Chapter 4

COMPETITION BETWEEN DESCRIBED AND LEARNED VALUE SIGNALS IN ECONOMIC CHOICE

4.1 Introduction

Imagine you are at a restaurant close to your home, one you have visited many times in the past, and you are trying to decide what to order for dinner. The menu lists several options, some of which you have tried before, and some of which are specials which are newly introduced to the menu every week. How do you choose between the different options available? In order to make a choice, you will need to consider foods that you have tried before, relying on your previous experience eating them, as well as foods that are completely new to you, in which case you may need to predict how much you will enjoy them based on the descriptions provided on the menu.

When making decisions based on individual preference, people may need to consider different types of information, some of which may be available in a descriptive format (such as the food specials from the restaurant example), and some of which may be retrieved from memory based on a history of previous rewards (such as the familiar foods in the restaurant menu). Moreover, some options may involve both types of information simultaneously: in the restaurant example, this could correspond, for instance, to a dish that is a slight, novel variation on another dish that has been tried before. In situations such as these, the values associated with the options presented in the two different formats, descriptive and experiential, will need to be represented in the decision making circuitry and somehow integrated in order to generate a single choice.

There is wide recognition that two distinct valuation systems might affect even simple economic choices such as the one described above (Kahneman and Tversky, 1979; Barron and Erev, 2003; Hertwig, Barron, et al., 2004; Jessup, Bishara, and Jerome R Busemeyer, 2008; Hertwig and Erev, 2009; FitzGerald et al., 2010). While the descriptive system relies on clearly described outcomes, the experiential system uses previous experience without relying on detailed information about the outcomes associated with the available options. Several studies have identified a phenomenon, referred to as the description-experience gap, according to which subjects' behavior is risk averse in the gain domain and risk seeking in the loss domain when options are presented in a descriptive format, whereas the opposite is true when they are presented in an experience-based format (Kahneman and Tversky, 1979; Barron and Erev, 2003; Hertwig, Barron, et al., 2004; Hertwig and Erev, 2009). In addition, subjects tend to overweight low probability events and underweight high probability events when those are fully described, but the reverse occurs when subjects learn about these events from experience. Evidence supporting this phenomenon suggests that two distinct valuation systems may be at work when subjects assign values to the options presented in each type of decision. However, most studies so far have focused on the differences between these systems as they operate separately, and therefore much remains unknown about how they interact at the time of choice, within the course of a decision, in order to consistently generate choices in real world situations.

In this chapter we present a study of the interactions between the descriptive and the experiential valuation systems. Our aim was to understand how these systems interact and compete for control when both are simultaneously recruited to produce a single choice, and to investigate how this interaction changes with contextual variables, such as cues about the relevance of each system to the current decision. For this purpose, we designed a task where each decision required subjects to make use of information retrieved from memory, based on their experience in previous trials, as well as information available only at the time of the decision. This design allowed us to investigate how control was allocated between the two valuation systems during the decision process. Furthermore, the inclusion of an experimental variable that changed the relative relevance of these two types of information in each trial allowed us to examine the influence of contextual variables on the interactions between the two systems.

Our approach differs from most previous work in this literature in that it directly addresses the interaction and competition between the descriptive and experiential systems when both are relevant to the choice at hand. Moreover, our experimental design exogenously and randomly changed the relative importance of each system throughout trials, which is useful in understanding this interaction and which has not been used in previous studies.

We found that a simple computational model of arbitration between the two systems could reasonably describe subjects' choices in our task. We tested a few variations of this model, allowing for both linear and non-linear effects of the exogenous relative relevance of the two systems on choice. Additionally, we modified the model to include an effect of either the relative strength of preference (i.e., to what degree one option was better than the other in each valuation system), or the uncertainty in the estimates from the experiential system. We found only a small influence of the relative strength of preference on choice and no influence of relative uncertainty, suggesting that the weights given to the two valuations systems during choice computation can be mostly accounted for by a non-linear transformation of the experimental variable that controlled relative system relevance.

4.2 Related Literature

Growing evidence from behavioral and neuroimaging studies suggests the existence of two distinct valuation systems in the human brain which come into play during the decision making process: a descriptive system relying on explicitly stated variables and an experiential system that evaluates options on the basis of experience (Hertwig, Barron, et al., 2004; Jessup, Bishara, and Jerome R Busemeyer, 2008; Hertwig and Erev, 2009; FitzGerald et al., 2010; Glöckner et al., 2012). Hertwig et al. found evidence for this dichotomy in a behavioral study showing that people tend to overweight the probability of rare events when making decisions from description, and to underweight that same probability when deciding based on experience (Hertwig, Barron, et al., 2004). In an fMRI study with humans, FitzGerald et al. found differential sensitivity to learned and described values and risk in brain regions typically associated with reward processing (FitzGerald et al., 2010). Later, Glöckner and colleagues used eye-tracking and physiological arousal measures to study the differences between descriptive and experiential valuation systems (Glöckner et al., 2012). They found that different computational models are better at predicting choices in the two types of decision, and that arousal and attention measurements also differ between them, providing further qualitative evidence for a distinction between the two systems.

Collectively, these results suggest that decisions based on described information and those based on previous experience exhibit different behavioral patterns and may stem from distinct neural systems. But this hypothesis leads to an important open question related to arbitration: when multiple valuation or decision systems are in operation, how is control allocated to each of them? Do they compete or cooperate? Several studies have attempted to understand this arbitration process (Doya et al., 2002; N. D. Daw, Niv, and Dayan, 2005; N. D. Daw, Gershman, et al., 2011; Beierholm et al., 2011; Wunderlich, Dayan, and Raymond J Dolan, 2012; Lee, Shimojo, and J. P. O'Doherty, 2014; Economides et al., 2015; Kool, F. A. Cushman, and Gershman, 2016; Kool, Gershman, and F. A. Cushman, 2017; Miller, Botvinick, and Brody, 2017; Russek et al., 2017). Notably, Daw et al. proposed a Bayesian principle of arbitration between model-based and model-free decision systems based on uncertainty, relying on the trade-off between the flexibility of the first system and the computational simplicity of the second (N. D. Daw, Niv, and Dayan, 2005). Wunderlich et al. identified two distinct areas of human striatum relating to forward planning and to values learned during extensive training, as well as functional coupling between these areas and a region in ventromedial prefrontal cortex that suggests a mechanism of value comparison (Wunderlich, Dayan, and Raymond J Dolan, 2012). More recently, Lee and colleagues found neuroimaging evidence for an arbitration mechanism in the human brain that allocates control over behavior to model-based and model-free systems as a function of their reliability (Lee, Shimojo, and J. P. O'Doherty, 2014), while Kool et al. studied the same arbitration process under a cost-benefit framework, suggesting that humans perform on-line cost-benefit analysis of effort and reliability in order to switch between systems (Kool, Gershman, and F. A. Cushman, 2017).

Our study differs from the ones listed above in several ways. First, while FitzGerald et al. looked for potential differential representations of the two classes of values (FitzGerald et al., 2010), here we investigated the competition between the valuation systems identified in this literature. We focused on understanding how the two systems interact and compete for control when both are relevant to the same decision, which could not be addressed through the task design used by FitzGerald et al. Furthermore, we carried out a qualitative comparison of broad classes of models that might drive the competition between the two valuation systems, which has not been addressed in previous studies. Finally, our task design included a built-in experimental variation of the relative importance of the two systems that allowed us to further investigate the nature of the arbitration mechanism.

4.3 Materials and Methods

Subjects

In this experiment we tested 27 healthy subjects (15 female, mean age 20), all of whom were Caltech students. Subjects reported no history of psychiatric or neurological disorders, and no current use of any psychoactive medications. Of the initial set of 27 subjects, 25 were able to complete the experiment (2 subjects could not be scanned due to discomfort and/or claustrophobia in the scanner). Each of the 25 remaining subjects completed 300 trials, split into 5 consecutive fMRI sessions. Subjects received a $30 show-up fee, as well as additional earnings based on performance, as described below. The experiment was approved by Caltech's IRB and all subjects provided informed consent prior to participation.

