Neural and behavioral probability weighting function

Agnieszka Tymula[1]**, Hiroshi Yamada**[2]

Recent theoretical models challenge the existence of a probability weighting function as it was traditionally conceived in Prospect Theory in ways that are not straightforward to test using choice data. This study transcends these constraints by directly observing probability distortions in the brain, free from utility confounds. Utilizing a unique dataset comprising 64,175 decision trials and 78,067 neural measurement trials, we pinpoint neural activity (a basic biological decision processing unit) that exclusively encodes probability, independent of payoff magnitudes. Our results demonstrate that neural probability weighting functions diverge from those estimated behaviorally under conventional assumptions. Furthermore, incorporating a biologically realistic utility function enhances our ability to reconstruct neural probability weighting from observed choices, offering direct biological evidence on the bases of economic decision-making.

Keywords: neuroeconomics, monkey, decision making under risk, probability

[1] corresponding author: School of Economics, University of Sydney, Sydney, 2006 NSW, Australia; agnieszka.tymula@sydney.edu.au
[2] Division of Biomedical Science, Institute of Medicine, University of Tsukuba
1-1-1 Tenno-dai, Tsukuba, Ibaraki, 305-8577 Japan

Introduction

The probability weighting function, alongside the reference-dependent utility and discount function, is a cornerstone of the behavioral economics toolkit. In their seminal paper on Prospect Theory, Kahneman & Tversky (1979) proposed that in decisions involving small stakes and risk, people do not perceive probabilities objectively. Instead, they assign "decision weights" to probabilities using an inverse S-shaped probability weighting function. This explanation rapidly got adopted into both theoretical and empirical social science research. Probability weighting has attracted significant and sustained attention – a vast amount of social science research utilizes the concepts of probability weighting to inform theoretical and empirical frameworks, influencing behavior-altering policies in finance, environmental conservation, health, social sectors, and beyond. This key behavioral phenomenon is taught to all students in economics, psychology, finance, and marketing.

Despite much progress, obtaining consensus regarding the shape of the probability weighting function remains challenging. A major obstacle preventing researchers from precisely describing the probability weighting function is that the probability weighting and utility functions are not directly and separately observable in standard economic data. Instead, they are typically estimated from observable decisions. A typical task to elicit probability weighting involves people choosing between lotteries with varying payoffs and probabilities. Then the probability weighting and utility functions are usually jointly estimated from choice. Since the two functions are assumed to multiply, the specific functional assumptions about one function can dramatically affect parameter estimates of both the utility and probability weighting functions. Moreover, it is usually assumed that utility of a given monetary payoff is fixed and independent of the other payoffs available to the chooser in the choice set. This assumption contradicts both neural and behavioral evidence that shows that utility is influenced by the choice set (Khaw et al., 2017a; Louie & Glimcher, 2012b; Padoa-Schioppa, 2009), and could bias both the utility and probability weighting function estimates. Finally, several recent theoretical papers demonstrated that behaviors that were originally attributed by Kahneman and Tversky (1979) to probability weighting, can be explained by context-dependent utility instead (Glimcher & Tymula, 2023; Kontek & Lewandowski, 2018; Schneider & Day, 2018). Remarkably, probability weighting functions estimated from observed choice, vary considerably between the individuals. A sizeable proportion of people (20% or more) seem to perceive probabilities objectively (Bruhin et al., 2010; Conte et al., 2011; Hey & Orme, 1994; Tversky & Kahneman, 1992). The remaining 80% can be far from being universally well-described by the inverse S-shaped probability weighting function (Abdellaoui, 2000;

Abler et al., 2006; Bruhin et al., 2010; Harbaugh et al., 2002; Wilcox, 2015). Moreover, one challenge for human experimental data is that it may not always contain enough decisions to separate probability weighting and utility to provide reliable estimates on the subject level.

In this study, we observe probability weighting directly in the brain without confounding it with the utility function, and then compare it to the probability weighting function estimated from behavior using a large number of decisions. Our unique dataset consists of 64,175 decisions made by two monkeys and 78,067 single-lottery trials in which monkeys did not make any decisions and their brain activity was recorded while they anticipated the receipt of a lottery. Our experimental design overcomes the shortcomings of previous studies in several ways. First, in our study, brain activity is recorded using electrodes implanted directly into our subjects' brains. This means that we observe neural response to our task stimuli with zero noise.[1] Second, on each trial, we randomly select what lotteries to present to the subject from a set of 100 lotteries that were constructed by crossing ten payoffs and ten probability magnitudes. This allows us to eliminate any correlation in lotteries across and within trials. This also means that the payoff and probability magnitudes in our study are orthogonal eliminating unintentional correlation between payoff and probability magnitudes introduced by design (e.g. by pairing large amounts with small probabilities). Additionally, we only record brain activity when a single lottery is presented to the subject, meaning that choice set effects do not bias the neurally observed probability weighting. This means that we observe probability distortions perceived by the subjects in their pure form rather than a signal that could be explained by a context-dependent utility or some unaccounted for distributional features of payoffs in our experiment. Fourth, and the most importantly, our design allows us to identify neurons that encode information about probability but not about the payoff magnitude, thus providing a clean measurement of probability weighting function in the brain. Finally, the combination of neural measurement trials with behavioral choice trials allows us to stress check whether the neural signal that we use to identify probability weighting is meaningful in the sense that it can, out-of-sample, predict behavior.

While the use of non-human subjects is not standard in economics, many studies provided evidence that non-human primates are a good model for human behavior (Kagel et al., 1995; Platt & Glimcher, 1999; Tymula et al., 2023; Yamada et al., 2013). Conducting studies with monkeys offers several unique advantages. Firstly, many decisions can be obtained for reliable and precise estimates of

---

[1] Electrophysiology is superior to fMRI in the quality of neural recordings due to higher temporal and spatial resolution. fMRI is also susceptible to artifacts related to motion, physiological processes, and scanner noise.

probability weighting from behavior, which is not feasible with human participants in a laboratory setting. Consequently, human studies tend to produce noisier estimates than desired. Monkeys in our study perform the task over many months, yielding a dataset of 64,175 decisions allowing us to estimate structural models with high precision. Secondly, experimental economists often question whether the relatively modest laboratory incentives motivate participants enough to elicit their true preferences as well as about the external factors unobservable to the experimenter. In our study, we have full control over the monkeys' economy and manage their hydration status (i.e., physical wealth in water content) throughout the study. The subjects are rewarded with beverage lotteries, and the liquid they earn while performing the task is their only source of hydration. They are also compensated for every decision they make. Overall, this places them in an environment with more consequential decisions spread over a longer period compared to a traditional economic experiment. Finally, the direct neural recordings that provide a much better signal-to-noise ratio are not available for human studies except in very specific patient populations performed by a neurosurgeon as a part of medical treatment (Kandel et al., 2021). Although previous studies have estimated probability weighting functions from the behavior of non-human primates (Farashahi et al., 2018; Ferrari-Toniolo et al., 2019; Fujimoto & Minamimoto, 2019; Imaizumi et al., 2022; Stauffer et al., 2015; Tymula et al., 2023), no study to date has attempted to isolate it from neural activity. In our previous study (Imaizumi et al., 2022), we have shown that by aggregating neural activity throughout the reward circuitry, we could reconstruct from this neural activity the risk preferences estimated from monkeys' behavior, however, in that study, we did not attempt to measure probability distortions separately from utility and instead we estimated them jointly, as has typically been done in the economics literature. This study leverages the methodological benefits of non-human primate studies, to precisely isolate probability weighting function in neural activity.

