Exploring the bounded rationality in human decision anomalies through an assemblable computational framework

Yi-Long Lu[1], Yang-Fan Lu[2,3], Xiangjuan Ren[1,3,4,7], Hang Zhang[1,3,4,5,6]*

[1] School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China

[2] Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China

[3] Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China

[4] PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China

[5] Key Laboratory of Machine Perception, Ministry of Education, Peking University, Beijing, China

[6] Chinese Institute for Brain Research, Beijing, China

[7] Max Planck Research Group NeuroCode, Max Planck Institute for Human Development, Berlin, Germany

*: corresponding author.

* Correspondence should be addressed to: Hang Zhang, School of Psychological and Cognitive Sciences, Peking University, Beijing, China. E-mail: hang.zhang@pku.edu.cn

Short title: Assemblable Resource-Rational Modules Framework

Abstract

Some seemingly irrational decision behaviors (anomalies), once seen as flaws in human cognition, have recently received explanations from a rational perspective. The basic idea is that the brain has limited cognitive resources to process the quantities (e.g., value, probability, time, etc.) required for decision making, with specific biases arising as byproducts of the resource allocation that is optimized for the environment. This idea, as a variant of bounded rationality, has grown into a fast-evolving subfield. However, the following issues may limit its development: the assumptions of different models lack consistency, each model typically focuses on one single environmental factor, and the covered decision anomalies are still limited. To address these issues, here we develop a computational framework—the Assemblable Resource-Rational Modules Framework (ARRM)—that integrates ideas from different lines of boundedly-rational decision models as freely assembled modules. The framework can accommodate the joint functioning of multiple environmental factors, and allow new models to be built and tested along with the existing ones, potentially opening a wider range of decision phenomena to bounded rationality modeling. We further apply ARRM to modeling the “peanuts effect” in decision under risk—people are more risk-seeking for gambles with smaller values, an anomaly that is pervasive and hard to explain. For one new and three published datasets that cover two different task paradigms and both the gain and loss domains, our boundedly-rational models reproduce two characteristic features of the peanuts effect and outperform previous models in fitting human decision behaviors.

Keywords: resource rationality, efficient coding, rate-distortion theory, decision under risk, probability distortion, peanuts effect

Author Summary

In daily life, we often make decisions that may seem illogical or irrational at first glance. These decisions were once viewed as mere flaws in our thinking. But what if our brains, with their limited processing power, are actually making the best possible choices given the constraints they face? Past work has introduced the concept of “bounded rationality”, suggesting that our seemingly illogical decisions are in fact rational, considering our brain's constraints and the environment we operate in. Building on this foundation, we offer a fresh perspective. We have developed a tool, the Assemblable Resource-Rational Modules Framework (ARRM). Imagine this as a Lego set, where each piece represents different theories about our decision-making process, many of which have been proposed by earlier researchers. Our innovation is in how we can combine these pieces, enhancing our understanding of complex decisions. We applied this to the “peanuts effect”, where people risk more for smaller rewards. Our approach opens up new avenues for understanding why our brains make the choices they do in the vast, intricate landscape of daily life.

INTRODUCTION

Many decision problems in the real world can be framed as choice among gambles whose outcomes are probabilistic. What principles do human decisions follow? An early influential theory was the Expected Utility Theory (EUT; von Neumann & Morgenstern, 1944), which assumes that humans maximize expected utility, as any rational agents would do. However, following the pioneering experimental work of Allais (1953), significant violations of EUT were identified in human decisions. These anomalies motivated Kahneman and Tversky's Cumulative Prospect Theory (CPT; Kahneman & Tversky, 1979; Tversky & Kahneman, 1992), which retains EUT's assumption of expected utility maximization but introduces a distortion of probability in the computation of expected utility, typically an overestimation of small probability and underestimation of large probability (i.e., an inverted-*S*-shaped distortion curve). With probability distortion modeled by a one- or two-parameter function (Gonzalez & Wu, 1999; Kahneman & Tversky, 1979; Prelec, 1998), CPT provides a good fit to a variety of human decision behaviors, including those that challenge EUT, and thus becomes one of most widely used decision models in economics, psychology, and neuroscience (Glimcher & Fehr, 2014).

From a cognitive view, CPT is unsatisfying in its lack of mechanistic explanation for the observed value or probability distortions. Even as a descriptive model for human behaviors, it has some critical limitations. For example, it assumes a separability of value and probability, that is, neither does the utility function depend on probability, nor does the probability distortion function depend on value. This separability assumption turns out to be oversimplified, failing to capture the rich interaction of value and probability in human decisions (Peterson et al., 2021). A well-documented phenomenon is known as the "peanuts effect": People tend to take more risk when the value at stake is small (Estle et al., 2006; Fehr-Duda et al., 2010; Green et al., 1999; Jones & Oaksford, 2011; Kachelmeier & Shehata, 1992; Prelec & Loewenstein, 1991; Weber & Chapman, 2005).

These highly patterned violations of EUT and even CPT lead some researchers to consider an entirely different possibility: people may not compute or maximize expected utility at all, but instead rely on simple heuristic algorithms to make decisions (Gigerenzer et al., 2011), often discarding part of the available information. For example, the maximax heuristic compares the best outcome of each gamble and chooses the one whose best outcome is the best (Coombs et al., 1970); the similarity heuristic ignores the dimension that is similar for different gambles and focuses on the other dimension (Rubinstein, 1988). Numerous different heuristic decision models have been proposed (Birnbaum & LaCroix, 2008; Brandstätter et al., 2006; Leland, 1998). However, one problem with heuristic models is a lack of consensus on which heuristic to use in which situation (Callaway et al., 2022). An even more serious problem is that these heuristics may not produce the same types of decision biases as humans do (Birnbaum, 2008). Indeed, no known heuristic models have outperformed CPT in capturing the comprehensive patterns of human decision behaviors (Erev et al., 2010).

A third modeling approach for human decisions, alternative to CPT and heuristics, is bounded rationality (Anderson, 1990; Simon, 1955), where humans are modeled as agents who have limited computational power, but who use it efficiently to achieve adaptive behaviors, with decision biases arising as byproducts. An early model along this line is Decision by Sampling (Stewart et al., 2006), which explains the inverted-*S*-shaped probability distortion as the consequence of an adaptive representation of probability according to the event probabilities in the environment. More recently, such bounded rationality has been formulated within the framework of information theory (Bates et al., 2019; C. A. Sims, 2003; C. R. Sims, 2016), providing normative explanations for the subjective distortions of value (Bates et al., 2019; Gershman & Bhui, 2020; C. R. Sims, 2016), probability (Bhui & Gershman, 2018; Zhang et al., 2020), numerosity (Heng et al., 2020), and temporal delay (Gershman & Bhui, 2020). What lies at the heart of these models is the assumption of limited cognitive resources and a rational use of these resources. This approach, as a branch of bounded rationality, is thus called resource

rationality (Bhui et al., 2021; Lewis et al., 2014; Lieder & Griffiths, 2020).

The resource rationality approach is appealing, because it can provide normative accounts for how human decisions may be determined by both the environmental and cognitive constraints. Meanwhile, it is faced with three challenges. First, most of its applications focus on rational resource allocation controlled by one single dimension of the environment, which could be an oversimplification of human decision problems. A full treatment needs to deal with the joint influence of multiple environmental factors. Second, it is common for different alternative assumptions to be used in the resource-rational modeling of different phenomena. Because most studies do not compare across alternative resource-rational models, it is often unknown whether a specific assumption is necessary for a specific phenomenon and which assumption best describes the optimalization problem posed on and solved by human cognition. Third, only a small portion of known decision anomalies have been addressed by the approach (Frydman & Jin, 2021; Gershman & Bhui, 2020; Summerfield & Parpart, 2022). It can hardly explain a wider range of decision phenomena unless its model space is significantly expanded.

In this article, we developed a framework—the Assemblable Resource-Rational Modules (ARRM) framework—to meet these challenges and pave the way for the future development of resource-rational decision models. Following previous work (Bhui et al., 2021; Dayan & Abbott, 2001; C. R. Sims, 2018; Wei & Stocker, 2015), we modeled cognitive processing as information transmission through a constrained communication channel. The ARRM framework specifies the modules of resource-rational models (cognitive resources, prior in memory, encoding and decoding schemes, etc.) and provides a taxonomy for each module to classify its different variants in the literature. The framework allows us to formulate the joint effects of multiple environmental factors on resource allocation and ultimately on decision biases. More generally, it allows us to assemble assumptions from different models to create new resource-rational models. By comparing different models against human data, we could further discriminate between different assumptions of resource rationality and identify those that best describe

human cognition.

We further applied the ARRM framework to decision under risk, exemplifying its explanatory power for the peanuts effect—a pervasive anomaly that was hard to explain. In particular, we constructed 10 new models by combining previous resource-rational models with hypotheses for two different environmental factors and tested these models on one new and three published datasets that span two different task paradigms and both the gain and loss domains. The winning model was the bounded log-odds model (Zhang et al., 2020) enhanced with the rational inattention hypothesis (Gershman & Bhui, 2020; C. A. Sims, 2003) and the structural prior hypothesis (Pleskac & Hertwig, 2014), which provides a complete description of human decision behaviors including two characteristic features of the peanuts effect. We close with a discussion of ARRM's other potential applications and its relationship to heuristic-based or attention-dependent decision models.

Results

The Assemblable Resource-Rational Modules Framework

One bottleneck of human cognitive processing is its limited resources which may be framed as information transmission through a constrained communication channel (Bhui et al., 2021; Dayan & Abbott, 2001; C. R. Sims, 2018; Wei & Stocker, 2015). Following previous work, we developed the Assemblable Resource-Rational Modules (ARRM) framework (see Figure 1A for an overview) to model resource-rational agents who use their limited communication channel adaptively to process the information needed for decision making.

Without loss of generality, the processing of a specific stimulus can be decomposed into an encoding process (from stimulus s to a noisy internal representation m) and a decoding process (from representation m to a subjective estimate $\hat{s}$). The encoding and decoding schemes chosen by a resource-rational agent may depend on the agent's cognitive resources, prior beliefs about the environment, and optimization goal (loss function). These "modules" together constitute the ARRM framework. The framework has multiple variants for each module, and as the "assemblable" in its name suggests, allows the variants from different modules to be combined in an unconventional way to form new resource-rational models. Below we specify the possible variants for each module, focusing on how the model space may be systematically expanded to cover a broader range of decision phenomena.

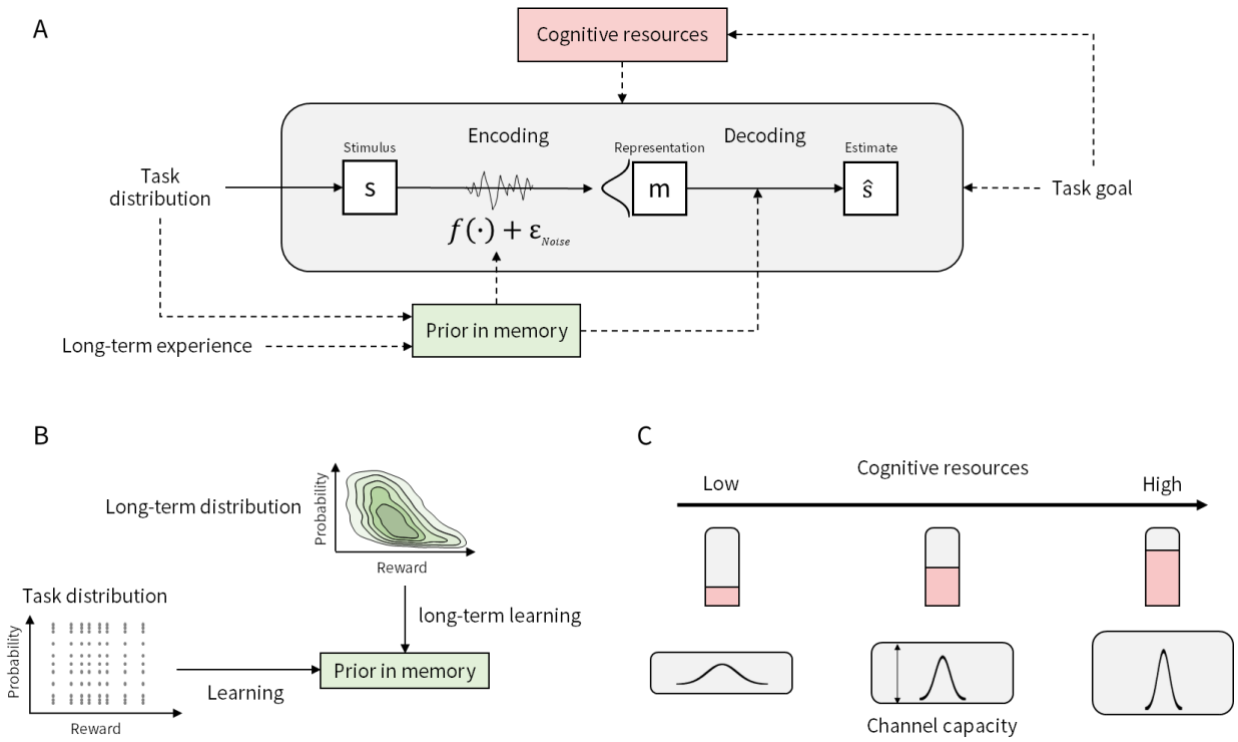


Figure 1. The ARRM framework of stimulus processing. (**A**) Schema of the framework, where a resource-rational agent's encoding and decoding processes are constrained by limited cognitive resources and optimized under specific prior in memory and optimization goal (loss function). The allocation of cognitive resources may also be optimized. (**B**) Illustration of two possible origins of the prior information in memory: the stimulus distribution in the current task (task distribution) and the distribution from long-term life experience (long-term distribution). The prior distribution can either be a marginal distribution over the stimulus in question (e.g., the probability of an outcome), or a conjunctive distribution of multiple attributes (e.g., the conjunctive distribution of value and probability), as implied by the saying "the higher the risk, the greater the return" (illustrated in the plot for long-term distribution). (**C**) Illustration of channels with different levels of cognitive resources. Higher cognitive resources lead to higher representational precision (i.e., lower noise).

Cognitive Resources: Three Allocation Problems

All resource-rational models of decision making share a common assumption that the brain tries to optimize information transmission through a constrained communication channel, but differ in their assumptions about the resource constraints or optimization goal. They descend from two theoretical lines—efficient coding (Attneave, 1953; Barlow, 1961; Simoncelli & Olshausen, 2001; Wei & Stocker, 2015) and the rate-distortion theory (Bates et al., 2019; Gershman & Bhui, 2020; MacKay, 2003; C. R. Sims, 2016). The resource assumptions in these two lines of models are analogous to the charges in two types of mobile data plans: Efficient-coding models resemble plans with a fixed monthly fee, where each bit of transmitted information counts as a gain; in contrast, rate-distortion models are like plans charged by data usage, where each bit of information adds to the cost.

Depending on whether information flow is viewed as cost or gain in a resource-rational model, there are two different ways to measure the resources available to a specific cognitive process. From the cost view, cognitive resources may be measured based on information theory as the rate of information flow through the communication channel. According to a classic measure, the information rate of cognitive processing (e.g., mapping stimulus to response) is only 2 to 3 bits per stimulus (Miller, 1956); in judgment or decision-making tasks involving multiple stimulus dimensions, the information rate for a specific stimulus (e.g., probability) may be even lower, up to 1 to 2 bits (Zhang et al., 2020). From the gain view, cognitive resources may be defined by the parameters of the (possibly hypothetical) encoding device, such as the precision of a bounded scale with homogeneous Gaussian noise (Wei & Stocker, 2015; Zhang et al., 2020) or the number of internal states used for representation (Bhui & Gershman, 2018; Heng et al., 2020).

No matter how cognitive resources are defined, the following assumption applies: The higher the cognitive resource used to encode a stimulus, the more precisely it can be represented (Figure 1C). What controls the cognitive resources available to the stimulus at a

specific moment? Below we specify three different problems of resource allocation that a resource-rational agent may face, starting with a classic problem and then extending to more dynamic or multiple-dimensional resource allocation problems.

The first problem is a classic coding problem in information theory: how to allocate cognitive resources across the different possible values of the to-be-encoded stimulus. This is also the problem most of previous efficient coding models (Bhui & Gershman, 2018; Polanía et al., 2019; Wei & Stocker, 2015) focus on, that is, to adjust the encoding scheme according to the prior distribution (or information source) of the stimulus to fulfill specific goals. As a special case, for fixed-length internal scale with homogeneous Gaussian noise, if the goal is to maximize the mutual information between stimulus and representation, the rational agent should use a stimulus-to-representation mapping that leads to uniformly distributed representations on the internal scale (Dayan & Abbott, 2001). Intuitively, the more frequently a value (or value range) occurs for the stimulus, the more resources should be invested. But note that this optimal solution (and its intuition) may not apply to encoding problems with different resource constraints or optimization goals. For example, if the resource constraint is imposed on the amount of information transmitted and the goal is to minimize the squared error between the stimulus and its subjective estimate, as typical of rate-distortion models, then the optimal encoding scheme for a Gaussian information source would allocate resources homogeneously across possible stimulus values (Bates et al., 2019; Berger & Gibson, 1998), instead of assigning more resources to more frequent stimulus values.

Cognitive resources used to represent a stimulus are not necessarily stationary but may change over time partly under one's volition. From this more *dynamic view* comes the second resource-allocation problem: How should a resource-rational agent allocate cognitive resources available across time? In 2003, C. A. Sims proposed the hypothesis of *rational inattention* that people should invest more of their limited cognitive resources in choices that are more important (C. A. Sims, 2003), which has been used to model a variety of decision phenomena (Gershman

& Bhui, 2020; Grujic et al., 2022; Mikhael et al., 2021; C. A. Sims, 2003). The rational inattention hypothesis is also consistent with qualitative observations that people tend to take more cognitive effort when the task incentives increase (Kleinsorge & Rinkenauer, 2012; Otto & Vassena, 2021).

One difference between the first (classic-coding) and the second (dynamic-view) resource-allocation problems is that the former is only about the to-be-encoded stimulus itself, whereas the latter may allow influences from other stimulus dimensions, such as allocating more resources to a probability stimulus when it is in a gamble associated with a more valuable outcome. Besides, the solution to the former often requires remapping of the whole stimulus distribution, while that of the latter can be an adjustment of encoding precision for individual stimuli.