Task

The structure of a typical trial is depicted in **Figure 4.1**. Each trial began with a central fixation, for which the duration is a random inter-trial interval (ITI) between 4 and 7 seconds plus a variable amount as described below. In each trial, the subject had to choose between two pairs, one on the left and one on the right side of the screen. Each pair contained a fractal (at the top) and a lottery (at the bottom). The two fractals shown to the subject were randomly sampled from a set of 25, and remained the same in every trial throughout the experiment. Each fractal was associated with a probability of a fixed payout of $1: $p_{\text{fractal left}}$ and $p_{\text{fractal right}}$. These probabilities were not shown on the screen. They drifted slowly and independently throughout the experiment between 0.25 and 0.75, according to a Gaussian random walk with $\sigma = 0.025$ (subject were only told that the probabilities drifted slowly and independently, but were given no information about the bounds or the drift rate). **Figure 4.2** shows the evolution of fractal reward probabilities for a sample sequence of trials. The initial value of each probability was uniformly sampled between 0.25 and 0.75. The change in each probability was then sampled at every trial, and when the value of a probability went beyond one of the two thresholds, it bounced back into the allowed interval by moving the same amount in the opposite direction.

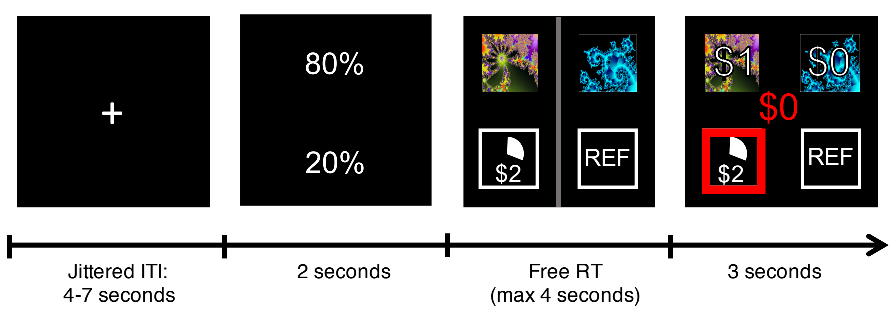

Figure 4.1: Task structure.

The lottery on the left changed from trial to trial and was represented by a probability $p_{\text{lottery left}}$, shown as a pie chart, and a magnitude $V_{\text{lottery left}}$. The pair $(p_{\text{lottery left}}, V_{\text{lottery left}})$ was taken from the set {(1, \$0.50), (0.25, \$2), (0.2, \$2.50),

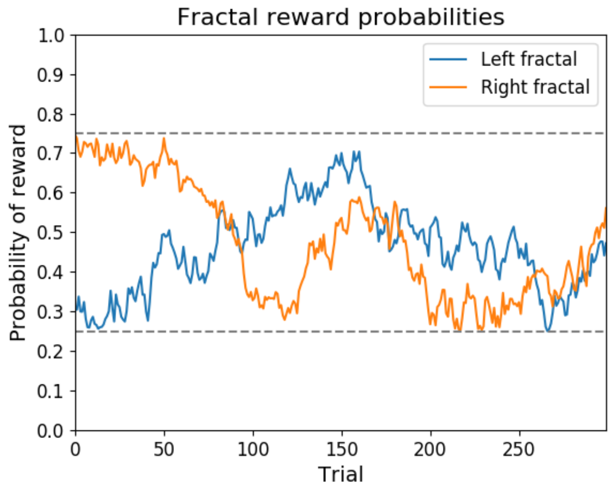

Figure 4.2: Example of fractal reward probabilities. The probabilities of receiving a reward associated with the fractals drifted slowly and independently over time between 0.25 and 0.75, according to a Gaussian random walk with $\sigma = 0.025$.

(0.1, \$5), (1, \$0.10), (0.1, \$1), (0.05, \$2), (0.01, \$10), (1, \$0.30), (0.3, \$1), (0.15, \$2), (0.1, \$3), (1, \$0.70), (0.7, \$1), (0.35, \$2), (0.1, \$7), (1, \$0.90), (0.9, \$1), (0.45, \$2), (0.1, \$9)}. Each of these 20 pairs occurred 3 times in each 60-trial session, in randomized order. The lottery on the right was the reference lottery and remained the same in every trial: a 50% probability of winning \$1; this was represented by the symbol "REF" on the bottom right corner of the screen in every trial.

Before each choice, the subject saw the probabilities associated with a fractal draw or a lottery draw, which were displayed for 2 seconds. The number at the bottom, which we call π, corresponded to the probability that the reward for the trial would be drawn from the lotteries, while the number at the top, $1 - \pi$, corresponded to the probability that the reward would be drawn from the fractals. The value of π in each trial was taken from the set {0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1}; each value below 0.3 or above 0.7 occurred 5 times within a 60-trial session, whereas all other values occurred 6 times each within the session.

The subject had a maximum of 4 seconds to respond in each trial. The subject reported their choice using their right hand, by pressing the "2" key for left and the "3" key for right. If the subject did not make a choice within that time, they saw the message "No response recorded!", and the experiment continued to the next trial. In trials where no response was recorded, no reward was added to the final payout.

In trials where the subject took less than 4 seconds to respond, the remaining time was added to the following ITI.

After the subject had made a choice, the computer drew the lottery associated with the option chosen: for instance, if the subject chose left, a reward would be drawn from the left lottery with probability π, and from the left fractal with probability $(1 - \pi)$. The subject saw a red box around the selected option (fractal or lottery) as well as the resulting reward drawn for that trial in the center of the screen in red (reward screen was shown for 3 seconds). In addition, the subject saw the reward drawn from each fractal in that trial overlaid on top of the fractals. However, only the amount shown in red in the center of the screen actually counted as the reward for that trial. At the end of the experiment, 175 trials were randomly selected, and the subject received the amount corresponding to the sum of rewards on those selected trials. This amount was then adjusted so that the minimum amount the subject could receive after completing the experiment was \$80, and the maximum amount was \$120.

Behavioral Models

We tested several different computational models with our behavioral data. In all models, we used π to describe the probability of a lottery draw in each trial, which was displayed to the subject on the screen. In each trial, the subject had to choose between a left and a right option. The value of each option was computed as:

$$V_i = w(\pi)V_i^D + (1 - w(\pi))V_i^E, \quad i \in \{\text{left}, \text{right}\}, \tag{4.1}$$

where V_i^D corresponds to the *described* value of the lottery, V_i^E corresponds to the *experienced* value of the fractal, and $w(\pi) \in [0, 1]$ is the relative weight given to the described value.

We modeled choices as a logistic function of the value difference between the two options:

$$P(\text{choice} = \text{left}) = \frac{1}{1 + e^{-\beta(V_{\text{left}} - V_{\text{right}})}}, \tag{4.2}$$

where β is the inverse temperature, a free parameter of the model which was fitted per subject.

The described value of a lottery at trial t was given by its expected value, which can be computed as:

$$V_{i,t}^{D} = EV_{i,t} = p_{\text{lottery } i,t} \times V_{\text{lottery } i,t}. \tag{4.3}$$

Since the subjects could not observe the probabilities of reward for the fractals directly, we used a Q-learning model to estimate the experienced value associated with each fractal at each trial t:

$$V_{i,t}^{E} = Q_{i,t}. \tag{4.4}$$

The Q-learning update was calculated at each trial as:

$$Q_{i,t} = Q_{i,t-1} + \alpha(R_{i,t-1} - Q_{i,t-1}), \tag{4.5}$$

where α is the subject's learning rate and $R_{i,t}$ is the reward sampled from fractal i on trial t, which was observed by the subject for both fractals in every trial. Also note that the Q values for both fractals were set to zero at the beginning of the experiment.

In the linear model, we set the relative weight given to the lottery as $w(\pi) = \pi$, such that this model contained only two free parameters: the learning rate α and the inverse temperature β.

In the non-linear model we used a two-parameter weighting function for w, which was given by:

$$w(\pi) = \frac{\delta\pi^{\gamma}}{\delta\pi^{\gamma} + (1-\pi)^{\gamma}}, \tag{4.6}$$

where δ and γ are two additional free parameters fitted per subject, leading to four free parameters total.

Finally, we also tested a nested model in which we used the absolute difference between the left and right lottery expected values as a measure of relative strength of preference in the descriptive system, and the absolute difference between the left and right fractal Q values as a measure of relative strength of preference in the experiential system. Using these two metrics, we computed a weight adjustment variable B which varied per trial, defined as:

$$B = \frac{|EV_{\text{left}} - EV_{\text{right}}|^{\kappa}}{|EV_{\text{left}} - EV_{\text{right}}|^{\kappa} + |Q_{\text{left}} - Q_{\text{right}}|^{\kappa}}, \tag{4.7}$$

where κ is an additional free parameter which controls the shape of the curve B as a function of the absolute lottery expected value difference. Using B, we computed the relative weight given to the descriptive system, u, as:

$$u = \mu w + (1 - \mu)B, \tag{4.8}$$

where μ is an additional free parameter controlling the relative contributions of the weight w and of the weight adjustment variable B to the weight u. The nested model contained a total of six free parameters: the learning rate α, the inverse temperature β, the weighting function parameters γ and δ, plus κ and μ.