In this paper, we use single neuron activity, which is the biological basis of the brain function, and ask whether the brain is capable of encoding probability separately from utility when subjects evaluate the utility of lotteries. If the information about both payoff and probability magnitudes is always coded together even at the level of single neurons, any attempt to separate the two functions using single-neuron recordings would be futile and would not offer much advantage over behavioral estimation. Fortunately, we find that some neurons in our dataset encode probability only (and not payoff), some encode payoff only (and not probability), and some encode both probability and payoff magnitudes. Importantly, these different types of neurons are distributed across all brain regions we record from and co-exist in relative proximity to one another. This indicates that using fMRI, another method frequently used to record human brain activity that lacks the resolution of single-neuron recordings, it

would be nearly impossible to isolate the probability weighting function from the utility function and to gain a new understanding of how our neural valuation system processes probabilities.

Next and most importantly, we use our neural data to illustrate the shape of the probability weighting function. We find that for both subjects, the neural probability weighting is S-shaped. This contrasts with the probability weighting function estimated from the same subjects' behavior. We investigate two possible explanations for this discrepancy. First, we check whether the neural signal is predictive of choice. If it is not, then it is not surprising that the neural and behaviorally estimated probability weighting functions are distinct. Our analysis does not support this explanation because our neural data accurately predicts choice. Second, we examine whether the restrictive assumptions on the functional form of the utility function in the behavioral estimation might contaminate the estimated probability weighting functions. Recent theoretical papers (Bordalo et al., 2012; Glimcher & Tymula, 2023; Kontek & Lewandowski, 2018; Schneider & Day, 2018) have demonstrated that when the utility function depends on the current and past choice sets this can lead to behaviors traditionally explained using probability weighting functions, even when the modeled choosers perceive probabilities objectively without distortions. As a result, we find that using a more realistic context-dependent utility function indeed improves the match between the neural and behaviorally estimated probability weighting and enhances the accuracy of predicting our subjects' choices, highlighting the importance of having more accurate models of behavior.

In the following sections, we describe the experimental design, present the results, and then conclude.

Experimental design

The subjects in our study were two rhesus monkeys, named SUN and FU. They participated in two types of tasks, both involving lotteries but differentiated by the presence or absence of choice. In the "single cue no choice task," a lottery was presented without an alternative, whereas in the "choice task," the subjects chose between two lotteries presented on the screen. The single cue no choice task was used for examining neural probability weighting function. Since the subjects had no use for money, their rewards were given in the form of juice/water, delivered after each trial. We designed one hundred lotteries by crossing ten payoff magnitudes (ranging from 0.1 to 1.0 mL in 0.1 mL increments) with ten probability levels (from 0.1 to 1.0 in 0.1 increments). The lotteries' payoff and probability magnitudes were visually communicated using pie charts: the payoff magnitude was

indicated by the number of green segments, and the probability level by the number of blue segments. Figure 1 illustrates an example trial, where the subject is choosing between a lottery on the left that pays 0.5 mL of water or nothing each with a 50% chance and a lottery on the right that pays 0.2 mL of water with a 90% chance and nothing with 10% chance. After ten months of intensive training, the animals demonstrated a robust understanding of the tasks and stimuli (Imaizumi et al., 2022; Tymula et al., 2023). The data analyzed in this study was collected only after the completion of the training phase, ensuring the subjects' proficient comprehension of the stimuli. The monkeys performed these tasks five days a week to earn liquid rewards.
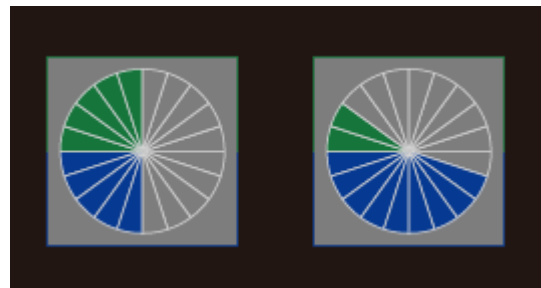
Figure 1. Example trial in choice task. The number of green pie segments represents payoff size in (each green pie = 0.1 mL) and the number of blue segments represents the probability (each blue pie = 10%).

Single cue no-choice task. To initiate each trial, monkeys had two seconds to align their gaze to the gray central fixation target. After fixation for one second, one pie chart providing information on the probability and magnitude of one lottery was presented for 2.5 seconds at the same location as the central fixation target. During these 2.5 seconds, brain activity was recorded. We calculate each neuron's activity (firing rate, FR) as the sum of the occurrences of action potentials[2] that it produced during the 2.5-second presentation of the lottery before the payoff is received (i.e., the frequency of the single neuron activity). The pie chart was then removed, and subjects received a liquid payoff, as indicated by the number of green pie chart segments, with the probability indicated by the number of blue pie chart segments. After 4–6 seconds, the next trial began. On each trial, one lottery was randomly selected, with replacement, from the set of one hundred possible lotteries. Overall, SUN and FU completed 38,678 and 39,389 single cue no-choice task trials respectively during which we measured the brain activity in 686 neurons, located across four brain regions associated with the

---

[2] An action potential is a brief electrical impulse that serves as the primary means of communication throughout the nervous system. It's a fundamental element of neural activity, allowing neurons (nerve cells) to transmit signals over long distances within the body.

computation of value: dorsal (DS) and ventral (VS) striatum, central part of orbitofrontal cortex (cOFC), and medial orbitofrontal cortex (mOFC). See Table 1 for the number of neurons whose activity was recorded in each region in each subject. Single cue no choice task trials were presented in blocks of 100-120 trials.

Table 1. The number of neurons whose activity was recorded by brain region and subject.

	monkey		
brain region	SUN	FU	Total
DS	98	96	**194**
VS	89	55	**144**
cOFC	98	92	**190**
mOFC	64	94	**158**
Total	**349**	**337**	**686**

Choice task. As in the single cue task, at the beginning of each trial, subjects had two seconds to align their gaze to the gray central fixation target. After fixation for one second, two lotteries, randomly selected from the set of one hundred, were presented to monkeys as pie charts for 2.5 seconds. After 2.5 seconds, two smaller gray circles appeared at the same location as lotteries together with a fixation target in the middle. After a 0.5-second delay, the fixation target disappeared. Subjects were then allowed 2 seconds to make their choice by shifting their gaze to the lottery that they preferred. Then, they received a reward according to the payoff and probability magnitudes they selected. After 4–6 seconds another trial began. Subjects completed approximately 30 to 60 trials of the choice task before switching to the single cue no-choice task or ending the experimental session. Our dataset includes 44,883 decisions made by monkey SUN (obtained in 884 blocks spread over 242 days) and 19,292 decisions made by monkey FU (obtained in 571 blocks spread over 127 days). During choice task neural activity was not recorded. Further details on methods, including details that would be particularly of interest to neuroscientists, are provided in (Imaizumi et al., 2022).