The two problems discussed above are both about resource allocation *within the stimulus dimension*. In real-world choices where multiple dimensions are involved, different dimensions may compete for a shared pool of cognitive resources, analogous to multiple items in working memory (van den Berg & Ma, 2018). This is our third resource-allocation problem: how to allocate resources *among different dimensions* (e.g., value versus probability in a gamble). Cross-dimension resource allocation may also vary from time to time, extending the dynamic view discussed above. For homogeneous dimensions, a rational agent should, at each moment, allocate more resources to the dimension that is more uncertain (Dewan, 2020). If the dimensions are heterogeneous, a rational agent should allocate more resources also to dimensions that contribute more to the task goal (Bates et al., 2019; Bates & Jacobs, 2021). Though such cross-dimension resource allocation is an important part of real-world decisions, how it works is still largely unknown.

The ARRM framework can be used to model the resource allocation in all these three types of problems.

Priors in Memory: Two Sources

A resource-rational agent must also consider the stimulus distribution of the environment, under which any optimization goal is defined. More precisely, it is the agent's subjective beliefs about the distribution instead of the objective distribution that defines the optimization problem. To emphasize that these prior beliefs are maintained in memory and their representations may also be resource-constrained, we call them *priors in memory*.

Different studies have used different sources of prior distributions to model human behavior: Some studies equate the agent's prior with the stimulus distribution in the current task environment (Heng et al., 2020; Jazayeri & Shadlen, 2010), others use the long-term environmental statistics (Grujic et al., 2022; Stewart et al., 2006; Wei & Stocker, 2015), and still others simply assume a Gaussian distribution for the convenience of analytical derivation (Gershman & Bhui, 2020).

In our framework, we refer to the stimulus distribution in the current task environment and in the long-term environment respectively as the task distribution and the long-term distribution (Figure 1B). For a rational agent who aims to maximize performance in a specific task, using priors learned from the task distribution seems to be self-evident. Meanwhile, using priors based on the long-term distribution is also justified, because learning the stimulus distribution separately for each task might be costly and even implausible when only a few samples are available (Binz et al., 2022; Dasgupta et al., 2020). In such situations, it is thus reasonable to assume that people may base their priors on the more stable (i.e., more slowly changing) long-term distribution, especially when similar stimuli are repeatedly encountered in real life. Indeed, the long-term prior has accounted for a wide range of human behaviors in judgment and decision making (Girshick et al., 2011; Griffiths & Tenenbaum, 2006; Stewart et al., 2015; Wei & Stocker, 2015).

We illustrate these two sources of priors using the example of the conjunctive distribution of reward and probability in gambles (Figure 1B). This conjunctive distribution is used to highlight

the possible covariation between different stimulus dimensions, which may influence a resource-rational agent's choice of encoding and decoding schemes even when the coding is only concerned with one of the dimensions. In decision under risk experiments using a factorial design for reward and probability, the two dimensions are independent of each other in the task distribution. In contrast, the long-term distribution from life experience may have certain structure, and echo "the higher the risk, the greater the return", exhibiting a negative correlation between reward and probability (Pleskac & Hertwig, 2014).

The resulting priors in memory can be a mixture of the two sources of distributions, with the task distribution carrying the immediate task information and the long-term distribution for prior knowledge acquired over years. There can even be different updating speeds for the former and the latter, so that people may learn the task environment well after hundreds of trials but are still subject to the influence from the long-term distribution (See Supplemental Text S1 for two possible updating algorithms for priors in memory).

Encoding Schemes: the flexibility and speed of adaptiveness

An encoding scheme is a mapping from stimulus to internal representation that is often stochastic. The stochastic mapping may be separated into two components—a deterministic mapping rule and a noise distribution. The former determines the mean of the internal values to which the stimulus is mapped, while the latter controls the encoding precision. Depending on how flexible the mapping can adapt to cognitive resources and priors in memory, the encoding schemes used in previous rational models may be classified into three levels of adaptiveness: fixed, boundedly-adaptive, and fully-adaptive.

Fixed encoding. For the fixed encoding scheme, neither the encoding mapping $f(\cdot)$ nor the encoding noise ε may change with cognitive resources or priors in memory. The representation m for a specific stimulus s can be written as

$$m = f(s) + \varepsilon\,. \tag{1}$$

Such fixed encoding schemes are assumed in traditional Bayesian observer models (Petzschner et al., 2015), where the observer can still be adaptive to the environment during the decoding process through Bayesian inference (i.e., integrating priors about environmental statistics with the noisy representation of the stimulus). Besides its application in perception, fixed encoding with Bayesian inference has also been used to explain decision phenomena, such as the inverted-*S*-shape of probability distortion (Khaw et al., 2021; Zhang & Maloney, 2012).

Boundedly-adaptive encoding. A boundedly-adaptive encoding scheme may adapt to cognitive resources or priors, but in a rather constrained way. An example is the Bounded Log-Odds (BLO) model (Ren et al., 2021; Zhang et al., 2020), which assumes that probability is encoded on a bounded log-odds scale and the mapping function from stimulus to representation is a linear-in-log-odds function with truncations. The functional form of the mapping is fixed but the truncation parameters may be adaptive. The precision of the encoding scale (or equivalently, the length of the scale if precision is held constant) is determined by the available cognitive resources. Such boundedly-adaptive encoding may be formalized as:

$$m = f\left(s; \theta_{g_s,R}\right) + \varepsilon_R \,, \tag{2}$$

where $\theta_{g_s,R}$ denotes parameters that are optimized for stimulus prior distribution g_s and resource R, and ε_R denotes the encoding noise under resource R.

Fully-adaptive encoding. Different from BLO, many other efficient coding models assume a more flexible encoding scheme, without constraining the functional form of the encoding mapping (Bhui & Gershman, 2018; Heng et al., 2020; Wei & Stocker, 2015). In these models, encoding mapping can be changed flexibly to minimize the expected loss under any specific prior distribution of the stimulus. Though these efficient coding models usually do not specify the effects of varying precision, they have the implicit assumption that the encoding precision of a stimulus varies with the available cognitive resource. For such encoding schemes, the

representation can be written as

$$m = f_{g_s,R}(s) + \varepsilon_R \,, \tag{3}$$

where $f_{g_s,R}(\cdot)$ denotes an encoding function that is optimized for the prior distribution g_s and resource R.

The encoding schemes in rate-distortion models (Bates et al., 2019; Gershman & Bhui, 2020) can be even more flexible, which allow both mapping and precision to be changed:

$$m = f_{g_s,R}(s) + \varepsilon_{g_s,R}(s), \tag{4}$$

where $\varepsilon_{g_s,R}(s)$ denotes the encoding noise for stimulus s changes with the prior distribution g_s and resource R.

Slow versus fast adaptiveness. The adaptation of encoding schemes to the environment may occur on multiple different time scales. Many of the encoding schemes that we categorize as fully-adaptive in previous efficient coding models (Stewart et al., 2006; Wei & Stocker, 2015) are actually adaptations to the real-world statistics (i.e., long-term prior in our term), which may take years to happen. Some seemingly fixed encoding schemes (e.g., logarithmic representation of magnitude; Dehaene et al., 2008; Petzschner et al., 2015) may come from a slow adaptation to the long-term environmental statistics (e.g., the power-law distribution of magnitudes; Stewart et al., 2006). Boundedly-adaptive encoding schemes can be seen as fast but minor adjustments within the functional form shaped by long-term adaptation (Payzan-LeNestour & Woodford, 2022).

The encoding scheme may change with priors and cognitive resources at different speeds. How fast the encoding mapping can adapt to the stimulus distribution is still largely unknown. Though the adaptation to the task distribution is found in the representation of value (Frydman & Jin, 2021; Louie et al., 2015) and relative frequency (Ren et al., 2018), the effects are usually not as large as the prediction of efficient coding theories, with some effects even being controversial (Alempaki et al., 2019; Stewart et al., 2015). In contrast, there is stronger evidence

that the precision of encoding may change on a trial-by-trial basis (Bates & Jacobs, 2021; Gershman & Bhui, 2020).

Decoding Schemes

Some efficient coding models skip the decoding process and directly use the internal representation m for further judgment or decision making (Bhui & Gershman, 2018; Heng et al., 2020; Stewart et al., 2006). The decoding process may not be necessary when the task is to compare whether one quantity is greater or less than the other, in which case the representation can be used directly for the task.

However, in the more general case where a stimulus is to be explicitly estimated or to be integrated with stimuli on other dimensions (e.g., probability to be integrated with value), a decoding process must be assumed (Polanía et al., 2019; Wei & Stocker, 2015; Zhang et al., 2020). For a rational agent, a statistically principled decoding method is Bayesian inference (Petzschner et al., 2015; Wei & Stocker, 2015), which combines the noisy representation m with the prior information to calculate the posterior distribution of the stimulus. To further obtain a point estimate $\hat{s}$, one needs to specify the loss function to optimize.

For the rate-distortion models, one may directly obtain the mapping from s to $\hat{s}$ based on the rate-distortion theory, without going into the details of encoding and decoding. Even so, a Bayesian decoding process is implicitly assumed, because it constitutes the only optimal solution to the optimization problem.

Optimization Goal (Loss Function)

The optimization goal (loss function) is indispensable for the definition of any rational agent. Different from Bayesian ideal observers who only care about decoding, a resource-rational agent would optimize both the encoding and the decoding schemes (denoted by $T(\cdot)$ as a whole), subject to cognitive constraints, to minimize the expected loss. For stimulus s, prior distribution g_s and loss function $L(\cdot)$, the optimal $T(\cdot)$ is defined by

$$T^* = \underset{T}{\operatorname{argmin}}\, \mathbb{E}_{g_s}[L(s, T(s))\,]\,. \tag{5}$$

Here the loss function may be a function of the deviation between stimulus s and its subjective estimate $\hat{s}$, or refers to the loss that is explicitly defined by the task (e.g., error rate or monetary loss). It may also be defined separately for encoding (without decoding). For example, in previous efficient coding models, the goal is often to maximize the mutual information between stimulus and representation (Bhui & Gershman, 2018; Wei & Stocker, 2015). Using the explicit task goal to define loss function is common in previous rate-distortion models (Bates & Jacobs, 2021) but less popular in previous efficient coding models (see Heng et al., 2020 for an exception), though it is applicable to both lines of resource-rational models.

"Bounded" Resource Rationality

Solving the exact optimization problem can be difficult and even intractable for the brain, especially on a fast time scale in a complex environment. Humans may adapt to the environment in a more bounded way than perfect resource-rational agents would behave. The potential boundedness in resource rationality could be modeled by the ARRM framework, which was discussed earlier for individual modules and is summarized below.

First, the encoding scheme may be boundedly adaptive instead of fully adaptive to the environment, such as assumed in the bounded log-odds model (Zhang et al., 2020). There is also evidence that an approximate implementation of resource rationality like the Decision by Sampling model (Stewart et al., 2006) may behave closer to humans than an exact resource-rational agent does (Heng et al., 2020).

Second, the encoding scheme adapts to the environment often in the long run (Stewart et al., 2006; Wei & Stocker, 2015) instead of on a time scale as short as minutes or hours (Alempaki et al., 2019; Bujold et al., 2021; Frydman & Jin, 2021; Ren et al., 2018). In contrast, the adaptation of the decoding scheme may be much faster, according to the Bayesian literature (Maloney & Zhang, 2010; Petzschner et al., 2015),.

What is explored in the ARRM framework above are theoretical possibilities, while how humans may behave like resource-rational agents is an empirical question. Next, we apply the framework to decision under risk and examine which set of theoretical possibilities best fits human behaviors.

Understanding the Peanuts Effect in Decision Under Risk

We now use the ARRM framework to understand the behavior pattern of human decision under risk. Here we focus on a decision anomaly known as the peanuts effect, because it is well documented but can hardly be explained by previous decision models, including those from the resource rationality approach. Our goal is two-fold: first, to show how the different modules of the ARRM framework can be assembled to extend the model space of resource rationality; second, to explain and predict the peanuts effect in human decisions. Before diving into the resource-rational modeling of the peanuts effect, we briefly review the history of the anomaly and why it would be hard to explain.

The Peanuts Effect

The anomaly was first presented by Markowitz (1952) as an (informal) observation that people may reverse their preferences between a sure reward and a gamble when the rewards of the two are proportionally enlarged. The set of choices constructed by Markowitz have the common format “X with certainty or one chance in ten of getting Y?”[1], where Y=10X and X can be 10 cents, $1, $10, $100, $1,000, or $1,000,000. In all these choice pairs, the sure reward and the gamble have the same expected gain. The puzzling phenomenon is that most people

[1] Markowitz (1952) constructed a parallel set of choices for the loss domain in a similar format as “X with certainty or take one chance in ten of owing Y?”. Contrary to the gain domain, most people seem to prefer the sure loss for small X and Y but the gamble for large X and Y. As we will see, the peanuts effect in both the gain and loss domains can be explained by the same model developed in the ARRM framework.

seem to prefer the gamble for small X and Y but the sure reward for large X and Y. It receives the name “peanuts effect” in Prelec and Loewenstein (1991), defined as an elevated risk aversion for large gambles. The phenomenon has been repeatedly replicated in later studies on decision under risk (Binswanger, 1981; Estle et al., 2006; Fehr-Duda et al., 2010; Kachelmeier & Shehata, 1992; Weber & Chapman, 2005), for rating as well as choice tasks (Weber & Chapman, 2005), for real as well as virtual monetary rewards (Binswanger, 1981; Fehr-Duda et al., 2010; Kachelmeier & Shehata, 1992). As we will soon see, this phenomenon also widely exists in previous datasets that were not intended to test it, often unnoticed in the original publications.

The theoretical difficulty in explaining the peanuts effect. Although the peanuts effect has been documented for decades, it has no satisfactory explanation yet. In his original paper, Markowitz (1952) attributes the peanuts effect to utility functions that are convex in small values but concave in large values. According to such utility functions, the higher value of a small gamble is much more appealing than its paired sure reward, while for large gambles the appeal is weaker. However, there is little empirical evidence for such utility functions; what are widely accepted (Tversky & Kahneman, 1992) and agree with the non-parametric measurement (Gonzalez & Wu, 1999) of human decisions are power utility functions. Fehr-Duda et al. (2010) examined human choices for gambles with a range of different probabilities and argued that no utility functions may fully explain the peanuts effect.

The peanuts effect poses an anomaly not only for the expected utility theory, but also for CPT (Tversky & Kahneman, 1992) and many other theories (Prelec & Loewenstein, 1991; Wakker, 2010), because it challenges the *separability assumption* that value and probability are independently processed before their integration, that is, the utility function $u(x)$ is determined only by value, and the probability weighting function $w(p)$ only by probability. If such separability held, scaling up the values proportionally for the sure reward and gambles should not have influenced human preference between the two, at least under the widely accepted power utility

function. In our Study 1, we will present an even stronger objection to Markowitz's (1952) explanation, proving that no utility or probability weighting functions may explain the observed peanuts effect, unless the separability assumption is loosened.

What is difficult to explain in the peanuts effect is not just the existence of an interaction between value and probability (which issue may be readily fixed by loosening the assumption of independence), but the particular direction of the interaction. Parallel to decision under risk, there is an interaction effect between value and delay in temporal discounting, where humans discount future values less for higher values (Prelec & Loewenstein, 1991). This magnitude effect is explained by Prelec and Loewenstein's (1991) decision theory, which assumes that higher values on one dimension increase the relative importance of the dimension in decision making. It may also be understood from the resource-rational perspective (Gershman & Bhui, 2020), where discounting is assumed to be consequence of noisy representation of future value, with higher values attracting more cognitive resources and thus associated with less discounting. However, similar logic cannot explain the peanuts effect, because it runs in the opposite direction as the magnitude effect in temporal discounting, if probability weighting function is treated as an analog of temporal discounting function.

That is why Prelec and Loewenstein's (1991) decision theory could explain 9 out of 10 anomalies summarized in their paper but not the peanuts effect. This failure led Prelec and Loewenstein (1991) to resort to a distinctively different perspective to explain the peanuts effect, that people may maximize their own emotional experience instead of the expected utility, as formulated in the decision affect theory (Mellers et al., 1997). They reasoned that when the value of the gamble is higher, people would face a higher disappointment for choosing the gamble but losing it. As a result, people may reverse their preference between gamble and sure reward when the value of the gamble increases. A similar idea was revisited in Weber and Chapman (2005). Though intuitively appealing, this reasoning holds only if the disappointment function is an increasingly accelerating function of the utility difference, which does not receive

empirical evidence from experiments measuring affective responses to gamble outcomes (Mellers et al., 1997). The empirical measurement of the affective response function even suggests a reverse effect.

The two components of the peanuts effect. Though most previous studies (Binswanger, 1981; Estle et al., 2006; Kachelmeier & Shehata, 1992; Prelec & Loewenstein, 1991; Weber & Chapman, 2005) frame the peanuts effect as something about risk attitude, Fehr-Duda et al. (2010) find it to be more adequately described as a modulation of outcome value on the probability distortion function. A closer examination of the effect reveals two sub-effects (illustrated in Figure 2): (1) the sensitivity (slope) of subjective probability increases with value, and (2) the elevation (cross-over point) of subjective probability decreases with value.

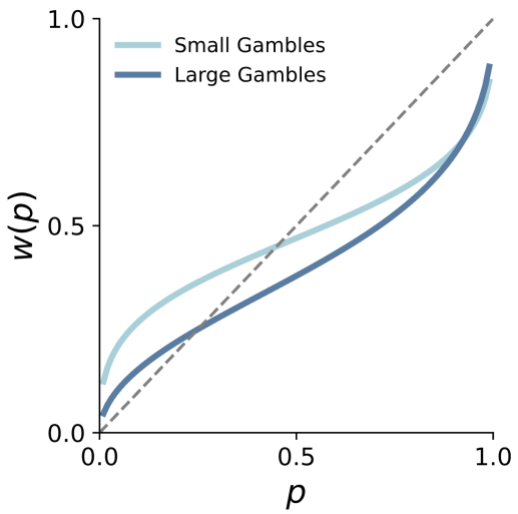


Figure 2. The Peanuts Effect on Probability Distortion. Illustration of the peanuts effect in decision under risk. Subjective probability (aka decision weight) $w(p)$ is plotted as a function of objective probability p. Compared to that of small gambles (light blue curve), the subjective probability of large gambles (dark blue curve) has a steeper slope and a lower crossover point (i.e., where $w(p) = p$).