Bayesian Update Learning Model

A Bayesian update model was used to estimate, in each trial, a posterior distribution over all possible values of $p_{\text{fractal } i}$ for the left and right fractals.

In this model we assumed that the subject knew the boundaries and drift rate of the fractal reward probabilities (even though subjects were not told the specifics of how these probabilities drifted, only that they drifted slowly and independently over time). Under this assumption, the subject began the experiment with a flat prior $P^{\text{in}}_{t=0}(p_{\text{fractal } i})$ over the interval $[0.25, 0.75]$ for each of the two fractals.

After observing the fractal rewards in each trial, the posteriors (for the values within the interval $[0.25, 0.75]$) were updated according to the rule:

$$P^{\text{out}}_t(p_{\text{fractal } i}) = \frac{P_t(R_i|p_{\text{fractal } i})P^{\text{in}}_t(p_{\text{fractal } i})}{\int P_t(R_i|\hat{p}_{\text{fractal } i})P^{\text{in}}_t(\hat{p}_{\text{fractal } i})d\hat{p}_{\text{fractal } i}}, \tag{4.9}$$

where $P_t(R_i|p_{\text{fractal } i})$ is the likelihood of observing reward R_i at trial t, which is equal to $p_{\text{fractal } i}$ when $R_i = 1$, and to $1 - p_{\text{fractal } i}$ when $R_i = 0$; and $P^{\text{in}}_t(p_{\text{fractal } i})$ is the prior at trial t.

In addition, the prior was updated in each trial according to:

$$P^{\text{in}}_t(p_{\text{fractal } i}) = P^{\text{out}}_t(p_{\text{fractal } i}) + \mathcal{N}(0, \sigma = 0.025), \tag{4.10}$$

to account for the drift in the fractal reward probabilities. A diagram of this model is shown in **Figure 4.3**.

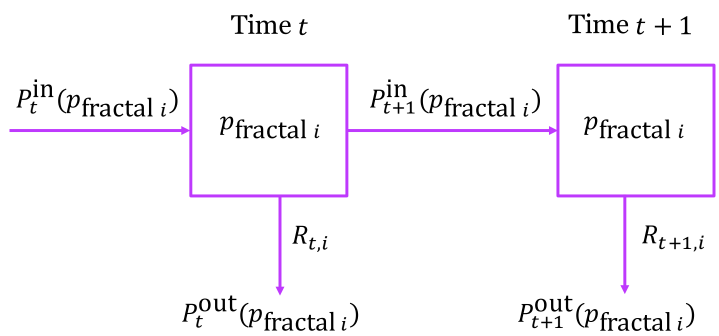

Figure 4.3: Diagram of the Bayesian update model for the fractal probabilities. $P_t^{\text{out}}(p_{\text{fractal } i})$ is the posterior distribution at time t for the probability associated with fractal i, where i is either left or right. $P_t^{\text{in}}(p_{\text{fractal } i})$ is the prior distribution at time t for that same probability.

fMRI Data Acquisition

Functional imaging data was acquired from all subjects while they participated in the task. Functional imaging was performed on a 3T Siemens (Erlangen) Trio scanner located at the Caltech Brain Imaging Center (Pasadena) with a 32-channel radio frequency coil for all the MR scanning sessions. Each subject was scanned in 5 consecutive sessions, and 882 volumes were obtained for each session.

4.4 Results

We used the binary choice task described above to investigate the interactions between the descriptive and the experiential valuation systems. Our analysis, which we describe in more detail in the subsections below, was structured as follows. First, we estimated various computational models of the choices from each subject through a maximum likelihood estimation approach. Then, we performed several tests to quantify the influence of each valuation system on the subjects' behavior, and the interactions between the two systems under different experimental conditions.

The current section is structured as follows. We begin by validating the efficacy our experimental paradigm. We checked if subjects' performance indicated that they understood the task, whether they were able to reasonably estimate the expected values of the lotteries and the probabilities associated with the fractals, and whether they correctly adjusted their choices in the special trial conditions where only one system was relevant (i.e., where $\pi = 0$ or 1). We then carried several additional analyses, to test: 1. if the probability of a lottery draw, π, influenced choices, and whether this influence was linear; 2. whether the weight given to each system was dependent on their individual relative strength of preference; 3. whether the relative influence of the two systems was affected by changes in the relative uncertainty of the experience values; and 4. whether the relative influence of the two systems was affected by the presence of conflict. More details about each of these analyses are provided below.

Basic Paradigm Validation

An important feature of our experimental paradigm was the probability π of a lottery draw, which randomly changed from trial to trial. On trials where π was equal to one, subjects should only have taken into consideration the difference between the values assigned to the two lotteries; conversely, when π was equal to zero, they should only have considered the values assigned to the fractals. In order to validate the ability of our task to study descriptive versus experiential valuation, we first looked at the psychometrics of trials where $\pi = 1$ or 0. This allowed us to verify that subjects understood the logic of the experiment and that they managed to learn a reasonable estimate for the values assigned to the fractals.

We defined the value assigned to each lottery, left or right, at trial t (*described* value, $V^D_{i,t}$) as the expected value of the lottery, i.e., its probability multiplied by its magnitude:

$$V^D_{i,t} = EV_{i,t} = p_{\text{lottery } i,t} \times V_{\text{lottery } i,t}. \tag{4.11}$$

We use V^D_i and EV_i exchangeably throughout the chapter to signify the described value of lottery i, i.e., its expected value. Moreover, the value assigned to each fractal, left or right, at trial t (*experienced* value, $V^E_{i,t}$) was given by the Q value of that fractal at that trial:

$$V^E_{i,t} = Q_{i,t}. \tag{4.12}$$

The Q values were obtained from a myopic Q-learning model fitted per subject, in which the values were updated as follows:

$$Q_{i,t} = Q_{i,t-1} + \alpha(R_{i,t-1} - Q_{i,t-1}), \tag{4.13}$$

where α is the learning rate, $R_{i,t-1}$ is the reward drawn from fractal i at trial $t-1$, and the Q values for both fractals are equal to zero at the beginning of the experiment. **Figure 4.4** shows a histogram of the individual learning rates, which were fitted separately for each subject through maximum likelihood. In the remainder of this chapter, we sometimes omit the trial indicator t to simplify the notation. More details about the computation of the described and experienced values, the Q-learning model, and the model fitting procedure are given in the Materials and Methods section.

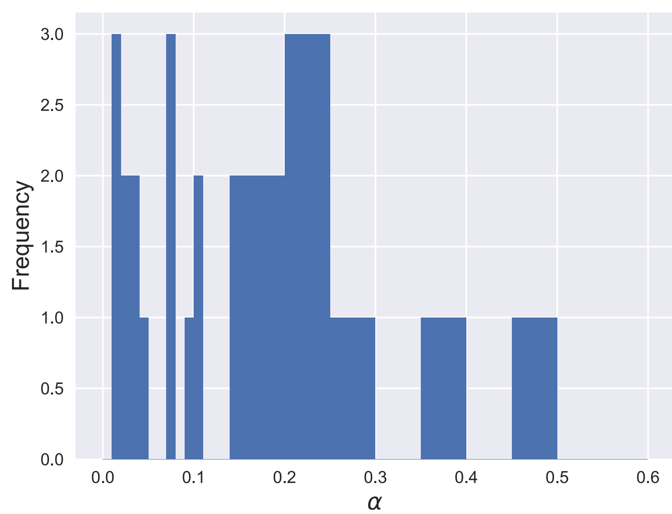

Figure 4.4: Histogram of the learning rate α used in the Q-learning model, fitted per subject.

Looking at subjects' choices, we found that they consistently used the lottery expected value difference, $EV_{\text{left}} - EV_{\text{right}}$, and not the fractal Q value difference, $Q_{\text{left}} - Q_{\text{right}}$, in trials where $\pi = 1$, and the opposite was true when $\pi = 0$. This is shown in the psychometric choice curves in **Figure 4.5**, and in the estimated regression coefficients for a mixed effects model using both value differences, as shown in **Table 4.1**.