Results

Identification of probability and payoff neurons

To identify neurons that respond to probability and/or payoff magnitude, we divided our one hundred

lotteries into categories of high (>50%) and low (≤50%) probability of receiving the payoff, and high (>0.5ml) versus low (≤0.5ml) payoff magnitude. For each neuron, we conducted two t-tests. The first test compared its activity in trials featuring lotteries with high versus low probability, and the second test compared its activity in trials with high versus low payoff lotteries. We defined neurons as probability neurons and payoff neurons if their respective t-tests were significant at the 5% level. This method was chosen over linear regression as it does not assume a specific functional form for probability representation in the brain[3]. It only requires that the average activity between low and high-probability trials is significantly different, which should hold under any pattern of probability coding. Probability-only neurons, if they exist, are particularly interesting as their activity in response to probability is not confounded by payoff magnitude.

Our initial findings, presented in Figure 2, show that both probability and payoff neurons are distributed across all four valuation regions of the brain. None of these regions is exclusively dedicated to either probability or payoff magnitude. This suggests that if the human brain shares this property, fMRI studies, which measure aggregated activity per voxel (a unit typically consisting of about one million neurons), may not be able to isolate the neural representation of probability weighting independent of the utility function. In our sample, 22.6% of neurons encode probability exclusively, 18.7% encode payoff only, and 14.0% encode both probability and payoff. The discovery of neurons sensitive exclusively to probability is crucial because it allows us to observe how probability is encoded in these neurons without interference from value or expected value signals. Thus, direct recordings from single neurons provide a unique opportunity to precisely measure how a chooser processes objective probability information. In the main paper, we present findings from analyses that include only neurons significantly responsive to probability but not to payoff magnitude. Like prior studies (Enomoto et al., 2020; Yamada et al., 2013, 2018), we observe that some neurons respond positively and others negatively to payoffs. Similarly, across all four brain regions, we identify neurons that are significantly more active, as well as those that are significantly less active, in high versus low probability trials. In the paper, we include probability only neurons that respond positively and negatively to probability. In the appendix, we further validate the robustness of our results through supplementary analyses that incorporate all neurons that positively respond to probability.

---

[3] Previous studies in neuroeconomics, often used a linear regression to identify neurons sensitive to payoff magnitude. In the appendix, we show that this type of selection makes little difference to our results. For our purposes, our selection procedure allows us to remain agnostic about the functional form.
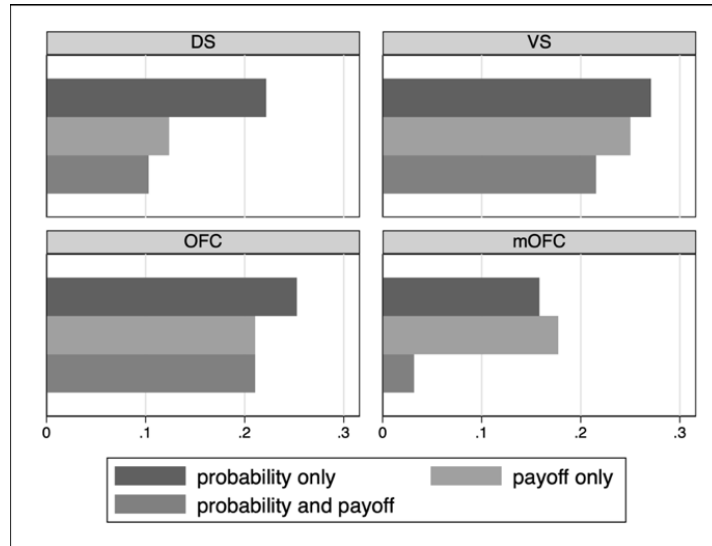
Figure 2. **Proportion of neurons in our sample that encode probability only (dark gray), payoff only (light gray), and both probability and payoff (bottom bar)** across dorsal striatum (DS), ventral striatum (VS), central orbitofrontal cortex (cOFC) and medial orbitofrontal cortex (mOFC).

Our probability and payoff neurons are interesting to economists only if they predict choice. We find that their activity correctly predicts about 74-87% of decisions. To calculate this percentage, for each of the 100 lotteries we calculated its subjective value. Since neural activity is observable, there is no need for estimation. Simply, for each of the 100 lotteries, and for each subject separately, we calculate the average neural activity across all instances when this lottery was presented. We base this measure on the activity of all neurons that either coded probability or payoff magnitude (not both). The activity of neurons that have a negative relationship with probability and/or payoff magnitudes is subtracted. To make sure that all neurons whose activity was measured in the study have the same impact on the subjective value, we first calculate the average activity for each lottery within a neuron and then average it across neurons. Using the calculated subjective values, we predicted that subjects would choose the lottery with the higher subjective value in each trial. The prediction accuracy was remarkably high, with subjective values predicting 87.22% of choices correctly for SUN and 87.74% for FU. This accuracy is notable for several reasons. Firstly, we are only recording from a subset of neurons in the brain's value regions. The ability of this subset to predict out-of-sample behaviors so accurately is impressive. Secondly, the context of presenting two lotteries at a time likely influences the subjective value in ways not detectable in neural data from single-lottery presentations. Lastly, these predictions are out-of-sample.
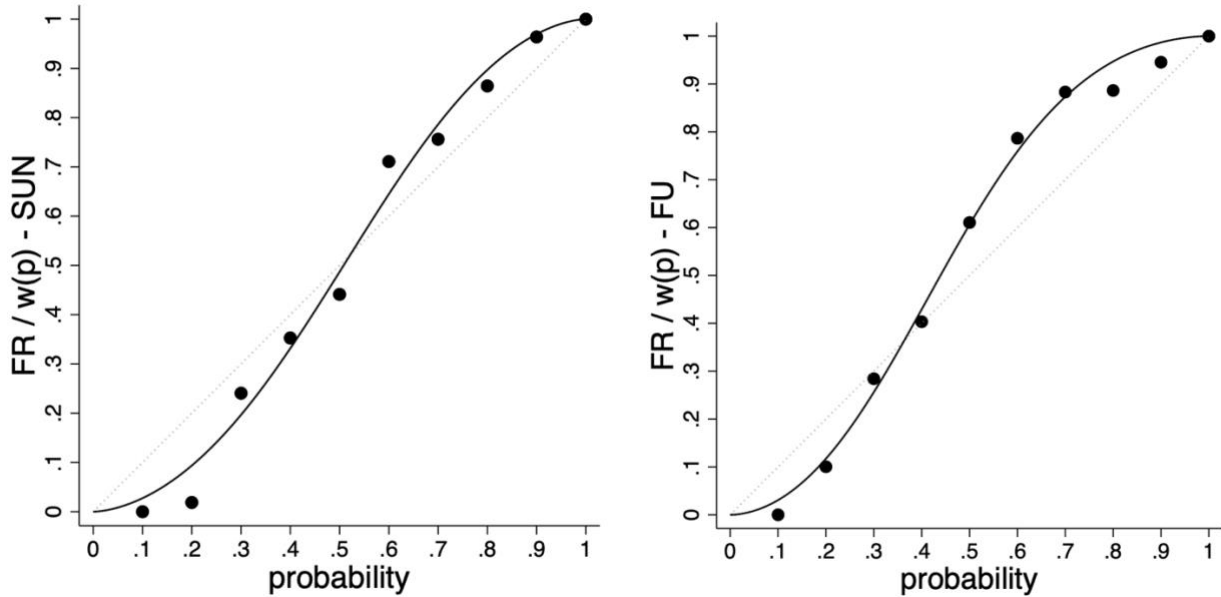
Neural probability weighting function

