On the Source and Instability of Probability Weighting*

Cary Frydman and Lawrence J. Jin

January 30, 2023

ABSTRACT

We propose and experimentally test a new theory of probability distortions in risky choice. The theory is based on a core principle from neuroscience called efficient coding, which states that perception is more accurate for those stimuli that the agent expects to encounter more frequently. We show that when the agent holds a prior that attaches greater likelihood to extreme probabilities compared to intermediate probabilities, efficient coding generates the inverse S-shaped probability weighting function from prospect theory. More important, as the agent's prior varies, the model predicts that probability distortions change systematically. Across two experiments, we provide novel support for the efficient coding hypothesis. In the first experiment, we show that perception of probability is more precise for those values near the boundaries of 0 and 1. These data validate the key model assumption about the shape of the prior, which generates the steeper slope of the weighting function for low and high probabilities compared to intermediate probabilities. In the second experiment, we provide evidence that the shape of the weighting function can be modulated through changes in the prior. Our theory generates additional novel predictions regarding heterogeneity and time variation in probability distortions.

JEL classification: G02, G41
Keywords: efficient coding, perception, prospect theory, probability weighting

*Frydman is at the Marshall School of Business, University of Southern California. Jin is at the SC Johnson College of Business, Cornell University. We are grateful for helpful comments from Nicholas Barberis and audience members at New York University and the 2023 JDM Winter Symposium. Frydman acknowledges financial support from the NSF and USC Marshall School of Business iORB fund. *E-mail addresses:* cfrydman@marshall.usc.edu (C. Frydman), lawrence.jin@cornell.edu (L. Jin).

I. Introduction

Consider the prospect of winning \$25 with 5% probability and winning \$0 otherwise. Under expected utility theory, valuation of the lottery should be linear in the probability of winning the \$25 outcome. For decades, however, experimental evidence points to decision-making that is nonlinear in probability: for example, subjects report a larger increase in valuation as the experimenter increases the probability from 5% to 10%, compared with an increase from 30% to 35% (Tversky and Kahneman, 1992; Camerer and Ho, 1994; Gonzalez and Wu, 1999; Bernheim and Sprenger, 2020). This nonlinearity gives rise to probability distortions, which in turn, have been used to explain fundamental puzzles in risky choice such as the simultaneous demand for gambling and insurance. Probability distortions have also been invoked to explain a wide array of anomalies in financial markets (Barberis, 2018).

In this paper, we ask a basic question: *why* do humans distort probabilities? We explore whether the distortion can be traced to fundamental properties about how the brain processes information. Our hypothesis is that valuation is linear in *perceived* probability, which differs from objective probability due to information processing constraints. Furthermore, we propose that this hypothesis can microfound the probability weighting function from prospect theory—which is the leading behavioral theory of choice under risk (Kahneman and Tversky, 1979). We are by no means the first to investigate the psychological underpinning of probability weighting (Gonzalez and Wu, 1999; Bordalo, Gennaioli, and Shleifer, 2012; Khaw, Li, and Woodford, 2021, 2022; Enke and Graeber, 2022).[1] But the mechanism we test—which is called efficient coding—is novel in the sense that it generates new testable hypotheses about the source and instability of probability weighting. To assess the empirical validity of our proposed mechanism, we develop a theory and present two experimental tests. The data largely confirm the idea that probability weighting derives, at least in part, from the efficient allocation of cognitive resources.

The basic premise of our theory is that, when faced with a risky prospect, the decision maker (henceforth DM) bases her decision on her perception of state probabilities; in general, the DM's perception does not coincide with the objective probability. This wedge between reality and per-

[1] See also important earlier work by Viscusi (1989), Tversky and Kahneman (1992), Rottenstreich and Hsee (2001), Stewart, Chater, and Brown (2006), Zhang and Maloney (2012), Steiner and Stewart (2016), and Zhang, Ren, and Maloney (2020).

ception stems from limited cognitive resources. That is, we assume that the DM does not have the capacity to form a precise cognitive representation of each probability that she could potentially be presented. But, what she can do is to efficiently allocate cognitive resources towards accurately representing those probabilities that she *expects* to encounter in a given class of situations.

For example, consider someone who is often faced with binary lotteries that have probabilities which fall in the range from say, 45% to 55%. We argue that, when faced with two novel lotteries, one with 50-50 odds and another with 51-49 odds, this person will be better able to discriminate between the two lotteries, compared to someone who is often faced with skewed lotteries that have extreme odds such as 90-10 or 95-5. Our model formalizes how the DM's prior belief about the probabilities she expects to encounter will optimally distort her perception of the probabilities that she actually encounters.

Our theoretical contribution builds on our own previous work in which we demonstrated that risk taking depends on the prior distribution of monetary outcomes to which the DM is adapted (Frydman and Jin, 2022). In that model, we made the simplifying assumption that probabilities were encoded without noise. Here, we show how noisy and efficient coding can generate signature patterns of probability distortions. In this regard, our work is closely related to recent theory by Steiner and Stewart (2016), Zhang et al. (2020), and Khaw et al. (2021, 2022). However, an important difference is that in our model, we show how the shape of the prior distribution over probabilities pins down the degree of probability distortion.

The model works as follows. The DM encodes a single probability p subject to a capacity constraint. We assume her objective is to choose an encoding function that maximizes the mutual information between the true probability p and the encoded signal R_p. The model predicts that the agent will encode with more precision those probabilities that she expects to encounter more often—according to her prior belief.

Thus, a key input to the model is the DM's prior belief about p. We show theoretically that when the DM holds a U-shaped prior—one that puts more mass on probabilities farther from 50%—she will encode these "extreme" probabilities with more precision. In fact, this type of prior, when coupled with efficient coding, generates the inverse S-shaped probability weighting function that is often assumed in prospect theory (Gonzalez and Wu, 1996; Prelec, 1998). The endogenous weighting function exhibits high sensitivity for extreme probabilities. In contrast, the weighting

function exhibits low sensitivity over intermediate probabilities, where the DM optimally chooses to encode probabilities with the least precision.

An obvious question is whether the U-shaped prior is a reasonable specification for prior expectations about probabilities. This is important, because as we show, efficient coding without a U-shaped prior is insufficient to generate an inverse S-shaped probability weighting function. Interestingly, Stewart et al. (2006) argue that people's experiences with probabilities are, in fact, likely to be concentrated near the extreme values of 0 and 1; for example, Stewart et al. (2006) document that the frequency of probability phrases in the English language is skewed towards extreme probabilities. Brown and Qian (2004) and Zou, Brown, Zhao, and Dong (2008) provide evidence where subjects report that extreme probabilities occur more frequently than intermediate probabilities. Moreover, in economics experiments that investigate probability distortions, experimenters tend to use designs that oversample extreme probabilities (Tversky and Kahneman, 1992; Gonzalez and Wu, 1999; Bernheim and Sprenger, 2020).

Given how crucial the role of the prior is in generating theoretical predictions in our model, we design our first experiment to provide a joint test of the U-shaped prior and efficient coding. The design is motivated by the model prediction that perception should be noisier for intermediate probabilities relative to extreme probabilities. Our test is simple: on each of several trials, we present subjects with a pair of probabilities that are displayed as two fractions, and we then ask them to classify which fraction is larger. The crucial aspect of the design is that we fix the numerical distance between fractions, and only vary the *level* of the fractions across trials. If perception is noisier for intermediate probabilities, we should observe a U-shaped accuracy function.

We recruit 800 subjects from Prolific, an online data collection platform. We find that accuracy is substantially higher at extreme values of probability—below 0.2 or above 0.8. Moreover, response times are the shortest when the probabilities to be discriminated lie near the extremes. Our evidence is therefore consistent with the joint hypothesis that (i) human subjects hold a U-shaped prior about probabilities and (ii) encoding of such probabilities is efficient given cognitive constraints. Interestingly, the encoding function that endogenously arises in our model provides a qualitative match to the encoding function that is assumed in Khaw et al. (2021, 2022). Moreover, the U-shaped accuracy function that we uncover in our data is also consistent with the hump-shaped "cognitive uncertainty" function that Enke and Graeber (2022) document when eliciting certainty

equivalents across the unit interval.

We then move on to a second experiment in which we test whether biases in perception of probability feed into subjective valuation of risky lotteries. To investigate the link between perception of probability and valuation, we exogenously manipulate a subject's prior belief about probabilities, and then elicit certainty equivalents for lotteries of the form: $(\$25, p; \$0, 1-p)$. Efficient coding predicts that when the DM is adapted to a distribution of probabilities that is concentrated near intermediate probabilities, she should have less difficulty discriminating probabilities over this range, compared to someone who is adapted to extreme probabilities. Crucially, because the DM does not have access to the objective probability p, she must rely only on her perceived value of p when forming a valuation. These observations lead to a testable prediction: conditional on a given risky lottery, valuation will depend systematically on the prior distribution to which a subject is adapted.

We find strong evidence in favor of this prediction. The slope of the measured probability weighting function over intermediate probabilities is significantly higher for subjects who are adapted to intermediate probabilities, compared to those who are adapted to extreme probabilities. Specifically, for subjects who are adapted to intermediate probabilities, increasing an intermediate probability of winning \$25 by 1% causes a \$0.26 increase in their certainty equivalent. In contrast, for subjects who are adapted to extreme probabilities, the same increase of an intermediate probability of winning \$25 by 1% leads to only a \$0.17 increase in their certainty equivalent. Thus, the marginal effect of a 1% increase in probability on valuation of the *same* lottery is about 50% larger for subjects who are adapted to intermediate probabilities. We find that the effect of the prior on valuation carries a similar magnitude for lotteries associated with small probabilities of winning \$25. Finally, the effect is of a smaller magnitude, but remains statistically significant, for lotteries associated with high probabilities of winning \$25.

Taken together, our results suggest that probability weighting stems from people having an imprecision in forming the cognitive representation of probability. At the same time, other candidate explanations for probability weighting may also be active. The salience theory by Bordalo et al. (2012) proposes that probability weighting arises from limited attention. In their theory, attention is disproportionately drawn to salient payoffs that are precisely defined. As in our framework, salience theory also delivers a context-dependent probability weighting function. One distinction,

however, is that salience theory's weighting function depends on the payoffs of the risky lottery, whereas our theory of efficient coding does not take these payoff amounts as input.[2]

Another candidate explanation for probability weighting is the framework of cognitive uncertainty proposed by Enke and Graeber (2022). In their framework, agents have a precise perception of probability, but cognitive noise corrupts the valuation process. As in efficient coding and salience theory, cognitive uncertainty also generates a weighting function that depends on the environment. For example, Enke and Graeber (2022) find that when displaying a lottery in a more complex format (i.e., a compound lottery), the inferred probability distortions become more pronounced. Thus, our efficient coding-based explanation joins a growing class of models that microfounds the probability weighting function, and importantly, makes predictions about how the shape of the weighting function changes with context. One distinguishing feature of our model, however, is that it is based on an optimization process that delivers a nonlinear probability weighting function.[3]

Zhang et al. (2020) develop a model that is closely related to ours—it is also an efficient coding model of probabilities. An important distinction, however, is that our model is able to accommodate any shape of a prior distribution, whereas the model in Zhang et al. (2020) only takes as input the range of probabilities to be considered. Their model predicts that, as the range of probabilities becomes wider, the slope of the weighting function should decrease. The authors find suggestive correlational evidence for this effect: in a meta-analysis on the slope of the weighting function across twelve different experimental studies, the range of probabilities used in a given experimental design negatively predicts the slope of the weighting function. Our model takes a substantial step forward by showing that the *shape* of the DM's prior distribution has a systematic and causal effect on valuation.

Our model also shares similarities with the decision by sampling model from Stewart et al. (2006). In that model, when the DM is presented with a probability, she draws samples from memory and compares the currently presented probability with the randomly drawn samples. This

[2] Khaw et al. (2022) propose a model in which the imprecision in representing a monetary amount is tied to the probability of winning that amount. In their model, less resources are devoted towards processing payoffs that are less likely to be delivered.

[3] Steiner and Stewart (2016) provide an alternative explanation of probability weighting that is also based on noisy perception of probability. In particular, they show that a nonlinear probability weighting function can arise as an optimal response to perceptual noise. Khaw et al. (2021, 2022) also propose a model that delivers an endogenous probability weighting function based on an optimal response to noisy coding of probabilities. One difference—which we discuss further in Section VI.2—is that the encoding function in our model also arises from an optimization process.

cognitive operation enables the DM to rank the current probability in the context of probabilities that are recalled from memory. Thus, memory plays a similar role to the prior in our model. Bhui and Gershman (2018) show theoretically that efficient coding can implement decision by sampling. Indeed, the Heng, Woodford, and Polanía (2020) model that we use is flexible enough to implement decision by sampling with a specific encoding function, one that also depends on the DM's prior. Our experimental results are qualitatively consistent with both the decision by sampling model and efficient coding; at the same time, our results reject a broad class of models that predict a stable probability weighting function.

Finally, it is worth noting that one of the earliest psychological explanations for the probability weighting function was tied to the idea that decision makers cannot perfectly discriminate between all probabilities. Tversky and Kahneman (1992) propose that the inverse S-shaped weighting function arises from the principle of diminishing sensitivity from two separate reference points, namely 0 and 1. Gonzalez and Wu (1999) point out that the concept of diminishing sensitivity is tied closely to the psychophysical notion of discriminability—which itself is related to the basic premise of our efficient coding model. Gonzalez and Wu (1999) propose that the weighting function can be characterized by a discriminability "slope" parameter and an attractiveness "level" parameter. Our model can be interpreted as providing a microfoundation both for the lack of perfect discriminability and its malleability across contexts.

The rest of this paper is organized as follows. Section II presents a theory to show that efficient coding and a U-shaped prior together lead to an inverse S-shaped probability weighting function. Section III presents an experiment that provides a joint test of efficient coding and the U-shaped prior. Section IV shows theoretically that the slope of different portions of the probability weighting function changes as the DM's prior changes. Section V presents an experiment which uses data on lottery valuations to confirm that the weighting function is indeed malleable in the manner predicted by efficient coding. Section VI provides additional discussions, and Section VII concludes.