Table 4.1 describes the results of a mixed effects logistic regression on trials with π equal to either 1 or 0, where the dependent variable was the probability of choosing left, and the independent variables were the lottery expected value difference, the fractal Q value difference, the lottery expected value difference modulated by an

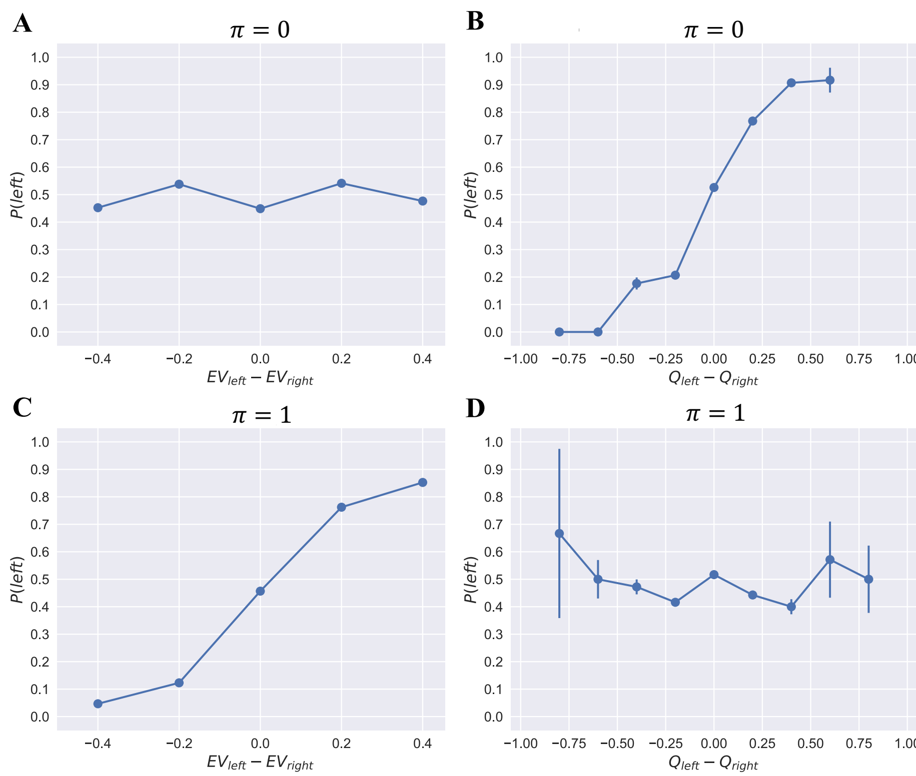

Figure 4.5: Psychometric choice curves as a function of the lottery expected value difference (A and C) and the fractal Q value difference (B and D). The top row shows only trials where $\pi = 0$ (A and B), whereas the bottom row shows only trials where $\pi = 1$ (C and D). Data was aggregated from all subjects. Error bars show 95% confidence intervals for the data pooled across all subjects.

indicator function for $\pi = 1$, and the fractal Q value difference modulated by an indicator function for $\pi = 0$. The regression allowed for random effects in these variables. We can see from the table that the lottery value difference has a significant effect on choice when $\pi = 1$ ($p < 10^{-16}$) but not when $\pi = 0$, and that the fractal value difference has a significant effect when $\pi = 0$ ($p = 1.01^{-15}$) but not when $\pi = 1$. We found no significant difference between the influence of the lottery value difference in $\pi = 1$ trial choices and the influence of the fractal value difference in $\pi = 0$ trial choices (paired t-test, $p = 0.25$).

We carried out an analogous analysis for trial response times (RTs). We found that the absolute lottery expected value difference, $|EV_{\text{left}} - EV_{\text{right}}|$, significantly affected RTs in trials where $\pi = 1$ ($p = 3.93 \times 10^{-16}$) but not in trials where $\pi = 0$,

Table 4.1: Choice logit mixed effects model: trials where $\pi = 1$ vs. trials where $\pi = 0$.

Regressor	Estimate	Std. Error	z value	p-value
Intercept	-0.006159	0.191351	-0.032	0.9743
$\pi = 1$ indicator	-0.373910	0.204786	-1.826	0.0679
$\pi = 1$ indicator $\times(EV_L - EV_R)$	7.022994	0.592529	11.853	< 2e-16
$\pi = 1$ indicator $\times(Q_L - Q_R)$	0.021057	0.552671	0.038	0.9696
$\pi = 0$ indicator $\times(EV_L - EV_R)$	0.428643	0.371517	1.154	0.2486
$\pi = 0$ indicator $\times(Q_L - Q_R)$	7.616962	0.949029	8.026	1.01e-15

whereas the absolute fractal Q value difference, $|Q_{\text{left}} - Q_{\text{right}}|$, significantly affected RTs in trials where $\pi = 0$ ($p = 1.19 \times 10^{-12}$) but not in trials where $\pi = 1$. This is illustrated in the plots in **Figure 4.6**, and in the regression coefficients for a mixed effects model using both absolute value differences, as shown in **Table 4.2**. The mixed effects model used for the RTs was similar to the choice logit mixed effects model described above, except that here we used the RT as the dependent variable, and took the absolute value of the differences in values for both lotteries and fractals.

Table 4.2: RT mixed effects model: trials where $\pi = 1$ vs. trials where $\pi = 0$.

Regressor	Estimate	Std. Error	t value
Intercept	1.09955	0.07233	15.201
$\pi = 1$ indicator	0.19478	0.07032	2.770
$\pi = 1$ indicator $\times\|EV_L - EV_R\|$	-0.40933	0.14183	-2.886
$\pi = 1$ indicator $\times\|Q_L - Q_R\|$	-0.05352	0.14869	-0.360
$\pi = 0$ indicator $\times\|EV_L - EV_R\|$	0.11285	0.14630	0.771
$\pi = 0$ indicator $\times\|Q_L - Q_R\|$	-0.82095	0.17008	-4.827

It is interesting to note that, in **Table 4.2**, the negative coefficient for the absolute fractal value difference in $\pi = 0$ trials is twice as large as the negative coefficient for the absolute lottery value difference in $\pi = 1$ trials (-0.82 vs. -0.41). We hypothesize this is due to the different ways in which these two types of values are computed. Since the individual fractal values were retrieved from memory and required no explicit computation, the preference for one fractal over the other could be decided during the probability screen, before the subject even saw the choice options. This means that, when $\pi = 0$, the subject could make a choice between

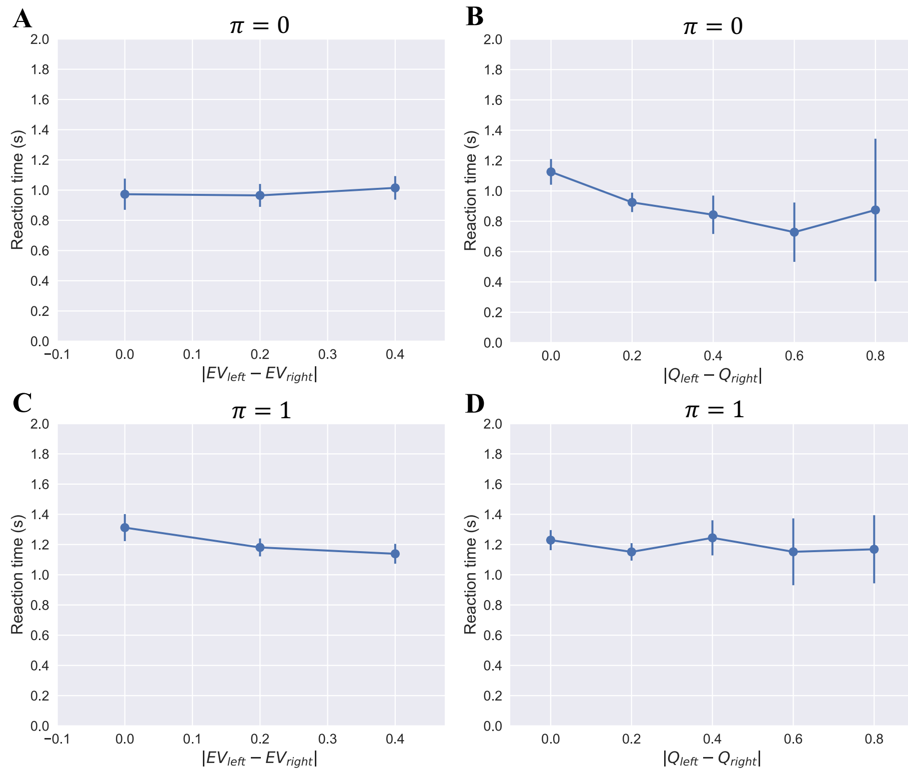

Figure 4.6: Response times as a function of the absolute lottery expected value difference (A and C) and the absolute fractal Q value difference (B and D). The top row shows only trials where $\pi = 0$ (A and B), whereas the bottom row shows only trials where $\pi = 1$ (C and D). Data was aggregated from all subjects. Error bars show 95% confidence intervals for the data pooled across all subjects.

left and right before they saw the choice screen, since the lotteries presented were irrelevant, leading to faster RTs. On the other hand, the lottery expected values always had to be computed explicitly once the choice screen appeared, and because in $\pi = 1$ trials they were especially relevant, RTs in those trials tended to be longer.