We have established that our neural measurements can predict choices and that there are neurons specifically sensitive to probability whose activity is not confounded by value signals. In Figure 3, we present the neural probability weighting function. This figure is based on the activity of neurons that solely encode probability and do not encode payoff, providing the cleanest measure of brain's response to probability. This dataset comprises 17,958 neural measurements (9,304 for SUN and 8,654 for FU) collected from 155 neurons (82 for SUN and 73 for FU). To construct Figure 3, we first calculate the average activity of each neuron at various probability levels. We then compute the overall average activity at each probability level across all neurons, incorporating neurons that negatively respond to probability with a negative sign. This method ensures that all neurons contribute equally, regardless of the number of measurements recorded from each. To facilitate a direct comparison between probability weighting functions measured behaviorally and in the brain function, we transform this neural measure to a scale of [0,1] using the formula for the neurobiological probability response ratio (NPRR) previously used by Berns (2008):

$$FR_{norm} = \frac{FR - FR_{min}}{FR_{max} - FR_{min}}$$

where FR is the actual average firing rate at a given probability level, FR_{min} is the lowest and FR_{max} is the highest average firing rate across all probability levels. We find that the neural probability weighting in each monkey is slightly S-shaped. We confirm this by fitting FR_{norm} with a two-parameter probability weighting function [4] , $w(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma}$, using nonlinear least-squares estimation via nl command in Stata (see Table 2 "neural" for the parameter estimates). The γ parameter, which gives the function S-shape does not statistically differ between the subjects. The δ parameter is significantly higher for FU ($p < 0.001$) which captures that the fixed point, $w(p) = p$, occurs for lower probability for FU.

---

[4] In Tymula et al. (2023), we conducted a horserace between the expected value, expected utility, and prospect theory models with four different probability weighting functions using the same behavioral dataset. We used power utility function and various probability weighting functions (Goldstein & Einhorn, 1987; Kahneman & Tversky, 1979b; Prelec, 1998). We found that prospect theory models with a two-parameter probability weighting function provided the best fit to behavioral data according to the Bayesian Information Criterion (BIC) (see Fig 2 in Tymula et al. (2023)). Following this finding, here we use the two-parameter probability weighting function (Goldstein & Einhorn, 1987).

Figure 3. **Neural probability weighting function.** Includes data from all neurons that significantly respond to probability but do not respond to payoff magnitude. The dots represent FR_{norm} and the solid curves are Goldstein and Einhorn (1987) two-parameter probability weighting functions that best fit the neural data.

In the Appendix, we conduct several robustness checks to validate our findings through three alternative approaches. First, we analyze the data using only those probability-only neurons that show a positive response to increasing probability, excluding any neurons whose activity decreases with probability magnitude (Figure S1). Second, we broaden our dataset to include all neurons that react to probability magnitude, encompassing those that also respond to both probability and payoff magnitudes (Figure S2). Third, we employ a different classification method for neurons; here, each neuron is classified as coding for probability and/or payoff magnitude based on whether its activity significantly ($p < 0.05$) correlates with either lottery probability or payoff magnitude. This classification is determined by a linear regression model ($FR = b_0 + b_1 * payoff + b_2 * probability$) run separately for each neuron (Figure S3). The results from these varied approaches are qualitatively similar, reinforcing the consistency of our original findings (Figures S1 - S3). These additional analyses enhance our confidence in the robustness of our results and that our original observations are not an artifact of the specific subset of neurons, or the method of analysis used in the main study.

Comparison of neural and behavioral probability weighting functions

Our second key question investigates whether the neural probability weighting functions are similar to those estimated using standard behavioral economics methods. To address this, we use data from the choice task that contains the same subjects' decisions but made in different blocks of the same experimental sessions. Using this data, we estimate probability weighting functions from behavior within a random utility framework. A lottery, denoted as $L(m,p)$ represents a gamble that pays m with probability p, and 0 otherwise. We adopt a popular two-parameter probability weighting function (Goldstein & Einhorn, 1987):

$$w(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma}$$

For the utility function, we use a power utility function, commonly utilized in economic literature:

$$u(m) = m^r$$

The expected utility of the lottery is then calculated as $U(m,p) = w(p)u(m)$. The probability that participants choose the lottery on the right side (L_R) over the one on the left (L_L) is estimated using a logistic choice function, $P(L_R) = 1/(1 + e^{-Z})$ where $Z = \frac{V(L_R)-V(L_L)}{\beta}$ and free parameter β captures the degree of stochasticity observed in choice. We fit the data by maximizing the log-likelihood of the observed choices. All estimations are done in Stata 18.

In Figure 4, just as in Figure 3, we replot the observed neural probability weighting with black dots, and the two-parameter probability weighting function that best fits the neural data is depicted as a black solid line. The black longdash line represents the probability weighting function estimated from behavior using the above model with power utility function. Visually, Figure 4 indicates that the behavioral probability weighting function (longdash) substantially differs from both the observed neural activity (dot) and the estimated (solid line) neural probability weighting functions. One can further validate these differences by comparing each subject's probability weighting parameters estimated from the neural data against those derived from behavioral measurements, see Table 2 under the headings 'neural' versus 'power'. For SUN, δ and γ parameters estimated from neural activity and behavior are significantly different at 0.001 and 0.1 significance levels respectively. For FU, δ and γ parameters estimated from neural activity and behavior are significantly different at 0.01 and 0.001 significance levels respectively. These findings suggest that the observed differences in neurally measured and behaviorally estimated probability weighting necessitate exploring alternative models.
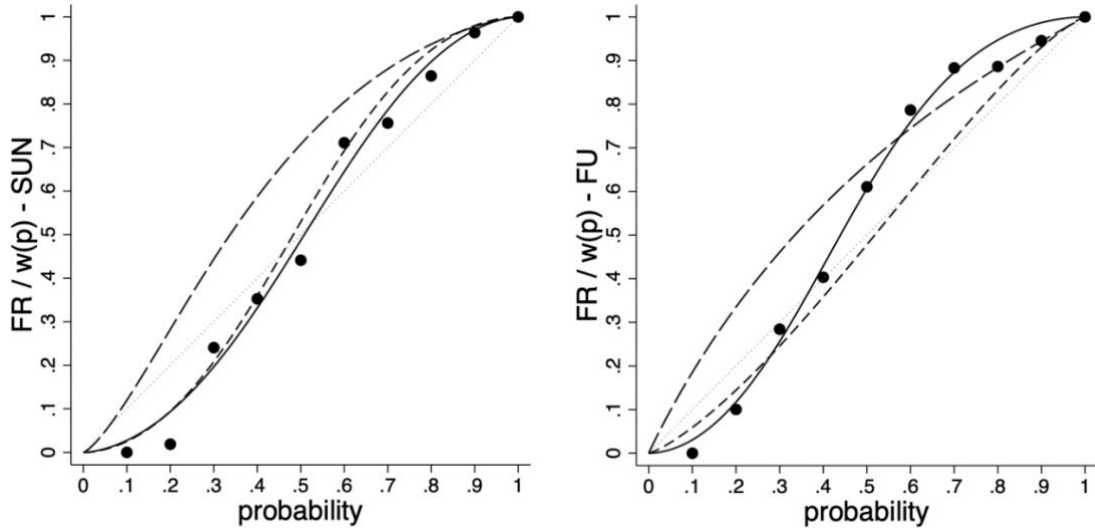
Figure 4. **Neural and behavioral probability weighting functions.** Includes data from neurons that respond to probability but not to payoff magnitude. The dots are FR_{norm} and the solid curve is Goldstein and Einhorn (1987) probability weighting function fit to neural data. The long dash and short dash curves are Goldstein and Einhorn (1987) probability weighting functions that best fit behavioral data estimated with the power and divisive normalization utility function, respectively.