II. The Model: Generating the Inverse S-shaped Probability Weighting Function

In this section, we apply the efficient coding model of Heng et al. (2020) to study the perception of probability. The model works as follows. Before the DM is presented with a probability p, she holds a prior belief about the probability, denoted by $f(p)$. Then, upon presentation of the probability p, the DM's perceptual system generates a noisy signal R_p; specifically, R_p is randomly drawn from a conditional distribution—or, in the language of Bayesian inference, a likelihood function—denoted by $f(R_p|p)$. The likelihood function captures the idea that the DM encodes information about probability with cognitive noise, even if the probability p is clearly presented to the DM. Alternatively, the likelihood function can reflect noise in retrieving information about p from memory when the DM is asked to make a decision.

Following Heng et al. (2020), we assume that the DM encodes probability p through a finite number of n "neurons," where the output state of each neuron takes the value of 0 or 1. The output states of these n neurons are assumed to be mutually independent, and each neuron takes the value 1 with probability $\theta(p)$ and 0 with the remaining probability $1-\theta(p)$. The encoded value of p is therefore represented by an output vector of $0s$ and $1s$, with length n. Given that the neurons are mutually independent, a sufficient statistic for the output vector is the sum across the n output values, which gives rise to the noisy signal R_p. Taken together, the likelihood function is given by

$$f(R_p|p) = \binom{n}{R_p} (\theta(p))^{R_p} (1-\theta(p))^{n-R_p}, \tag{1}$$

where the noisy signal R_p can take on integer values from 0 to n.

A signature feature of efficient coding is that the DM endogenously chooses the likelihood function $f(R_p|p)$ as a function of the prior $f(p)$. We assume the DM chooses $\theta(p)$ in (1) to maximize mutual information between probability p and its noisy signal R_p

$$\max_{\theta(p)} I(p, R_p), \tag{2}$$

where the mutual information $I(p, R_p)$ is defined as the difference between the marginal entropy of R_p and the entropy of R_p conditional on p. Intuitively, the objective function in (2) leads

the DM to better discriminate between probability values that she expects to encounter more frequently given her prior belief. A large literature in sensory perception documents strong support for this objective function (Laughlin, 1981; Girshick, Landy, and Simoncelli, 2011; Wei and Stocker, 2015). Heng et al. (2020) show that the optimal coding rule that maximizes $I(p, R_p)$ is given by

$$\theta(p) = \left(\sin\left(\frac{\pi}{2}F(p)\right)\right)^2, \tag{3}$$

where $F(p)$ is the cumulative distribution function of the prior belief $f(p)$.

Given the prior belief and the noisy signal, the DM follows Bayes' rule to generate posterior belief about p

$$f(p|R_p) = \frac{f(R_p|p)f(p)}{\int_0^1 f(R_p|p)f(p)dp}. \tag{4}$$

For an objective probability value p, the DM draws the noisy signal R_p according to the likelihood function $f(R_p|p)$ from equation (1). For each noisy signal R_p, the DM forms a posterior distribution about p given by equation (4). Together, these two steps imply that the average perception of p is

$$v(p) = \sum\nolimits_{R_p=0}^{n} f(R_p|p) \cdot \mathbb{E}[\tilde{p}|R_p], \tag{5}$$

where

$$\mathbb{E}[\tilde{p}|R_p] \equiv \int_0^1 f(p|R_p)pdp = \frac{\int_0^1 f(R_p|p)f(p)pdp}{\int_0^1 f(R_p|p)f(p)dp} \tag{6}$$

is the posterior mean of p conditional on R_p. Note that $v(p)$ in (5) represents the DM's subjective valuation of p averaged across different values of R_p, and in general, $v(p)$ is a nonlinear function that maps objective probabilities into decision-relevant probabilities—similar to the decision weights from prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992).[4] At the same time, unlike the deterministic weighting function exogenously assumed in prospect theory, $v(p)$ in our model is derived from stochastic probability perception. To formally define the stochastic nature

[4] As we later discuss in Section VI.1, our theory and prospect theory differ in an important way: our theory computes $v(p)$ as the DM's average *belief* about probability p; by contrast, prospect theory posits that the DM's belief about p is correct, and the overweighting or underweighting of p arises from her *preferences*.

of probability perception, we compute the standard deviation of the DM's subjective valuation,

$$\sigma(p) = \left[\sum\nolimits_{R_p=0}^{n} f(R_p|p) \cdot (\mathbb{E}[\tilde{p}|R_p] - v(p))^2\right]^{1/2}. \tag{7}$$

Equation (7) makes it clear that the randomness in probability perception comes from the fact that a given objective value of p randomly generates one of the n possible values of R_p; each R_p then generates a different posterior distribution $f(p|R_p)$ and hence a different valuation $\mathbb{E}[\tilde{p}|R_p]$.

It is important to note that, in an efficient coding model, the DM's subjective valuation of probability $v(p)$ depends on her prior belief $f(p)$ through *two* channels. First, the posterior belief $f(p|R_p)$ is directly affected by the prior belief $f(p)$, as shown in equation (4). Second and more importantly, $f(p)$ indirectly affects the posterior belief $f(p|R_p)$ through the coding rule $\theta(p)$. In particular, equation (3) shows that the coding rule takes as input the prior, so that the same value of p is encoded differently depending on the prior. Thus, the prior continues to play its traditional and direct role in Bayesian inference, but it also affects information processing through the endogenously determined likelihood function given by equations (1) and (3).

The above framework demonstrates that the DM's prior belief, when combined with efficient coding, governs the shape of the weighting function $v(p)$. Here we pick a particular shape of the prior distribution and illustrate how the various components of the model pin down the nonlinearities in probability perception. We assume that the DM holds a U-shaped prior of the following form

$$f(p) = \frac{we^{-\lambda_1 p} + (1-w)e^{-\lambda_2(1-p)}}{w\lambda_1^{-1}(1-e^{-\lambda_1}) + (1-w)\lambda_2^{-1}(1-e^{-\lambda_2})}, \tag{8}$$

where the parameters w, λ_1, and λ_2 control for the shape of the function. This prior puts more weight on extreme probabilities compared to intermediate probabilities. Some motivation for this type of prior comes from previous work which shows that, in both daily life and economics experiments, extreme probabilities are encountered more often than intermediate probabilities (Stewart et al., 2006). Because of the crucial role the U-shaped prior assumption plays in our theoretical analysis, in the next section, we conduct an experiment which tests this assumption.

[Place Figure I about here]

We now illustrate the implications of such a U-shaped prior. Recall that the prior influences the

posterior in part through the optimal choice of the likelihood function. Thus, here we first examine the properties of the likelihood function implied by the U-shaped prior. Figure I Panel A plots the U-shaped prior from equation (8); the parameter values are: $\lambda_1 = \lambda_2 = 8$, and $w = 0.5$. Given this prior, Panel B plots the corresponding coding rule $\theta(p)$. It shows that, under the U-shaped prior, $\theta(p)$ is steep for low and high values of p; it is much flatter for intermediate values of p. This nonlinear pattern is a consequence of the DM's objective to maximize the mutual information between p and R_p; importantly, the nonlinear pattern of the coding function will be passed on to the weighting function $v(p)$, which will form the basis of our experimental tests.

Panel C plots the implied likelihood function for four different probability values: $p = 0.02$, $p = 0.04$, $p = 0.49$ and $p = 0.51$. We note that $p = 0.02$ and $p = 0.04$ generate distinctive likelihood functions; when p is low, the slope of $\theta(p)$ is high, and a small change in p leads to a large shift in the likelihood function. As such, the DM is able to easily discriminate between low probabilities. By the symmetry of the likelihood function—that is, $f(R_p|p) = f(n - R_p|1 - p)$—the DM is also able to easily discriminate between high probabilities. On the contrary, $p = 0.49$ and $p = 0.51$ give rise to nearly identical likelihood functions; when p is intermediate, the slope of $\theta(p)$ is low and changes in p cause almost no change in the likelihood function. As such, the DM finds it difficult to discriminate between intermediate probabilities.

[Place Figure II about here]

Given the prior and the implied likelihood functions, we next turn to the DM's subjective valuation of probability $v(p)$. We substitute the U-shaped prior from equation (8) into the coding rule in equation (3) and explicitly compute $v(p)$ according to equation (5). Figure II demonstrates that, with efficient coding and with the U-shaped prior, the resulting weighting function $v(p)$ is inverse S-shaped, consistent with prospect theory. To understand this shape, we first note from (5) that $v(p)$ is constructed by combining the likelihood function $f(R_p|p)$ with the posterior mean $\mathbb{E}[\tilde{p}|R_p]$. We also note that the posterior mean $\mathbb{E}[\tilde{p}|R_p]$ does not depend directly on p. As a result, when p varies, changes in $v(p)$ depend only on changes in the likelihood function. As shown in Figure I Panel B, the slope of $\theta(p)$ is steep both for high and low values of p; for these probability values, a small change in p leads to a large shift in the likelihood function and hence also a large change in $v(p)$; as such, the slope of $v(p)$ is steep. On the other hand, the slope of $\theta(p)$ is low for

intermediate values of p; in this case, changes in p cause little change in the likelihood function and hence also little change in $v(p)$, resulting in a shallow slope of $v(p)$. Taken together, the inverse S-shaped coding function $\theta(p)$ in Figure I Panel B, which is implied by the U-shaped prior, gives rise to an inverse S-shaped probability weighting function $v(p)$.

[Place Figure III about here]

It is worth noting that both efficient coding and a U-shaped prior are fundamentally important for generating an inverse S-shaped probability weighting function. To illustrate how the weighting function depends on both components, we examine two scenarios. The first scenario keeps the U-shaped prior but replaces the efficient coding rule in equation (3) by an inefficient one

$$\theta(p) = \left(\sin\left(\frac{\pi}{2}p\right)\right)^2. \tag{9}$$

An important implication is that this alternative coding rule does not connect the prior belief $f(p)$ with the likelihood function $f(R_p|p)$.[5] Figure III Panel A plots the weighting function $v(p)$ when the coding rule takes the form of (9). It shows that, in the absence of efficient coding, the U-shaped prior alone is insufficient to generate an inverse S-shaped probability weighting function.

The second scenario keeps the efficient coding rule but replaces the U-shaped prior by a uniform prior. That is, we replace the U-shaped prior in equation (8) by a uniform prior $f(p) = 1$. Figure III Panel B shows that, with the uniform prior, efficient coding alone is insufficient to generate an inverse S-shaped weighting function. The results in Figure III therefore highlight that it is the particular set of likelihood functions that are optimized under a U-shaped prior that are crucial for generating the inverse S-shaped probability weighting function. In the next section, we conduct an experiment which is aimed precisely at testing whether subjects' perception is consistent with the likelihood functions that are efficient under a U-shaped prior.

[5] Note that there are an infinite number of inefficient coding rules; equation (9) is just one of them that we use to illustrate that a noisy coding model with a U-shaped prior is not sufficient to generate an inverse S-shaped probability weighting function.

III. Experiment 1: Discrimination task

We have shown how efficient coding and a U-shaped prior can theoretically generate an inverse S-shaped probability weighting function. A key assumption driving our theoretical result is that the DM has a prior which puts more mass on extreme probabilities compared to intermediate probabilities. In this section, we test an important implication of the theory, namely that perception is noisier for intermediate probabilities compared to extreme probabilities.

III.1. Background on numerical cognition of fractions and probabilities

The foundation of our experimental design comes from the literature on numerical cognition. There are numerous papers documenting empirical evidence consistent with the hypothesis that humans have an "approximate number system" for integers (Dehaene, 2011). The approximate number system facilitates an intuitive but noisy representation of number that uses brain circuits which are distinct from those invoked in symbolic representation and manipulation of numerical quantities. Recently, the approximate number system has been shown to have important consequences for the way that humans process monetary amounts in simple risky choice (Khaw et al., 2021; Frydman and Jin, 2022).

It is more controversial, however, to argue that the same approximate number system also facilitates representation of ratios and fractions—which in turn could support an intuitive representation of probability. Clarke and Beck (2021) put forth the hypothesis that the number sense supports rational numbers—those that can be expressed as a ratio between two integers—but not irrational numbers.[6] In particular, the authors offer an evolution-based explanation for why the approximate number system would support rational numbers (pg 22): "...organisms need to reason under uncertainty...The capacity to engage in fast and efficient probabilistic reasoning would, thus, be enormously advantageous...To do so, they need to represent probabilities; and the most straightforward to do this is to represent rational numbers."

Some evidence in favor of the hypothesis that the approximate number system supports fractions comes from Matthews and Chesney (2015). Those authors show that humans exhibit a "distance effect" when comparing two fractions: error rates decrease as the difference between the two frac-

[6]The controversy of such a hypothesis was highlighted when Clarke and Beck (2021) triggered 26 published commentaries.

tions increases.[7] Crucially, Matthews and Chesney (2015) demonstrate the distance effect even when one fraction is displayed symbolically (as a ratio of Arabic numerals) and the other is displayed non-symbolically (as two different clouds of dots). The fact that a distance effect obtains in this "cross-modal" format suggests that humans convert the clouds of dots into an abstract and intuitive notion of ratio, which is then compared with the cognitive representation of the symbolically presented fraction.[8]

An obvious question is whether the ratios and fractions that have been used in the previous experimental literature are interpreted by subjects as probabilities. An alternative hypothesis is that subjects instead interpret the ratios in a manner that is detached from any probabilistic connotation. Some recent evidence from developmental psychology, however, suggests that ratio perception can indeed be regarded as probability perception. Szkudlarek and Brannon (2021) give children an experimental task in which they are asked to choose a machine that will give the highest chance of delivering a specific colored ball. The probability is shown symbolically in one condition, but non-symbolically in the other condition. Interestingly, accuracy in the non-symbolic version of the probability discrimination task was highly correlated with accuracy in the symbolic version of the task, where probabilities were displayed with symbolic fractions. This evidence suggests that probabilities can be represented both with symbolic and non-symbolic stimuli, and the representation of each type may share a common cognitive underpinning. Further evidence of this effect from adult experimental subjects comes from Eckert, Call, Hermes, Herrmann, and Rakoczy (2018).