Taken collectively, these results show that subjects were able to ignore the valuation signals that were irrelevant to each trial, indicating that they understood the task, succeeded in computing the expected values of the lotteries, and estimated the values associated with the fractals in a manner consistent with the Q-learning model.

Role of Exogenous Changes on the Relative Relevance of the Two Systems

Next, we investigated how choices and RTs were affected by the relative relevance of the descriptive and the experiential valuation systems, which we were able to exogenously manipulate by randomly changing the probability π at every trial. We developed a computational model to describe how choices were affected by π, and fitted this model to our data to check whether the relative relevance of the two systems was reflected in the weights given to them in the model. Additionally, we examined whether this effect was linear or non-linear, and checked for the presence of biases in favor of one system over the other.

We defined the non-linear choice model as follows. The value of each option, left or right, was given by:

$$V_i = w(\pi)V_i^D + (1 - w(\pi))V_i^E, \quad i \in \{\text{left}, \text{right}\}, \tag{4.14}$$

where V_i^D corresponds to the described value of the lottery, V_i^E corresponds to the experienced value of the fractal, and $w(\pi) \in [0, 1]$ is the relative weight given to the described value, which is a function of π.

We modeled choices as a logistic function of the value difference between the two options:

$$P(\text{choice} = \text{left}) = \frac{1}{1 + e^{-\beta(V_{\text{left}} - V_{\text{right}})}}, \tag{4.15}$$

where β is the inverse temperature, a free parameter of the model which was fitted per subject.

Finally, we used a two-parameter weighting function to obtain the relative weight w given to the lottery, which was given by:

$$w(\pi) = \frac{\delta\pi^\gamma}{\delta\pi^\gamma + (1 - \pi)^\gamma}, \tag{4.16}$$

where γ and δ are two additional free parameters fitted per subject, leading to four free parameters total (α, β, γ and δ). The γ parameter primarily controls the curvature of the weighting function, while δ primarily controls its elevation. Examples of curves obtained for different values of γ and δ and shown in **Figure 4.7**.

Several aspects of the two-parameter weighting function are worth highlighting. First, note that when both γ and δ are equal to 1, we obtain $w(\pi) = \pi$, and this model reduces to a linear model, i.e., the relative weight given to the descriptive system is a linear function of the probability π. Second, the use of this weighting function provides us with a natural notion of bias: for instance, when $\pi = 0.5$, no bias means $w = 0.5$, a bias towards the lotteries means $w > 0.5$, and a bias towards the fractals means $w < 0.5$. Finally, note that when $\delta > 1$ the weight is biased towards the lotteries, whereas when $\delta < 1$ it is biased towards the fractals.

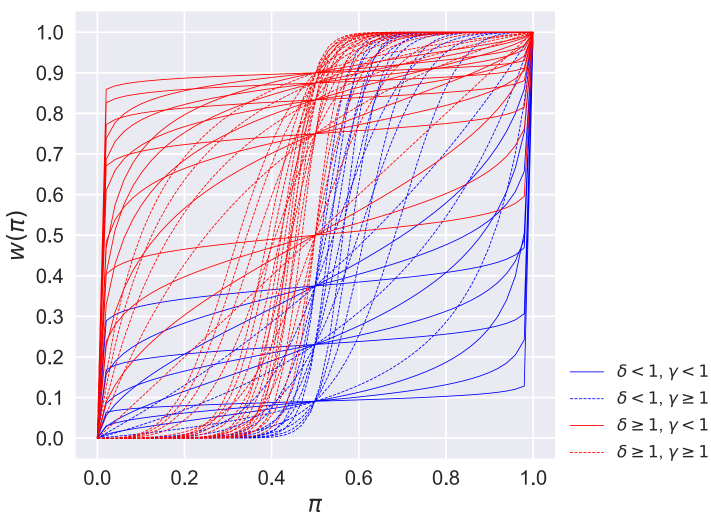

Figure 4.7: Different curves of the two-parameter weighting function. Curves were obtained with varying values of γ and δ, which were taken from the set $\{0.1, 0.3, 0.6, 1, 3, 5, 7, 9\}$. Curves are displayed in four groups according to whether or not γ is less than 1, and to whether or not δ is less than 1.

An interesting aspect of the choice model used here is the fact that it describes the choice prediction of a Drift-Diffusion Model (DDM), where choices are made based on the sequential integration between a left value and a right value. These values can be obtained through our model by computing a weighted sum of the lottery expected value and the fractal Q value corresponding to the same side, using w as the weight.

We fitted the above model for all subjects individually, through a maximum likelihood estimation procedure. For each subject, we performed a grid search over the four-parameter space, and chose the combination of parameters that yielded the largest likelihood value. The results of this fitting procedure are shown in **Figure 4.8**. Summary statistics are provided in **Table 4.3**.

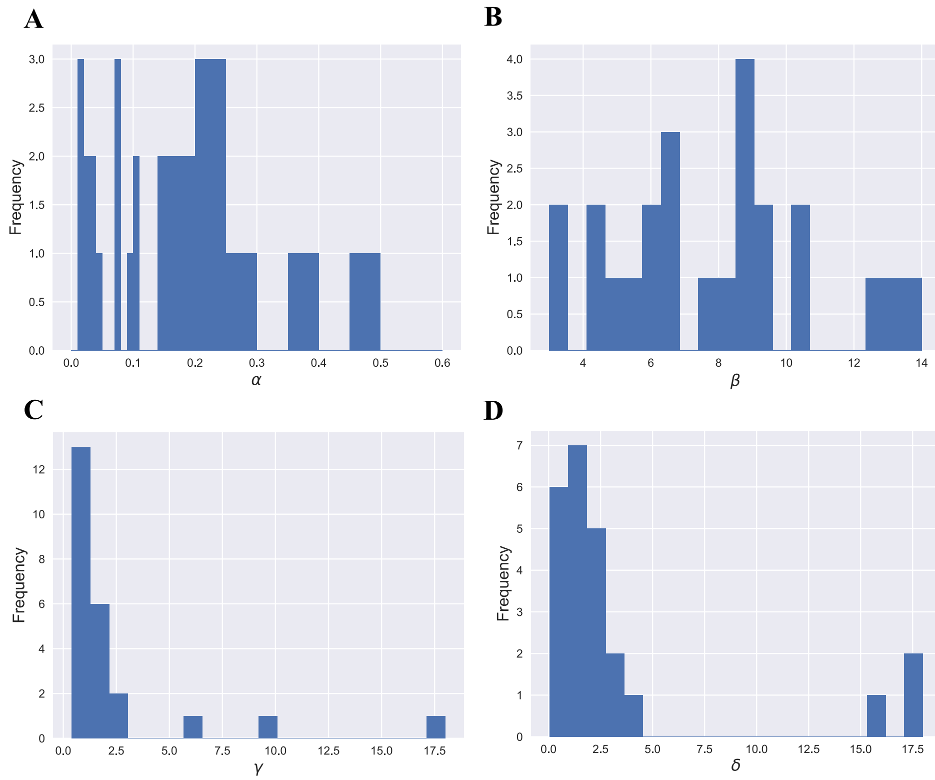

Figure 4.8: Histograms for the non-linear model parameters, fitted per subject. (A) α. (B) β. (C) γ. (D) δ.

Table 4.3: Non-linear model fitting summary statistics.

Parameter	**Mean**	**SD**
α	0.13	0.12
β	7.79	2.88
γ	2.42	3.87
δ	3.57	5.25

Using the values fitted for parameters γ and δ, we obtained the weighting curve for each subject, as shown in **Figure 4.9**.

We used the results from model fitting to exclude any subjects whose learning rate α was equal to zero, indicating that they were not able to reasonably learn the probabilities associated with the fractals and therefore effectively perform the task.