Though several resource-rational models have treated the distortion of probability (Bhui & Gershman, 2018; Zhang et al., 2020) or value (Heng et al., 2020; Polanía et al., 2019), the previous treatments are about probability or value alone and cannot explain why probability distortion may be influenced by the associated value. In the following, we apply the ARRM framework to modeling human decision under risk. By combining the rational inattention hypothesis of cognitive resource allocation (Gershman & Bhui, 2020; Grujic et al., 2022; C. A. Sims, 2003) with a hypothesis of value-probability covariation in prior beliefs (Pleskac & Hertwig, 2014), we are able to provide a full explanation for the peanuts effect in probability distortion.

ARRM Models for Probability Distortion

We use $g: (x_1, p; x_2)$ to denote a monetary gamble with two outcomes, that is, receiving x_1 with probability p, or x_2 otherwise, where $|\mathrm{x}_1| > |\mathrm{x}_2|$. A sure payoff (gain or loss) is a special case of this two-outcome gamble with $p = 1$. Following CPT (Tversky & Kahneman, 1992), we assume that the expected utility of gamble g can be computed as[2]

$$U(g) = w(p)u(x_1) + \big(1 - w(p)\big)u(x_2), \tag{6}$$

where $u(\cdot)$ denotes the utility function and $w(\cdot)$ denotes the probability weighting (distortion) function. What differs our following modeling from standard applications of CPT is that we replace the probability weighting function $w(\cdot)$ of p with the subjective estimate $\hat{p}$ decoded from noisy representation m_p of p, which may vary from gamble to gamble as the consequence of resource-rational encoding and decoding.

The Rational Inattention (RI) Hypothesis. As discussed earlier for the module of cognitive resources in the ARRM framework, the second resource-allocation problem can be understood as a tradeoff between task performance and cognitive effort: A higher encoding precision would

[2] This equation holds when both outcomes are gains or losses, i.e., $x_1 > x_2 \geq 0$ or $x_1 < x_2 \leq 0$. When one is gain and the other is loss, they need two separate weighting functions.

lead to a more accurate estimation of the stimulus and thus better task performance, but at a cost of higher demand for cognitive resources. To balance task performance and cognitive effort, a resource-rational agent should allocate cognitive resources among different stimuli according to their potential rewards. In behavioral economics, this is known as the rational inattention (RI) hypothesis (Grujic et al., 2022; Matějka & McKay, 2015; Mikhael et al., 2021; C. A. Sims, 2003). That is, stimuli associated with a larger reward receive more attention, while those that are worthless are ignored (Kool & Botvinick, 2018; Matějka & McKay, 2015).

When applied to the representation of probability, the RI hypothesis implies that the cognitive resource increases with the potential gain of the gamble. For gamble $(x_1, p; x_2)$, the resource allocated to the probability p should increase with its associated potential gain (or saved loss) $|\mathrm{x}_1 - \mathrm{x}_2|$:

$$R_{\mathscr{g}} = R_0 + \beta \ln(|\mathrm{x}_1 - \mathrm{x}_2| + 1) \ , \tag{7}$$

where $R_0 \geq 0$ is a parameter for the baseline resource (i.e., when $\mathrm{x}_1 - \mathrm{x}_2 = 0$), and $\beta \geq 0$ is a slope parameter controlling how steep the allocated resource changes with value. Similar linear-log forms of rational inattention assumptions have been used in previous studies (Gershman & Bhui, 2020; Mikhael et al., 2021).

The Structural Prior (SP) Hypothesis. According to the module of priors in memory of the ARRM framework, one source for prior beliefs is the long-term experience in everyday life. Such everyday statistics have been used by previous resource-rational models to explain human distortion of value, probability or time (Bhui & Gershman, 2018; Stewart et al., 2006, 2015), but the priors used in these practices are limited to one single dimension of stimuli.

However, as the saying “the higher the risk, the greater the return” suggests, a structural relationship of reward and probability—the two dimensions of a gamble—widely exists in everyday life, ranging from financial investment to the submission of academic papers (Pleskac & Hertwig, 2014). That is, when the potential value is large, the probability of success tends to

be small. Indeed people may use value as a contextual cue to infer the associated probability (Leuker et al., 2018, 2019). In the structural prior (SP) hypothesis, we hypothesize that people have prior beliefs about the probabilistic interdependence between value and probability.

Following previous studies (Khaw et al., 2021), we model the agent's prior belief for p as a Gaussian distribution on the log-odds scale:

$$\lambda(p) \sim N\left(\lambda_{g}, \sigma_s^2\right), \tag{8}$$

where $\lambda(p) = \ln\frac{p}{1-p}$ is the log-odds of p. The λ_{g} and σ_s^2 respectively denote the mean and variance of the Gaussian prior, with the subscript g of λ_{g} to emphasize its dependence on the gamble in question. When transformed back to the probability scale, the prior would be *U*-shaped for large variance and inverted-*U*-shaped for small variance, resembling the beta distribution used to model the real-world statistics of probability (Stewart et al., 2006; Zhu et al., 2020).

The SP hypothesis for gamble g: $(x_1, p; x_2)$, can then be formulated as

$$\lambda_{g} = \lambda_0 - \theta \ln(|\mathrm{x}_1 - \mathrm{x}_2| + 1)\ , \tag{9}$$

where λ_0 is a baseline parameter, $|\mathrm{x}_1 - \mathrm{x}_2|$ (as before) is the potential gain (or saved loss) of the gamble, and θ is a slope parameter controlling the direction and extent that probability is expected to change with value , with $\theta > 0$ indicating an assumed negative correlation between probability and value (i.e., "the higher the risk, the greater the return"), $\theta = 0$ for no correlation, and $\theta < 0$ for positive correlation.

Study 1: Modeling the peanuts effect in the ARRM framework

To see how well the resource-rational models we constructed in the ARRM framework can explain human decision behaviors, we first need a model-free quantification of the phenomenon in question. For the peanuts effect, which reflects the influence of outcome value on probability distortion (Fehr-Duda et al., 2010), the ideal experimental design shall allow us to measure the

probability distortion curve separately for different values. Fortunately, two previously published datasets satisfy our requirement, one from Gonzalez and Wu (1999) and the other from Zhang et al. (2020), which used an almost identical design of 165 different two-outcome gambles from a factorial combination of 11 different probabilities and 15 different sets of values (see Methods). For each of the 10 or 75 participants, the authors of the studies measured the certainty equivalent (CE) for each gamble, that is, the sure reward that is perceived as equally preferable as the gamble. Based on these certainty equivalents, we elicited the probability distortion curve for each set of values and quantified the two aspects of the peanuts effect—how the sensitivity and elevation parameters of probability distortion change with value.

The peanuts effect in the GW99 and ZRM20 datasets

Non-parametric estimation of probability distortions had been performed for the GW99 (Gonzalez & Wu, 1999) and ZRM20 (JDA and JDB) datasets (Zhang et al., 2020) in their original studies. Such estimation, however, had collapsed data across gambles with different outcome values.

The normalized certainty equivalent, $\mathrm{NCE} = (\mathrm{CE} - x_2)/(x_1 - x_2)$, was used as a measure of subjective probability, which could be computed for each gamble based on the estimated CE for the participant (denoted $\widehat{NCE}$ as estimation). Though $\widehat{NCE}$ is not strictly equal to subjective probability unless the utility function is linear, it served as a useful measure to visualize the influence of outcome value on probability distortion. Figure 3 plots the mean $\widehat{NCE}$ across participants for each dataset (GW99, n=10) or sub-dataset (JDA, n=51; JDB, n=24), separately for gambles with different x_1 (Only gambles with $x_2 = 0$ is plotted for visualization purposes).

The observations in all these datasets confirmed the peanuts effect and its two sub-effects, whose descriptions had been scattered in different previous studies (Fehr-Duda et al., 2010; Green et al., 1999; Prelec & Loewenstein, 1991). Other than exhibiting the typical inverted *S*-shape (linear in the log-odds scale) as in the original analyses of the datasets, the probability

distortion curve had (1) *greater sensitivity* (i.e., steeper slope) and (2) *lower elevation* (i.e., smaller crossover point) for larger gambles (darker curves in Figure 3). The sensitivity effect and elevation effect were supported by linear mixed-effect model (LMM) analysis for participants' $\widehat{NCE}$ (211 in Methods). When the potential gain ($D = |x_1 - x_2|$) was greater, participants were more sensitive to the differences between different probabilities (D by probability interaction, GW99: b = 0.084, 95% CI [0.055, 0.114], P < 0.001; JDA: b = 0.067, 95% CI [0.060, 0.074], P < 0.001; JDB: b = 0.050, 95% CI [0.019, 0.081], P = 0.002), and had lower $\widehat{NCE}$ (main effect of D, GW99: b = −0.211, 95% CI [−0.270, −0.153], P < 0.001; JDA: b = −0.090, 95% CI [−0.114, −0.066], P < 0.001; JDB: b = −0.093, 95% CI [−0.131, −0.055], P < 0.001).

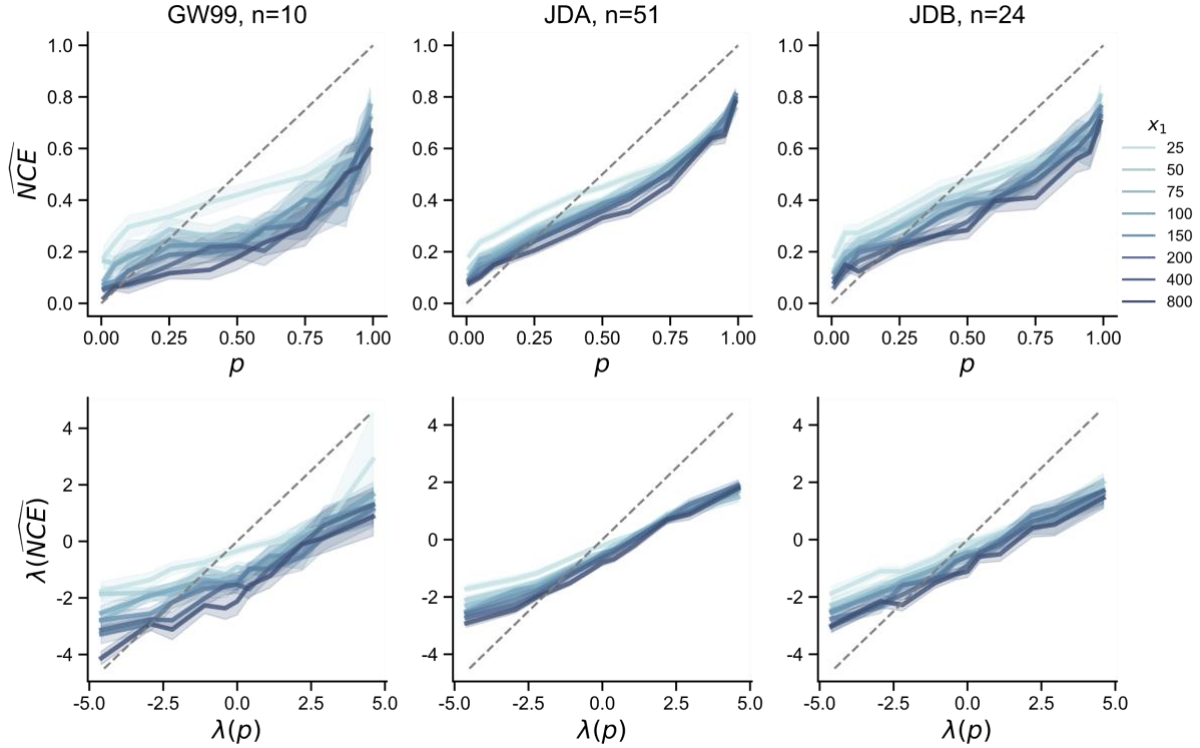


Figure 3. The peanuts effect in the GW99 and ZRM20 (JDA and JDB) datasets. For each dataset or sub-dataset (one column), participants' normalized certainty equivalent ($\widehat{NCE}$, a measure of subjective probability) is plotted against objective probability (p), separately visualized for gambles $(x_1, p; 0)$ with different x_1 (color-coded in different darkness). The upper

and lower rows are respectively for probability and log-odds plots. Shadings denote standard error. Both sub-effects of the peanuts effect are visible in all the datasets: the larger the gamble (darker curves), (1) the greater the sensitivity of subjective probability, and (2) the lower the elevation of subjective probability.

Exchangeability test: rejecting the separability assumption of value and probability

If participants' choices had followed the prospect theory with power utility function $u(x) = x^{\alpha}$ and any probability weighting function $w(p)$, for gamble $(x_1, p; x_2)$ with $x_2 = 0$, we would have $NCE = w(p)^{\alpha}$. If so, NCE would simply be a function of p and have nothing to do with outcome value. However, as shown in the sensitivity and elevation effects above, NCE changed systematically with outcome value. In other words, the peanuts effect violates the separability assumption widely accepted in previous decision models—utility is only determined by value and subjective probability only by probability. This has motivated us to develop resource-rational models to capture the cross-dimensional influence of outcome value on probability distortion.

But one may wonder whether some non-power utility function, such as the four-fold function proposed by Markowitz (1952), could explain the peanuts effect without modifying the separability assumption. To exclude this possibility, we developed a model-free test for separability, termed the *exchangeability test*, whose conclusion can apply to any forms of utility and probability weighting functions.

The basic idea of the test is to navigate in a space of gambles through two different paths, analogous to first going west then north vs. first north then west in the physical space. If the separability between value and probability held, the two paths would lead to the same destination. In particular, the test (illustrated in Figure 4A) is based on the CE measured in the GW99 and ZRM20 datasets for different gambles (Figure 3). For gamble $(x_1, p; 0)$, the separability assumption implies

$$U(x_1, p) = w(p)u(x_1) = U(CE, 1), \tag{10}$$

where $U(x_1, p)$ is the expected utility of the gamble, and $U(CE, 1)$ is the utility of the CE. For any $x_a > x_b > x_c$, we could find probabilities p_b and p_c so that the CE of gamble $(x_a, p_b; 0)$ is x_b and the CE of gamble $(x_a, p_c; 0)$ is x_c. Using these two probabilities we can construct two new gambles $(x_c, p_b; 0)$ and $(x_b, p_c; 0)$. It is easy to prove that these two gambles have the same CE (by applying Eq. 10). Significant differences between the two gambles would allow us to reject separability (see Methods for more details).

The GW99 and ZRM20 datasets allowed us to test 40 pairs of equivalence, across which the measured CE for $(x_c, p_b; 0)$ and $(x_b, p_c; 0)$ were significantly different from each other (paired *t*-test, *t*(39) = 2.79, *p*=0.008), thus rejecting the separability assumption (Figure 4B). That is, no utility or probability weighting functions would explain the observed peanuts effect, unless value and probability are allowed to influence the processing of each other.

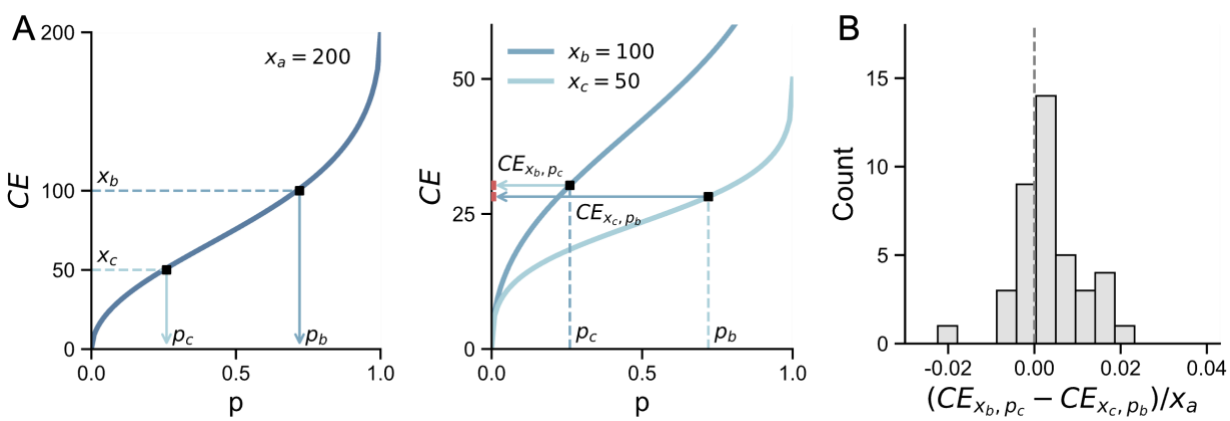


Figure 4. The exchangeability test in the GW99 and ZRM20 datasets. (**A**) Illustration of the rational of the exchangeability test. The test was based on the 8 certainty equivalents (CE) curves estimated from the datasets, separately for gambles with $x_1 = 25,\ 50,\ 75,\ 100,\ 150,\ 200,\ 400,\ 800$. From these 8 different values of x_1, we drew 40 combinations of x_a, x_b, x_c that satisfy $(x_a, p_b; 0) \sim x_b$, $(x_a, p_c; 0) \sim x_c$, and $0.01 \leq p_b, p_c \leq 0.99$, where $\sim$ denotes "equally preferred" (**left panel**). We then tested whether the equivalence pair predicted by the

separability of value and probability—$(x_c, p_b; 0)$ and $(x_b, p_c; 0)$—were equivalent in CE (**right panel**). (**B**) Results of the exchangeability test: the histogram of the 40 equivalence pairs' CE differences. The mean $(CE_{x_b,p_c} - CE_{x_b,p_c})/x_a$ of the 40 equivalence pairs were significantly greater than 0.

Testing 10 resource-rational models in the ARRM framework

Next, we will show resource-rational models could provide a full explanation for the peanuts effect. Meanwhile, by comparing models with or without a specific resource-rational hypothesis, we can see whether the hypothesis holds for human decision behavior.

We constructed 10 resource-rational models in the ARRM framework, using three different base models with or without the RI or SP hypothesis, and fit the models to each participant's $\widehat{NCE}$ using maximum likelihood estimates (see Methods and Supplemental Methods for details). The three base models were the Bounded Log-Odds model (BLO, Zhang et al., 2020), the Decision by Sampling model (DbS, Bhui & Gershman, 2018; Stewart et al., 2006), and the Rate-Distortion Model (RD, Bates et al., 2019; Gershman & Bhui, 2020), which have different resource-rational assumptions (see Supplemental Methods). As shown in Figure 5, the BLO+RI+SP model was the best-fitting model, achieving the lowest summed ΔAIC (the lower the better) in the JDA and JDB datasets and the highest protected exceedance probability (the probability of outperforming the other models in goodness-of-fit) in all the datasets.