Table 2. **Neural and behavioral estimates of two-parameter Goldstein and Einhorn (1987) probability weighting functions**. For the probability weighting parameters, r, and alpha, stars indicate significant differences from 1. For the rest of the parameters from 0. Neural is the best fit for the neural data, power are the behavioral estimates using the power utility function, and ESVT are behavioral estimates using the ESVT utility function.

	SUN			FU		
	neural	power	ESVT	neural	power	ESVT
			probability weighting			
δ	0.9488	2.4187***	1.1245+	1.5392***	1.9634***	0.9155
	(0.0909)	(0.0549)	(0.0648)	(0.1182)	(0.0754)	(0.0851)
γ	1.5990**	1.3137***	1.7152***	1.7677***	0.9818	1.2191***
	(0.1524)	(0.0220)	(0.0452)	(0.1248)	(0.0232)	(0.0434)
			utility parameters			
r		0.8787***			0.5776***	
		(0.0156)			(0.0186)	
α			1.6571***			1.0455
			(0.0709)			(0.0689)
ω			0.0928***			0.0704***
			(0.0072)			(0.0091)
β/ε		0.0647***	0.0149***		0.0636***	0.0220***
		(0.0008)	(0.0020)		(0.0011)	(0.0034)
N	10	44883	44883	10	19292	19292
BIC		19677	19493		8680	8621

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Next, we investigate whether using a neurobiologically realistic utility function could improve the accuracy of recovering neural probability weighting from subjects' decisions and improve choice prediction. Empirical studies (e.g., Yamada et al. (2023)) and theoretical work (Glimcher & Tymula, 2023; Kontek & Lewandowski, 2018; Schneider & Day, 2018) suggest that the assumptions about utility functions significantly influence estimates of probability weighting, and vice versa.[5] We chose to model utility using divisive normalization, a well-established neural computation found throughout the neural pathway from the retina to the cortex (Carandini & Heeger, 2012; Louie et al., 2011; Yamada et al., 2018). Divisive normalization has been theoretically demonstrated to be an efficient computation for encoding value (Steverson et al., 2019), especially with Pareto-distributed payoffs (Bucher & Brandenburger, 2022). Divisive normalization is interesting because it can account for many behaviors associated with Prospect Theory but without the probability weighting function or the discontinuity created by the loss aversion parameter (Glimcher & Tymula, 2023). It also effectively captures various choice set effects (Guo & Tymula, 2021; Khaw et al., 2017b; Landry & Webb, 2021; Louie et al., 2013; Webb, 2020; Webb et al., 2021), and recent research suggests that our brains may use this encoding mechanism even when it is not efficient (Kurtz David et al., 2023). A comprehensive summary of this model and its historical context is provided elsewhere (Carandini & Heeger, 2012; Glimcher, 2022; Glimcher & Tymula, 2023; Louie et al., 2015). Following Glimcher & Tymula (2023), we refer to the model with the ESVT acronym which stands for expected subjective value theory. In our ESVT model, the expected utility of lottery L_1 when choosing between L_1 and L_2 is:

$$EU(m_1, p_1; m_2, p_2) = \frac{w(p_1)m_1^{\alpha}}{\varepsilon + \omega(w(p_1)m_1^{\alpha} + w(p_2)m_2^{\alpha})}$$

where α is a parameter called predisposition. We follow Webb et al. (2021) and use parameter ω to nest the power utility model in ESVT. If we estimate that $\omega = 0$, we are back in the prospect theory framework with a power utility function where ε becomes the noise level (β in our previous behavioral model). If $\omega \neq 0$, then we conclude that the utility of the lottery L_1 is normalized by itself and the other elements of the choice set. The remaining details of the structural estimation are the same as before.

For both SUN and FU, we find that the parameter ω is significantly different than zero (Table 2). Moreover, the Bayesian Information Criterion is lower for the ESVT model than for the traditional power utility function model, indicating a better fit. Based on these results—significant parameter

---

[5] Abdellaoui (2000) designed a clever estimation method that overcomes this problem. Unfortunately, the method relies on the utility being independent of other alternatives in the choice set.

values of ω, slightly lower BIC—we conclude that behavioral data is more accurately estimated using the neurobiologically realistic ESVT utility function. Next, we check which model better approximates the probability weighting function estimated in the brain.

In Figure 4, we plot the probability weighting function estimated with the ESVT utility function as a curve with short-dash. Visually, for SUN, this curve closely aligns with the average neural probability weighting signal represented by black dots, indicating a near-perfect match. To formally assess how accurately the probability weighting functions estimated with ESVT and traditional power utility models approximate the neural measurements, we calculate residuals for each model. These residuals are determined by comparing each model's estimate of the probability weight to the averaged and normalized neural measurement at each probability level. We then compute the sum of square errors for these residuals. The results show that for both monkeys, the sum of square errors is lower for the ESVT model than for the power utility model, suggesting a better fit with the neural data. (0.028 vs. 0.002, p=0.015 for SUN and 0.016 vs. 0.009, p=0.468 for FU). These findings support the ESVT model's superior ability to replicate the neural probability weighting, particularly evident in the SUN dataset.

Discussion

The most striking finding of this paper is the persistence of the distortion in the neural signature of probability. Even though in our experimental design we eliminated some suggested reasons for why probability weighting occurs, we find that the probability weighting function is not linear. For instance, (Frydman & Jin, 2023) argue that individuals are more likely to encounter small or large probabilities, leading to an argument for an inverse S-shaped probability weighting function that helps to distinguish between more frequently occurring probabilities. However, in our experimental design, all probabilities are uniformly drawn from a range of 10% to 100%, in 10% increments, thus eliminating any advantage for steepening the probability weighting function within specific probability ranges. One might argue that perhaps life events external to the experiment could have influenced probability weighting measured in the experiment. Yet, this would contradict Frydman and Jin (2023)'s observation that individuals in natural settings are more inclined to experience extreme probabilities rather than 50-50 events. Herold & Netzer (2023) proposed that probability weighting arises as a secondary fitness-maximizing solution for individuals with nonlinear utility functions. However, in our design, the payoff magnitudes, just like probabilities, were drawn from a uniform distribution thus

eliminating fitness-maximizing reasons for non-linear utility. Furthermore, by pairing each payoff magnitude with each probability level, we ensure their orthogonality and eliminate the reasons for strategic adjustments of probability weighting to utility. Yet, despite these precautions, we consistently observe S-shaped neural probability weighting. Finally, the S-shaped probability weighting sometimes found in decisions from experience could be a result of a sampling error and thus imperfect understanding of probability information in studies that rely on small samples. We find that the S-shaped distortion persists even though our subjects had an extensive experience with the stimuli which included 10-months-long training before the data was collected for analysis and then performed the task daily for another couple of months. The S-shaped distortions that we document offer a new outlook on probability weighting which is particularly relevant given that the existing theoretical explanations of probability weighting largely focused on figuring out under what conditions probability weighting should be inverse S-shaped (Blavatskyy, 2007; Enke & Graeber, 2023; Herold & Netzer, 2023; Steiner & Stewart, 2016; Zhang et al., 2020) while it may be additionally worthwhile to consider the flexibility in the shapes of probability weighting as well (Bordalo et al., 2012; Frydman & Jin, 2023; Glimcher & Tymula, 2023).

