), 233-238.
Platt, M. L., & Plassmann, H. (2014). Multistage valuation signals and common neural currencies. *Neuroeconomics*, 237-258.
Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural computation*, *20*(4), 873-922.
Regenwetter, M., & Davis-Stober, C. P. (2012). Behavioral variability of choices versus structural inconsistency of preferences. *Psychological review*, *119*(2), 408.
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.
Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionst model of decision making. *Psychological review*, *108*(2), 370.
Russo, J. E., & Dosher, B. A. (1983). Strategies for multiattribute binary choice. *J Exp Psychol Learn Mem Cogn*, *9*(4), 676-696. https://doi.org/10.1037//0278-7393.9.4.676
Rustichini, A., Conen, K. E., Cai, X., & Padoa-Schioppa, C. (2017). Optimal coding and neuronal adaptation in economic decisions. *Nature communications*, *8*(1), 1208.
Savage, L. J. (1972). *The foundations of statistics*. Courier Corporation.
Shadlen, M. N., & Kiani, R. (2013). Decision making as a window on cognition. *Neuron*, *80*(3), 791-806.
Shafir, E. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory & cognition*, *21*(4), 546-556.
Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-Based Choice. *Cognition*, *49*(1-2), 11-36. https://doi.org/Doi 10.1016/0010-0277(93)90034-S

Siems, M., Cao, Y., Tohidi-Moghaddam, M., Donner, T. H., & Tsetsos, K. (2023). Rhythmic sampling of multiple decision alternatives in the human brain. *bioRxiv*, 2023.2012.2008.570734. https://doi.org/10.1101/2023.12.08.570734

Simoncelli, E. P. (2003). Vision and the statistics of the visual environment. *Current opinion in neurobiology*, *13*(2), 144-149.

Slovic, P. (1995). The Construction of Preference. *American Psychologist*, *50*(5), 364-371. https://doi.org/Doi 10.1037/0003-066x.50.5.364

Soltani, A., De Martino, B., & Camerer, C. (2012). A range-normalization model of context-dependent choice: a new model and evidence. *PLoS computational biology*, *8*(7), e1002607.

Spektor, M. S., Bhatia, S., & Gluth, S. (2021). The elusiveness of context effects in decision making. *Trends in Cognitive Sciences*, *25*(10), 843-854.

Srivastava, N., & Schrater, P. R. (2012). Rational inference of relative preferences. *Advances in neural information processing systems*, *25*.

Stearns, S. C. (2000). Daniel Bernoulli (1738): evolution and economics under risk. *Journal of biosciences*, *25*(3), 221-228.

Summerfield, C., & Tsetsos, K. (2012). Building Bridges between Perceptual and Economic Decision-Making: Neural and Computational Mechanisms. *Front Neurosci*, *6*, 70. https://doi.org/10.3389/fnins.2012.00070

Summerfield, C., & Tsetsos, K. (2015). Do humans make good decisions? *Trends Cogn Sci*, *19*(1), 27-34. https://doi.org/10.1016/j.tics.2014.11.005

Summerfield, C., & Tsetsos, K. (2020). Rationality and efficiency in human decision-making. *The Cognitive Neurosciences*, *6*, 427-439.

Talluri, B. C., Urai, A. E., Tsetsos, K., Usher, M., & Donner, T. H. (2018). Confirmation Bias through Selective Overweighting of Choice-Consistent Evidence. *Curr Biol*, *28*(19), 3128-3135 e3128. https://doi.org/10.1016/j.cub.2018.07.052

Tan, K., Dong, S., Liu, X., Chen, W., Wang, Y., Oldroyd, B. P., & Latty, T. (2015). Phantom alternatives influence food preferences in the eastern honeybee A pis cerana. *Journal of Animal Ecology*, *84*(2), 509-517.

Tentori, K., Osherson, D., Hasher, L., & May, C. (2001). Wisdom and aging: Irrational preferences in college students but not older adults. *Cognition*, *81*(3), B87-B96.

Teodorescu, A. R., & Usher, M. (2013). Disentangling decision models: from independence to competition. *Psychological review*, *120*(1), 1.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Tohidi-Moghaddam, M., & Tsetsos, K. (2024). The timescale and functional form of context-dependence during human value-learning. *bioRxiv*, 2024.2002. 2001.578398.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, *315*(5811), 515-518.

Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychol Rev*, *121*(2), 179-205. https://doi.org/10.1037/a0036137

Tsetsos, K. (2012). *Information integration in perceptual and value-based decisions* (Publication Number U600209) [Ph.D., University of London, University College London (United Kingdom)]. ProQuest Dissertations & Theses Global. England. https://bris.idm.oclc.org/login?url=https://www.proquest.com/dissertations-theses/information-integration-perceptual-value-based/docview/1442492478/se-2?accountid=9730 https://bris.on.worldcat.org/atoztitles/link?sid=ProQ:&issn=&volume=&issue=&title=Information+integration+in+perceptual+and+value-based+decisions&spage=&date=2012-01-01&atitle=Information+integration+in+perceptual+and+value-based+decisions&au=Tsetsos%2C+K.&id=doi:

Tsetsos, K., Chater, N., & Usher, M. (2012). Salience driven value integration explains decision biases and preference reversal. *Proc Natl Acad Sci U S A*, *109*(24), 9659-9664. https://doi.org/10.1073/pnas.1119569109

Tsetsos, K., Gao, J., McClelland, J. L., & Usher, M. (2012). Using Time-Varying Evidence to Test Models of Decision Dynamics: Bounded Diffusion vs. the Leaky Competing Accumulator Model. *Front Neurosci*, *6*, 79. https://doi.org/10.3389/fnins.2012.00079

Tsetsos, K., Moran, R., Moreland, J., Chater, N., Usher, M., & Summerfield, C. (2016). Economic irrationality is optimal during noisy decision making. *Proc Natl Acad Sci U S A*, *113*(11), 3102-3107. https://doi.org/10.1073/pnas.1519157113

Tsetsos, K., Usher, M., & Chater, N. (2010). Preference reversal in multiattribute choice. *Psychol Rev*, *117*(4), 1275-1293. https://doi.org/10.1037/a0020580

Turner, B. M., Schley, D. R., Muller, C., & Tsetsos, K. (2018). Competing theories of multialternative, multiattribute preferential choice. *Psychol Rev*, *125*(3), 329-362. https://doi.org/10.1037/rev0000089

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological review*, *79*(4), 281.

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, *211*(4481), 453-458. https://doi.org/DOI 10.1126/science.7455683

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, *5*, 297-323.

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management Science*, *39*(10), 1179-1189.

Urai, A. E., de Gee, J. W., Tsetsos, K., & Donner, T. H. (2019). Choice history biases subsequent evidence accumulation. *Elife*, *8*. https://doi.org/10.7554/eLife.46331

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological review*, *108*(3), 550.

Usher, M., & McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychol Rev*, *111*(3), 757-769. https://doi.org/10.1037/0033-295X.111.3.757

Usher, M., Tsetsos, K., Glickman, M., & Chater, N. (2019). Selective integration: an attentional theory of choice biases and adaptive choice. *Current Directions in Psychological Science*, *28*(6), 552-559.

Vlaev, I., Chater, N., Stewart, N., & Brown, G. D. (2011). Does the brain calculate value? *Trends Cogn Sci*, *15*(11), 546-554. https://doi.org/10.1016/j.tics.2011.09.008

Von Neumann, J., & Morgenstern, O. (2007). Theory of games and economic behavior: 60th anniversary commemorative edition. In *Theory of games and economic behavior*. Princeton university press.

Wald, A. (2004). *Sequential analysis*. Courier Corporation.

Wald, A., & Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *The Annals of Mathematical Statistics*, 326-339.

Webb, R., Glimcher, P. W., & Louie, K. (2021). The normalization of consumer valuations: Context-dependent preferences from neurobiological constraints. *Management Science*, *67*(1), 93-125.

Wollschläger, L. M., & Diederich, A. (2012). The 2 N-ary choice tree model for N-alternative preferential choice. *Frontiers in psychology*, *3*, 189.

Yang, T., & Shadlen, M. N. (2007). Probabilistic reasoning by neurons. *Nature*, *447*(7148), 1075-1080.