By comparing models that only differ in one dimension of assumptions, we can see which assumption best agrees with human data. In most cases, models based on the coding scheme of BLO were better than those based on DbS or RD, though the AIC difference was moderate. Regardless of the choice of the base models (DbS, RD, or BLO), the models enhanced with the RI or SP hypothesis always fit better than those without them. The structural prior parameter θ (Eq. 9) in the winning model (BLO+RI+SP) was significantly greater than 0 (one-sample *t*-test,

one-sided: $t(84) = 2.506$, $p = 0.007$), implying a prior belief consistent with the saying "the higher the risk, the greater the return".

Figure 5. Model comparison results of the GW99 and ZRM20 (JDA and JDB) datasets. The ΔAIC (with the BLO+RI+SP model as the reference) summed across participants is plotted as the metric of goodness-of-fit for each of the 10 models constructed in the ARRM framework. Smaller summed ΔAIC indicates better fit. The protected exceedance probability (P_{exc}, the probability that a specific model outperforms the other models), a group-level Bayesian measure of goodness-of-fit, is marked for the best two models. To highlight the differences between models with and without the RI or SP hypothesis, models are color coded into four categories: base models (without RI or SP), RI models (with RI, without SP), SP models (with SP, without RI), and full models (with both RI and SP). Each panel is for one dataset or sub-dataset. In all the datasets, full models (orange) performed the best, among which the BLO+RI+SP model had the highest exceedance probability.

The BLO+RI+SP model well captured both sub-effects of the peanuts effect as well as the inverted-*S*-shaped probability distortion (Figure 6). Intuitively, the sensitivity effect—the increase of sensitivity with gamble outcome—is explained by the RI hypothesis, which assumes that people would invest more cognitive resources to gambles of higher values. The elevation effect—the decrease of crossover point with gamble outcome—is explained by the SP hypothesis that assumes a negative correlation between probability and outcome value.

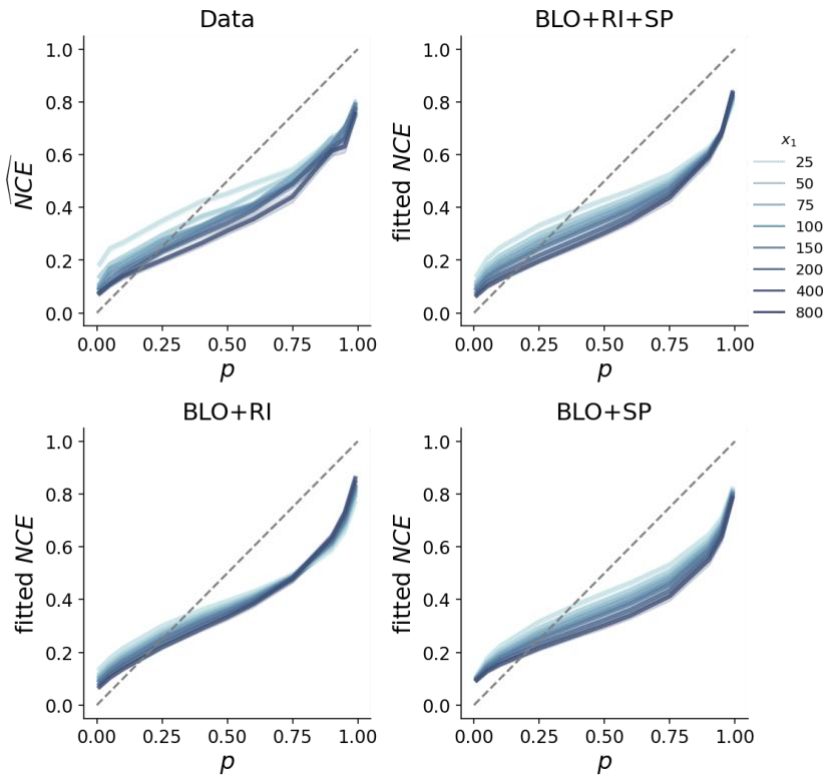


Figure 6. Data versus model predictions, pooled over the GW99 and ZRM20 datasets. Participants' probability distortion curves ($\widehat{NCE}$, upper-left panel) versus model-predicted distortion curves (NCE, other three panels), separately visualized for gambles $(x_1, p; 0)$ with different x_1. The prediction of the best-fit BLO+RI+SP model well captured both the sensitivity and elevation effect. Those of the BLO+RI and BLO+SP respectively captured the sensitivity and elevation effect, but not both, illustrating the necessity of both hypotheses. Conventions follow Figure 3.

To summarize, the BLO model enhanced by the RI and SP hypotheses provided a full explanation for the peanuts effect in decision under risk, with the hypotheses respectively accounting for the two sub-effects. Though the 10 models we considered above were all resource-rational models, a comprehensive set of non-resource-rational models were tested on the GW99 and ZRM20 (JDA and JDB) datasets by Zhang et al. (2020), according to which BLO outperformed all of a large set of decision models they tested, including the conventional CPT model and several heuristic-based models.

Study 2: Replicating the Peanuts Effect and its ARRM Modeling in a Binary Choice Task

Both the GW99 and ZRM20 datasets we reanalyzed above were collected with the certainty equivalent task, where the options were presented in a tabular form, contrasting one gamble with a list of sure rewards in an ascending or descending order. This tabular presentation is efficient in eliciting certainty equivalents but is not the most typical scenario in decision making. To exclude the possibility that our findings only apply to the certainty equivalent task, we performed a new experiment of decision under risk using a binary choice task, where participants chose between a two-outcome gamble $(x_1, p; 0)$ and a sure reward on each trial. There were 8 levels of value x_1 (25, 50, 75, 100, 150, 200, 400 or 800 Chinese Yuan) and 5 levels of winning probability p (0.05, 0.25, 0.5, 0.75 of 0.95), combining into 40 different gambles. Each gamble was repeated for 10 times, paired with different sure rewards that were selected by an adaptive procedure (see Methods for details).

In this binary choice task (Figure 7A, $n = 21$), we replicated the peanuts effect found in the certainty equivalent task of Study 1. According to the LMM analysis on participants' $\widehat{NCE}$, when the potential gain D of the gamble was greater, participants were more sensitive to the

differences between different probabilities (D by probability interaction, b = 0.039, 95% CI [0.006, 0.071], P = 0.020), and had lower subjective estimates of the probabilities (main effect of D, b = −0.084, 95% CI [−0.128, −0.041], P < 0.001).

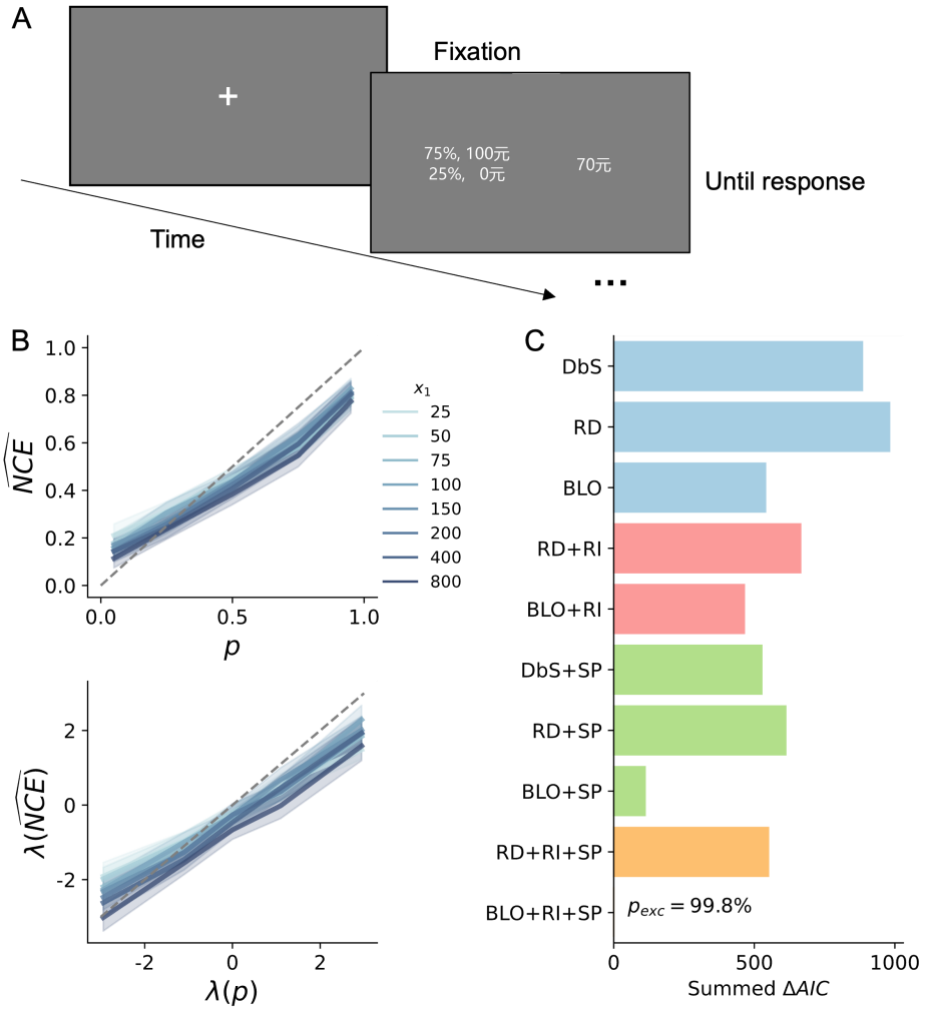


Figure 7. Task and results of the binary choice experiment (*n*=21). (**A**) The binary choice task. On each trial, participants chose whether to receive a gamble or a sure reward. The "元"in the display is the Chinese character for Yuan. (**B**) The measured peanuts effect. To facilitate comparison with the results of the certainty equivalent task (Figure 3), participants' choices were used to derive certainty equivalents, which were then transformed into normalized certainty equivalent ($\widehat{NCE}$, upper panel) and its log-odds ($\lambda(\widehat{NCE})$, lower panel) for visualization.

Conventions follow Figure 3. Similar to the results of the certainty equivalent task, gambles with a higher potential gain were associated with higher sensitivity and lower elevation in subjective probability (i.e., the two sub-effects of the peanuts effect). (**C**) Model comparison results. Conventions follow Figure 5. Again, the BLO+RI+SP model outperformed all other models in goodness-of-fit.

Similar to Study 1, the BLO+RI+SP model was the best model among the 10 resource-rational models in fitting participants' binary choices (Figure 7B). Again, the BLO base (coding scheme) outperformed the other two, and the models with the RI or SP hypothesis outperformed those without them, other things being the same. The structural prior parameter θ in the winning model (BLO+RI+SP) was again significantly greater than 0 (one-sample *t*-test, one-sided: $t(19) = 3.597$, $p < 0.001$).

Study 3: Using ARRM to Measure Probability Distortions Under Gain and Loss

We have shown that the peanuts effect in decision under risk can be fully explained by resource-rational models developed in the ARRM framework, as the consequence of allocating more resources to higher-valued gambles (the RI hypothesis) and compensating for the covariation of value and probability in life experience (the SP hypothesis). But the datasets we analyzed in Studies 1 and 2 were limited to the gain domain (i.e., participants chose between risky and sure rewards). In Study 3, we extend our findings to the loss domain, reanalyzing the FD05 dataset from Fehr-Duda et al. (2010), where 153 participants were tested on both gain gambles and loss gambles. The dataset was also among the few decision studies that involved "real money" for participants instead of imaginary or small monetary payoffs.

As we introduced earlier, Fehr-Duda et al. (2010) were the first to propose that the peanuts effect reflects the effect of outcome value on probability distortion; they further measured the

change of the probability distortion curve under different levels of values, separately for the gain and loss domains. In our reanalysis, we fit their dataset using the winning ARRM model, the BLO model enhanced with the RI and SP hypotheses. On one hand, we could see how well the model can explain probability distortion patterns in the loss domain. On the other hand, the model parameters would allow us to understand why probability distortion may differ between gain and loss, that is, whether the allocation of cognitive resources or compensation for prior beliefs may differ between the two domains.

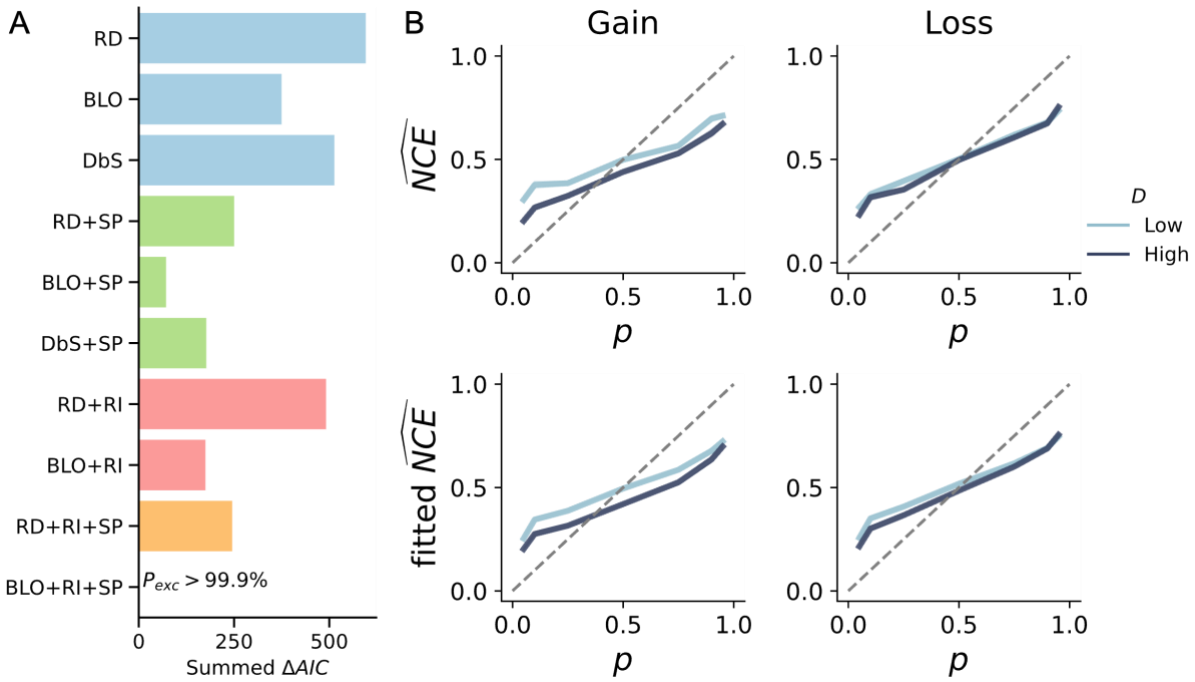


Figure 8. Modeling results for the FD05 Dataset (n=153). (**A**) Model comparison results for the FD05 dataset (153 participants, each of whom completed 56 trials). Different utility functions were used for gain and loss. Smaller ΔAIC indicates better fit. (B) Participants' probability distortion curves (upper row) versus model-predicted distortion curves (lower row). The left and right columns are respectively for the gain and loss domains. Similar to Figure 6, participants' normalized certainty equivalent ($\widehat{NCE}$) is plotted against objective probability (p). For visualization purposes, the gambles in the FD05 dataset were divided into two categories

according to whether the potential gain $D = |x_1 - x_2|$ of the gamble $(x_1, p; x_2)$ was greater than 65 Yuan (darker curve) or not (lighter curve).

Similar to Studies 1 and 2, we visualized probability distortions separately for gambles with different outcomes (Figure 8B) and ran an LMM analysis to verify the possible main and interaction effects of potential gain (D), probability, and outcome domain (gain or loss) on participants' $\widehat{NCE}$. As to the elevation effect, participants' subjective estimates of the probabilities decreased with the potential gain D (*b* = −0.141, 95% CI [−0.165, −0.117], *P* < 0.001), with an attenuated decreasing trend in the loss domain (D by IsLoss interaction, *b* = 0.121, 95% CI [0.090, 0.151], *P* < 0.001). Though participants' sensitivity to probability did not change with D (D by probability interaction, *b* = −0.011, 95% CI [−0.028, 0.007], *P* = 0.239), a three-way interaction with the outcome domain (D by probability by IsLoss interaction, *b* = 0.048, 95% CI [0.023, 0.073], *P* < 0.001) indicates that the sensitivity effect might exist in the loss domain. The lack of sensitivity effect in the gain domain might reflect a lack of statistical power, because a significant sensitivity effect was identified in our following modeling analysis.

Again, the best-fitting model to participants' choices was the BLO+RI+SP model (Figure 8A). To obtain reliable parameter estimates for further analysis, we also performed a hierarchical Bayesian fit for the BLO+RI+SP model. The fitted model can well capture participants' behaviors, including the peanuts effect, in both the gain and loss domains (Figure 8B). Consist with our findings in Studies 1 and 2, participants tended to invest more cognitive resources to gambles involving larger stakes, with the group-level rational-inattention parameters (Eq.7) in both the gain (μ_{β_G}) and loss (μ_{β_L}) domains to be greater than 0 (95% HDI > 0; see Figure 9A, left panel). The group-level structural-prior parameters (Eq.9) μ_{θ_G} and μ_{θ_L} were also both greater than 0 (95% HDI > 0; see Figure 9B, left panel), as if participants held a prior belief that higher risk shall accompany higher gain or loss.