Both Szkudlarek and Brannon (2021) and Eckert et al. (2018) provide evidence consistent with the hypothesis that imprecision of ratio perception follows Weber's law. The evidence for this hypothesis comes from the fact that error rates in probability discrimination decrease as the ratio of ratios increases. However, at least in these two studies, the distance between probabilities on each trial is confounded with the ratio of ratios. In the next subsection, we design a task in which we fix the distance between ratios and vary only the *level* of ratios across the unit interval. This

[7] Evidence for a distance effect from earlier experiments where subjects compare integers has been taken as support for the existence of an approximate number system for integers.

[8] The cross-modal evidence complements other demonstrations of the distance effect when both stimuli are represented as symbolic fractions (DeWolf, Grounds, Bassok, and Holyoak, 2014; Jacob and Nieder, 2009; Kallai and Tzelgov, 2009, 2012; Meerta, Grégoire, and Noël, 2010; Schneider and Siegler, 2010; Siegler, Thompson, and Schneider, 2011; Meerta, Grégoire, Seron, and Noël, 2012).

design enables us to investigate how the degree of imprecision in perception of ratios shifts across the unit interval.

III.2. Design

Our design builds on symbolic numerical discrimination tasks that began with Moyer and Landauer (1967). Such tasks can be used to elucidate the patterns of imprecision that experimental subjects display when asked to classify which of two symbolic numbers is larger (e.g., "1" vs. "8"). A classic result from this literature is that subjects exhibit a "distance effect," whereby error rates and response times become larger as the numerical distance between quantities becomes smaller.

Here, we seek to understand how the degree of imprecision varies systematically over the unit interval. To do so, we incentivize subjects to accurately and quickly classify which of two fractions is larger. We elicit responses across multiple trials, and crucially, we fix the distance between fractions on each trial at 2%. Thus, we control for the distance effect and look for systematic variation in imprecision depending on the location of the two fractions in the unit interval. Using the theory from the previous section as a guide, we expect that the degree of imprecision should be largest for fractions near 50%, and it should be smallest for those fractions near the extremes (0% and 100%).

[Place Table I and Figure IV about here]

Each trial contains a pair of fractions. We build the pairs by applying several restrictions that follow from the design in Matthews and Chesney (2015). First, we use only reduced fractions to eliminate any cognitive processing involved with reducing fractions. For simplicity, the numerator and denominator in each fraction have at most 2 digits. Second, we avoid fraction pairs where the numerator is the same for each fraction in the pair; we similarly avoid fraction pairs where the denominator is the same for each fraction in the pair. This is done so that subjects cannot simply use the heuristic of comparing only the numerators when the denominators are the same, or of comparing only the denominators when the numerators are the same. These restrictions, along with our main goal of varying fraction levels across the entire unit interval lead to a set of 28 fraction pairs that are presented in Table I. We define the trials on which we show these 28 fraction pairs as "test trials." We show each subject each of the 28 fraction pairs twice, and we counterbalance the location of each fraction on the right or left part of the screen. This generates

a total of 56 test trials per subject; Table I shows that, on each test trial, the denominator for each fraction takes the value of either 25 or 50. Figure IV provides a screenshot of one sample trial. Subjects choose either the left or right fraction by pressing the "Q" or "P" key on their keyboard.[9]

Finally, to reduce the chance that subjects develop a strategy where they simply double the numerator associated with the denominator of 25 and then just compare numerators, we add a set of 28 "catch" trials to the experiment. These catch trials have random numerators and denominators. Moreover, the distance between fractions on catch trials is not fixed at 2%. The 28 catch trials are also presented in Table I.

When constructing our design, we made a conscious choice to use fractions rather than percentages. Our justification for this design choice is that the fractional notation is likely to help foster the perception of a probability. If we were to instead use percentages, subjects may employ a heuristic where they ignore the percent sign (%) and simply compare integers. Our goal here is to assess whether imprecision has any features that are unique to the unit interval, compared to other number systems that have already been well-studied (e.g., integers).

III.3. Procedures

As described above, the experiment consists of 56 test trials and 28 catch trials, for a total of 84 trials. Subjects were paid $2.25 for completing all 84 trials. Additionally, 10% of subjects were randomly selected to receive a bonus payment. The bonus payment follows the incentive scheme in Frydman and Jin (2022), in which subjects are paid as a function of both accuracy and speed. Specifically, if a subject is selected to receive a bonus, we pay $0.10 for each correct answer and we subtract $0.01 for each second it takes the subject to respond. We impose a minimum bonus amount of $0, and subjects are aware of this minimum. For those subjects who were randomly selected to receive a bonus, the average total earning was $6.06. We recruit 800 subjects from Prolific, an online data collection platform. The sample size, main analyses, and data exclusion criteria are pre-registered on Aspredicted.org.[10]

[9] The experimental instructions are provided in Online Appendix A.

[10] The pre-registration document for this experiment is at: https://aspredicted.org/blind.php?x=YKS_FMK.

III.4. Results

As outlined in the pre-registration document, we exclude trials for which the response time is less than 0.2 seconds or greater than 10 seconds. After this exclusion, we have 65,511 trials, of which 43,833 are test trials. The average accuracy among test trials is 73.3% and the average response time is 2.44 seconds. The substantial error rate—in the presence of time pressure—suggests that the task is not trivial.

[Place Figure V about here]

Figure V shows how the accuracy and response times vary across the unit interval. Each point in the figure summarizes data from one pair of fractions used on test trials; recall that each of the 28 fraction pairs used on a test trial is presented to a subject twice, as we counterbalance the left-right display of the two fractions. For both panels in Figure V, the "average probability" on the x-axis denotes the average of the two fractions on a given test trial. One can see, visually, that accuracy is U-shaped while response times are hump-shaped. This specific pattern of non-uniform imprecision across the unit interval is a key prediction of efficient coding when the subject's prior is U-shaped.

To provide a more formal test of the systematic imprecision, we partition the unit interval into "extreme" and "intermediate" probabilities. As outlined in pre-registration, we define a fraction to be *extreme* if it is less than 0.2 or greater than 0.8; all remaining probabilities are defined to be *intermediate*. Table II provides results from mixed effects regressions where the dependent variables of interest are accuracy and response times. The key explanatory variable is a dummy that takes the value of one if both fractions on the test trial are extreme.[11]

[Place Table II about here]

Table II Column (1) shows that average accuracy among intermediate test trials is 72.3%, while accuracy on extreme test trials is significantly higher by 2.6% (p-value < 0.001). Moreover, Column (2) shows that response times are significantly shorter by 0.112 seconds when both fractions are

[11] Our design precludes the scenario in which, within a test trial, one fraction is extreme and the other is intermediate; either both fractions are extreme or both are intermediate.

extreme compared to intermediate (p-value < 0.001). Thus, subjects not only exhibit higher accuracy for fractions near the endpoints of the unit interval, but they also implement these decisions significantly more quickly. This result is important as it provides novel evidence for a systematic imprecision across the unit interval in a design that controls for the difference between fractions.

One natural concern, however, with the above results is that smaller fractions in our design tend to have only one digit in the numerator, whereas larger fractions have two digits in the numerator.[12] Thus, it may be that our tests are just picking up the extra difficulty that subjects have with perceiving a two-digit numerator relative to a one-digit numerator. To address this concern, we re-run our analyses restricting to trials where both fractions have two digits. This restriction implies that all extreme fractions are now above 0.8. Our tests therefore rely on comparisons of behavior among fractions in the intermediate range with those extreme fractions near one.

Columns (3) and (4) of Table II show that the regression results are robust to removing trials with one-digit numerators. These subsample results are additionally important because they highlight that imprecision over the unit interval fails to conform to the pattern of imprecision often found among integers. Specifically, it is often found that as the distance between integers remains fixed, imprecision grows monotonically in the *level* of the integers (Moyer and Landauer, 1967; Buckley and Gillman, 1974; Dehaene, 2011). We instead find evidence that subjects are *more* accurate and respond more quickly when discriminating between large fractions compared to when discriminating between intermediate fractions.

We conclude this section by providing evidence on the distance effect from the 28 catch trials. These results serve as a replication of previous experimental work on symbolic fraction comparisons, except that in our design, subjects are financially incentivized. As outlined in the pre-registration document, we test whether error rates and response times decrease in the distance between fractions. We run mixed effects linear regressions where the dependent variable is either error rates or response times and the explanatory variable is the distance between the two fractions. We also include a random intercept and a random slope. Consistent with previous work, we find that both error rates and response times are negatively correlated with the distance between fractions; each negative correlation is significant at the 0.1% level.

[12] All fractions in our design have two-digit *denominators*.

IV. The Model: Instability of Probability Weighting

The experimental data presented in the previous section are consistent with the efficient coding hypothesis and with human subjects holding a U-shaped prior belief about probability. As we emphasized in Section II, the prior belief is a key input to the efficient coding model. Thus, in principle, a change in the DM's prior belief should affect her perception through a change in the likelihood functions. In this section, we formally show how a shift in the prior affects perception and hence, subjective valuation.

In particular, we highlight how efficient coding generates a probability weighting function that is *malleable*. We begin by assuming that the DM's prior belief takes the following form

$$f_c(p) = \xi \cdot \underbrace{f_1(p)}_{\text{a stable component}} + (1-\xi) \cdot \underbrace{f_2(p)}_{\text{a fast-moving component}}. \tag{10}$$

We refer to $f_c(p)$ as a "mixed" prior, which contains two components. The first component is stable, and is denoted as $f_1(p)$. This component represents the "natural" prior that the DM has formed based on encountering different probability values with different frequencies over her lifetime. The second, fast-moving component is denoted as $f_2(p)$; it captures transient changes in statistics of the local environment. For example, $f_2(p)$ can capture the empirical distribution of probabilities that we present in the lab experiment in the next section.[13] The parameter ξ in equation (10) represents the weight the DM puts on the stable component; this weight may depend on the rate at which the statistics of the environment change.

In light of our experimental results from the previous section—as well as the arguments outlined in Stewart et al. (2006)—it is reasonable to assume that the natural prior $f_1(p)$ is U-shaped. We operationalize this assumption by setting $f_1(p)$ equal to the U-shaped prior from equation (8)

$$f_1(p) = \frac{we^{-\lambda_1 p} + (1-w)e^{-\lambda_2(1-p)}}{w\lambda_1^{-1}(1-e^{-\lambda_1}) + (1-w)\lambda_2^{-1}(1-e^{-\lambda_2})}. \tag{11}$$

For the second, fast-moving component, we consider two specifications. Our main goal is to demonstrate how shocks to the fast-moving component drive differences in valuation. As such, it is useful

[13] A growing literature has argued that people adapt, at least in part, towards the local context that changes over the course of an experimental session (Burke, Baddeley, Tobler, and Schultz, 2016; Zimmermann, Glimcher, and Louie, 2018; Conen and Padoa-Schioppa, 2019; Payzan-LeNestour and Woodford, 2022).

to choose two specifications of the fast-moving component that have very different shapes. We investigate an "intermediate" prior component, which has density only in the center of the unit interval and thus balances out the U shape of the stable component. Formally, for this "intermediate" prior component, we assume

$$f_2(p) = \begin{cases} \dfrac{1}{p_h - p_l} & p_l \leq p \leq p_h \\ 0 & \text{otherwise} \end{cases}, \tag{12}$$

where p_l and p_h take intermediate values and $0 < p_l < p_h < 1$.

We also investigate an "extreme" prior component, which has density only near the the endpoints of the unit interval; as such, this prior component accentuates the U-shape of the stable component. For this "extreme" prior component, we assume

$$f_2(p) = \begin{cases} \dfrac{1}{(p_{h,2} - p_{h,1}) + (p_{l,2} - p_{l,1})} & p_{l,1} \leq p \leq p_{l,2} \\ \dfrac{1}{(p_{h,2} - p_{h,1}) + (p_{l,2} - p_{l,1})} & p_{h,1} \leq p \leq p_{h,2} \\ 0 & \text{otherwise} \end{cases}, \tag{13}$$

where $p_{l,1}$ and $p_{l,2}$ take low values, $p_{h,1}$ and $p_{h,2}$ take high values, and $0 \leq p_{l,1} < p_{l,2} < p_{h,1} < p_{h,2} \leq 1$.

The two different specifications of the fast-moving component give rise to two different mixed priors, which we refer to as the intermediate prior and the extreme prior. We specify the parameter values for these two priors as follows. For the stable component $f_1(p)$ of each mixed prior, we set $\lambda_1 = \lambda_2 = 8$ and $w = 0.5$. For the fast-moving intermediate prior component, we set $p_l = 0.38$ and $p_h = 0.62$. For the fast-moving extreme prior component, we set $p_{l,1} = 0.1$, $p_{l,2} = 0.21$, $p_{h,1} = 0.79$, and $p_{h,2} = 0.9$. Finally, we set ξ, the weight on the stable component in the mixed prior, to 0.5.

Armed with two specific mixed priors, we now use them as inputs to our efficient coding model and assess the differences in perception of probability. For each mixed prior $f_c(p)$, we substitute it into the coding rule in equation (3). The two different coding rules—one for each specification of the mixed prior—give rise to a distinct set of efficient likelihood functions. Figure VI summarizes these

results: it plots, for both the intermediate prior and the extreme prior, the prior distribution $f_c(p)$, its coding rule $\theta(p)$, and the likelihood functions. Figure VII then plots the probability weighting functions $v(p)$ implied by these two mixed priors.

[Place Figures VI and VII about here]

Figures VI and VII demonstrate the *malleability* of probability weighting: as the DM's prior beliefs changes (the upper graph of Figure VII), the implied probability weighting function $v(p)$ changes significantly (the lower graph of Figure VII). More important, the way in which the weighting function changes is governed by efficient coding. Specifically, for intermediate probabilities, the slope of the weighting function is steeper when the fast-moving prior component "fills in" some density for intermediate probabilities; however, for extreme probabilities, the slope of the weighting function is steeper when the fast-moving prior component puts extra density in the extreme portions of the unit interval, near 0 and 1.