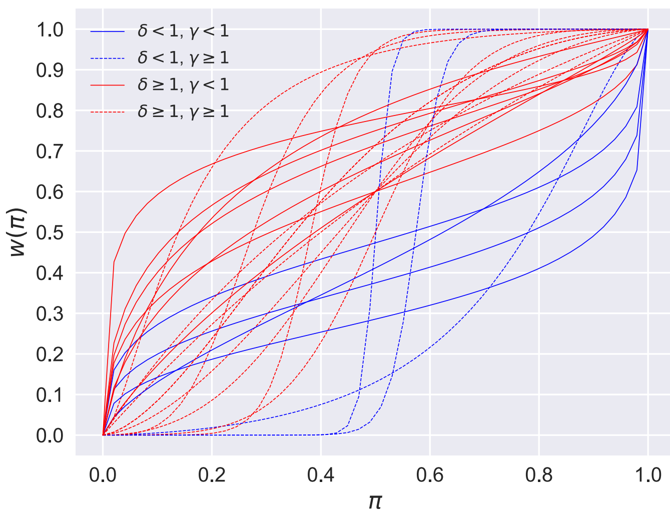

Figure 4.9: Curves obtained for the two-parameter weighting function for each subject. Curves are grouped into 4 groups according to whether or not γ is less than 1, and to whether or not δ is less than 1.

Using this criterion, we excluded one subject from all further analyses.

We also fitted the same computational model under linear constraints, i.e., setting $\gamma = 1$ and $\delta = 1$. **Table 4.4** shows a comparison between the linear and non-linear models using two evaluation metrics: negative log-likelihood (NLL) and Bayesian information criterion (BIC). Note that smaller numbers for these metrics correspond to better fittings. The table shows that, using the NLL as the evaluation metric, the non-linear model generates a better fit for all subjects; using the BIC, the non-linear model generates a better fit for 6 out of the 24 subjects, which is likely due to the fact that the BIC also takes into account the number of free parameters in the model (which is four in the non-linear model and only two in the linear model).

As an additional comparison between the non-linear and the linear models, we simulated each of them using the trial conditions from the experiment and the best fitting parameters obtained for each subject, then compared the choices generated by each model with the subjects' actual choices. This process was repeated 100 times. The non-linear model was able to correctly predict, on average, 67% of choices (mean = 0.67, SD = 0.0041), while the linear model correctly predicted 65% of them (mean = 0.65, SD = 0.005), and this difference was statistically significant

Table 4.4: Comparison between linear and non-linear model fittings.

Subject	NLL non-linear	NLL linear	BIC non-linear	BIC linear
1	156.42	164.59	335.65	340.58
2	171.64	177.69	366.07	366.77
3	127.58	130.86	277.97	273.13
4	154.63	156.34	332.07	324.08
5	144.38	145.19	311.57	301.78
6	119.13	128.63	261.08	268.66
7	165.6	167.11	353.89	345.57
8	121.67	125.15	266.16	261.7
10	155.14	159.82	333.06	331.03
11	144.95	147.94	312.63	307.25
12	143.77	146.45	310.35	304.31
13	137.61	151.74	298.02	314.88
14	189.69	190.25	402.14	391.88
15	146.82	148.07	316.46	307.55
16	169.53	174.44	361.88	360.28
17	163.58	168.55	349.97	348.5
18	117.06	118.77	256.87	248.91
19	168.92	170.94	360.65	353.29
20	122.9	125.96	268.62	263.34
22	136.74	143.92	296.26	299.24
23	175.56	177.37	373.87	366.12
24	169.69	171.27	362.18	353.94
25	136.27	137.13	295.3	285.64
27	124.38	142.8	271.57	297.0
Mean:	148.49	152.96	319.76	317.31
SD:	20.05	19.03	40.08	38.07

(two-sample t-test, $t = 18.92$, $p < 10^{-16}$).

We found a moderately significant difference between the α parameter fitted in the non-linear and the linear models (mean $\alpha = 0.13$ vs. 0.09; paired t-test, $p = 0.01$), and no significant difference for the β parameter (mean $\beta = 7.79$ vs. 8.08; paired t-test, $p = 0.47$). Importantly, we found that the δ parameter fitted in the non-linear model was significantly greater than 1 ($t = 2.35$, $p = 0.01$), indicating an overall tendency to overweight the lottery expected values (red curves in **Figure 4.7** and **Figure 4.9**).

We now present an analysis of the psychometrics of the task using the fitted weights w. As discussed above, the values of the left and right options can be computed as the weighted sum of the lottery expected values and the fractal Q values. The group-level choices and RTs are plotted as a function of the difference between the resulting left and right values in **Figure 4.10**. These results validate our model by showing that subjects' behavior in the task was compatible with the values generated by the model. Note that, although we show RTs in this analysis, we performed the model fitting using only choice data, and RTs were not included in the fitting procedure.

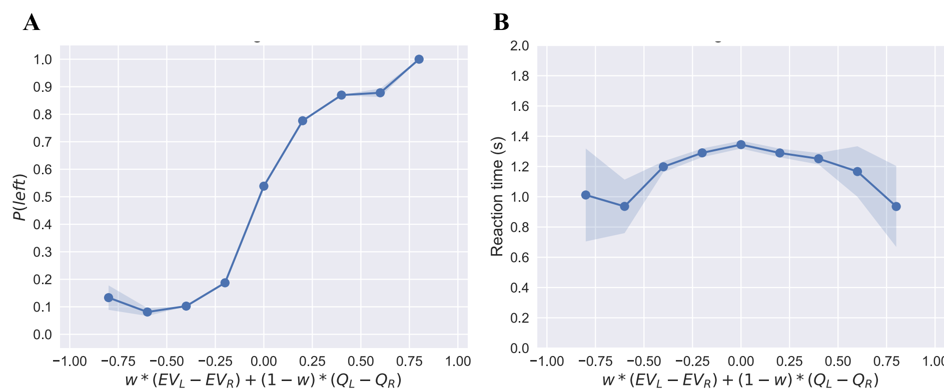

Figure 4.10: Basic psychometrics. (A) Choice curve as a function of value difference. (B) Response time as a function of value difference. Data was aggregated from all subjects. Shaded error bars show 95% confidence intervals for the data pooled across all subjects.

We also looked at the influence of w on choices. **Figure 4.11** shows the choice curves as a function of lottery expected value difference, with curves grouped by the relative weight of the descriptive system, w, and as a function of fractal Q value difference, with curves grouped by the relative weight of the experiential system, $1 - w$. **Figure 4.12** shows the coefficients obtained from two mixed effects models, one for choices and one for RTs, which grouped trials based on the value of w (the same grouping used in **Figure 4.11**). The choice mixed effects model was a logistic regression where the dependent variable was choice, and the independent variables were the lottery expected value difference and the fractal Q value difference. The RT mixed effects model was equivalent, but we used the absolute value of the differences for both lotteries and fractals. We can see from these results that subjects' choices and RTs were modulated by the model's weight w. Additionally, when $0.4 \leq w \leq 0.6$, both the lotteries and the fractals had a similar impact on choices and RTs, as evidenced by the similar coefficients for that trial group in both mixed effects models.

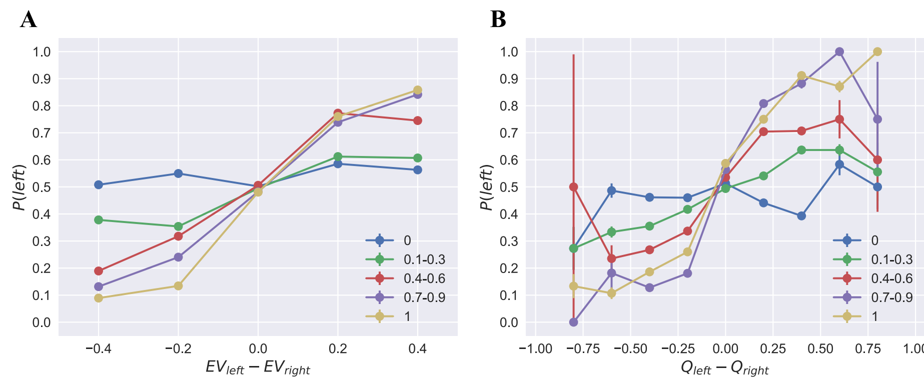

Figure 4.11: Choice curves grouped by weights. (A) Choice curves as a function of lottery expected value difference, grouped by w. (B) Choice curves as a function of fractal Q value difference, grouped by $1-w$. Data was aggregated from all subjects. Error bars show 95% confidence intervals for the data pooled across all subjects.