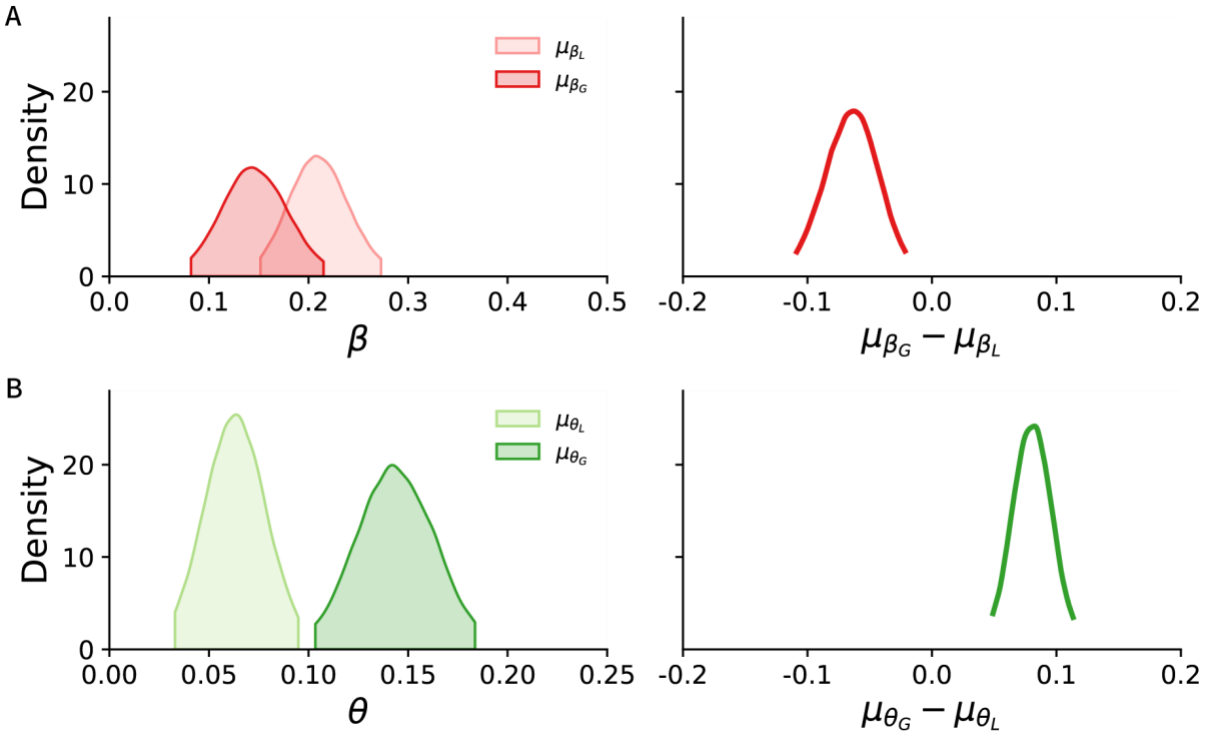


Figure 9. Gain versus loss: group-level rational inattention and structural prior parameters in FD05. The posterior distribution of each parameter (with 95% HDI) was estimated from hierarchical Bayesian modeling for the group of participants, separately for the gain and loss domains. (**A**) The posterior distribution of the rational inattention parameters μ_{β_G}, μ_{β_L} and their difference. Participants' resource allocation was less sensitive to the change of outcome value for gain than for loss. (**B**) The posterior distribution of the structural prior parameters μ_{θ_G}, μ_{θ_L} and their difference. Participants' adjustment to the structural prior was more sensitive to the change of outcome value for gain than for loss.

The sizes of these effects were different for gain and loss gambles. Participants increased their resource investment less quickly for one unit of potential gain than for one unit of potential loss (the 95% HDI of $\mu_{\beta_G} - \mu_{\beta_L} < 0$, see Figure 9A, right panel). The extent to which participants compensated for a structural prior also differed between gain and loss, with a stronger negative

relationship assumed for probability and gain than for probability and loss (the 95% HDI of $\mu_{\theta_G} - \mu_{\theta_L} > 0$, see Figure 9B, right panel).

Discussion

Herbert Simon had a famous metaphor that the structure of the environment and the computational capabilities of the mind are the two blades of the scissors, together shaping human rational behavior (Simon, 1990). This idea of bounded rationality has motivated numerous studies on human heuristics and biases (Anderson, 1990; Lieder & Griffiths, 2020), but only recently been applied to modeling classic decision-making tasks such as decision under risk (Zhang et al., 2020) or intertemporal choice (Gershman & Bhui, 2020), tasks not traditionally perceived to be cognitively demanding.

Extending this approach, we developed ARRM as a computational framework to model rational agents who have limited cognitive resources to process the stimuli used for decision making but who use these resources appropriately to maximize task performance. In one new and three published datasets, we further demonstrated that ARRM can provide a full and normative explanation for human behavioral patterns in decision under risk, including a decision anomaly (the peanuts effect) that was difficult to be explained by previous decision theories. The seemingly inconsistency in human risk attitude turns out to reflect the allocation of more resources to more important probabilities and the compensation for the believed covariation of value and probability in the environment. The ARRM models can also serve as useful computational measures, providing insights into the differences in risk attitude between the gain and loss domains. The resource allocation is more sensitive to the change in loss than in gain, while the believed value-probability covariation is stronger in gain than in loss.

Below we discuss our work and its implications in the context of a broader range of decision theories and problems.

Relationship of attention-dependent decision models to resource rationality

Some decision theories model human decision making as an attention-modulated valuation process, with the amount of attention allocated to a specific outcome determining its decision

weight. For example, the transfer of attention exchange model assumes that the lower-valued outcomes in an option receive a larger proportion of attention than the higher-valued outcomes in the same option, when their probabilities are identical (Birnbaum, 2005). In contrast, the salience theory assumes that the same attributes of different options compete for attention in the decision process, during which higher-valued outcomes receive more attention due to their saliency and are thus endowed with higher decision weights (Bordalo et al., 2012; Mormann, 2016). Similarly, the parallel constraint satisfaction theory proposes that gambles with higher value and winning probability would lead to higher neural activation (Glöckner & Herbold, 2011). Further examples may include the decision field theory, which assumes decisions based on a dynamic sampling process, during which attention shifts among outcomes according to their associated probabilities (Busemeyer & Townsend, 1993; Johnson & Busemeyer, 2005).

Although attention is often viewed as a limited cognitive resource, these attention-dependent decision models are different from resource-rational models in two important ways. First, they often do not have normative motivations for why attention should be allocated in a particular way. As shown in the examples above, the attention allocation strategies assumed in different models often conflict with each other (Fiedler & Glöckner, 2012; Glöckner & Herbold, 2011; Johnson & Busemeyer, 2005). Second and more important, attention-dependent models often assume that attention directly influences decision weight, instead of influencing the precision of representation or information capacity as in resource-rational models. With more attention for larger gambles, attention-dependent models would not explain the peanuts effect but rather counterfactually predict that people would be more risk-seeking for larger gambles.

Empirical findings are also more consistent with the assumption of our models that attention may influence the representational precisions of probability and value. On the individual level (Pachur et al., 2018), individuals who devote a higher proportion of attention to the probabilities of gambles have a steeper slope of probability distortion, implying more precise representation of probability, whereas those who allocate higher attention to the outcomes of gambles have

higher sensitivity to the differences in outcomes. On the option level (Zilker & Pachur, 2021), people have a shallower slope of distortion for the probability in the risky option when they spend a larger proportion of time fixating the safe option.

Relationship of heuristic-based decision models to resource rationality

Heuristic-based models, which assume that human decision behaviors follow simple decision rules instead of maximizing expected utility, have been used to explain human decision biases that are unexplainable to CPT. Heuristic-based decisions often ignore part of the available information (Pachur et al., 2017). This can be viewed as a scenario where a resource-limited rational agent allocates resources exclusively to dimensions perceived as most crucial. For example, the similarity heuristic (Rubinstein, 1988) ignores the similar dimension of gamble options and focuses on their dissimilar dimension, which can be viewed as the extreme case of a rational cross-dimension resource allocation that invests more cognitive resources to the dimension with higher variance (Bates et al., 2019; Dewan, 2020). The more extreme maximax heuristic (Coombs et al., 1970) simply ignores the probability dimension and focuses on the option with the highest value, which yields suboptimal choices in general but might be rational when the contribution of probability to expected reward is small or the cognitive resource is extremely limited. Indeed, these heuristic-based decision algorithms are claimed to satisfy *ecological rationality* (Hertwig et al., 2022; Mata et al., 2012), according to which the heuristics are crystallized adaptive solutions of the brain to specific decision problems in the environment.

But unlike resource-rational solutions, these heuristic algorithms *per se* cannot vary with decision problems. In other words, each of them may approximate a resource-rational agent under specific resource constraints and task goals, but is not equivalent to the latter in general. Indeed, compared with heuristic-based decision models, resource-rational models can better explain human behaviors across different task environments (Callaway et al., 2022).

What were treated as heuristics in previous studies may enter the ARRM framework as

prior beliefs. When people are found to judge the unknown probability of winning a larger reward to be smaller (Leuker et al., 2018; Pleskac & Hertwig, 2014) and respond faster in environments with negative correlations between outcome value and probability (Leuker et al., 2019), these behaviors were explained as a use of heuristics—higher risk often accompanies higher return. In our models, the same structural relationship is defined as a two-dimensional joint prior that contributes to a more precise estimation of probability. The prior, according to our taxonomy, belongs to a long-term prior that is likely learned through one's life experience.

In our work, we did not directly compare the ARRM models with any specific heuristic-based models, but Zhang et al. (2020) have shown that the BLO model outperforms a representative set of heuristic-based models in fitting two of the datasets we reanalyzed here. Our best model (BLO+RI+SP), in turn, performed much better than the BLO model in fitting these datasets, and thus would outperform the set of heuristic-based models by transitivity.

Biological plausibility of resource rationality

It is unclear whether or how an exact resource-rational agent may be implemented in the nervous system. As we argued earlier, approximate (or "bounded") solutions to the resource allocation problems are more feasible than exact ones. For example, consistent with the hypothesis of rational inattention, larger incentives lead to higher arousal of the nervous system, which further modulates the encoding precision (Grujic et al., 2022; Mikhael et al., 2021).

Approximate resource-rational solutions may also be implemented by divisive normalization, a canonical neural computation originally found in the perceptual system (Carandini & Heeger, 2012) and later extended to higher-level neural circuits (e.g., valuation) as well (Louie et al., 2015). According to divisive normalization, the firing rate of a neuron to a specific stimulus is normalized by a population of neurons responding to similar stimuli, which allows neuron responses to adapt to the environmental statistics and achieve efficient coding (Simoncelli & Olshausen, 2001; Valerio & Navarro, 2003). In decision making, divisive

normalization has been used to account for how the valuation of an option may change with the simultaneously (Louie et al., 2013) or recently (Khaw et al., 2017) presented options. Given that higher firing rate implies higher signal-to-noise ratio and thus higher representational precision, divisive normalization may be understood as allocation of cognitive resources (a total number of firing rates) among different options.

Other decision problems where ARRM may apply

When introducing the ARRM framework, we have summarized three different resource allocation problems and two different types of prior beliefs. These different modules, when assembled as needed, would allow us to provide normative explanations for a wider range of decision phenomena than previous resource-rational decision models did. Our modeling of the peanuts effect in decision under risk is just an example. Below we briefly discuss a few decision problems where ARRM may apply.

One potential application concerns the decoy effects, that is, the probability of choosing one option against another may be dramatically changed by the presence of a third option (decoy) even when the decoy itself is rarely chosen (Louie et al., 2015). A memorizable example from Dan Ariely's book "Predictably Irrational" (Ariely, 2010) is as follows. The reader of the journal *The Economist* was offered three subscription options: inexpensive e-copy, expensive hard-copy, or a bundle of hard-copy and e-copy of the same price as the hard-copy alone. Most people would choose the last option, though they would probably find the first option more attractive if the middle option did not exist. The decoy effects have diverse variants, depending on how the attributes of the decoy are related to those of the major options. Louie and colleagues (Louie et al., 2013, 2015) proposed that the decoy effects reflect context-dependent encoding of option values. From the perspective of resource rationality, the decoy effects may arise from the agent's solutions to two resource allocation problems: how to allocate resources across options and across different attribute dimensions. These two problems can be

interrelated; just as in our modeling of the peanuts effect, the value of one dimension (outcome) may influence the resource allocation on the other dimension (probability).

The ARRM framework may also be applied to decision problems that involve both delay and risk, such as choosing among delayed gambles (e.g., 10% probability to receive $1M after 10 months). As we reviewed earlier, the magnitude effects in temporal discounting and probability distortion have opposite directions (Estle et al., 2006; Prelec & Loewenstein, 1991). Both now have normative explanations, with the former modeled in Gershman and Bhui's (2020) resource-rational model and the latter (i.e., the peanuts effect) modeled by us in the present article. However, when delay and risk are both involved, there turn out to be more sophisticated interactive patterns (Luckman et al., 2020; Vanderveldt et al., 2015, 2017), whose explanation is still elusive. The difficulty of modeling the effect, similar to the decoy effect discussed above, lies in the necessity to model the interrelated resource allocations across options and across attribute dimensions. ARRM can serve as a framework for developing such resource-rational models.

Limitations and Future Directions

Similar to existing resource-rational studies, we have only tested our models on tasks that are simple enough to be precisely predicted—choosing between a two-outcome gamble and a (or a series of) sure payoff(s). For gambles with more than two outcomes or tasks with multiple gambles, the resource-rational models are still to be worked out, due to the increased complexity of the resource allocation problems in more complex scenarios.

As a second limitation, we have only considered the influence of outcome value on the probability distortion process while keeping the utility function fixed, which follows the theoretical convention in the literature but is not necessarily the only possibility. Similar ARRM models may be constructed for the processing of value, with the value-probability interaction being the other way around: the decreased risk preference for larger outcome may be interpreted as increased

sensitivity of value (i.e., greater exponent for the utility function) for larger probability. In which direction the interaction actually occurs will be a question for future studies.

We assume people have limited cognitive resources and allocate their resources rationally for decision making. Though the behavior of such a resource-rational agent is derivable as the solution to an optimization problem, how to define the optimization problem is an empirical question. What are the cognitive resources, what are people's prior beliefs for the environment, and how flexible the encoding and decoding schemes can be are all fundamental empirical questions that are still largely unknown. The ARRM framework will enable us to develop new resource-rational models as well as to test them along with existing models to explore alternative theoretical possibilities. Further behavioral and neurophysiological measures to test these theoretical possibilities are also what the field calls for.

METHODS

Ethics statement

The experiment for the new dataset (Study 2) had been approved by the Institutional Review Board of School of Psychological and Cognitive Sciences at Peking University. All participants provided written informed consent and were paid for their time.

Methods of Study 1

Experimental design of the datasets

Gonzalez & Wu's dataset (GW99). The GW99 dataset (Gonzalez & Wu, 1999) includes the certainty equivalents (CEs) for 165 different virtual gambles from each of 10 US participants. In their experiment, each gamble was in the form of $(x_1, p; x_2)$, that is, a probability of p to win $\$x_1$ and otherwise $\$x_2$, where $x_1 > x_2 \geq 0$. There were 15 pairs of (x_1, x_2): (25, 0), (50, 0), (75, 0), (100, 0), (150,0), (200, 0), (400, 0), (800, 0), (50, 25), (75, 50), (100, 50), (150, 50), (150, 100), (200, 100) and (200, 150). These outcome pairs were combined with 11 probabilities (0.01, 0.05, 0.1, 0.25, 0.4, 0.5, 0.6, 0.75, 0.9, 0.95 and 0.99) to generate $15 \times 11 = 165$ gambles. For each gamble, the participant chose between the gamble and 12 different sure rewards (presented as two successive tables) in an adaptive two-stage procedure (see Gonzalez & Wu, 1999 for details), based on which the CE was estimated.

Zhang, Ren, & Maloney's dataset (ZRM20: JDA and JDB). This decision dataset (Zhang et al., 2020) contains the CEs for 165 different virtual gambles from each of 75 Chinese participants from two sub-experiments (JDA: 51 participants; JDB: 24 participants). The JDA and JDB had almost identical procedures and designs as GW99, except that the outcomes were in Chinese Yuan instead of US dollars and participants were also tested on a separate task of relative frequency judgment (not analyzed here) in the experiment. The two sub-experiments differed from each other mainly in whether the trials of the decision and judgment tasks were

interleaved (JDA) or in different blocks (JDB). Our purpose of presenting the two sub-experiments separately was to demonstrate the robustness of the results.

Statistical Analysis

For each participant and gamble, we calculated the normalized certainty equivalent ($\widehat{NCE}$) based on the measured CE, and then obtained its log-odds $\lambda(\widehat{NCE})$. The $\widehat{NCE}$ and $\lambda(\widehat{NCE})$ were used as nonparametric measures of the subjective probability. For each dataset, we performed the following linear mixed-effects model (LMM) analysis on $\lambda(\widehat{NCE})$ to test the potential influence of outcome values on probability distortion (as suggested by the peanuts effect):

$$\lambda(\widehat{NCE}) \sim 1 + \lambda(p) * D + (1 + \lambda(p) * D | \text{paticipant}) \quad (11),$$

where $\lambda(p)$ denotes the log-odds of the probability p in gamble $(x_1, p; x_2)$, $D = |\mathrm{x}_1 - \mathrm{x}_2|$ denotes the potential gain, 1 denotes the intercept term, $\lambda(p) * D$ denotes the main effects of $\lambda(p)$ and D as well as their interaction, and $(\cdot\,|\text{paticipant})$ denotes the random-effects terms. In the case the full model could not converge (which rarely occurred), we simplified the random effects structure following statistical guidelines (Barr et al., 2013; D. Bates et al., 2018; Brauer & Curtin, 2018). Data were normalized on the group level before entering LMMs. All the LMMs were implemented in python 3.8 with the statsmodels (v0.13.1) package and then replicated in R 4.0.3.

Model Fitting and Comparison

We fit the 10 models (described earlier) to the $\widehat{NCE}$ of each participant using maximum likelihood estimates. The Akaike Information Criterion (AIC; Akaike, 1974) was used to compare the goodness-of-fit among different models. We used the AIC of model BLO+RI+SP as a reference and calculated each model's ΔAIC as its difference from the reference. Smaller ΔAIC indicates better fit. For each dataset or sub-dataset, we also performed the group-level Bayesian model selection (Rigoux et al., 2014; Stephan et al., 2009) and calculated the protected exceedance probability for each model.

The exchangeability test

The GW99 and ZRM20 datasets measured the CE for each gamble, this enabled us to conduct the model free exchangeability test. The measured CE curve for each gamble value x_1 only has 11 probabilities, thus we first conduct a linear interpolation to approximate the full CE cure in probability range 0.01~0.99. Given these CE curves, for any $x_a > x_b > x_c$, we could find probabilities p_b and p_c so that the CE of gamble $(x_a, p_b; 0)$ is x_b and the CE of gamble $(x_a, p_c; 0)$ is x_c. By applying Eq. 10, we have

$$U(x_b, p_c) = w(p_c)u(x_b) = w(p_c)w(p_b)u(x_a) = U(x_c, p_b) \quad (12),$$

thus the gambles $(x_c, p_b; 0)$ and $(x_b, p_c; 0)$ should have the same CE.