Our paper represents a unique contribution. This is the first study to measure the brain's response to probability through single-neuron activity dedicated solely to probability, and not payoff magnitude. This is the cleanest and most direct measurement of how the brain encodes probability that is possible using current technology. Consequently, the probability weighting functions we document are directly observed, rather than estimated from behavior, and are not confounded by or entangled with utility. Demonstrating that such signature of probability weighting can be measured in the brain opens the door to answering more complex questions in the future, such as exploring how efficient coding constraints (Frydman & Jin, 2021; Glimcher, 2022; Glimcher & Tymula, 2023; Louie & Glimcher, 2012a; Polanía et al., 2019), perceptual factors (Oprea, 2022), and salience (Bordalo et al., 2012) relate to probability encoding in the brain. This direct neural measurement has enabled us to empirically demonstrate the importance of using a correct utility function when recovering probability weighting from behavior. Notably, for both subjects, the probability weighting estimated from behavior closely aligns with (and for one subject, essentially mirrors) the neural measurement when we model utility using the canonical, context-dependent utility model from neuroscience and neuroeconomics, rather than the traditionally used power utility function. This suggests that as our understanding of how the brain encodes value improves, we will become increasingly better at recovering probability weighting from behavior.

Obtaining such high quality neural (and behavioral) measurements has been possible because the subjects in our study are macaques. While rhesus macaques are not human, we share roughly 93% of our DNA sequences (Wang et al., 2014). Many studies demonstrated a close parallel between human and monkey behaviors and brain function. Economists many years before argued how economics can learn from animal studies (Kagel et al., 1995). This study is a perfect example that illustrates these advantages – we directly observe a noiseless probability weighting signal in the brain that is not accessible with choice data. We also observe a lot of decisions, ensuring that our estimated probability weighting from behavior is more reliable. Previous monkey studies estimated probability weighting from behavior (Farashahi et al., 2018; Ferrari-Toniolo et al., 2019; Fujimoto & Minamimoto, 2019; Imaizumi et al., 2022; Stauffer et al., 2015; Tymula et al., 2023) but none of them measured or reported probability weighting in single neuron activity. Additionally, our lotteries were constructed with the highest number of different probability levels in any monkey study so far, which enhances the precision of our measurement in both brain activity and behavior. The existing studies with monkey subjects were not conclusive about the shape of the probability weighting estimated from behavior but demonstrated that some features of how choices are presented affect the estimates, reinforcing the need for studies that establish how the perception of probabilities changes based on environment or experience (Ferrari-Toniolo et al., 2019; Tymula et al., 2023).

Kahneman and Tversky, when they conceived prospect theory in the 1970s, could only rely on observed choices to guide their assumptions about how probability and payoff value are encoded. Since then, science has made remarkable progress. We can now directly observe neuronal activity that guides choice using a range of tools. A series of studies in the 2000s set out to measure and describe the functional properties of the probability weighting function using non-invasive functional magnetic resonance imaging (fMRI) techniques that measure brain activity by tracking blood oxygenation levels throughout the brain (Abler et al., 2006; Berns et al., 2008; Hsu et al., 2009; Preuschoff et al., 2006; Tobler et al., 2008). Remarkably, these neuroeconomic studies generally agree that value-coding areas of the brain are capable of encoding probability, both with and without distortions. Consistent with our finding that probability is encoded by neurons across different value regions in the brain, none of these earlier fMRI studies pointed to a specific brain region that encodes probability alone. This means that measuring the probability signal separately from value is challenging. We overcome this shortcoming by using a direct and completely noiseless measurement of brain activity using electrodes implanted in the neurons of the brain while monkeys evaluate probabilistic lotteries, employing the largest number of probability levels used to date with monkey subjects.

Bibliography

Abdellaoui, M. (2000). Parameter-Free Elicitation of Utility and Probability Weighting Functions. *Management Science*, *46*(11), 1497–1512.

Abler, B., Walter, H., Erk, S., Kammerer, H., & Spitzer, M. (2006). Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *NeuroImage*, *31*(2), 790–795. https://doi.org/10.1016/j.neuroimage.2006.01.001

Berns, G. S., Capra, C. M., Chappelow, J., Moore, S., & Noussair, C. (2008). Nonlinear neurobiological probability weighting functions for aversive outcomes. *NeuroImage*, *39*(4), 2047–2057. http://linkinghub.elsevier.com/retrieve/pii/S1053811907009615

Blavatskyy, P. R. (2007). Stochastic expected utility theory. *Journal of Risk and Uncertainty*, *34*(3). https://doi.org/10.1007/s11166-007-9009-6

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience Theory of Choice Under Risk. *The Quarterly Journal of Economics*, *127*(3), 1243–1285. https://doi.org/10.1093/qje/qjs018

Bruhin, A., Fehr-Duda, H., & Epper, T. (2010). Risk and Rationality: Uncovering Heterogeneity in Probability Distortion. *Econometrica*, *78*(4), 1375–1412.

Bucher, S., & Brandenburger, A. (2022). Divisive Normalization is an Efficient Code for Multivariate Pareto-Distributed Environments. *Proceedings of the National Academy of Sciences*, *119*(40).

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

Conte, A., Hey, J., & Moffatt, P. (2011). Mixture Models of Choice Under Risk. *Journal of Econometrics*, *162*(1), 79–88.

Enke, B., & Graeber, T. (2023). Cognitive Uncertainty. *Quarterly Journal of Economics*, *138*(4). https://doi.org/10.1093/qje/qjad025

Enomoto, K., Matsumoto, N., Inokawa, H., Kimura, M., & Yamada, H. (2020). Topographic distinction in long-term value signals between presumed dopamine neurons and presumed striatal projection neurons in behaving monkeys. *Scientific Reports*, *10*(1). https://doi.org/10.1038/S41598-020-65914-0

Farashahi, S., Azab, H., Hayden, B., & Soltani, A. (2018). On the Flexibility of Basic Risk Attitudes in Monkeys. *Journal of Neuroscience*, *38*(18), 4383–4398. https://doi.org/10.1523/JNEUROSCI.2260-17.2018

Ferrari-Toniolo, S., Bujold, P. M., & Schultz, W. (2019). Probability Distortion Depends on Choice Sequence in Rhesus Monkeys. *Journal of Neuroscience*, *39*(15), 2915–2929. https://doi.org/10.1523/JNEUROSCI.1454-18.2018

Frydman, C., & Jin, L. J. (2021). Efficient coding and risky choice. *Quarterly Journal of Economics*.