To understand these results, we first note that, by construction, intermediate probabilities occur much more frequently under the intermediate prior, compared to the extreme prior. Thus, under the intermediate prior, the coding rule $\theta(p)$ has a much steeper slope for intermediate probabilities, which causes the likelihood function $f(R_p|p)$ to shift substantially as p varies over this intermediate range. To see this visually, Figure VI shows a substantial change in the likelihood function as p increases from 0.49 to 0.51, under the intermediate prior. By contrast, the same increase in p from 0.49 to 0.51 leads to almost no change of the likelihood function under the extreme prior. This greater separation of likelihood functions for nearby probabilities gives rise to the higher slope of the probability weighting function under the intermediate prior.

The same intuition holds when examining perception of probabilities closer towards the boundaries of the unit interval. By construction, low or high probabilities occur more frequently under the extreme prior, compared to the intermediate prior. Therefore, as p varies over the low or high range of probability, the likelihood function shifts to a greater extent under the extreme prior. Figure VI shows a substantial change in the likelihood function when p increases from 0.12 to 0.14 under the extreme prior. By contrast, the same increase in p from 0.12 to 0.14 causes a much smaller change in the likelihood function under the intermediate prior. These patterns of overlap in the likelihood functions in turn generate a steeper slope of the probability weighting function

under the extreme prior.

In the next section, we design an experiment that exogenously manipulates the fast-moving component of a subject's prior. The particular specifications of the fast-moving component follow closely from the intermediate and extreme priors we have theoretically considered above. Using data on lottery valuations, the experiment will allow us to test whether the slope of different portions of the probability weighting function is malleable in the manner predicted by our theory of efficient coding.

V. Experiment 2: Main task

V.1. Design

Our experimental design is intended to test whether subjects' prior belief about probability causally affects their valuations of risky lotteries. As in the theory outlined in the previous section, here we assume that subjects in our experiment only have noisy perception of probability; their perception of lottery payoffs is noiseless. To justify such an assumption, we consider only binary lotteries with two possible payoffs, and crucially, the two possible payoffs are constant across all trials. As such, noisy perception of lottery payoffs should be limited.

Before describing the details of the experimental design, we should make clear that testing for the causal effect of prior belief on valuation is not a new endeavor. For example, in the theory and subsequent experimental tests of Kőszegi and Rabin (2007), lottery valuations are affected by the DM's reference point, which is equated with her (rational) expectations. The important distinction with previous literature is that valuation in our setting depends on prior belief about a feature of the choice set that has *already* been presented to the DM. In particular, it is the DM's prior belief about the probability p that optimally shapes perception and drives valuation.

We now turn to the details of our design. On each trial, a subject is presented with a risky lottery of the following form

$$(\$25, p; \$0, 1-p). \tag{14}$$

The subject is then asked to provide her certainty equivalent for the lottery by using a slider bar that appears below a graphical representation of the risky lottery. Figure VIII provides a screenshot

of an example trial. We choose to use the slider method rather than the commonly used multiple price list, as it allows us to elicit an exact valuation instead of a switching interval. For similar reasons, the slider method is also adopted by Khaw et al. (2022).

[Place Figure VIII about here]

Each subject in the experiment completes a total of thirty trials, and thus submits a total of thirty certainty equivalents for various values of probability p. Subjects are incentivized to report their certainty equivalents using a Becker-Degroot-Marschak mechanism (Becker, DeGroot, and Marschak, 1964). We define the first twenty-four trials as "adaption trials" and the final six trials as "test trials." The core of our design is to exogenously manipulate the distribution of p on the adaptation trials and test the impact on valuation in the test trials. The distribution of p on adaptation trials plays the role of the fast-moving component of the subject's mixed prior, as we previously described in Section IV. Importantly, for different sets of adaptation trials, we hold constant the set of test trials. In this way, the only feature of the experiment that varies is the distribution of p that subjects experience before they provide their valuations on the six test trials.

When constructing our design, we face a tradeoff in choosing the number of test trials per subject. On the one hand, more test trials provide us with a greater number of data points on which to estimate the treatment effect of the prior. On the other hand, the more test trials we use, the weaker the treatment effect will be for the final few test trials. This is because early test trials will begin to contaminate the prior that we attempt to induce with the adaptation trials. Thus, while we would ideally like to construct a large set of test trials that sample p across the entire unit interval, we are constrained by the number of test trials that we can use *per subject*.

Because this design constraint operates only at the subject level, we choose to construct three separate test trial ranges, and randomly assign subjects to one of the three ranges. In this manner, we can span a large range of the unit interval while minimizing the number of test trials needed per subject. Specifically, we define a *low* range for which $p \in \{0.11, 0.15, 0.19, 0.23, 0.27, 0.31\}$, an *intermediate* range for which $p \in \{0.38, 0.42, 0.47, 0.53, 0.58, 0.62\}$, and a *high* range for which $p \in \{0.69, 0.73, 0.77, 0.81, 0.85, 0.89\}$.

After subjects are randomized into one of the three test trial ranges, we further randomize them into an adaptation condition. Each test trial range is associated with one of two possible adaptation

conditions. Table III provides the exact values of p that we use in each adaptation condition. For the *low* test trial range, we adapt subjects to either a *low* or *intermediate* set of adaptation trials. The intuition is that, for a subject who is adapted to low values of p, she should be able to easily discriminate between low values of probability once the low range test trials arrive. In contrast, for a subject who is adapted to intermediate values of p, she should be less able to discriminate between low values of probability. Thus, conditional on being randomized into the *low* test trial range, our model predicts that valuation will respond more strongly to changes in p among subjects in the *low* adaptation condition, compared to those in the *intermediate* adaptation condition.

[Place Table III about here]

A similar intuition applies for the other two test trial ranges. For those subjects randomized into the *high* test trial range, we adapt them to either a *high* or *intermediate* set of adaptation trials. For a subject who is adapted to high values of p, she should be better able to discriminate between probabilities in this high range once the test trials arrive, compared to a subject who is adapted to intermediate values of p.

Finally, for those subjects randomized into the *intermediate* test trial range, we adapt them to either an *extreme* or *intermediate* set of adaptation trials. A subject who is adapted to extreme values of p—closer to the endpoints of the unit interval— should be less able to discriminate between probabilities in the intermediate range once the test trials arrive, compared to a subject who is adapted to intermediate values of p.

V.2. Procedures

We recruit 600 subjects from Prolific. Subjects are paid \$3.00 for completing the experiment. In addition, 10% of the subjects are randomly selected to receive a bonus whose amount is based on their decision in one randomly selected trial. Specifically, the bonus amount is the outcome of a Becker-Degroot-Marschak mechanism: on the randomly selected trial, we draw a monetary amount randomly from a uniform distribution between $[\$0, \$25]$. If the subject's certainty equivalent is below or equal to the randomly drawn monetary amount, the subject receives the monetary amount. If, however, the subject's certainty equivalent is above the monetary amount, the computer plays the lottery and pays the subject either \$25 or \$0 with the trial-specific probabilities.

At the beginning of the experiment, subjects are presented with instructions and then need to pass a two-question comprehension check; the experimental instructions and the comprehension check are provided in Online Appendix A. On average, subjects completed the experiment in 11 minutes. Conditional on being chosen to receive a bonus, they received an average total earning of $19.54. The sample size, main analyses, and data exclusion criteria are pre-registered on Aspredicted.org.[14]

V.3. Results

Following our pre-registration, we apply two exclusion criteria before analyzing the data. First, we exclude those subjects who violated a basic monotonicity property in their responses. In particular, for each subject, we estimate a linear regression of the certainty equivalent on p based on the first twenty-four adaptation trials. We then record the regression coefficient for each subject, and drop all subjects for whom the coefficient is negative; this restriction excludes 39 of the 600 subjects. Note that we only estimate the regression for each subject on the first twenty-four trials, so that we do not select on the dependent variable of interest (certainty equivalents on test trials). Next, we drop all observations for which the subject reported their certainty equivalent in less than one second; this restriction excludes an additional 28 observations. Taken together, we are left with a data set that contains 16,802 trials, of which 3,362 are test trials.

We now turn to testing our main hypothesis for each of the three different test trial ranges. In order to arrive at an estimate of the subject's perceived value of p, here we assume that the subject has a linear utility function. This implies that the subject's perception of p can be estimated simply as $\frac{CE}{25}$, where CE represents the certainty equivalent that the subject provides.[15]

Figure IX shows a graphical depiction of the effect of the prior on perception of probability. The figure has three panels; each panel plots the perceived probability against the objective probability for two sets of subjects. The left panel involves two sets of subjects who are both randomized into the low test trials; one set is adapted to the *low* adaptation trials, while the other set is adapted to the *intermediate* adaptation trials. The right panel involves two sets of subjects who are both

[14]The pre-registration document for this experiment is at: https://aspredicted.org/blind.php?x=W25_48V.

[15]Later in this section, we consider a more general case that allows the subject to have intrinsic risk aversion over the lottery payoff; for a payoff of X, the utility is $u(X) = X^{\alpha}$, where $\alpha \leq 1$. We find that, under nonlinear utility specifications, our experimental data continue to support the theoretical predictions from Section IV and Figure VII.

randomized into the high test trials; one set is adapted to the *high* adaptation trials, while the other set is adapted to the *intermediate* adaptation trials. Finally, the middle panel involves two sets of subjects who are both randomized into the intermediate test trials; one set is adapted to the *extreme* adaptation trials, while the other set is adapted to the *intermediate* adaptation trials.

[Place Figure IX about here]

Recall from Section IV and Figure VII our theoretical prediction: subjects who are adapted to intermediate probabilities should have greater difficulty discriminating between low probabilities, compared to those who are adapted to low probabilities. In other words, when the perceived probability is plotted against the objective probability, the slope should be lower for subjects who are adapted to intermediate probabilities; the left panel of Figure IX confirms this prediction. Similarly, subjects who are adapted to intermediate probabilities should also have greater difficulty discriminating between high probabilities, compared to those who are adapted to high probabilities; the right panel of Figure IX confirms this prediction. Finally, subjects who are adapted to intermediate probabilities should have *less* difficulty discriminating between intermediate probabilities, compared to those who are adapted to extreme probabilities; the middle panel of Figure IX confirms this prediction. In sum, the three panels in Figure IX provide experimental evidence for the theoretical prediction that the slope of different portions of the probability weighting function is malleable in the manner predicted by the efficient coding hypothesis.

[Place Table IV about here]

To formally test the efficient coding hypothesis, Table IV presents results from three mixed effects linear regressions, one for each test trial range. For these regressions, the dependent variable is our estimate of the subject's perceived valuation of p; the independent variables include the objective value of probability p, a dummy variable labelled *intermediate* which takes the value of one if the subject is randomized into the intermediate adaptation condition (and zero otherwise), and the interaction between p and *intermediate*. We include only test trials in the regressions, which allows us to estimate the effect of adaptation trials on valuation.

Column (1) reports the regression results for the low test trial range. The key variable of interest, $p \times intermediate$, has a significantly negative coefficient. This indicates that subjects who

are adapted to intermediate probabilities form perception of low probability that is less sensitive to changes in the objective probability, compared to subjects who are adapted to low probabilities. Column (3) reports the regression results for the high test trial range. Again, the interaction term, $p \times intermediate$, has a significantly negative coefficient. This indicates that subjects who are adapted to intermediate probabilities form perception of high probability that is less sensitive to changes in the objective probability, compared to subjects who are adapted to high probabilities. Finally, Column (2) reports the regression results for the intermediate test trial range. Here, we see that the interaction term has a significantly *positive* coefficient. Consistent with the prediction of efficient coding, this result indicates that subjects who are adapted to intermediate probabilities form perception of intermediate probability that is *more* sensitive to changes in the objective probability, compared to subjects who are adapted to extreme probabilities.

In the above analyses, our key dependent variable is the subject's perceived value of p. However, because this perceived value is unobservable, we estimate it using the certainty equivalent data—under the assumption that the subject has a linear utility function. If, instead, we relax this linear utility assumption and allow the subject to have intrinsic risk aversion over the lottery payoff X, where $X = 25$ with probability p and $X = 0$ with probability $1-p$, then, on each trial, the certainty equivalent that the subject provides is

$$CE = [\mathbb{E}(\tilde{p}|R_p) \cdot (25)^{\alpha}]^{1/\alpha} = [\mathbb{E}(\tilde{p}|R_p)]^{1/\alpha} \cdot 25, \tag{15}$$

where α is the risk aversion parameter and $\alpha \leq 1$. Equation (15) implies that $\mathbb{E}(\tilde{p}|R_p) = (\frac{CE}{25})^{\alpha}$.

We re-estimate our main regressions in Table IV, but now using $(\frac{CE}{25})^{\alpha}$ as the dependent variable. We have examined a variety of values for α, including $\alpha = 0.88$, the value reported in Tversky and Kahneman (1992), and $\alpha = 0.70$, the value used by Barberis, Jin, and Wang (2021) to match the behavior of real-world retail investors. We find that our regression results are robust to these alternative nonlinear utility specifications: the coefficient on $p \times intermediate$ remains significantly negative for the low test trial range and the high test trial range, and it remains significantly positive for the intermediate test trial range. We also note that, with higher risk aversion (i.e., as α decreases), the implied perceived value of p will change. For example, when $\alpha = 0.7$, we find that small probabilities are distorted upwards: $v(p) > p$ when p is small.

VI. Discussion

VI.1. Belief distortion as a source of probability weighting

Both of our experiments showcase the role that efficient coding plays in generating perceptual distortions of probability. Experiment 1 highlights that subjects' discrimination ability varies systematically over the unit interval, in a way that is predicted by efficient coding when subjects have a U-shaped prior. Experiment 2 highlights the malleability of probability weighting, as exogenous shifts in the prior lead to predictable changes in valuation.

The model that we present, which builds on earlier work by Khaw et al. (2021) and Frydman and Jin (2022), has only one free parameter n. This parameter measures the number of "neurons" that the DM uses for encoding probability, so it can be interpreted as a capacity constraint. Our assumption that n is finite—and thus information processing is constrained—is responsible for the prediction that the average perception of p does not equal p. As such, our model of probability weighting derives exclusively from a friction in information processing and thus points to a belief-based channel as an important driver of probability distortions.