For the GW99 and ZRM20 datasets, we draw 56 combinations from gamble values (x_1, x_2): (25, 0), (50, 0), (75, 0), (100, 0), (150,0), (200, 0), (400, 0), (800, 0). The value combinations $[x_a, x_b, x_c]$ were then passed into the interpolated CE curve of x_a to get the probability p_b and p_c. Combinations with p_b and p_c outside the probability range 0.01~0.99 (16 combinations) were excluded from analysis. The probabilities p_b and p_c were further passed into the CE curve of x_c and x_b respectively to get the measured CE_{x_c,p_b} and CE_{x_b,p_c}. The measure CEs were scaled by x_a and then passed into paired t-test.

Methods of Study 2

Participants

Twenty-one students (15 females, mean age = 20.6, SD = 2.2) in Peking University participated in our study. The experiment took approximately 1 hour and participants received 60 Chinese Yuan for their time. The study had been approved by the Committee for Protecting Human and Animal Subjects of School of Psychological and Cognitive Sciences at Peking University and all participants provided informed consent before the experiment.

Design and Procedure

On each trial, a two-outcome gamble, denoted $(x_1, p; x_2)$, and a sure reward were presented side by side on the screen. Participants need to choose between the two by pressing the “F” (left) or “J” (right) on the keyboard. The value of x_1 had 8 levels: 25, 50, 75, 100, 150, 200, 400 and 800 (in Chinese Yuan). The value of x_2 was always 0. The winning probability p had 5 levels: 0.05, 0.25, 0.5, 0.75 and 0.95. A combination of x_1 with p led to 40 different gambles in total.

The sure reward paired with a specific gamble was selected according to the following two-round adaptive procedure. In the first round of 200 trials, the sure rewards for each gamble were set to 5 different levels, which were 0.1, 0.3, 0.5, 0.7 or 0.9 times of the x_1 in the gamble. After the first round, for each gamble, the participant’s 5 choices were entered into a logistic regression against the sure-reward-to-x_1 ratio to estimate a location parameter where the participant would be indifferent between the gamble and the sure reward (i.e., with a probability of 0.5 to choose sure reward). The location parameter was rounded to one decimal and added by −0.08, −0.04, 0, 0.4 or 0.08. The resulting ratio was then used to generate the 5 sure rewards for the gamble in the second round of 200 trials. In both rounds, trials for different gambles and sure rewards were interleaved in random order. For each participant, this two-round procedure was repeated twice, thus leading to 800 trials in total and 20 trials for each

gamble.

Statistical Analysis

To make the results of this binary choice task comparable to those of the certainty equivalent task presented earlier, we estimated the certainty equivalent (CE) for each gamble using a logistic regression of choices against the sure rewards, based on all the 20 trials for the gamble. This estimated NCE ($\lambda(\widehat{NCE})$) was used for visualization and further statistical analysis—an LMM similar to that used for the certainty equivalent task (Eq. 11).

Model Fitting and Comparison

Instead of fitting the estimated NCE, we fit the 10 models directly to participants' binary choices (see the Decision Model section for details) using maximum likelihood estimates. As before, ΔAIC and protected exceedance probability were calculated for model comparison in goodness-of-fit.

The experiment design and statistical analyses were preregistered at OSF (https://osf.io/p76rg) for testing the peanuts effect. The modeling analysis (with the 10 models) were not preregistered[3].

Study 3

Fehr-Duda et al.'s dataset (FD05)

The experiment was performed in Beijing in 2005 on 153 Chinese participants, for each of whom the certainty equivalents for 56 different gambles (28 gain gambles and 28 loss gambles) were measured (Fehr-Duda et al., 2010). For each gamble, participants were asked to mark out

[3] When we preregistered the experiment, we had not known about the work by Fehr-Duda et al. (2010), who reasoned that the peanuts effect should be understood as the influence of outcome value on probability distortion. Instead, we had framed the peanuts effect as the influence of winning probability on the utility function, an alternative mathematical form to describe the peanuts effect (see Supplemental Text S2 and Figures S1 and S2). We later switched to the view of Fehr-Duda et al. (2010) to make a better connection to the decision literature.

their choices on a paper sheet where on the left side was the gamble and on the right side were 20 equally spaced sure payoffs. All gambles in FD05 had two outcomes, denoted $(x_1, p; x_2)$ as before, where $x_1 > x_2 \geq 0$ (gain gamble) or $x_1 < x_2 \leq 0$ (loss gamble). The probability p had 7 levels: 0.05, 0.10, 0.25, 0.50, 0.75, 0.9, or 0.95. Half of the 28 gain gambles offered outcomes ranging from 4 to 55 Yuan (low-stake) and the other half offered 65 to 950 Yuan (high-stake). The 28 loss gambles had the same settings as the gain gambles except for reversing the sign of the outcomes. The highest gain or loss in the gambles amounted to several times of participants' average monthly income at that time. Participants received monetary payoffs based on a random draw from their choices.

Statistical Analysis

Unlike the original paper (Fehr-Duda et al., 2010) that tested the effect of outcome value by dividing gambles into two categories according to the value of x_2, we computed the potential gain ($D = |\mathrm{x}_1 - \mathrm{x}_2|$) for each gamble and used it as a continuous regressor. We used the following linear mixed-effects model (LMM) to examine the effects of outcome value and domain on the estimated NCE ($\lambda(\widehat{NCE})$, calculated as before):

$$\lambda(\widehat{NCE}) \sim 1 + \lambda(p) * D * \mathrm{IsLoss} + (1 + \lambda(p) * D * \mathrm{IsLoss} | \mathrm{Participant}), \quad (13)$$

where (as before) $\lambda(p)$ denotes the log-odds of the probability p in gamble $(x_1, p; x_2)$, IsLoss denotes a dummy variable for outcome domain (0 for gain, 1 for loss), 1 denotes the intercept term, $\lambda(p) * D * \mathrm{IsLoss}$ denotes the main effects of $\lambda(p)$, D and IsLoss as well as their interactions, and $(\cdot\,|\mathrm{paticipant})$ denotes the random-effects terms.

Hierarchical Bayesian Modeling

Because the gamble set for each participant and domain (28 gambles) was much smaller than that of Study 1 (165 gambles) or Study 2 (40 gambles), we fit the BLO+RI+SP model to participants' NCEs using the hierarchical Bayesian method (Lee & Wagenmakers, 2013; Martin & Wiecki, 2018) instead of maximum likelihood estimates to obtain more reliable parameter

estimates.

To extend our model to loss domain, we assumed that RI parameter β (i.e., the sensitivity of resource allocation to potential gain) and SP parameter θ (i.e., the sensitivity of compensating for the covariation between value and probability) could be different for gain and loss. We assumed each participant had two RI parameters (β_G and β_L) and SP parameters (θ_G and θ_L), with the subscripts G and L respectively denoting gain and loss. Following CPT, we assumed two different exponents α_G and α_L respectively for the power utility functions (Eq. S15) in gain and loss.

The model parameters were estimated using the Markov Chain Monte Carlo method (See Supplemental Text S3 and Table S3 for details), implemented by PyMC package on python 3.8. Four independent chains were run, each had 10000 samples after an initial burn-in of 5000 samples. The model had a good convergence with $\hat{R} \leq 1.01$ for all estimated parameters. The 95% highest density interval (HDI) was calculated for the group-level RI and SP parameters as well as their differences between the gain and loss domains.

Acknowledgments

HZ was supported by National Natural Science Foundation of China grants 32171095 and funding from Peking-Tsinghua Center for Life Sciences. Part of the analysis was performed on the High-performance Computing Platform of Peking University and the Computing Platform of the Center for Life Sciences at Peking University.

Data and Code Availability Statement

The new data and codes that support the findings of this study will be available in the Open Science Framework (OSF) repository upon publication of this article.

References

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

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

Anderson, J. R. (1990). *The adaptive character of thought*. L. Erlbaum Associates.

Ariely, D. (2010). *Predictably irrational: The hidden forces that shape our decisions* (rev. and expanded ed., 1. Harper Perennial ed). Harper Perennial.

Attneave, F. (1953). Psychological probability as a function of experienced frequency. *Journal of Experimental Psychology*, *46*(2), 81–86. https://doi.org/10.1037/h0057955

Barlow, H. B. (1961). Possible Principles Underlying the Transformations of Sensory Messages. In W. A. Rosenblith (Ed.), *Sensory Communication* (pp. 216–234). The MIT Press. https://doi.org/10.7551/mitpress/9780262518420.003.0013

Bates, C. J., & Jacobs, R. A. (2021). Optimal attentional allocation in the presence of capacity constraints in uncued and cued visual search. *Journal of Vision*, *21*(5), 3. https://doi.org/10.1167/jov.21.5.3

Bates, C. J., Lerch, R. A., Sims, C. R., & Jacobs, R. A. (2019). Adaptive allocation of human visual working memory capacity during statistical and categorical learning. *Journal of Vision*, *19*(2), 11. https://doi.org/10.1167/19.2.11

Berger, T., & Gibson, J. D. (1998). Lossy Source Coding. *IEEE Trans. Inf. Theory*, *44*, 2693–2723. https://doi.org/10.1109/18.720552

Bhui, R., & Gershman, S. J. (2018). Decision by sampling implements efficient coding of psychoeconomic functions. *Psychological Review*, *125*(6), 985–1001. https://doi.org/10.1037/rev0000123

Bhui, R., Lai, L., & Gershman, S. J. (2021). Resource-rational decision making. *Current Opinion in Behavioral Sciences*, *41*, 15–21. https://doi.org/10.1016/j.cobeha.2021.02.015

Binswanger, H. P. (1981). Attitudes Toward Risk: Theoretical Implications of an Experiment in Rural India. *The Economic Journal*, *91*(364), 867. https://doi.org/10.2307/2232497

Binz, M., Gershman, S. J., Schulz, E., & Endres, D. (2022). Heuristics from bounded meta-learned inference. *Psychological Review*. https://doi.org/10.1037/rev0000330

Birnbaum, M. H. (2005). Three New Tests of Independence That Differentiate Models of Risky Decision Making. *Management Science*, *51*(9), 1346–1358. https://doi.org/10.1287/mnsc.1050.0404

Birnbaum, M. H. (2008). Evaluation of the priority heuristic as a descriptive model of risky decision making: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review*, *115*(1), 253–260. https://doi.org/10.1037/0033-295X.115.1.253

Birnbaum, M. H., & LaCroix, A. R. (2008). Dimension integration: Testing models without trade-offs. *Organizational Behavior and Human Decision Processes*, *105*(1), 122–133. https://doi.org/10.1016/j.obhdp.2007.07.002

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience Theory of Choice Under Risk. *The Quarterly Journal of Economics*, *127*(3), 1243–1285. https://doi.org/10/gf7965

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

Bujold, P. M., Ferrari-Toniolo, S., & Schultz, W. (2021). Adaptation of utility functions to reward distribution in rhesus monkeys. *Cognition*, *214*, 104764. https://doi.org/10.1016/j.cognition.2021.104764

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432–459. https://doi.org/10.1037/0033-295X.100.3.432

Callaway, F., van Opheusden, B., Gul, S., Das, P., Krueger, P. M., Griffiths, T. L., & Lieder, F. (2022). Rational use of cognitive resources in human planning. *Nature Human Behaviour*, *6*(8), 1112–1125. https://doi.org/10.1038/s41562-022-01332-8

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*(1), 51–62. https://doi.org/10.1038/nrn3136

Coombs, C. H., Dawes, R. M., & Tversky, A. (1970). *Mathematical psychology: An elementary introduction*. Prentice-Hall.

Dasgupta, I., Schulz, E., Tenenbaum, J. B., & Gershman, S. J. (2020). A theory of learning to infer. *Psychological Review*, *127*(3), 412–441. https://doi.org/10.1037/rev0000178

Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. Massachusetts Institute of Technology Press.

Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or Linear? Distinct Intuitions of the Number Scale in Western and Amazonian Indigene Cultures. *Science*, *320*(5880), 1217–1220. https://doi.org/10.1126/science.1156540

Dewan, A. (2020). *Costly Multidimensional Information* (p. 55).

Estle, S. J., Green, L., Myerson, J., & Holt, D. D. (2006). Differential effects of amount on temporal and probability discounting of gains and losses. *Memory & Cognition*, *34*(4), 914–928. https://doi.org/10/bbkw7x

Fehr-Duda, H., Bruhin, A., Epper, T., & Schubert, R. (2010). Rationality on the rise: Why relative risk aversion increases with stake size. *Journal of Risk and Uncertainty*, *40*(2), 147–180. https://doi.org/10/dgmr2s

Fiedler, S., & Glöckner, A. (2012). The Dynamics of Decision Making in Risky Choice: An Eye-Tracking Analysis. *Frontiers in Psychology*, *3*. https://doi.org/10/gf5sv5

Frydman, C., & Jin, L. J. (2021). Efficient Coding and Risky Choice. *The Quarterly Journal of Economics*, *137*(1), 161–213. https://doi.org/10.1093/qje/qjab031

Gershman, S. J., & Bhui, R. (2020). Rationally inattentive intertemporal choice. *Nature Communications*, *11*(1), 3365. https://doi.org/10.1038/s41467-020-16852-y

Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The foundations of adaptive behavior*. Oxford University Press.

Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: Visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience*, *14*(7), 926–932. https://doi.org/10.1038/nn.2831

Glimcher, P. W., & Fehr, E. (Eds.). (2014). *Neuroeconomics: Decision making and the brain* (Second edition). Elsevier/AP, Academic Press is an imprint of Elsevier.

Glöckner, A., & Herbold, A.-K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*, *24*(1), 71–98. https://doi.org/10.1002/bdm.684

Gonzalez, R., & Wu, G. (1999). On the Shape of the Probability Weighting Function. *Cognitive Psychology*, *38*(1), 129–166. https://doi.org/10.1006/cogp.1998.0710

Green, L., Myerson, J., & Ostaszewski, P. (1999). Amount of reward has opposite effects on the discounting of delayed and probabilistic outcomes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 418–427. https://doi.org/10/ckgjz8

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal Predictions in Everyday Cognition. *Psychological Science*, *17*(9), 767–773. https://doi.org/10.1111/j.1467-9280.2006.01780.x

Grujic, N., Brus, J., Burdakov, D., & Polania, R. (2022). Rational inattention in mice. *Science Advances*, *8*(9), eabj8935. https://doi.org/10.1126/sciadv.abj8935

Heng, J. A., Woodford, M., & Polania, R. (2020). Efficient sampling and noisy decisions. *ELife*, *9*, e54962. https://doi.org/10.7554/eLife.54962

Hertwig, R., Leuker, C., Pachur, T., Spiliopoulos, L., & Pleskac, T. J. (2022). Studies in

Ecological Rationality. *Topics in Cognitive Science*, *14*(3), 467–491. https://doi.org/10.1111/tops.12567

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, *13*(8), 1020–1026. https://doi.org/10.1038/nn.2590

Johnson, J. G., & Busemeyer, J. R. (2005). A Dynamic, Stochastic, Computational Model of Preference Reversal Phenomena. *Psychological Review*, *112*(4), 841–861. https://doi.org/10.1037/0033-295X.112.4.841

Jones, S., & Oaksford, M. (2011). Transactional problem content in cost discounting: Parallel effects for probability and delay. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 739–747. https://doi.org/10.1037/a0022219

Kachelmeier, S. J., & Shehata, M. (1992). Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence from the People's Republic of China. *The American Economic Review*, *82*(5), 1120–1141. JSTOR.

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

Khaw, M. W., Glimcher, P. W., & Louie, K. (2017). Normalized value coding explains dynamic adaptation in the human valuation process. *Proceedings of the National Academy of Sciences*, *114*(48), 12696–12701. https://doi.org/10/gcn9gg

Khaw, M. W., Li, Z., & Woodford, M. (2021). Cognitive Imprecision and Small-Stakes Risk Aversion. *The Review of Economic Studies*, *88*(4), 1979–2013. https://doi.org/10.1093/restud/rdaa044

Kleinsorge, T., & Rinkenauer, G. (2012). Effects of Monetary Incentives on Task Switching. *Experimental Psychology*, *59*(4), 216–226. https://doi.org/10.1027/1618-3169/a000146

Kool, W., & Botvinick, M. (2018). Mental labour. *Nature Human Behaviour*, *2*(12), 899–908. https://doi.org/10.1038/s41562-018-0401-9

Leland, J. W. (1998). Similarity Judgments in Choice Under Uncertainty: A Reinterpretation of the Predictions of Regret Theory. *Management Science*, *44*(5), 659–672. https://doi.org/10.1287/mnsc.44.5.659

Leuker, C., Pachur, T., Hertwig, R., & Pleskac, T. J. (2018). Exploiting risk–reward structures in decision making under uncertainty. *Cognition*, *175*, 186–200. https://doi.org/10.1016/j.cognition.2018.02.019

Leuker, C., Pachur, T., Hertwig, R., & Pleskac, T. J. (2019). Do people exploit risk–reward structures to simplify information processing in risky choice? *Journal of the Economic Science Association*, *5*(1), 76–94. https://doi.org/10/gmhksb

Lewis, R. L., Howes, A., & Singh, S. (2014). Computational Rationality: Linking Mechanism and Behavior Through Bounded Utility Maximization. *Topics in Cognitive Science*, *6*(2), 279–

311. https://doi.org/10.1111/tops.12086

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, e1. https://doi.org/10.1017/S0140525X1900061X

Louie, K., Glimcher, P. W., & Webb, R. (2015). Adaptive neural coding: From biological to behavioral decision-making. *Current Opinion in Behavioral Sciences*, *5*, 91–99. https://doi.org/10/gdq732

Louie, K., Khaw, M. W., & Glimcher, P. W. (2013). Normalization is a general neural mechanism for context-dependent decision making. *Proceedings of the National Academy of Sciences*, *110*(15), 6139–6144. https://doi.org/10.1073/pnas.1217854110

Luckman, A., Donkin, C., & Newell, B. R. (2020). An evaluation and comparison of models of risky intertemporal choice. *Psychological Review*, *127*(6), 1097–1138. https://doi.org/10/gnbd5x

MacKay, D. J. C. (2003). *Information theory, inference, and learning algorithms*. Cambridge University Press.