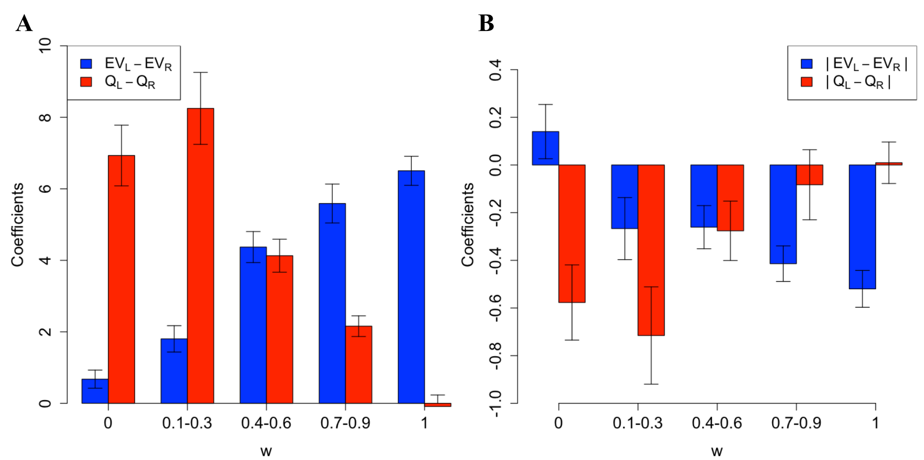

Figure 4.12: Coefficients of the mixed effects models, with trials grouped by w. (A) Coefficients for choice logit mixed effects model. (B) Coefficients for RT mixed effects model.

The results discussed so far indicate that our computational model provides a reasonable account of subjects' psychometrics in this task. We found that both descriptive and experiential values were taken into account by subjects, and that the relative influence of each type of value was modulated by a weight w which is a non-linear function of the probability π and which is different for each subject.

Role of Changes in the Relative Strength of Preferences in the Two Systems

In the previous subsection we looked into the influence of an exogenous experimental variable, π, on the relative importance given by subjects to the two decision systems, descriptive and experiential. We now investigate whether the relative strength of preference within each system also affects subjects' behavior. In our paradigm, it makes sense to take this measure into account, since the difference in values between lotteries or fractals may also impact the relative relevance of the two value systems. For instance, if the difference between the lottery expected values in a particular trial is zero, then an optimal decision maker would only consider the fractal value difference when making a choice for that trial.

We modified the non-linear model described above such that it contains a measure of bias relevant to our task that can flexibly affect choices, while allowing for the possibility of no existing bias by including the simpler model as a special case. It is important to note, however, that the nature of this test is qualitative, in that it can indicate whether or not the relative strength of preference in each system has an impact on choices, but cannot quantitatively describe this influence.

We used the absolute difference between the left and right lottery expected values as a measure of relative strength of preference in the descriptive system, and the absolute difference between the left and right fractal Q values as a measure of relative strength of preference in the experiential system. Using these two metrics, we computed a weight adjustment variable B which varied per trial, and which we defined as:

$$B = \frac{|EV_{\text{left}} - EV_{\text{right}}|^{\kappa}}{|EV_{\text{left}} - EV_{\text{right}}|^{\kappa} + |Q_{\text{left}} - Q_{\text{right}}|^{\kappa}}, \tag{4.17}$$

where κ is an additional free parameter which controls the shape of the curve B as a function of the absolute lottery expected value difference. **Figure 4.13** shows examples of curves obtained for B as a function of the absolute lottery expected value difference, for different values of κ and different values of the absolute fractal Q value difference. When $\kappa = 0$, we get $B = 0.5$, indicating no bias towards either system. On the other hand, when κ goes to infinity, we obtain a winner-takes-all setting, in which there is a full bias towards the system with the largest relative strength of preference (i.e., $B = 1$ when $|EV_{\text{left}} - EV_{\text{right}}| > |Q_{\text{left}} - Q_{\text{right}}|$, and $B = 0$ when $|EV_{\text{left}} - EV_{\text{right}}| < |Q_{\text{left}} - Q_{\text{right}}|$).

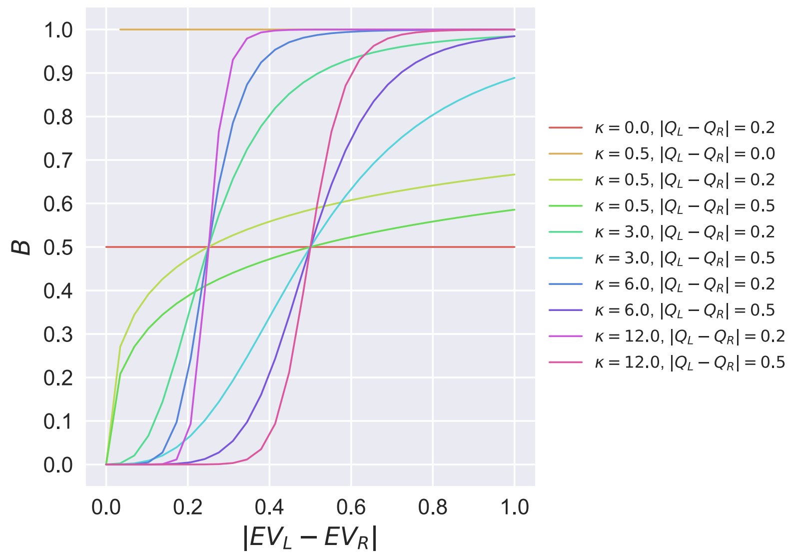

Figure 4.13: Examples of the weight adjustment curve B. Curves were obtained as a function of the absolute lottery expected value difference, for different values of κ and different values of the absolute fractal Q value difference.

Using B, we computed the relative weight given to the descriptive system, u, as:

$$u = \mu w + (1 - \mu)B \tag{4.18}$$

where μ is an additional free parameter which controls the relative contributions of the weight w and of the weight adjustment variable B to the weight u. We called this modified version of the non-linear model the nested model, which contained a total of six free parameters: the learning rate α, the inverse temperature β, the weighting function parameters γ and δ, plus κ and μ. Note that when $\mu = 1$ this model reduces to the non-linear model described above, and when $\mu = 0$ the weights are driven solely by the relative strength of preference. Since the two models are nested, they provide a natural qualitative way of testing the relative contribution of these two mechanisms to the integration and competition between the descriptive and experiential valuation systems.

We fitted the nested model for all subjects individually, through a maximum likelihood estimation procedure. For each subject, we performed a grid search over the six-parameter space, and chose the combination of parameters that yielded the largest likelihood value. The results of this fitting procedure are shown in **Figure 4.14**. Summary statistics are provided in **Table 4.5**.