Such a belief-based channel is worth emphasizing, given that the literature beginning with Kahneman and Tversky (1979) has largely interpreted probability weighting as an expression of preferences, rather than a misperception of beliefs. Specifically, our results suggest that the degree of probability weighting should be modulated by the agent's prior expectations, and thus information provision and past experience may play an important role in characterizing the shape of the weighting function. This interpretation is consistent with recent work by Enke and Graeber (2022), who show that the weighting function can be modulated by the complexity of the lottery under consideration. If the weighting function were instead exclusively driven by preferences, we should expect that the weighting function does not vary with complexity—which is counterfactual to the data presented by Enke and Graeber (2022). At the same time, our data do not rule out preferences as a source of probability weighting. That is, if we shut down our noisy encoding channel so that the objective and perceived probability coincide, it may well be that valuation is still passed through an exogenous nonlinear probability weighting function.

There are some intricacies of the mechanism that our experiments do not nail down. For example, consider the scenario in which a subject is presented with a risky lottery and is simply

asked to report the probability that is associated with the upside payoff. We speculate that subjects would be able to report back to the experimenter the exact probability that is displayed on screen. This raises the question: if the subject has conscious access to the objective probability, how could the distortion of probability still operate through a belief-based channel?

One hypothesis is that when information about p is retrieved from memory for the purpose of valuation, it is retrieved with errors of the type we have modeled in our efficient coding framework. Obviously, this finer hypothesis about the particular stages at which representations of probability are and are not corrupted with error remains untested. Put differently, the cognitive representation of probability may be different when a subject is asked to simply repeat back a probability to the experimenter, compared to when she is asked to use it as an input to a subjective valuation process. What our model attempts to capture is the cognitive representation that is relevant for valuation.[16]

VI.2. Endogenizing the encoding function

As we briefly discussed in the Introduction, there have been many earlier attempts to microfound the probability weighting function. Our model is closest in spirit to recent models of noisy and efficient coding of probability. In particular, Khaw et al. (2021, 2022) assume that the DM encodes probabilities using a log odds encoding (or likelihood) function. That is, the DM's noisy signal of p is drawn from a normal distribution with mean of $\log(\frac{p}{1-p})$.[17] Note that this likelihood function hard-wires a pattern whereby extreme probabilities are easier to discriminate compared to intermediate probabilities. In contrast, our efficient coding model endogenously derives a likelihood function that is qualitatively consistent with this pattern, but under the assumption that the DM has a U-shaped prior.[18] Of course, if the DM's prior changes, our model then produces a different likelihood function, presumably of a different functional form, whereas the likelihood function in Khaw et al. (2021, 2022) remains fixed. Our data are consistent with the hypothesis that the

[16]The triple-code model proposed in Dehaene (1992) suggests that the representation of number differs across different decision contexts.

[17]Zhang and Maloney (2012) propose a related encoding function whereby the DM's distorted probability is a convex combination of the log odds, namely $\log(\frac{p}{1-p})$, and a constant.

[18]One conceptual difference exists between the likelihood function that arises endogenously under our model and the exogenously assumed likelihood function where the noisy signal is drawn from a normal distribution with mean of $\log(\frac{p}{1-p})$. Specifically, in the latter case, encoding of p becomes infinitely precise as p approaches either 0 or 1. This in turn implies that $v(p) \to 0$ as $p \to 0$. In contrast, with the endogenous likelihood function from our model, the implied perceived value of $p = 0$ is strictly above 0; $v(p) > 0$ as $p \to 0$. In other words, our model generates a discontinuity of the weighting function at the boundary; similarly, this discontinuity is also derived from the efficient coding model of Zhang et al. (2020).

likelihood function indeed shifts with prior beliefs.

Another closely related model of efficient coding of probability is proposed in Zhang et al. (2020). Those authors highlight the role that the prior plays in determining the precision of the encoding function: as the range of probabilities considered becomes narrower, precision of each of those considered probabilities increases. However, Zhang et al. (2020) do not consider general shapes of the prior, and thus cannot produce the non-uniform discrimination thresholds that our model generates in the case of a U-shaped prior. At an empirical level, we provide novel evidence that is consistent with the conjecture by Zhang et al. (2020) that "...human representation of probability can adapt to the environment, in the spirit of efficient coding..." (pg 22032).

VI.3. Broader implications of imprecision of ratios

In Experiment 1, we use fractions as stimuli to represent probabilities. In Experiment 2, we use explicit representations of state probabilities, both numerically (displaying percentages) and visually (displaying the ratio of two colored bars). Throughout the paper, we have mainly been concerned with deriving and testing implications for risky choice when the DM has an imprecise representation of *probability*.

However, our theory can be applied more broadly to real numbers in the unit interval that do not necessarily have to represent probabilities. For example, our framework of noisy and efficient coding will also make predictions about perception of proportions when there is no intrinsic risk to consider. To make ideas concrete, consider the task in Oprea (2022) where the DM must report a valuation of a riskless—but disaggregated—set of monetary payments. In particular, subjects are asked to provide a valuation for an asset that is comprised of 100 disaggregated payments. A proportion p of these payments are each worth \$2.50 and the remaining proportion $1-p$ of the payments are worth \$0. Thus, the subject's task is simply to integrate the 100 payments into a single valuation.

If the subject has imprecision in perceiving the proportion p, this will drive a wedge between the subject's reported valuation and the "true" valuation—the latter is simply the sum of the 100 disaggregated payments. Indeed, Oprea (2022) finds that subjects report valuations that seem to accord with the four-fold pattern of risk-taking—even in the absence of any intrinsic risk. Our model would predict the four-fold pattern for what Oprea (2022) calls "simplicity equivalents"—as

long as the DM has a prior over p that is U-shaped, which is consistent with the realized values of p in Oprea (2022)'s experiment. Moreover, our model makes an additional untested prediction, which is that the pattern of reported simplicity equivalents can be modulated by the DM's prior belief about proportions.

VI.4. Applications to financial economics

While probability weighting has been applied in a variety of areas of economics, including insurance and betting markets, it has been particularly fruitful in explaining facts in financial economics (Barberis, 2013; O'Donoghue and Somerville, 2018). The core component of probability weighting that most researchers in financial economics have invoked is the overweighting of small probabilities. Models with this basic assumption can generate the high equity premium (De Giorgi and Legg, 2012), abnormal returns of individual stocks (Barberis and Huang, 2008; Barberis, Mukherjee, and Wang, 2016; Barberis et al., 2021), overpricing of out-of-the-money stock options (Baele, Driessen, Ebert, Londono, and Spalt, 2019), and time-inconsistent risk-taking (Barberis, 2012; Ebert and Strack, 2015; Heimer, Iliewa, Imas, and Weber, 2021). Importantly, in all these studies, researchers assume that probability distortions are "hard-wired" and do not shift over time.

Our results suggest that the degree of probability distortions depends critically on the DM's prior belief about probability, which can vary over time. It follows that the financial phenomena described above should also vary over time—to the extent that they are generated by probability weighting. This insight leads to two concrete paths forward for future empirical work. First, our model predicts *when* probability distortions should arise, and thus when we should observe anomalous asset prices and risk-taking behavior. For example, when combined with the theory in Barberis et al. (2021), our model predicts conditional moments by identifying periods when overweighting of small probabilities should be more pronounced. During these periods—when perception of probability is heavily distorted—we should observe average returns of individual stocks that are consistent with those returns documented in the idiosyncratic volatility and failure probability anomalies. But during other periods, when investors' prior beliefs induce less perceptual distortions, the asset pricing anomalies should be weaker.

Second, by connecting probability distortions to investors' prior beliefs, our model also makes predictions about individual heterogeneity. In particular, those investors who hold prior beliefs that

probabilities are skewed towards 0 and 1 will distort probabilities more than investors who believe that probabilities are uniformly distributed. One challenge, of course, in testing these predictions is to obtain accurate measures of investors' prior beliefs about probabilities.

VII. Conclusion

We have provided a new explanation for probability weighting based on a core principle from neuroscience: efficient coding. We first develop a theory to show that efficient coding and a U-shaped prior together generate the inverse S-shaped probability weighting from prospect theory. The theory also predicts that, when the prior changes, the slope of different portions of the probability weighting function will change predictably.

In a pair of experiments, we provide novel support for the efficient coding model. First, we provide support for the joint hypothesis that perception of probability is governed by efficient coding and a U-shaped prior. The key piece of evidence is that perception of probability is more accurate for those probabilities that have higher density under a U-shaped prior. Second, we measure the impact of the prior belief on valuation. To do so, we experimentally manipulate the prior and then assess whether valuation of a given lottery changes systematically. Indeed, we find strong support that the shape of the probability weighting function is unstable yet reacts in the manner predicted by efficient coding.

In closing, it is useful to highlight the connection between our experimental results and the patterns documented in recent experiments by Payzan-LeNestour and Woodford (2022) and Frydman and Jin (2022). Each of the two latter papers provides experimental support in favor of efficient coding as a driver of diminishing sensitivity, which is another core ingredient in prospect theory. Thus, when viewed against this broader perspective, our data add credence to the hypothesis that efficient coding serves as a common mechanism underlying multiple features of prospect theory (Woodford, 2012a,b).

REFERENCES

Baele, Lieven, Joost Driessen, Sebastian Ebert, Juan M Londono, and Oliver G Spalt, 2019, Cumulative prospect theory, option returns, and the variance premium, *Review of Financial Studies* 32, 3667–3723.

Barberis, Nicholas, 2013, Thirty years of prospect theory in economics: A review and assessment, *Journal of Economic Perspectives* 27, 173–196.

Barberis, Nicholas, 2018, Psychology-based models of asset prices and trading volume, in Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics* (North Holland, Amsterdam).

Barberis, Nicholas C., 2012, A model of casino gambling, *Management Science* 58, 35–51.

Barberis, Nicholas C., and Ming Huang, 2008, Stocks as lotteries: The implications of probability weighting for security prices, *American Economic Review* 98, 2066–2100.

Barberis, Nicholas C., Lawrence J. Jin, and Baolian Wang, 2021, Prospect theory and stock market anomalies, *Journal of Finance* 76, 2639–2687.

Barberis, Nicholas C., Abhiroop Mukherjee, and Baolian Wang, 2016, Prospect theory and stock returns: An empirical test, *Review of Financial Studies* 29, 3068–3107.

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak, 1964, Measuring utility by a single-response sequential method, *Behavioral Science* 9, 226–232.

Bernheim, Douglas, and Charles Sprenger, 2020, On the empirical validity of cumulative prospect theory: Experimental evidence of rank-independent probability weighting, *Econometrica* 88, 1363–1409.

Bhui, Rahul, and Samuel Gershman, 2018, Decision by sampling implements efficient coding of psychoeconomic functions, *Psychological Review* 125, 985–1001.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2012, Salience theory of choice under risk, *Quarterly Journal of Economics* 127, 1243–1285.

Brown, Gordon D. A., and Jing Qian, 2004, The origin of probability weighting: A psychophysical approach, Working paper.

Buckley, Paul B., and Clifford B. Gillman, 1974, Comparisons of digits and dot patterns, *Journal of Experimental Psychology* 103, 1131–1136.

Burke, Christopher J., Michelle Baddeley, Philippe N. Tobler, and Wolfram Schultz, 2016, Partial adaptation of obtained and observed value signals preserves information about gains and losses, *Journal of Neuroscience* 36, 10016–10025.

Camerer, Colin F., and Teck-Hua Ho, 1994, Violations of the betweenness axiom and nonlinearity in probability, *Journal of Risk and Uncertainty* 8, 167–196.

Clarke, Sam, and Jacob Beck, 2021, The number sense represents (rational) numbers, *Behavioral and Brain Sciences* 44, e178.

Conen, Katherine E., and Camillo Padoa-Schioppa, 2019, Partial adaptation to the value range in the macaque orbitofrontal cortex, *Journal of Neuroscience* 39, 3498–3513.

De Giorgi, Enrico G., and Shane Legg, 2012, Dynamic portfolio choice and asset pricing with narrow framing and probability weighting, *Journal of Economic Dynamics and Control* 36, 951–972.

Dehaene, Stanislas, 1992, Varieties of numerical abilities, *Cognition* 44, 1–42.

Dehaene, Stanislas, 2011, *The Number Sense* (Oxford University Press, Oxford, United Kingdom).

DeWolf, Melissa, Margaret A Grounds, Miriam Bassok, and Keith J Holyoak, 2014, Magnitude comparison with different types of rational numbers, *Journal of Experimental Psychology: Human Perception and Performance* 40, 71–82.

Ebert, Sebastian, and Philipp Strack, 2015, Until the bitter end: On prospect theory in a dynamic context, *American Economic Review* 105, 1618–1633.

Eckert, Johanna, Josep Call, Jonas Hermes, Esther Herrmann, and Hannes Rakoczy, 2018, Intuitive statistical inferences in chimpanzees and humans follow Weber's law, *Cognition* 180, 99–107.

Enke, Benjamin, and Thomas Graeber, 2022, Cognitive uncertainty, Working paper.

Frydman, Cary, and Lawrence J. Jin, 2022, Efficient coding and risky choice, *Quarterly Journal of Economics* 137, 161–213.

Girshick, Ahna R., Michael S. Landy, and Eero P. Simoncelli, 2011, Cardinal rules: Visual orientation perception reflects knowledge of environmental statistics, *Nature Neuroscience* 14, 926–932.

Gonzalez, Richard, and George Wu, 1996, Curvature of the probability weighting function, *Management Science* 42, 1676–1690.

Gonzalez, Richard, and George Wu, 1999, On the shape of the probability weighting function, *Cognitive Psychology* 38, 129–166.

Heimer, Rawley, Zwetelina Iliewa, Alex Imas, and Martin Weber, 2021, Dynamic inconsistency in risky choice: Evidence from the lab and field, Working paper.

Heng, Joseph, Michael Woodford, and Rafael Polanía, 2020, Efficient sampling and noisy decisions, *eLife* 1–49.

Jacob, Simon N., and Andreas Nieder, 2009, Notation-independent representation of fractions in the human parietal cortex, *Journal of Neuroscience* 29, 4652–4657.