Maloney, L. T., & Zhang, H. (2010). Decision-theoretic models of visual perception and action. *Vision Research*, *50*(23), 2362–2374. https://doi.org/10.1016/j.visres.2010.09.031

Markowitz, H. (1952). The Utility of Wealth. *Journal of Political Economy*, *60*(2), 151–158. https://doi.org/10.1086/257177

Mata, R., Pachur, T., von Helversen, B., Hertwig, R., Rieskamp, J., & Schooler, L. (2012). Ecological Rationality: A Framework for Understanding and Aiding the Aging Decision Maker. *Frontiers in Neuroscience*, *6*. https://doi.org/10.3389/fnins.2012.00019

Matějka, F., & McKay, A. (2015). Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model. *American Economic Review*, *105*(1), 272–298. https://doi.org/10.1257/aer.20130047

Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision Affect Theory: Emotional Reactions to the Outcomes of Risky Options. *Psychological Science*, *8*(6), 423–429. https://doi.org/10.1111/j.1467-9280.1997.tb00455.x

Mikhael, J. G., Lai, L., & Gershman, S. J. (2021). Rational inattention and tonic dopamine. *PLOS Computational Biology*, *17*(3), e1008659. https://doi.org/10.1371/journal.pcbi.1008659

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158

Mormann, M. M. (2016). The Role of Salience and Attention in Choice under Risk: An Experimental Investigation. *SSRN Electronic Journal*. https://doi.org/10/gnnz8d

Otto, A. R., & Vassena, E. (2021). It's all relative: Reward-induced cognitive control modulation depends on context. *Journal of Experimental Psychology: General*, *150*(2), 306–313. https://doi.org/10.1037/xge0000842

Pachur, T., Schulte-Mecklenbeck, M., Murphy, R. O., & Hertwig, R. (2018). Prospect theory reflects selective allocation of attention. *Journal of Experimental Psychology: General*, *147*(2), 147–169. https://doi.org/10.1037/xge0000406

Pachur, T., Suter, R. S., & Hertwig, R. (2017). How the twain can meet: Prospect theory and models of heuristics in risky choice. *Cognitive Psychology*, *93*, 44–73. https://doi.org/10.1016/j.cogpsych.2017.01.001

Payzan-LeNestour, E., & Woodford, M. (2022). Outlier blindness: A neurobiological foundation for neglect of financial risk. *Journal of Financial Economics*, *143*(3), 1316–1343. https://doi.org/10.1016/j.jfineco.2021.06.019

Peterson, J. C., Bourgin, D. D., Agrawal, M., Reichman, D., & Griffiths, T. L. (2021). Using large-scale experiments and machine learning to discover theories of human decision-making. *Science*, *372*(6547), 1209–1214. https://doi.org/10.1126/science.abe2629

Petzschner, F. H., Glasauer, S., & Stephan, K. E. (2015). A Bayesian perspective on magnitude estimation. *Trends in Cognitive Sciences*, *19*(5), 285–293. https://doi.org/10.1016/j.tics.2015.03.002

Pleskac, T. J., & Hertwig, R. (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General*, *143*(5), 2000–2019. https://doi.org/10.1037/xge0000013

Polanía, R., Woodford, M., & Ruff, C. C. (2019). Efficient coding of subjective value. *Nature Neuroscience*, *22*(1), 134–142. https://doi.org/10.1038/s41593-018-0292-0

Prelec, D. (1998). The Probability Weighting Function. *Econometrica*, *66*(3), 497. https://doi.org/10.2307/2998573

Prelec, D., & Loewenstein, G. (1991). Decision Making Over Time and Under Uncertainty: A Common Approach. *Management Science*, *37*(7), 770–786. https://doi.org/10.1287/mnsc.37.7.770

Ren, X., Luo, H., & Zhang, H. (2021). Automatic and Fast Encoding of Representational Uncertainty Underlies the Distortion of Relative Frequency. *The Journal of Neuroscience*, *41*(16), 3692–3706. https://doi.org/10.1523/JNEUROSCI.2006-20.2021

Ren, X., Wang, M., & Zhang, H. (2018). Context Effects in the Judgment of Visual Relative-Frequency: Trial-by-Trial Adaptation and Non-linear Sequential Effect. *Frontiers in Psychology*, *9*, 1691. https://doi.org/10.3389/fpsyg.2018.01691

Rubinstein, A. (1988). Similarity and decision-making under risk (is there a utility theory resolution to the Allais paradox?). *Journal of Economic Theory*, *46*(1), 145–153. https://doi.org/10.1016/0022-0531(88)90154-8

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

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural Image Statistics and Neural Representation. *Annual Review of Neuroscience*, *24*(1), 1193–1216. https://doi.org/10.1146/annurev.neuro.24.1.1193

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665–690. https://doi.org/10.1016/S0304-3932(03)00029-1

Sims, C. R. (2016). Rate–distortion theory and human perception. *Cognition*, *152*, 181–198. https://doi.org/10.1016/j.cognition.2016.03.020

Sims, C. R. (2018). Efficient coding explains the universal law of generalization in human perception. *Science*, *360*(6389), 652–656. https://doi.org/10.1126/science.aaq1118

Stewart, N., Chater, N., & Brown, G. D. A. (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

Summerfield, C., & Parpart, P. (2022). Normative Principles for Decision-Making in Natural Environments. *Annual Review of Psychology*, *73*(1), 53–77. https://doi.org/10.1146/annurev-psych-020821-104057

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

Valerio, R., & Navarro, R. (2003). Optimal coding through divisive normalization models of V1 neurons. *Network: Computation in Neural Systems*, *14*(3), 579–593. https://doi.org/10.1088/0954-898X_14_3_310

van den Berg, R., & Ma, W. J. (2018). A resource-rational theory of set size effects in human visual working memory. *ELife*, *7*, e34963. https://doi.org/10.7554/eLife.34963

Vanderveldt, A., Green, L., & Myerson, J. (2015). Discounting of monetary rewards that are both delayed and probabilistic: Delay and probability combine multiplicatively, not additively. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 148–162. https://doi.org/10/gfv884

Vanderveldt, A., Green, L., & Rachlin, H. (2017). Discounting by Probabilistic Waiting: Discounting by Probabilistic Waiting. *Journal of Behavioral Decision Making*, *30*(1), 39–53. https://doi.org/10/f9sjf4

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.

Wakker, P. P. (2010). *Prospect Theory: For Risk and Ambiguity* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511779329

Weber, B. J., & Chapman, G. B. (2005). Playing for peanuts: Why is risk seeking more common for low-stakes gambles? *Organizational Behavior and Human Decision Processes*, *97*(1), 31–46. https://doi.org/10.1016/j.obhdp.2005.03.001

Wei, X.-X., & Stocker, A. A. (2015). A Bayesian observer model constrained by efficient coding can explain “anti-Bayesian” percepts. *Nature Neuroscience*, *18*(10), 1509–1517. https://doi.org/10.1038/nn.4105

Zhang, H., & Maloney, L. T. (2012). Ubiquitous Log Odds: A Common Representation of Probability and Frequency Distortion in Perception, Action, and Cognition. *Frontiers in Neuroscience*, *6*. https://doi.org/10.3389/fnins.2012.00001

Zhang, H., Ren, X., & Maloney, L. T. (2020). The bounded rationality of probability distortion. *Proceedings of the National Academy of Sciences*, *117*(36), 22024–22034. https://doi.org/10.1073/pnas.1922401117

Zhu, J.-Q., Sanborn, A. N., & Chater, N. (2020). The Bayesian sampler: Generic Bayesian inference causes incoherence in human probability judgments. *Psychological Review*, *127*(5), 719–748. https://doi.org/10.1037/rev0000190

Zilker, V., & Pachur, T. (2021). Nonlinear probability weighting can reflect attentional biases in sequential sampling. *Psychological Review*. https://doi.org/10.1037/rev0000304

Supplemental Methods

ARRM Models for Probability Distortion

Probability estimation models

To implement the ARRM framework for probability estimation, we adopt three previous resource-rational decision models as the "base" models and combine them with the RI and SP hypotheses ("modifiers") to produce a set of 10 different models. Below we briefly review each of the base models and how they may be enhanced by the RI and SP hypotheses (see Supplemental Tables S1 and S2 for a summary of model notations).

Decision by Sampling (DbS) Model. As its name suggests, the DbS model (Bhui & Gershman, 2018; Stewart et al., 2006) starts with sampling from memory. It assumes that the agent compares the stimulus with each sample and calculates the proportion of times the former is larger than the latter as the subjective estimation of the stimulus. When sample size is sufficiently large, this subjective estimation approximates the cumulative distribution function of the prior distribution of the samples (Bhui et al., 2021; Bhui & Gershman, 2018; Stewart et al., 2015). Given prior $\lambda(\mathrm{p}) \sim N(\lambda_s, \sigma_s^2)$ (Eq. 8, with λ_g set to a constant λ_s), the subjective estimation of probability p can be written as:

$$\hat{p} = \Phi\left(\frac{\lambda(p) - \lambda_s}{\sigma_s}\right), \tag{S1}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard Gaussian distribution.

There is no obvious way that the DbS model can be enhanced by the RI hypothesis, because its subjective estimation (Eq. S1) is not influenced by the available cognitive resource (in fact, the model assumes a sufficiently large sample size to approximate the cumulative probability function). Therefore, we only consider the SP hypothesis for DbS, which is simply to replace the λ_s in Eq. S1 with the λ_g in Eq. 9. In the special case of $\theta = 0$, it reduces to the original DbS without the SP hypothesis. The original DbS model has two free parameters: λ_s,

and σ_s. The DbS+SP model replaces the parameter λ_s with two parameters: λ_0 and θ. Unless otherwise specified, our count of model parameters does not include the parameters of the decision model (introduced later), which are shared across all 10 models.

Rate-Distortion (RD) Model. The RD model assumes a constraint on the rate of transmitted information and optimizes encoding and decoding schemes under this constraint to minimize the expected loss incurred by distorted subjective estimation of the stimulus. Similar to previous RD models (Bates et al., 2019; Gershman & Bhui, 2020), we assume a Gaussian information source and a squared-error loss function, for which the optimal encoding scheme is a linear mapping with homogeneous Gaussian noise (Berger & Gibson, 1998). We further assume that the encoding of probability is on the log-odd scale (Khaw et al., 2021; Zhang et al., 2020; Zhang & Maloney, 2012), so the representation can be written as

$$m = \lambda(p) + \varepsilon. \tag{S2}$$

Here $\lambda(p)$ is the log-odds of probability p, and $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ is a noise term with mean 0 and variance σ_ε^2.

The optimal decoding scheme is Bayesian decoding. Given the Gaussian prior $\lambda(p) \sim N(\lambda_s, \sigma_s^2)$ and representation m, the posterior estimate of $\lambda(p)$ follows a Gaussian distribution:

$$\lambda(p)|m \sim N\left(\frac{\sigma_s^2 m + \sigma_\varepsilon^2 \lambda_s}{\sigma_s^2 + \sigma_\varepsilon^2}, \frac{\sigma_s^2 \sigma_\varepsilon^2}{\sigma_s^2 + \sigma_\varepsilon^2}\right). \tag{S3}$$

Given m, the optimal estimate for $\lambda(p)$ that minimizes squared error on the log-odds scale is the posterior mean,

$$\widehat{\lambda_m} = \frac{\sigma_s^2 m + \sigma_\varepsilon^2 \lambda_s}{\sigma_s^2 + \sigma_\varepsilon^2}\ . \tag{S4}$$

Marginalizing out m, we have the expected estimate for $\lambda(p)$:

$$\begin{aligned}\widehat{\lambda_p} &= \mathbb{E}_{m|\lambda(p)}\frac{\sigma_s^2 m + \sigma_\varepsilon^2\lambda_s}{\sigma_s^2+\sigma_\varepsilon^2}\\ &= \frac{\sigma_s^2\mathbb{E}_{m|\lambda(p)}m + \sigma_\varepsilon^2\lambda_s}{\sigma_s^2+\sigma_\varepsilon^2}\\ &= \frac{\sigma_s^2\lambda(p) + \sigma_\varepsilon^2\lambda_s}{\sigma_s^2+\sigma_\varepsilon^2}\\ &= (1-\frac{1}{\frac{\sigma_s^2}{\sigma_\varepsilon^2}+1})\lambda(p) + \frac{1}{\frac{\sigma_s^2}{\sigma_\varepsilon^2}+1}\lambda_s. \end{aligned} \tag{S5}$$

According to information theory, for a Gaussian source (variance σ_s^2) transmitted through a Gaussian channel (variance σ_ε^2), the rate of information flow, R, equals the channel capacity, which satisfies

$$R = \frac{1}{2}\ln\left(1+\frac{\sigma_s^2}{\sigma_\varepsilon^2}\right). \tag{S6}$$

Thus $\frac{\sigma_s^2}{\sigma_\varepsilon^2}$ can be written as

$$\frac{\sigma_s^2}{\sigma_\varepsilon^2} = \mathrm{e}^{2R} - 1\,, \tag{S7}$$

which expresses how the internal noise may be constrained by cognitive resources. Substituting Eq. S7 into Eq. S5, we have

$$\widehat{\lambda_p} = (1-e^{-2R})\lambda(p) + e^{-2R}\lambda_s\,. \tag{S8}$$

This form is consistent with Linear in Log-Odds—a two-parameter descriptive model of probability distortion widely used in previous studies (Gonzalez & Wu, 1999; Zhang & Maloney, 2012)—and provides a cognitive explanation for its two parameters. The cognitive resource R controls the sensitivity parameter (slope) of the distortion curve, with higher resource associated with higher sensitivity (steeper slope); the prior λ_s controls the elevation parameter (intercept) of the curve.

The combinations of the RD model with the RI and SP hypotheses are straightforward, which respectively replace the R and λ_s in Eq. S8 with the R_g in Eq. 7 and the λ_g in Eq. 9. The original RD model has two free parameters: R and λ_s. The RD+RI and RD+RI+SP models

replace the parameter R with two parameters: R_0 and β. The RD+SP and RD+RI+SP models replace the parameter λ_s with two parameters: λ_0 and θ.

Bounded Log-Odds (BLO) Model. The BLO model (Ren et al., 2021; Zhang et al., 2020) assumes an agent who encodes probability in log-odds on an internal scale with finite length and homogeneous Gaussian noise. Define a bounded transformation of log-odds

$$\Gamma(p) = \begin{cases} \Delta^-, & \lambda(p) < \Delta^- \\ \lambda(p), & \Delta^- \leq \lambda(p) \leq \Delta^+ \\ \Delta^+, & \lambda(p) > \Delta^+ \end{cases}, \tag{S9}$$

where $\lambda(p)$ (as before) is the log-odds of p, and Δ^- and Δ^+ are the bounds parameters specifying the encoded stimulus range. A mapping of the transformed $\Gamma(p)$ to the internal representation scale $[-\Psi, \Psi]$ is:

$$\Lambda(p) = \frac{2\Psi}{\Delta^+ - \Delta^-}[\Gamma(p) - (\Delta^+ + \Delta^-)/2]. \tag{S10}$$

The noisy internal representation is

$$m = \Lambda(p) + \varepsilon. \tag{S11}$$

where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ is a noise term with mean 0 and variance σ_ε^2.

A subjective log-odds $\widehat{\lambda_m}$ can be decoded from the representation m and further transformed back to the probability scale to obtain $\widehat{p_m} = \lambda^{-1}(\widehat{\lambda_m})$. The decoding goal of the BLO model is not to minimize the squared error for $\widehat{\lambda_m}$ on the log-odds scale, but to minimize the squared error for $\widehat{p_m}$ on the probability scale, whose Bayesian inference has no known analytical form. The BLO model adopts the following form to compensate for the representational noise:

$$\widehat{\lambda_m} = \frac{1}{1 + \kappa V(\widehat{p_m})} m + \left(1 - \frac{1}{1 + \kappa V(\widehat{p_m})}\right)\lambda_s, \tag{S12}$$

where λ_s, as before, is the mean of the Gaussian prior, κ is a free parameter, $V(\widehat{p_m})$ is the variance of $\widehat{p_m}$ due to Gaussian noise on the log-odds scale, whose approximation has the form:

$$V(\widehat{p_m}) \approx \frac{3\sigma_\varepsilon^2}{2\pi^2} p(1-p). \tag{S13}$$

The agent is assumed to have a separate estimate of this variance without knowing the exact value of p. The parameter σ_ε is redundant with κ and thus omitted in model fitting.

Similar to the proof for Eq. S5, we can write the expected estimate of λ_m given p as

$$\widehat{\lambda_p} = \frac{1}{1+\kappa V(\widehat{p_m})}\Lambda(p) + (1 - \frac{1}{1+\kappa V(\widehat{p_m})})\lambda_s\,, \tag{S14}$$

To apply the RI hypothesis to BLO, we replace Ψ with the $R_{\mathscr{g}}$ in Eq. 7. Similar to the enhancement of the DbS and RD models, the SP hypothesis is applied to BLO through replacing the λ_s in Eq. S14 with the $\lambda_{\mathscr{g}}$ in Eq. 9.

Decision model

All the probability estimation models above are combined with the same decision model to produce responses in a specific task. We assume a power utility function

$$u(x) = x^\alpha, \tag{S15}$$

where $\alpha > 0$ is a free parameter, which may be different for gain and loss. Substituting it into Eq. 6 and substituting the probability weighting function $w(p)$ in Eq. 6 with the subjective estimate $\hat{p}$, the expected utility of gamble $\mathscr{g}$ can be written as

$$U(\mathscr{g}) = \hat{p}x_1^\alpha + (1-\hat{p})x_2^\alpha. \tag{S16}$$

where $\hat{p}$ follows Eq. S1 for the DbS model and computed as

$$\hat{p} = \lambda^{-1}(\widehat{\lambda_p}) \tag{S17}$$

for the RD and BLO models.

Two typical tasks of decision under risk, certainty equivalent task and binary choice task, are modeled as the following.

Certainty Equivalent Task. On each trial, participants choose between a gamble and an array of sure payoffs of varying values so that their certainty equivalent (CE)—the value of sure payoffs that is equally preferred—to each gamble can be estimated for individual participants.

The decision model above predicts

$$CE = [\hat{p}x_1^{\alpha} + (1-\hat{p})x_2^{\alpha}]^{1/\alpha}. \tag{S18}$$

We define normalized certainty equivalent

$$\mathrm{NCE} = (\mathrm{CE} - x_2)/(x_1 - x_2) \tag{S19}$$

as a measure of subjective probability and compute a log-odds transformation of the measure for both participants' and model-predicted NCEs. Participants' $\lambda(\widehat{NCE})$ is modeled as model prediction plus a Gaussian noise:

$$\lambda(\widehat{NCE}) = \lambda(NCE) + \varepsilon_{NCE}\ , \tag{S20}$$

where $\varepsilon_{NCE} \sim N(0, \sigma_{NCE}^2)$ and σ_{NCE} is a free parameter.