Frydman, C., & Jin, L. J. (2023). On the Source and Instability of Probability Weighting. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4546440

Fujimoto, A., & Minamimoto, T. (2019). Trait and state-dependent risk attitude of monkeys measured in a single-option response task. *Frontiers in Neuroscience*, *13*(JUL), 467953. https://doi.org/10.3389/FNINS.2019.00816/BIBTEX

Glimcher, P. (2022). Efficiently irrational: illuminating the riddle of human choice. *Trends in Cognitive Sciences*, *26*(8), 669–687. https://doi.org/10.31234/OSF.IO/ZHYMG

Glimcher, P., & Tymula, A. (2023). Expected Subjective Value Theory (ESVT): a representation of decision under risk and certainty. *Journal of Economic Behavior and Organisation*, *207*, 110–128.

Goldstein, W. M., & Einhorn, H. J. (1987). Expression Theory and the Preference Reversal Phenomena. *Psychological Review*. https://doi.org/10.1037/0033-295X.94.2.236

Guo, J., & Tymula, A. (2021). Waterfall illusion in risky choice – exposure to outcome-irrelevant gambles affects subsequent valuation of risky gambles. *European Economic Review*, *139*, 103889. https://doi.org/10.1016/J.EUROECOREV.2021.103889

Harbaugh, W. T., Krause, K., & Vesterlund, L. (2002). Risk attitudes of children and adults: Choices over small and large probability gains and losses. *Experimental Economics*, *5*, 53–84.

Herold, F., & Netzer, N. (2023). Second-best probability weighting. *Games and Economic Behavior*, *138*. https://doi.org/10.1016/j.geb.2022.12.005

Hey, J., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica: Journal of the Econometric Society*, *62*(6), 1291–1326.

Hsu, M., Krajbich, I., Zhao, C., & Camerer, C. F. (2009). Neural response to reward anticipation under risk is nonlinear in probabilities. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(7), 2231–2237.

Imaizumi, Y., Tymula, A., Tsubo, Y., Matsumoto, M., & Yamada, H. (2022). A neuronal prospect theory model in the brain reward circuitry. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-33579-0

Kagel, J. H., Battalio, R. C., & Green, L. (1995). *Economic choice theory: An experimental analysis of animal behavior*. Cambridge University Press.

Kahneman, D., & Tversky, A. (1979a). Prospect Theory. *Econometrica*, *47*(2), 311. https://doi.org/10.2307/3791683

Kahneman, D., & Tversky, A. (1979b). Prospect Theory - analysis of decision under risk. *Econometrica*, *47*(2), 263–292. http://gateway.webofknowledge.com/gateway/Gateway.cgi?GWVersion=2&SrcAuth=mekentosj&SrcApp=Papers&DestLinkType=FullRecord&DestApp=WOS&KeyUT=A1979GR75700001

Kandel, E. R., Koester, J. D., Mack, S. H., & Siegelbaum, S. A. (2021). *Principles of Neural Science* (6th ed.). McGraw Hill / Medical.

Khaw, M. W., Glimcher, P. W., & Louie, K. (2017a). Normalized value coding explains dynamic adaptation in the human valuation process. *Proceedings of the National Academy of Sciences*, *114*(48), 12696–12701. https://doi.org/10.1073/pnas.1715293114

Khaw, M. W., Glimcher, P. W., & Louie, K. (2017b). Normalized value coding explains dynamic adaptation in the human valuation process. *Proceedings of the National Academy of Sciences*, *114*(48), 12696–12701. https://doi.org/10.1073/pnas.1715293114

Kontek, K., & Lewandowski, M. (2018). Range-Dependent Utility. *Management Science*, *64*(6), 2473–2972. https://doi.org/10.2139/ssrn.2307858

Kurtz David, V., Alladi, V., Sinha, S., Bucher, S., Brandenburger, A., Louie, K., Dewan, A., Glimcher, P., & Tymula, A. (2023). *Inefficient Divisive Normalization: Human Choosers Employ Divisive Normalization Even When They Should Not*. 544.

Landry, P., & Webb, R. (2021). Pairwise Normalization: A Theory of Multi-Attribute Choice. *Journal of Economic Theory*, *193*.

Louie, K., & Glimcher, P. W. (2012a). Efficient coding and the neural representation of value. *Annals of the New York Academy of Sciences*, *1251*, 13–32. https://doi.org/10.1111/j.1749-6632.2012.06496.x

Louie, K., & Glimcher, P. W. (2012b). Set-size effects and the neural representation of value. In R. J. Dolan & T. Sharot (Eds.), *Neuroscience of Preference and choice: Cognitive and neural mechanisms* (pp. 143–169). Academic Press.

Louie, K., Glimcher, P. W., & Webb, R. (2015). Adaptive neural coding: From biological to behavioral decision-making. *Current Opinion in Behavioral Sciences*, *5*, 91–99. https://doi.org/10.1016/j.cobeha.2015.08.008

Louie, K., Grattan, L. E., & Glimcher, P. W. (2011). Reward value-based gain control: Divisive normalization in parietal cortex. *The Journal of Neuroscience*, *31*(29), 10627–10639. https://doi.org/10.1523/JNEUROSCI.1237-11.2011

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

Oprea, R. (2022). Simplicity Equivalents. *Mimeo*.

Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. *The Journal of Neuroscience*, *29*(44), 14004–14014. https://doi.org/10.1523/JNEUROSCI.3751-09.2009

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, *400*(6741), 233–238. https://doi.org/10.1038/22268

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

Prelec, D. (1998). The probability weighting function. *Econometrica*, 497–527. http://www.jstor.org/stable/10.2307/2998573

Preuschoff, K., Bossaerts, P., & Quartz, S. R. (2006). Neural Differentiation of Expected Reward and Risk in Human Subcortical Structures. *Neuron*, *51*(3), 381–390. https://doi.org/10.1016/j.neuron.2006.06.024

Schneider, M., & Day, R. (2018). Target-adjusted utility functions and expected-utility paradoxes. *Management Science*, *64*(1). https://doi.org/10.1287/mnsc.2016.2588

Stauffer, W. R., Lak, A., Bossaerts, P., & Schultz, W. (2015). Economic Choices Reveal Probability Distortion in Macaque Monkeys. *Journal of Neuroscience*, *35*(7), 3146–3154. https://doi.org/10.1523/JNEUROSCI.3653-14.2015

Steiner, J., & Stewart, C. (2016). Perceiving prospects properly. *American Economic Review*, *106*(7). https://doi.org/10.1257/aer.20141141

Steverson, K., Brandenburger, A., & Glimcher, P. (2019). Choice-theoretic foundations of the divisive normalization model. *Journal of Economic Behavior & Organization*, *164*, 148–165.

Tobler, P. N., Christopoulos, G. I., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2008). Neuronal distortions of reward probability without choice. *The Journal of Neuroscience*, *28*(45), 11703–11711. http://www.jneurosci.org/content/28/45/11703.full

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. http://www.springerlink.com/content/lwr6176230786245/

Tymula, A., Wang, X., Imaizumi, Y., Kawai, T., Kunimatsu, J., Matsumoto, M., & Yamada, H. (2023). Dynamic prospect theory: Two core decision theories coexist in the gambling behavior of monkeys and humans. *Science Advances*, *9*(20). https://doi.org/10.1126/sciadv.ade7972

Wang, J., Xia, Y., Wang, G., Zhou, T., Guo, Y., Zhang, C., An, X., Sun, Y., Guo, X., Zhou, Z., & Sha, J. (2014). In-depth proteomic analysis of whole testis tissue from the adult rhesus macaque. *Proteomics*, *14*(11). https://doi.org/10.1002/pmic.201300149

Webb, R. (2020). The (neural) dynamics of stochastic choice. *Management Science*, *65*(1). https://doi.org/10.1287/mnsc.2017.2931

Webb, R., Glimcher, P. W., & Louie, K. (2021). The Normalization of Consumer Valuations: Context-Dependent Preferences from Neurobiological Constraints. *Management Science*, *67*(1). https://doi.org/10.2139/ssrn.2462895

Wilcox, N. (2015). *Unusual Estimates of Probability Weighting Functions*.