Kahneman, Daniel, and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, *Econometrica* 47, 263–291.

Kallai, Arava Y., and Joseph Tzelgov, 2009, A generalized fraction: an entity smaller than one on the mental number line, *Journal of Experimental Psychology: Human Perception and Performance* 35, 1845–1864.

Kallai, Arava Y., and Joseph Tzelgov, 2012, When meaningful components interrupt the processing of the whole: The case of fractions, *Acta Psychologica* 139, 358–369.

Khaw, Mel Win, Ziang Li, and Michael Woodford, 2021, Cognitive imprecision and small-stakes risk aversion, *Review of Economic Studies* 88, 1979–2013.

Khaw, Mel Win, Ziang Li, and Michael Woodford, 2022, Cognitive imprecision and stake-dependent risk attitudes, NBER working paper No. 30417.

Kőszegi, Botond, and Matthew Rabin, 2007, Reference-dependent risk attitudes, *American Economic Review* 97, 1047–1073.

Laughlin, Simon, 1981, A simple coding procedure enhances a neuron's information capacity, *Zeitschrift fur Naturforschung* 36, 9–10.

Matthews, Percival G., and Dana L. Chesney, 2015, Fractions as percepts? exploring cross-format distance effects for fractional magnitudes, *Cognitive Psychology* 78, 28–56.

Meerta, Gaëlle, Jacques Grégoire, and Marie-Pascale Noël, 2010, Comparing 5/7 and 2/9: Adults can do it by accessing the magnitude of the whole fractions, *Acta Psychologica* 135, 284–292.

Meerta, Gaëlle, Jacques Grégoire, Xavier Seron, and Marie-Pascale Noël, 2012, The mental representation of the magnitude of symbolic and nonsymbolic ratios in adults, *Quarterly Journal of Experimental Psychology* 65, 702–724.

Moyer, Robert S., and Thomas K. Landauer, 1967, Time required for judgments of numerical inequality, *Nature* 215, 1519–1520.

O'Donoghue, Ted, and Jason Somerville, 2018, Modeling risk aversion in economics, *Journal of Economic Perspectives* 32, 91–114.

Oprea, Ryan, 2022, Simplicity equivalents, Working paper.

Payzan-LeNestour, Elise, and Michael Woodford, 2022, Outlier blindness: A neurobiological foundation for neglect of financial risk, *Journal of Financial Economics* 143, 1316–1343.

Prelec, Drazen, 1998, The probability weighting function, *Econometrica* 66, 497–527.

Rottenstreich, Yuval, and Christopher K. Hsee, 2001, Money, kisses, and electric shocks: On the affective psychology of risk, *Psychological Science* 12, 185–268.

Schneider, Michael, and Robert S. Siegler, 2010, Representations of the magnitudes of fractions, *Journal of Experimental Psychology: Human Perception and Performance* 36, 1227–1238.

Siegler, Robert S., Clarissa A. Thompson, and Michael Schneider, 2011, An integrated theory of whole number and fractions development, *Cognitive Psychology* 62, 273–296.

Steiner, Jakub, and Colin Stewart, 2016, Perceiving prospects properly, *American Economic Review* 106, 1601–1631.

Stewart, Neil, Nick Chater, and Gordon Brown, 2006, Decision by sampling, *Cognitive Psychology* 53, 1–26.

Szkudlarek, Emily, and Elizabeth M. Brannon, 2021, First and second graders successfully reason about ratios with both dot arrays and arabic numerals, *Child Development* 92, 1011–1027.

Tversky, Amos, and Daniel Kahneman, 1992, Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty* 5, 297–323.

Viscusi, W Kip, 1989, Prospective reference theory: Toward an explanation of the paradoxes, *Journal of Risk and Uncertainty* 2, 235–263.

Wei, Xue-Xin, and Alan A. Stocker, 2015, A bayesian observer model constrained by efficient coding can explain 'anti-bayesian' percepts, *Nature Neuroscience* 18, 1509–1517.

Woodford, Michael, 2012a, Inattentive valuation and reference-dependent choice, Working paper.

Woodford, Michael, 2012b, Prospect theory as efficient perceptual distortion, *American Economic Review Papers and Proceedings* 102, 41–46.

Zhang, Hang, and Laurence T. Maloney, 2012, Ubiquitous log odds: a common representation of probability and frequency distortion in perception, action, and cognition, *Frontier in Neuroscience* 6, 1–14.

Zhang, Hang, Xiangjun Ren, and Laurence T. Maloney, 2020, The bounded rationality of probability distortion, *Proceedings of the National Academy of Sciences* 117, 22024–22034.

Zimmermann, Jan, Paul W. Glimcher, and Kenway Louie, 2018, Multiple timescales of normalized value coding underlie adaptive choice behavior, *Nature Communications* 9, 3206.

Zou, Deqiang, Gordon D. A. Brown, Ping Zhao, and Songting Dong, 2008, The shape of probability weighting function: Findings in a binary comparison, *Journal of Marketing Science* 4, 56–68.

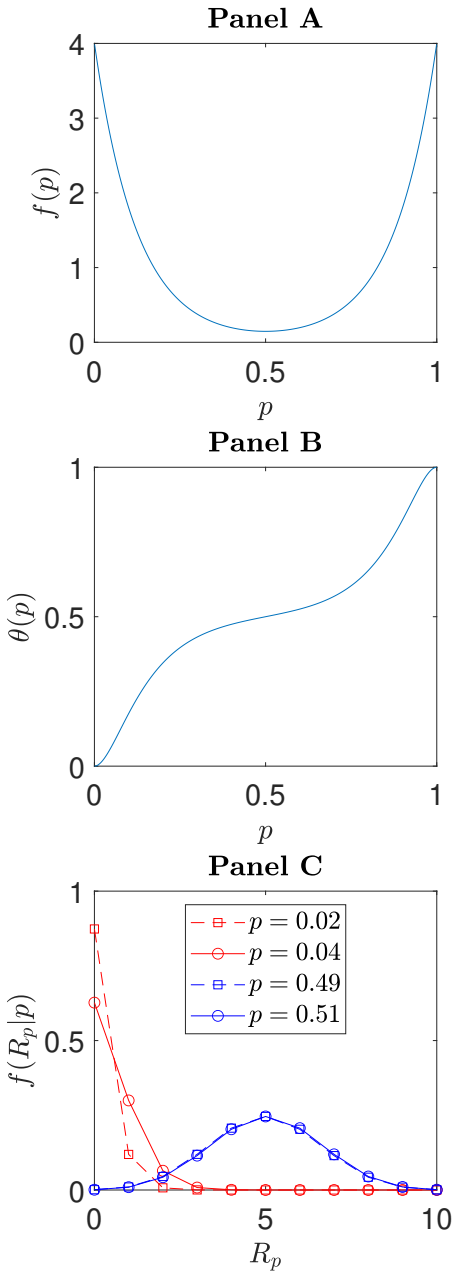


Figure I
Prior distribution, coding rule, and optimal likelihood functions

Panel A plots a U-shaped prior distribution in the form of (8) described in the main text; the parameter values are: $\lambda_1 = \lambda_2 = 8$ and $w = 0.5$. Panel B plots the coding rule $\theta(p)$, defined in equation (3), for the prior distribution from Panel A. Panel C plots the implied likelihood function $f(R_p|p)$, defined in equation (1), for four different probability values: $p = 0.02$, $p = 0.04$, $p = 0.49$, and $p = 0.51$; the parameter n is set to 10.

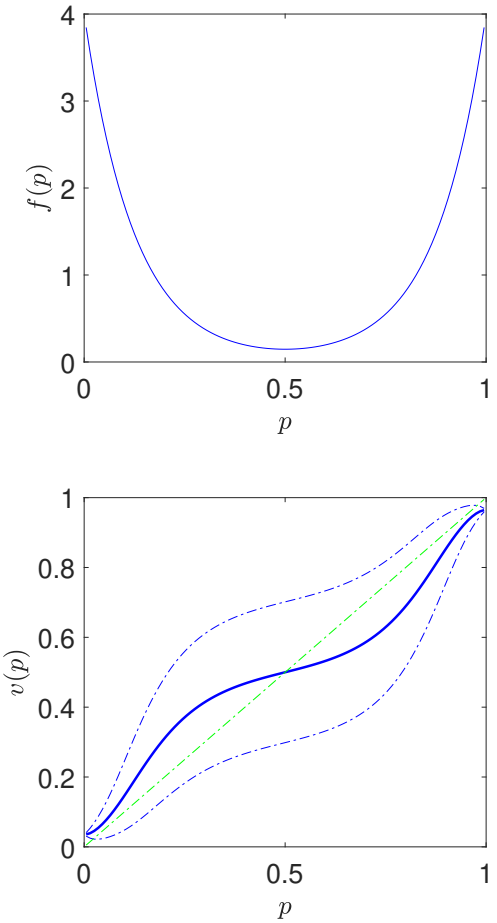


Figure II
Prior distribution and value function: Efficient coding and U-shaped prior

The upper graph plots a U-shaped prior distribution in the form of (8) described in the main text; the parameter values are: $\lambda_1 = \lambda_2 = 8$ and $w = 0.5$. The lower graph plots the subjective valuation implied by efficient coding, $v(p)$, and its one-standard-deviation bounds $v(p) \pm \sigma(p)$; when computing $v(p)$ and $\sigma(p)$, we use the coding rule from equation (3) in the main text and we set the parameter n to 10. The green dash-dot line is the forty-five degree line.

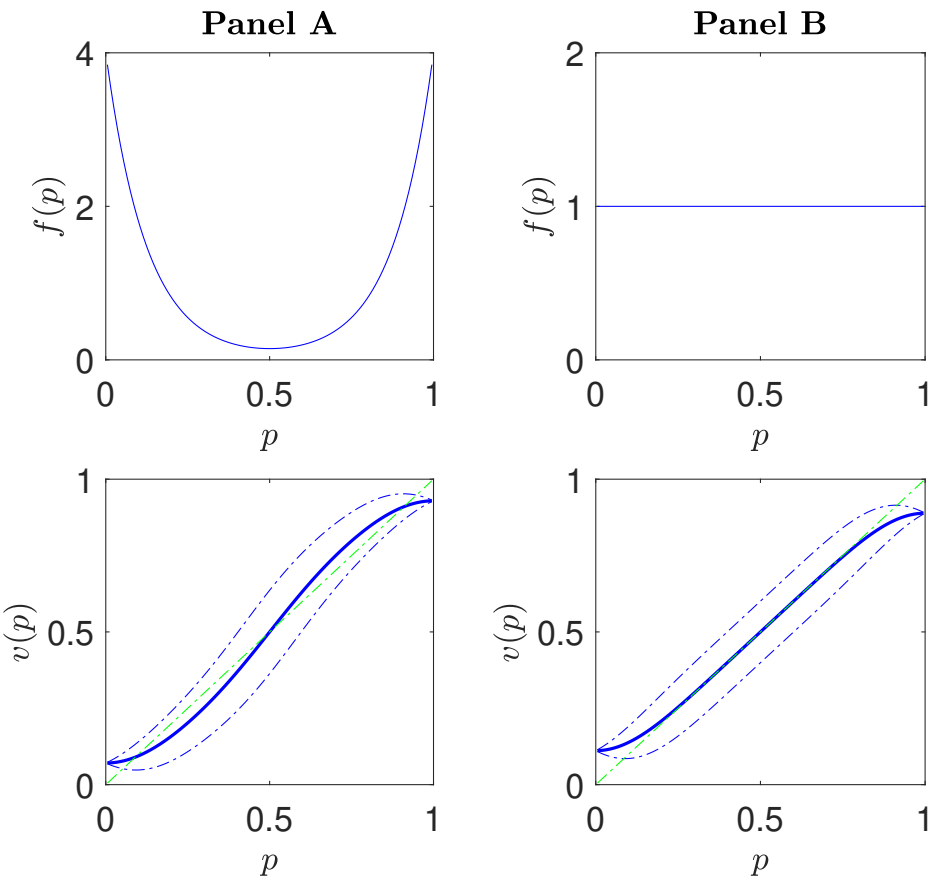


Figure III
Model implications when relaxing the efficient coding or U-shaped prior assumptions

Panel A: the upper graph plots a U-shaped prior distribution in the form of (8) described in the main text; the parameter values are: $\lambda_1 = \lambda_2 = 8$ and $w = 0.5$. The lower graph plots the subjective valuation implied by inefficient coding, $v(p)$, and its one-standard-deviation bounds $v(p) \pm \sigma(p)$; when computing $v(p)$ and $\sigma(p)$, we use the coding rule from equation (9) in the main text. Panel B: the upper graph plots a uniform prior distribution $f(p) = 1$. The lower graph plots the subjective valuation implied by efficient coding, $v(p)$, and its one-standard-deviation bounds $v(p) \pm \sigma(p)$; when computing $v(p)$ and $\sigma(p)$, we use the coding rule from equation (3) in the main text. For both panels, we set the parameter n to 10. In the lower graph of each panel, the green dash-dot line is the forty-five degree line.

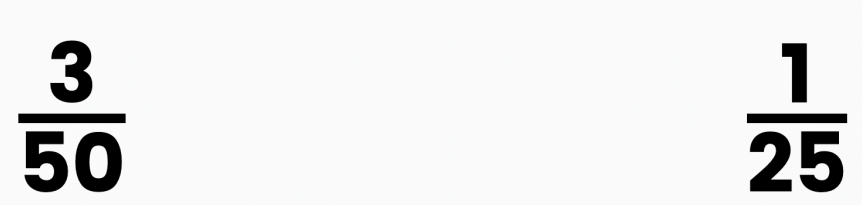


Figure IV
Screenshot from discrimination task

Subjects are incentivized to quickly and accurately select the greater of the two displayed fractions. Subjects input their response on the keyboard using either the "Q" key (for left) or the "P" key (for right).