Binary Choice Task. On each trial, participants choose between two gambles, or between one gamble and one sure payoff. We model the probability of choosing gamble g_i as a SoftMax function determined by its utility difference with the other gamble:

$$Pr(g_i) = \frac{\mathrm{e}^{bU(g_i)}}{\sum_{j=1}^{2} e^{bU(g_j)}}\ , \tag{S21}$$

where b is an inverse temperature parameter that controls the stochasticity of the choice. Participants' binary choice on each trial is modeled as a Bernoulli random variable based on $Pr(g_i)$.

Supplemental Tables S1 and S2

Table S1

General Notations

Symbol	
s	Stimulus to be represented.
m	Mental representation of the stimulus.
ε	Mental representation noise.
$g: (x_1, p; x_2)$	Monetary gamble with two outcomes, that is, receiving x_1 with probability p, or x_2 otherwise, where $\lvert x_1 \rvert > \lvert x_2 \rvert$.
$\widehat{CE}$	Measured certainty equivalent (CE) in the task, that is, the value of sure payoffs that is equally preferred to the gamble g.
$\widehat{NCE}$	Normalized certainty equivalent, $\widehat{NCE} = \left(\widehat{CE} - x_2\right)/(x_1 - x_2)$.
$u(\cdot)$	Utility function.
α	Exponent parameter of the power utility function.
$U(g)$	Expected utility of gamble g.
$w(\cdot)$	Probability weighting (distortion) function.
R_g	Cognitive resources allocated to represent the probability in gamble g.
R_0	Parameter indicating the baseline resource.
β	Slope parameter controlling the effect of value on resource allocation.
m_p	Noisy representation of probability p.
$\hat{p}$	Subjective estimate of probability p.
$\lambda(\cdot)$	Log-odds function, $\lambda(p) = \ln \frac{p}{1-p}$.
$\lambda^{-1}(\cdot)$	The inverse function of $\lambda(\cdot)$.
$\lambda(p)$	Log-odds of p.
$\widehat{\lambda_m}$	The estimate of $\lambda(p)$ given m.
$\widehat{\lambda_p}$	The expected estimate for $\lambda(p)$ over different values of m.
$N\left(\lambda_g, \sigma_s^2\right)$	The prior Gaussian distribution assumed for $\lambda(p)$, where λ_g denotes the mean and σ_s^2 the variance.
θ	Slope parameter controlling the relationship between prior mean probability λ_g and the associated potential gain.
λ_0	Parameter indicating the baseline prior mean.

Table S2

Model Notations

Decision by Sampling (DbS) Model	
$\Phi(\cdot)$	The cumulative distribution function of the standard Gaussian distribution.
λ_s	Parameter for the mean of the Gaussian prior for $\lambda(p)$, which is replaced by λ_0 and θ in the SP hypothesis.
σ_s^2	Parameter for the variance of the Gaussian prior for $\lambda(p)$.
Rate-Distortion (RD) Model	
R	Parameter for the channel capacity, which is replaced by R_0 and β in the RI hypothesis.
λ_s	Parameter for the mean of the Gaussian prior for $\lambda(p)$, which is replaced by λ_0 and θ in the SP hypothesis.
Bounded Log-Odds (BLO) Model	
Δ^+, Δ^-	Parameters controlling the bounds of the log-odds representation.
Ψ	Parameter for the half-range of the internal representation scale (resource), which is replaced by R_0 and β in the RI hypothesis.
$\widehat{p_m}$	The estimated probability given representation m, that is, $\widehat{p_m} = \lambda^{-1}(\widehat{\lambda_m})$.
$V(\widehat{p_m})$	Variance of $\widehat{p_m}$ due to Gaussian noise on the log-odds scale.
κ	Parameter that controls the effect of $V(\widehat{p_m})$.
λ_s	Parameter for mean of the Gaussian prior, which is replaced by λ_0 and θ in the SP hypothesis.
Decision Model for Certainty Equivalent Task	
NCE	Predicted normalized certainty equivalent of the model.
ε_{NCE}	Log-odds scale Gaussian estimation noise of $\lambda(NCE)$, $\varepsilon_{NCE} \sim N(0, \sigma_{NCE}^2)$.
σ_{NCE}^2	Variance parameter of estimation noise.
Decision Model for Binary Choice Task	
$Pr(g_i)$	Probability of choose gamble g_i.
b	Inverse temperature parameter of the choice.

Supplemental Text S1. Two possible updating algorithms for priors in memory

One possibility is to start with the long-term distribution acquired before the task, g_L^0, and keeps updating it across the task, for example, following delta-rule learning:

$$g_L^t = g_L^{t-1} + \alpha(s^t - g_L^{t-1}), \tag{S22}$$

where g_L^t denotes the long-term prior on trial t, s^t is the stimulus on trial t and α is a learning rate. The updating can also be parametric, constraining the functional form of the distribution and only updating its parameters. Then just set

$$g_M^t = g_L^t. \tag{S23}$$

An alternative way is to update a separate estimation of the task distribution:

$$g_T^t = g_T^{t-1} + \alpha(s^t - g_T^{t-1}), \tag{S24}$$

with $g_T^0 = g_L^0$. The memory prior at trial t, g_M^t, is a mixture of g_T^t and g_L^0:

$$g_M^t = \omega g_T^t + (1-\omega) g_L^0 , \tag{S25}$$

where ω is a weight parameter. Here we do not exclude the possibility that the stimuli in the current task will finally become part of the learned long-term distribution, but simply assume that the learning occurs much more slowly and the change would be negligible for tasks with a limited number of trials (Binz et al., 2022; Botvinick et al., 2019).

Supplemental Text S2. An alternative mathematical form to describe the peanuts effect

When we preregistered the experiment, we had not known about the work by Fehr-Duda et al. (2010), who reasoned that the peanuts effect should be understood as the influence of outcome value on probability distortion. Instead, we had framed the peanuts effect as the influence of winning probability on the utility function, an alternative mathematical form to describe the peanuts effect.

In the cumulative prospect theory (CPT, as well as other decision models), the subjective utility of a gamble g: $(x_1, p; x_2)$ is:

$$U(g) = w(p)u(x_1) + \big(1 - w(p)\big)u(x_2), \tag{S26}$$

where $u(\cdot)$ and $w(\cdot)$ respectively denote the utility and probability weighting functions. Suppose the utility function $u(x) = x^{\alpha}$. In the special case that $x_2 = 0$, the certainty equivalent (CE) of the gamble follows:

$$CE^{\alpha} = w(p)x_1^{\alpha}. \tag{S27}$$

Taking logarithms and then scaling by $1/\alpha$ on both sides lead to:

$$\ln(CE) = \ln(x_1) + \frac{\ln(w(p))}{\alpha}\ , \tag{S28}$$

That is, regardless of the functional form of $w(\cdot)$, CPT predicts that the logarithm of CE is a linear function of the logarithm of x_1, with slope equaling 1.

This form inspired us to use the following more general linear form (with slope parameter k and intercept parameter c)

$$\ln(CE) = k\ln(x_1) + c\ , \tag{S29}$$

to measure the relationship between $\ln(CE)$ and $\ln(x_1)$ in the human data, separately for each specific probability p. If the estimated k is not a constant but changes with p, the CPT prediction (Eq. S28) is violated, implying that the utility and probability weighting functions may not be

independent of each other as assumed in CPT.

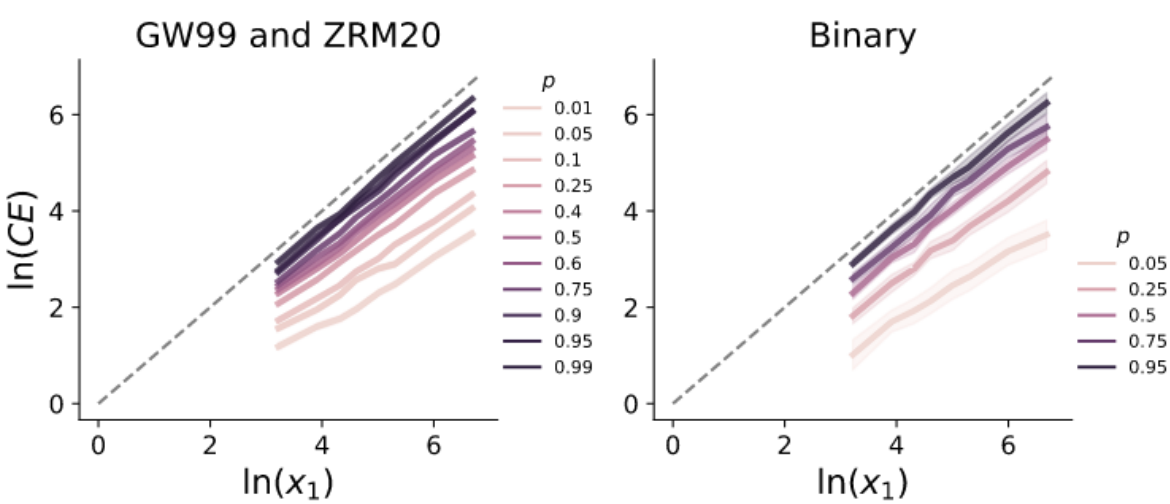
Indeed, by re-analyzing the measured CEs in published datasets (GW99 and ZRM20), we found a nice linear relationship between $\ln(CE)$ and $\ln(x_1)$ (for gambles with $x_2 = 0$), separately for each probability p (Figure S1, left). However, the slope k of the linear regression was less than 1, not constant but increased with p (Figure S2, left).

We later performed a pre-registered experiment (Study 2 reported in the main text) to test whether there were similar effects in a binary choice task. We replicated the findings above as expected (Figures S1 and S2, right).

The linear regression based on Eq. S29 thus constitutes an alternative mathematical form to describe the peanuts effect. The resulting findings are consistent with those reported in the main text, indicating an interaction between the utility function and the probability weighting function. As we discussed in the main text, we still do not know whether it reflects the influence of probability on the utility function, or the influence of value on the probability weighting function.

Figure S1

Relationship of CE to x_1 separately for different probabilities.

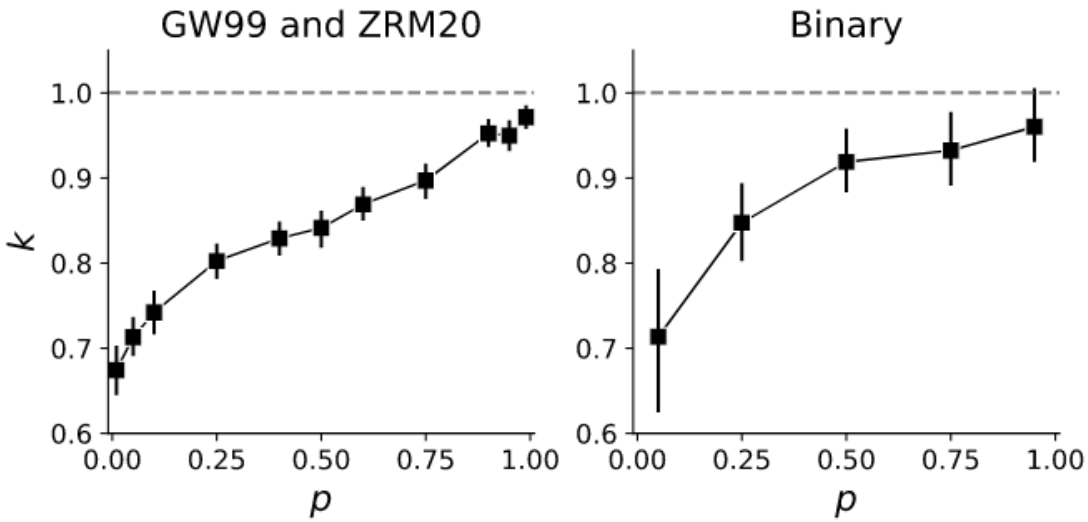


Note. Participants' CE for each gamble (with $x_2 = 0$) are plotted against the reward x_1 of the

gamble on the logarithmic scale, separately for different winning probabilities (in different colors). Each curve is for one probability. Shadings denote SE across participants. Left: The GW99 and ZRM20 datasets of the certainty equivalent task (corresponding to Study 1 in the main text). Right: The binary choice task (corresponding to Study 2 in the main text).

Figure S2

Regression slope k of $\ln(x_1)$ as a function of the winning probability p.



Note. Left: The GW99 and ZRM20 datasets of the certainty equivalent task (corresponding to Study 1 in the main text). Right: The binary choice task (corresponding to Study 2 in the main text). Error bars denote SE across participants.

Supplemental Text S3. Hierarchical Bayesian model for FD05 dataset

We used a hierarchical Bayesian version of BLO+RI+SP model to fit the FD05 dataset (Fehr-Duda et al., 2010). We assumed that the rational inattention (RI) parameter β, structural prior (SP) parameter θ could be different in the gain and loss domains. Following the Prospect Theory, we assumed two different exponents α_G and α_L for the power utility functions in gain and loss. To boost the sampling process, we reparametrized the Δ^-, Δ^+ and Ψ (Eq. S10) with B^-, B and τ:

$$\Delta^- = 3\mathrm{erf}(\mathrm{B}^-)\,, \tag{S30}$$

$$\Delta^+ = 3\mathrm{erf}(\mathrm{B}^- + \mathrm{B})\,, \tag{S31}$$

$$\tau = \frac{2\Psi}{\Delta^+ - \Delta^-} = \tau_0 + \beta_{G(or\ \mathrm{L})}\ln(|\mathrm{x}_1 - \mathrm{x}_2| + 1). \tag{S32}$$

where $\mathrm{erf}\,(.)$ is the Gaussian error function. Because the probabilities in the FD05 dataset were between 0.05 and 0.95, we bounded Δ^- and Δ^+ between −3 and 3.

Parameters and prior distributions are listed in Table S3.

Table S3

Parameters, priors and hyperpriors for the hierarchical Bayesian version of BLO+RI+SP model.

Parameters	Prior	Hyperpriors
$\kappa' = ln(\kappa)$	$N(\mu_{\kappa'}, t_{\kappa'})$	
B^{-}	$N(\mu_{\mathrm{B}^{-}}, t_{\mathrm{B}^{-}})$	
$\mathrm{B}' = \ln(\mathrm{B})$	$N(\mu_{\mathrm{B}'}, t_{\mathrm{B}'})$	
λ_0	$N(\mu_{\lambda_0}, t_{\lambda_0})$	
$\tau_0{}' = ln(\tau_0)$	$N(\mu_{\tau_0{}'}, t_{\tau_0{}'})$	$\mu_{(\cdot)} \sim N(0, 1/9)$
β_L	$N(\mu_{\beta_L}, t_{\beta_L})$	$t_{(\cdot)} \sim \Gamma(1, 1)$
$\Delta\beta = \beta_G - \beta_L$	$N(\mu_{\Delta\beta}, t_{\Delta\beta})$	
θ_L	$N(\mu_{\theta_L}, t_{\theta_L})$	
$\Delta\theta = \theta_G - \theta_L$	$N(\mu_{\Delta\theta}, t_{\Delta\theta})$	
$\alpha_G{}' = ln(\alpha_G)$	$N(\mu_{\alpha_G{}'}, t_{\alpha_G{}'})$	
$\alpha_L{}' = ln(\alpha_L)$	$N(\mu_{\alpha_L{}'}, t_{\alpha_L{}'})$	

Note. $\mu_{(\cdot)}$ and $t_{(\cdot)}$ respectively denote the mean and precision of the Gaussian prior $N(\mu_{(\cdot)}, t_{(\cdot)})$ for each parameter.

Supplemental References

Bates, C. J., Lerch, R. A., Sims, C. R., & Jacobs, R. A. (2019). Adaptive allocation of human visual working memory capacity during statistical and categorical learning. Journal of Vision, 19(2), 11. https://doi.org/10.1167/19.2.11

Berger, T., & Gibson, J. D. (1998). Lossy Source Coding. IEEE Trans. Inf. Theory, 44, 2693–2723. https://doi.org/10.1109/18.720552

Bhui, R., & Gershman, S. J. (2018). Decision by sampling implements efficient coding of psychoeconomic functions. Psychological Review, 125(6), 985–1001. https://doi.org/10.1037/rev0000123

Bhui, R., Lai, L., & Gershman, S. J. (2021). Resource-rational decision making. Current Opinion in Behavioral Sciences, 41, 15–21. https://doi.org/10.1016/j.cobeha.2021.02.015

Binz, M., Gershman, S. J., Schulz, E., & Endres, D. (2022). Heuristics from bounded meta-learned inference. Psychological Review. https://doi.org/10.1037/rev0000330

Botvinick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., & Hassabis, D. (2019). Reinforcement Learning, Fast and Slow. Trends in Cognitive Sciences, 23(5), 408–422. https://doi.org/10.1016/j.tics.2019.02.006

Fehr-Duda, H., Bruhin, A., Epper, T., & Schubert, R. (2010). Rationality on the rise: Why relative risk aversion increases with stake size. Journal of Risk and Uncertainty, 40(2), 147–180. https://doi.org/10/dgmr2s

Gershman, S. J., & Bhui, R. (2020). Rationally inattentive intertemporal choice. Nature Communications, 11(1), 3365. https://doi.org/10.1038/s41467-020-16852-y

Gonzalez, R., & Wu, G. (1999). On the Shape of the Probability Weighting Function. Cognitive Psychology, 38(1), 129–166. https://doi.org/10.1006/cogp.1998.0710

Khaw, M. W., Li, Z., & Woodford, M. (2021). Cognitive Imprecision and Small-Stakes Risk Aversion. The Review of Economic Studies, 88(4), 1979–2013.

https://doi.org/10.1093/restud/rdaa044

Ren, X., Luo, H., & Zhang, H. (2021). Automatic and Fast Encoding of Representational Uncertainty Underlies the Distortion of Relative Frequency. The Journal of Neuroscience, 41(16), 3692–3706. https://doi.org/10.1523/JNEUROSCI.2006-20.2021

Stewart, N., Chater, N., & Brown, G. D. A. (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

Zhang, H., & Maloney, L. T. (2012). Ubiquitous Log Odds: A Common Representation of Probability and Frequency Distortion in Perception, Action, and Cognition. Frontiers in Neuroscience, 6. https://doi.org/10.3389/fnins.2012.00001

Zhang, H., Ren, X., & Maloney, L. T. (2020). The bounded rationality of probability distortion. Proceedings of the National Academy of Sciences, 117(36), 22024–22034. https://doi.org/10.1073/pnas.1922401117