Yamada, H., Louie, K., Tymula, A., & Glimcher, P. W. (2018). Free choice shapes normalized value signals in medial orbitofrontal cortex. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-017-02614-w

Yamada, H., Tymula, A., Louie, K., & Glimcher, P. W. (2013). Thirst-dependent risk preferences in monkeys identify a primitive form of wealth. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(39). https://doi.org/10.1073/pnas.1308718110

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

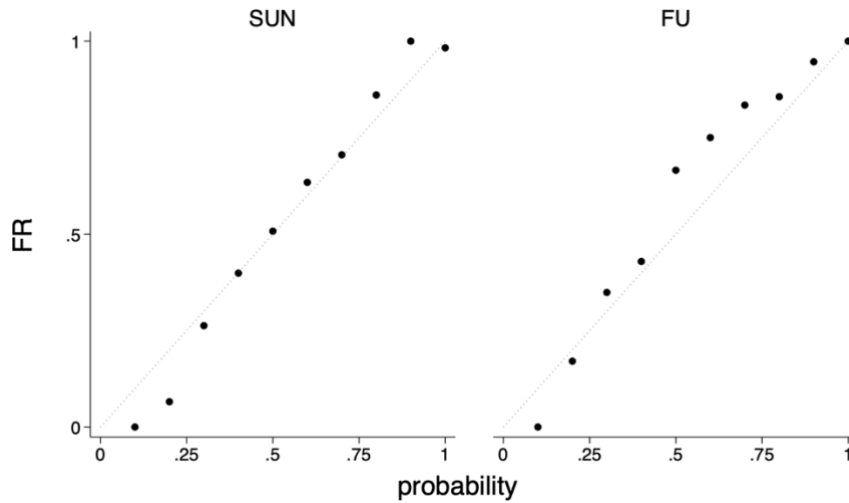# Appendix
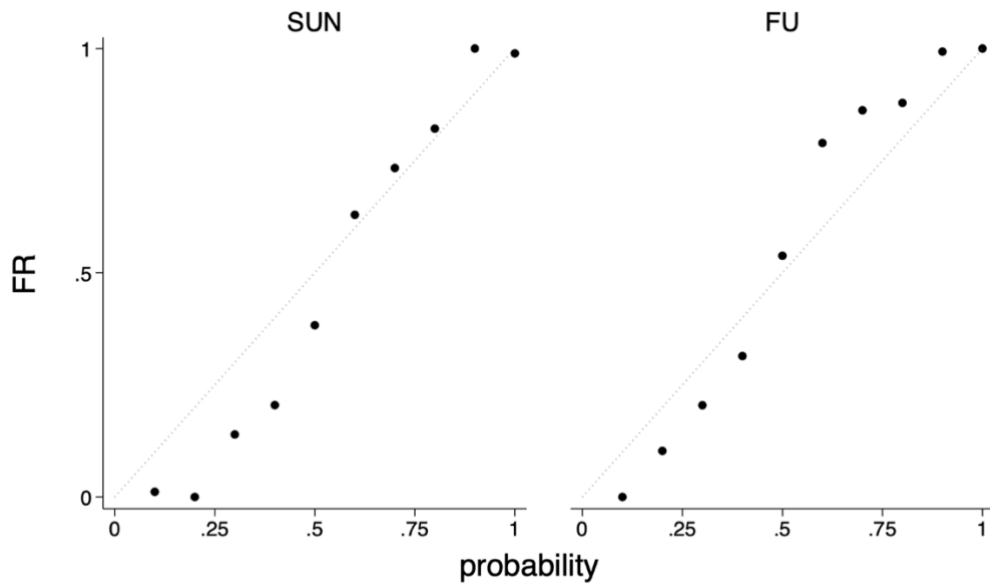
Figure S1. Neural probability weighting function. Includes data from all neurons that significantly and positively (i.e. higher probability higher activity) respond to probability and do not respond to payoff magnitude. Best fit probability weighting parameters are $\delta_{SUN} = 0.634$ with SE=0.059, $\delta_{FU} = 1.219$ with SE=0.106, $\gamma_{SUN} = 1.812$ with SE=0.153, and $\gamma_{FU} = 1.947$ with SE=0.160.
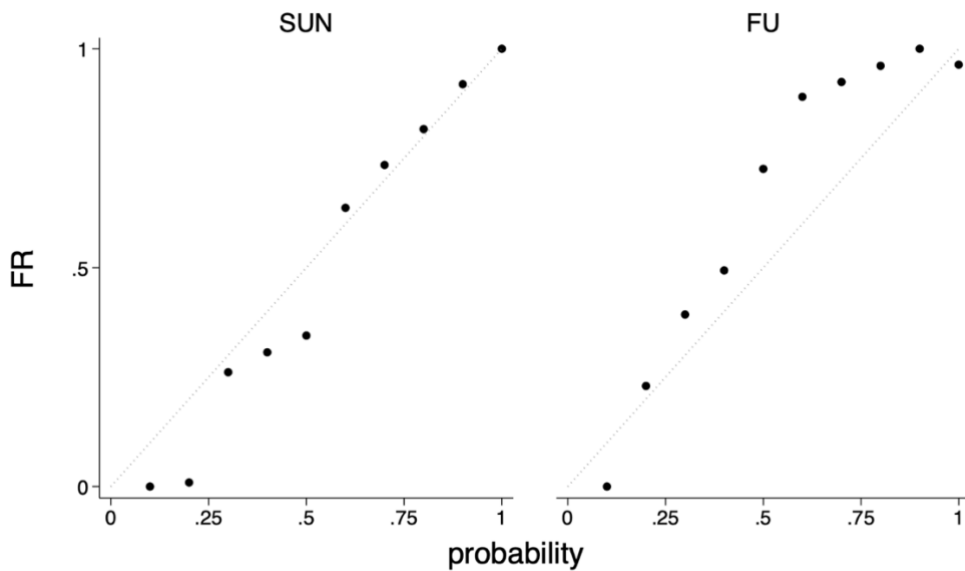
Figure S2. Neural probability weighting function. Includes data from all neurons that significantly respond to probability (positively and negatively) including those that respond to payoff magnitude. Best fit probability weighting parameters are $\delta_{SUN} = 0.747$ with SE=0.089, $\delta_{FU} = 2.627$ with SE=0.306, $\gamma_{SUN} = 1.452$ with SE=0.172, and $\gamma_{FU} = 1.775$ with SE=0.157.

Figure S3. Neural probability weighting function. Includes data from all neurons that significantly respond to probability and do not respond to payoff magnitude. Neurons are included if they positively respond to probability as determined by a linear regression. Best fit probability weighting parameters are $\delta_{SUN} = 0.983$ with SE=0.080, $\delta_{FU} = 1.582$ with SE=0.143, $\gamma_{SUN} = 1.384$ with SE=0.115, and $\gamma_{FU} = 1.437$ with SE=0.124.