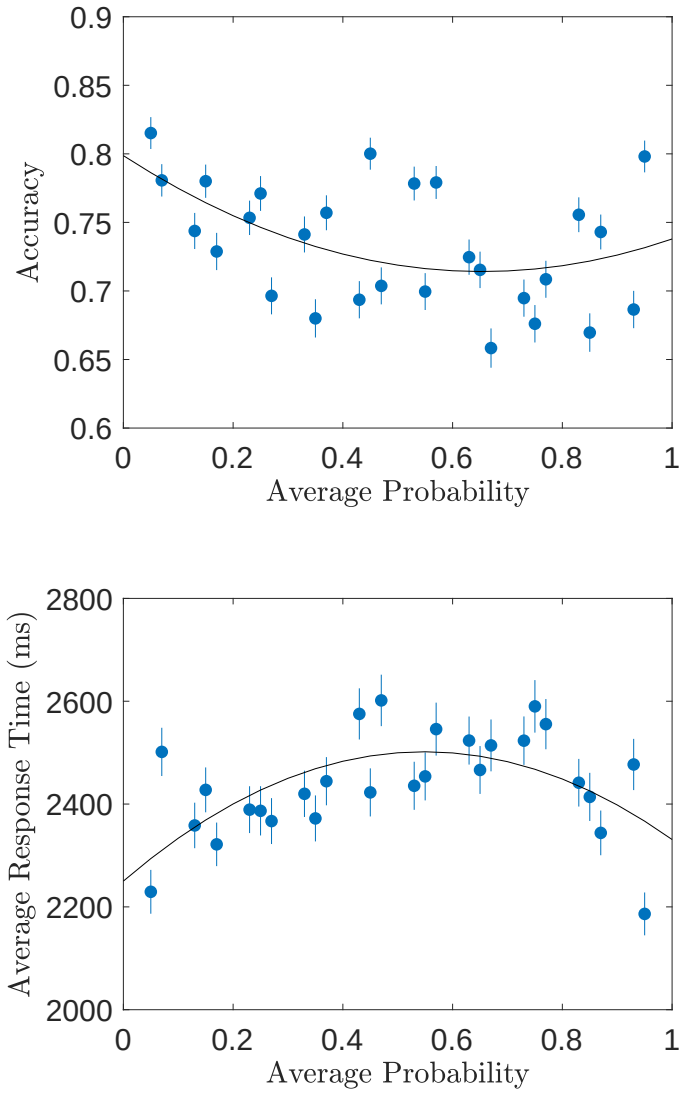


Figure V
Accuracy and response time data from the fraction discrimination task

The upper panel plots the average accuracy rate for classification among test trials. The lower panel plots the average response time among test trials. The x-axis denotes the average probability of the two fractions on a given trial. The smoothed black line represents the best fitting second order polynomial (fitted at the individual trial level). The blue vertical bars denote two standard errors of the mean for each level of average probability.

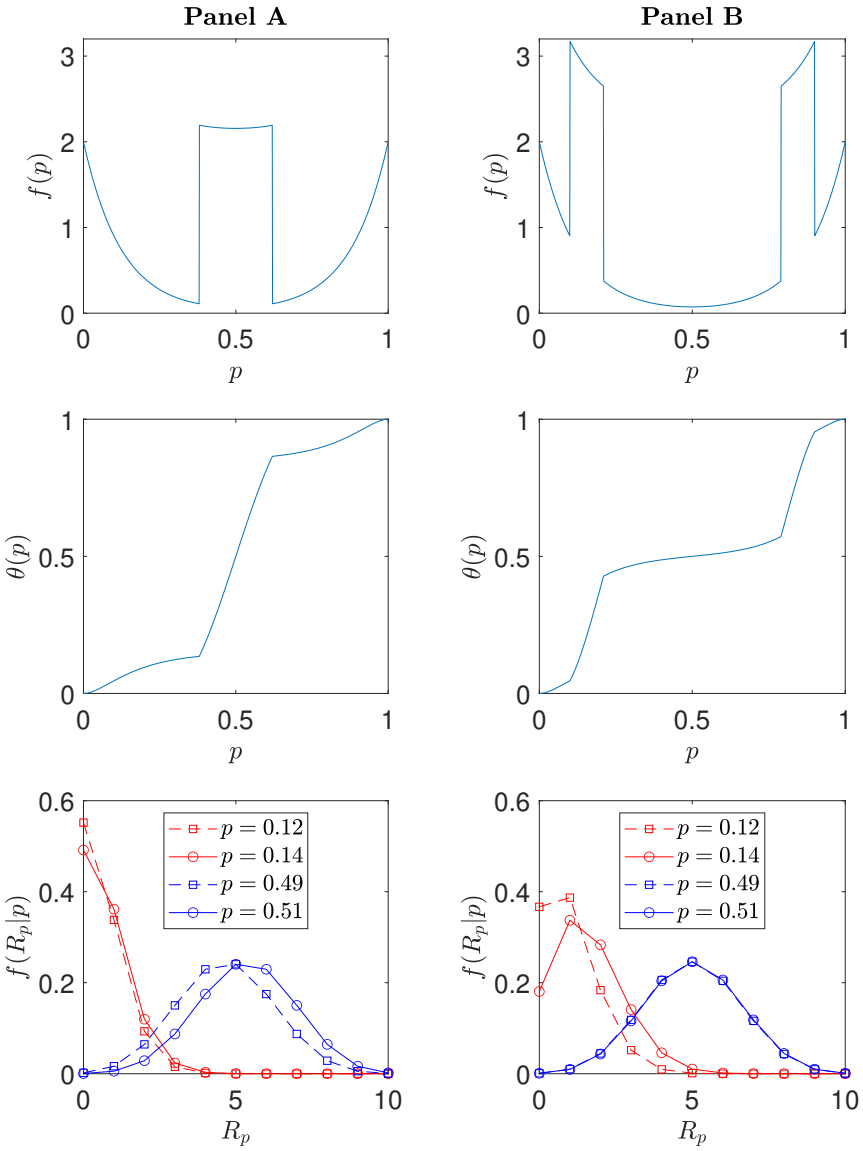


Figure VI

Prior distribution, coding rule, and optimal likelihood functions: Efficient coding and mixed prior

Panel A: the upper graph plots the mixed prior distribution in the form of (10) for the intermediate condition; the middle graph plots the corresponding coding rule $\theta(p)$; the lower graph plots the implied likelihood function, for $p = 0.12$, $p = 0.14$, $p = 0.49$, and $p = 0.51$. Panel B: the upper graph plots the mixed prior distribution in the form of (10) for the extreme condition; the middle graph plots the corresponding coding rule $\theta(p)$; the lower graph plots the implied likelihood function, for $p = 0.12$, $p = 0.14$, $p = 0.49$, and $p = 0.51$. For both conditions, the first component of the mixed prior is U-shaped in the form of (11); the parameter values are: $\lambda_1 = \lambda_2 = 8$ and $w = 0.5$. For the intermediate condition, we set $p_l = 0.38$ and $p_h = 0.62$; for the extreme condition, we set $p_{l,1} = 0.1$, $p_{l,2} = 0.21$, $p_{h,1} = 0.79$, and $p_{h,2} = 0.9$. For both panels, we set $\xi = 0.5$ and $n = 10$.

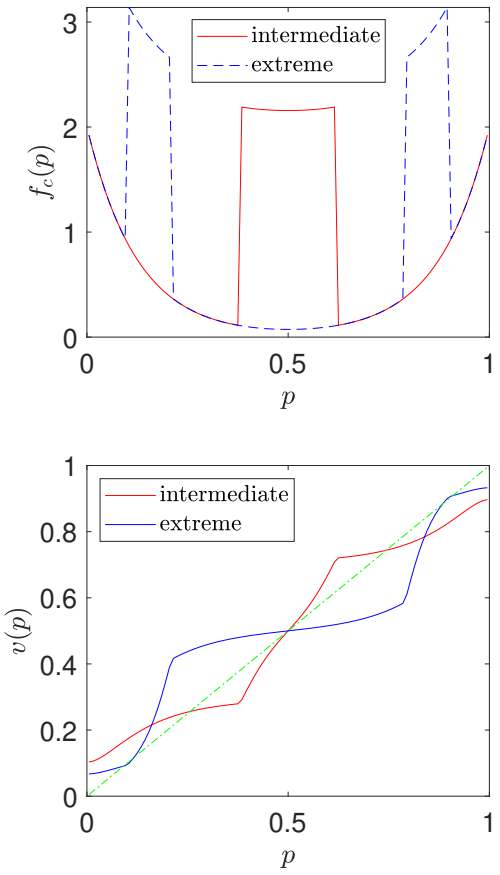


Figure VII
Prior distribution and value function: Efficient coding and mixed prior

The upper graph plots a mixed prior distribution $f_c(p)$ in the form of (10) described in the main text. The first, stable component takes the form of (11); the parameter values are: $\lambda_1 = \lambda_2 = 8$ and $w = 0.5$. The second, fast-moving component takes the form of (12) in the "intermediate" condition and takes the form of (13) in the "extreme" condition; the parameter values are: $p_l = 0.38$, $p_h = 0.62$, $p_{l,1} = 0.1$, $p_{l,2} = 0.21$, $p_{h,1} = 0.79$, and $p_{h,2} = 0.9$. The weight ξ the DM assigns to the stable component is set to 0.5. The lower graph plots, for both the "intermediate" condition and the "extreme" condition, the subjective valuation implied by efficient coding, $v(p)$; when computing $v(p)$, we set the parameter n to 10. The green dash-dot line is the forty-five degree line.

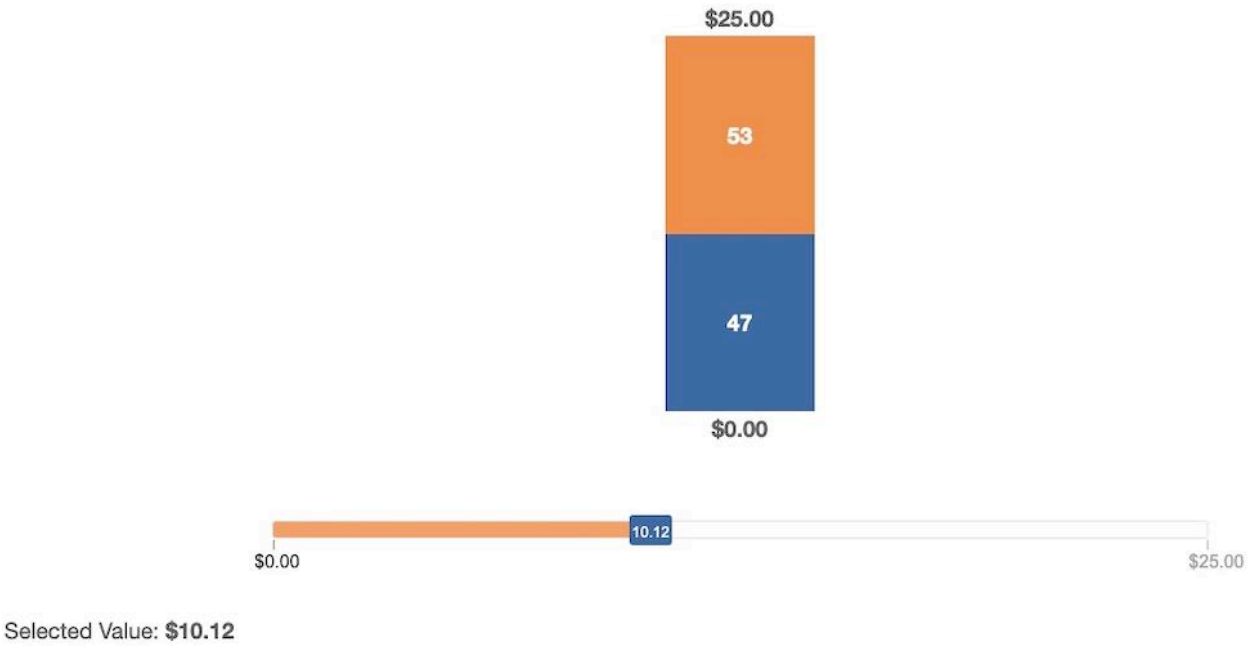


Figure VIII
Screenshot of trial from Experiment 2

In this example, the risky lottery pays $25 with 53% probability and $0 with 47% probability. The slider position indicates that the subject's reported certainty equivalent on this trial is $10.12.

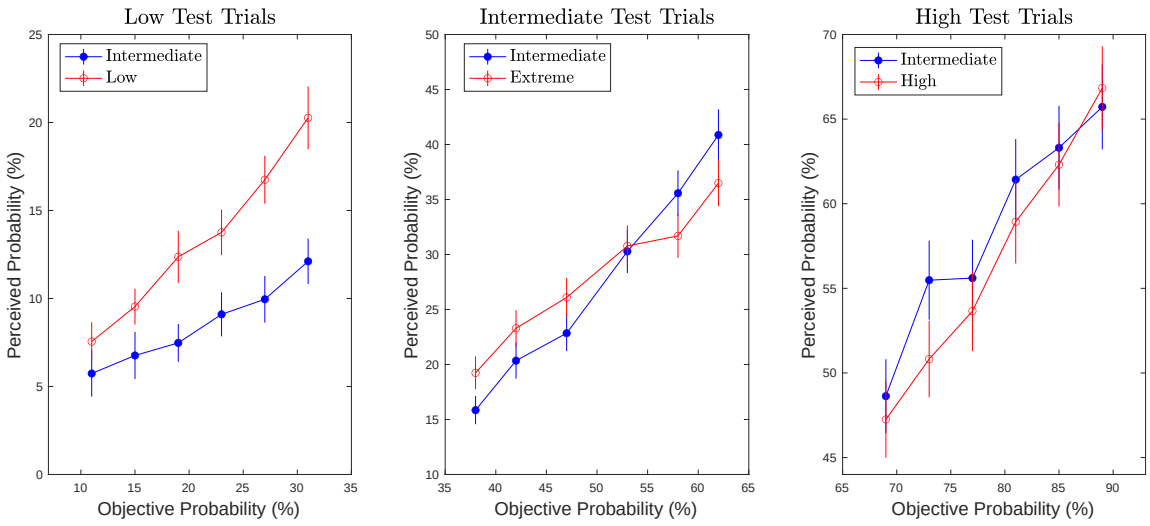


Figure IX
Perceived vs. objective probability

Each panel displays probabilities for one of the three sets of test trials. The legend in the upper left of each panel denotes the adaptation condition. Vertical bars denote two standard errors of the mean.

Table I
Fraction pairs in Experiment 1

Test trials				Catch trials			
Fraction 1 num	Fraction 1 denom	Fraction 2 num	Fraction 2 denom	Fraction 1 num	Fraction 1 denom	Fraction 2 num	Fraction 2 denom
1	25	3	50	3	45	1	20
2	25	3	50	2	9	3	8
3	25	7	50	7	55	3	20
4	25	7	50	5	6	3	8
4	25	9	50	9	46	4	21
6	25	11	50	3	7	2	7
6	25	13	50	13	55	6	21
7	25	13	50	5	9	3	7
8	25	17	50	17	45	8	21
9	25	17	50	5	9	2	7
9	25	19	50	19	55	9	20
11	25	21	50	3	5	3	8
11	25	23	50	23	45	11	20
12	25	23	50	7	8	2	9
13	25	27	50	27	55	13	20
14	25	27	50	5	6	7	8
14	25	29	50	29	45	14	21
16	25	31	50	2	9	5	9
16	25	33	50	33	55	16	21
17	25	33	50	5	6	3	4
18	25	37	50	37	45	18	19
19	25	37	50	3	7	5	6
19	25	39	50	39	55	19	20
21	25	41	50	7	8	3	7
21	25	43	50	43	45	21	25
22	25	43	50	5	6	3	7
23	25	47	50	47	55	23	25
24	25	47	50	2	7	7	8

Notes. Each row of the table characterizes a pair of fractions used in the discrimination task from Experiment 1. The left panel provides the 28 test trials; the right panel provides the 28 catch trials. "Fraction 1 num" and "Fraction 1 denom" denote the numerator and denominator of the first fraction. "Fraction 2 num" and "Fraction 2 denom" denote the numerator and denominator of the second fraction. We counterbalance the left-right ordering of the two fractions so that each subject is presented with a total of 56 test trials.

Table II
Accuracy and response times in Experiment 1

	All trials		Only trials with two-digit numerators	
	(1)	(2)	(3)	(4)
Dependent variable:	Correct	Response time (seconds)	Correct	Response time (seconds)
extreme	0.026***	–0.112***	0.011**	–0.152***
	(0.004)	(0.013)	(0.005)	(0.018)
Constant	0.723***	2.493***	0.718***	2.534***
	(0.007)	(0.040)	(0.007)	(0.041)
Observations	43,833	43,833	26,597	26,597

Notes. The table reports results from mixed effects linear regressions. *Correct* is a dummy variable that takes the value of one if the subject correctly classifies the larger fraction on a given trial. *extreme* takes the value of one if both fractions on the trial are less than 0.2 or greater than 0.8. Only test trials are included in the regression; catch trials are not included. We include a random intercept which allows accuracy to vary across subjects. Standard errors of the fixed effect estimates are clustered at the subject level and reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Table III
Design parameters for Experiment 2

Adaptation conditions				Test conditions		
Extreme	Low	Intermediate	High	Low	Intermediate	High
10	10	38	67	11	38	69
11	11	39	68	15	42	73
12	12	40	69	19	47	77
13	13	41	70	23	53	81
14	14	42	71	27	58	85
15	15	43	72	31	62	89
16	16	44	73			
17	17	45	74			
18	18	46	75			
19	19	47	76			
20	20	48	77			
21	21	49	78			
79	22	51	79			
80	23	52	80			
81	24	53	81			
82	25	54	82			
83	26	55	83			
84	27	56	84			
85	28	57	85			
86	29	58	86			
87	30	59	87			
88	31	60	88			
89	32	61	89			
90	33	62	90			

Notes. The table provides the specific values of probability used in each of the adaptation and test conditions. Each adaptation condition contains 24 trials, corresponding to the four columns on the left. Each test condition contains 6 trials, corresponding to the three columns on the right. Each entry in the table denotes a probability p (in percentage) associated with the upside payoff of \$25. Subjects who are randomized into the low test condition are further randomized into either the low or intermediate adaptation condition; subjects who are randomized into the high test condition are further randomized into either the high or intermediate adaptation condition; subjects who are randomized into the intermediate test condition are further randomized into either the extreme or intermediate adaptation condition.

Table IV
Perceptual biases of probability

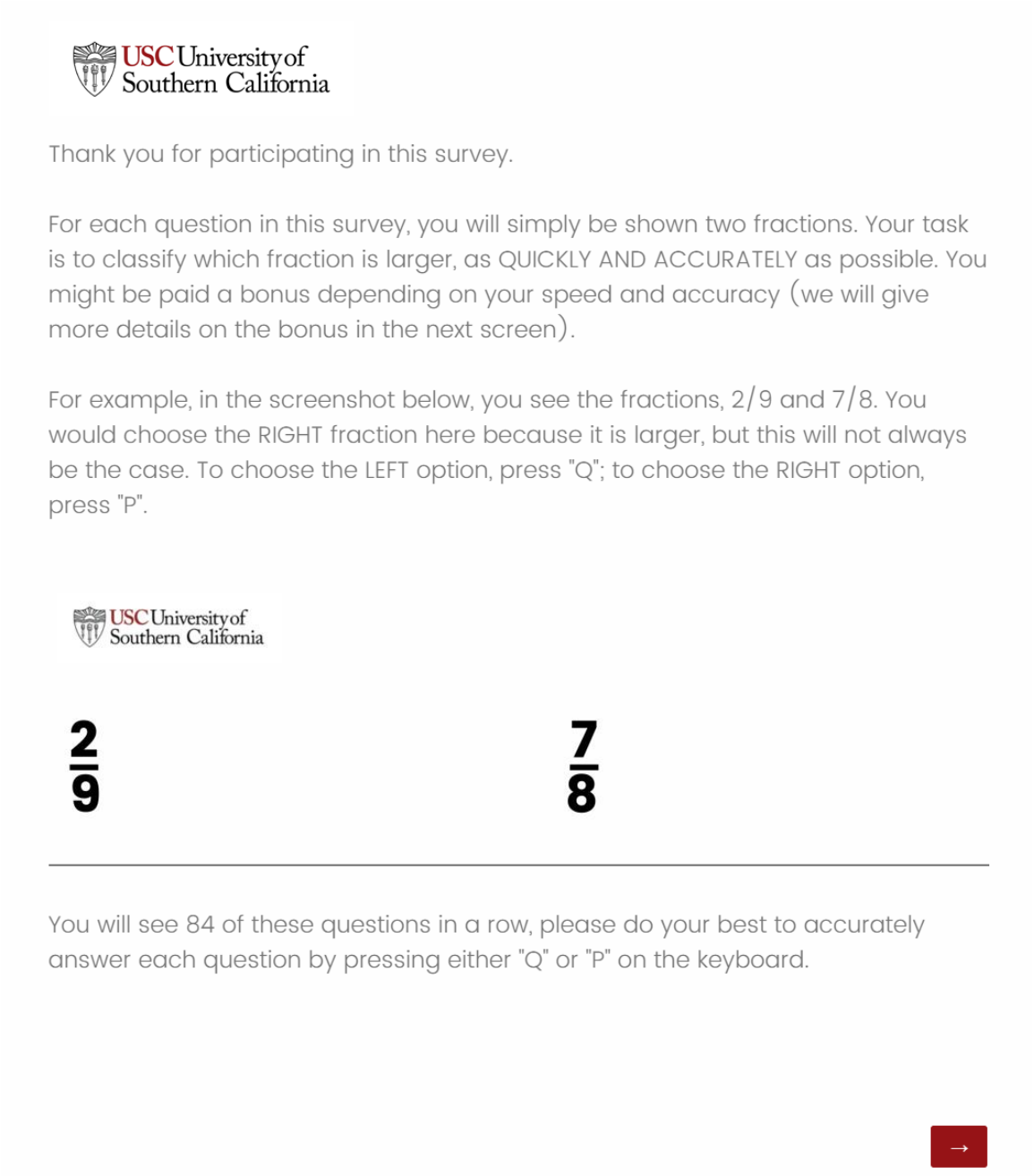
	(1)	(2)	(3)
Dependent variable: "Perceived probability"	Low test trials sample	Intermediate test trials sample	High test trials sample
p	0.620***	0.664***	0.983***
	(0.060)	(0.059)	(0.061)
intermediate	1.710	–18.531***	14.857**
	(1.901)	(4.092)	(6.745)
$p \times$*intermediate*	–0.312***	0.367***	–0.166**
	(0.070)	(0.094)	(0.084)
Constant	0.352	–5.279*	–21.057***
	(1.217)	(2.722)	(4.969)
Observations	1,127	1,096	1,133

Notes. The table reports results from mixed effects linear regressions in which the dependent variable is the perceived probability, estimated from the certainty equivalent on each trial. The dummy variable, *intermediate*, takes the value of one if the trial belongs to the intermediate adaptation condition, and zero otherwise. The variable p takes the objective value of the probability associated with the risky lottery's upside payoff. Only data from test trials are included. There are random effects on the independent variables p and the intercept. Standard errors of the fixed effect estimates are clustered at the subject level and reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Online Appendix

A. Experimental Instructions

A.1. Instructions for the discrimination task in Experiment 1



USC University of Southern California

Bonus Payment

1 out of every 10 survey respondents will be selected to receive a bonus payment. The bonus payment is $0.10 for every correct answer, but we will reduce this amount by $0.01 for every second it takes you to respond. Therefore, if you answered all 84 questions correctly, and it took you an average of 4 seconds to respond to each question, and you were randomly selected to receive the bonus payment, **you would receive a bonus of $8.40 - $3.36 = $5.04.** If instead, you only answered 50 out of the 84 questions correctly, and it took you an average of 4 seconds to respond, and you were randomly selected to receive a bonus, **you would receive a bonus of $5.00 - $3.36 = $1.64.** [Note that we impose a minimum bonus of $0.]

→

USC University of Southern California

You are now ready to begin. Remember, you will see 84 of these questions, and your task is to CHOOSE THE LARGER FRACTION.

Press Q for LEFT and P for RIGHT.

→

A.2. Instructions for the valuation task in Experiment 2

USC University of Southern California

Instructions (Page 1 of 3)

Please read these instructions carefully. There will be a short comprehension quiz at the end of the instructions, which you need to pass before continuing to the study.

The study consists of approximately 30 rounds. In each round, you will see a gamble which pays either $25 or $0, each with a different probability. For example, if you see the following gamble:



this means that the gamble pays $25 with 53% chance, and $0 with 47% chance. Importantly, there are no right or wrong answers in this study, please answer all questions depending on your personal opinions.

→

USC University of Southern California

Instructions (page 2 of 3)

On each round, you will be asked to tell us what amount of money -- in your opinion -- the gamble is worth to you. In order to help you form your opinions, think about the amount of money as the maximum price you'd be willing to pay to play the gamble. We will ask you to enter your decision on each round using a slider that goes from $0 (on the left) to $25 (on the right), like in the example below.



Selected Value: **$10.12**

In the above example, the person indicated that a 53% chance of winning $25 is worth $10.12 to them. In other words, the maximum amount they'd be willing to pay to play the gamble is $10.12. We emphasize that different people will give different answers to the same question, since answers depend on your own opinions. There are no right or wrong answers in this survey.

→

USC University of Southern California

Instructions (Page 3 of 3)

10% of all participants will be chosen to receive a bonus, in addition to the $3 for completing the survey.
We have developed the bonus payment method in a way that makes it optimal for you to report your true valuation of the gamble. If you don't care about the exact mechanism that we use to compute the bonus payment, you can feel free to skip the rest of the details on this screen (the comprehension quiz will not ask about how bonuses are computed). If you would like to know exactly how we determine the bonus, then please read below.

Details on how bonus is computed:

As you know, you will see a new gamble on each round and you will be asked to give a certain payment amount that is worth as much as the gamble to you. We will use your response to fill in the table below which actually contains 2,500 questions (for example, see the table below which corresponds to a round in which the gamble offers a 53% chance of winning $25 and 47% chance of winning $0). Question #1 asks if you'd rather have the gamble or $0.01 with certainty. Question #2 asks if you'd rather have the gamble or $0.02 for with certainty. Question #2,500 asks if you'd rather have the gamble or $25.00 with certainty.

Question #		Option A		Option B
1	Would you rather have:		or	$0.01 with certainty
2	Would you rather have:		or	$0.02 with certainty
3	Would you rather have:		or	$0.03 with certainty
4	Would you rather have:	With probability **53%: Get $ 25** With probability **47%: Get $ 0**	or	$0.04 with certainty
…	…		…	…
2,499	Would you rather have:		or	$24.99 with certainty
2,500	Would you rather have:		or	$25.00 with certainty

Once we get your response on the slider, we will fill in the answer to all 2,500 questions as follows. Suppose your response on the slider is denoted by "x". Then, we will assume you would have chosen Option A for all questions for which the certain amount in Option B is less than or equal to x. We will also assume you would have Chosen Option B for all questions for which the certain amount in Option B is greater than x.

If you are chosen to receive a bonus, then we will then randomly pick one of the 2,500 questions (from one randomly chosen round) and pay you according to what you chose on that one question. Each question and each round is equally likely to be chosen for payment. If you chose Option B on the randomly selected question, you would be paid the certain amount. If you chose Option A on the randomly selected question, then the gamble would be played out by the computer. Using the above table as an example, if you chose Option A then you would receive $25 with 53% chance and $0 with 47% chance.

Comprehension Quiz

The following two questions test your understanding of the instructions. Please answer them correctly in order to continue to the survey. Suppose that someone is presented with the gamble below and answers with the following slider location.



Selected Value: **$7.43**

1. Which one of the following statements is correct?

- The person values the gamble at exactly $7.43.
- The person values the gamble at more than $7.43
- The person values the gamble at less than $7.43

2. Which one of of the following statements is correct?

- The person would rather have $4.24 with certainty than play the gamble
- The person would rather have $6.43 with certainty than play the gamble
- The person would rather have $8.34 with certainty than play the gamble

→



You have successfully completed the comprehension quiz.

To begin the survey, please click the button below. Remember to answer each question carefully as you might be chosen to receive a bonus according to your answer on that question.

→