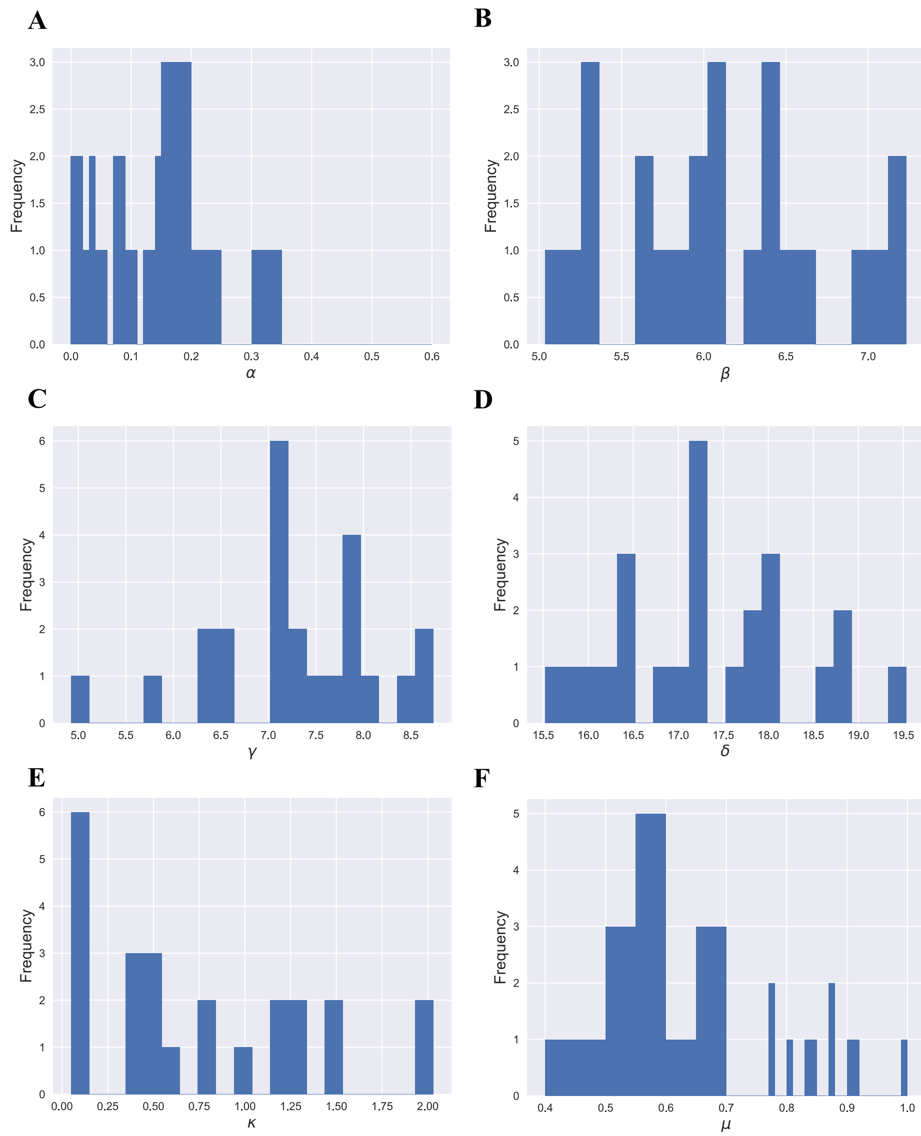

Figure 4.14: Histograms for the nested model parameters, fitted per subject. (A) α. (B) β. (C) γ. (D) δ. (E) κ. (F) μ.

Table 4.6 shows a comparison between the nested and the non-linear models using

Table 4.5: Nested model fitting summary statistics.

Parameter	**Mean**	**SD**
α	0.1	0.08
β	6.09	0.63
γ	7.31	0.9
δ	17.34	0.99
κ	0.76	0.6
μ	0.69	0.16

two evaluation metrics: negative log-likelihood (NLL) and Bayesian information criterion (BIC), where smaller numbers indicate better fittings. The table shows that, using the NLL as the evaluation metric, the non-linear model generates a better fit for all but one subject and, using the BIC, the same model generates a better fit for all subjects.

As an additional comparison between the non-linear and nested models, we simulated each of them using the trial conditions from the experiment and the best fitting parameters obtained for each subject, then compared the choices generated by each model with the subjects' actual choices. This process was repeated 100 times. The non-linear model was able to correctly predict, on average, 67% of choices (mean = 0.67, SD = 0.0041), while the nested model correctly predicted 64% of them (mean = 0.64, SD = 0.0046), and this difference was statistically significant (two-sample t-test, $t = 29.89$, $p < 10^{-16}$).

We found no significant difference between the α parameter fitted in the nested and the non-linear models (mean $\alpha = 0.1$ vs. 0.13; paired t-test, $p = 0.08$), a moderately significant difference for the β parameter (mean $\beta = 6.09$ vs. 7.79; paired t-test, $p = 0.01$), and significant differences for the γ (mean $\gamma = 7.31$ vs. 2.42; paired t-test, $p < 10^{-16}$) and δ (mean $\delta = 17.34$ vs. 3.57; paired t-test, $p < 10^{-16}$) parameters. Importantly, we found that the mean μ parameter fitted in the nested model was significantly less than 1 ($t = -9.295$, $p < 10^{-16}$), indicating an influence of the weight adjustment variable on subjects' choices.

Our qualitative results for the nested model fitting revealed that the strength of relative preference within the descriptive and experiential systems exerts a significant but relatively small influence on the relative weight that subjects assign to each system. Instead, subjects appear to heavily rely on the exogenous variable π to

Table 4.6: Comparison between non-linear and nested model fittings.

Subject	NLL nested	NLL non-linear	BIC nested	BIC non-linear
1	156.15	156.42	346.52	335.65
2	178.7	171.64	391.57	366.07
3	137.72	127.58	309.66	277.97
4	156.76	154.63	347.74	332.07
5	145.95	144.38	326.1	311.57
6	130.22	119.13	294.67	261.08
7	175.56	165.6	385.13	353.89
8	125.29	121.67	284.81	266.16
10	158.14	155.14	350.47	333.06
11	146.41	144.95	326.92	312.63
12	147.31	143.77	328.84	310.35
13	147.75	137.61	329.7	298.02
14	198.96	189.69	432.05	402.14
15	151.52	146.82	337.26	316.46
16	171.19	169.53	376.61	361.88
17	166.96	163.58	368.14	349.97
18	130.2	117.06	294.53	256.87
19	173.83	168.92	381.88	360.65
20	130.67	122.9	295.56	268.62
22	139.37	136.74	312.91	296.26
23	183.88	175.56	401.89	373.87
24	171.12	169.69	376.44	362.18
25	145.27	136.27	324.68	295.3
27	139.9	124.38	314.0	271.57
Mean:	154.53	148.49	343.25	319.76
SD:	18.96	20.05	37.9	40.08

compute these weights.

Role of Changes in the Relative Uncertainty of the Two Systems

Previous literature has found considerable evidence for competition between distinct decision and learning systems in decision making tasks with humans. Several modeling studies have explored competitive architectures, in which arbitration is performed based on some metric that gives preferential control to one of the two systems, such as relative uncertainty or accuracy. One example of this kind of architecture is that of uncertainty-based arbitration, in which the controller chooses the system with more accurate value estimates (Lee, Shimojo, and J. P. O'Doherty, 2014). A related architecture implements cost-benefit arbitration, i.e., the controller chooses the system more likely to optimize reward relative to cognitive effort (Kool, F. A. Cushman, and Gershman, 2016; Kool, Gershman, and F. A. Cushman, 2017). Here, we investigate the possibility of uncertainty-based competition between the descriptive and experiential systems in our study, using the variance from a Bayesian update model for the fractal probabilities as the relevant uncertainty measure.

We first applied a Bayesian update model to the probabilities of reward associated with each fractal. This model was implemented as an agent with full knowledge of the characteristics of the Gaussian random walk that we used to compute the fractal probabilities. We then used the sum of the standard deviations of the posterior distributions obtained for the left and right fractal probabilities as our trial-by-trial ex post uncertainty measure. Note that only the fractal probabilities could be used as a source of uncertainty in our paradigm, since there was no putative trial-by-trial variation stemming from the descriptive system.

The Bayesian update model was implemented as follows. We began by assuming full knowledge of the drift rate and of the upper and lower bounds applied to the fractal probability signals. The model contains no free parameters. At each trial, we generated a posterior distribution over all possible values for each of the two fractal probabilities. More details about the posterior computation are provided in the Materials and Methods section. **Figure 4.15** shows the mean and 95% confidence interval for the Bayesian estimates obtained for 3 different subjects throughout the experiment, as well the true values of the fractal probabilities and the Q values obtained from the Q-learning model. The figure shows that the Bayesian update model correctly tracks the fractal probabilities throughout the trials. In addition, note that the learning rate α, which was fitted per subject, strongly dictates how well subjects were able to estimate the values of the fractals: when α is very small, as shown for subject 17, the Q values provide a poor estimate for the probabilities; medium values of α, as with subject 22, lead to good estimates; and high values of α, as with subject 27, lead to high volatility in the Q values.

To validate the uncertainty measure we defined, we looked at the histogram of this metric across all trials from all subjects, and at the correlation between the difference between the mean of the Bayesian posteriors and the difference between the Q values, again for all trials from all subjects. Both results are shown in **Figure**

4.16. We found that there was a large spread of the uncertainty measure across trials and subjects (mean range across subjects = 0.17, SD range = 0.014, max range = 0.21, min range = 0.14), and that the estimated Q values were highly correlated with the mean estimates from the Bayesian update model. These results indicated that we could use the standard deviations from the Bayesian estimates as a proxy measure for uncertainty in the experiential system, in which subjects' valuations were represented through Q values.

In order to compare the extent to which the Q-learning model vs. the Bayesian estimates explained choices, we computed a logistic mixed effects regression applied to all trials where $\pi = 0$. We used the probability of choosing left as the dependent variable, and the difference between fractal Q values and the difference between Bayesian posterior means as the independent variables. The resulting coefficients for this model are shown in **Table 4.7**. We can see from the table that the Q value difference absorbs most of the variance in choices and is highly significant, whereas the coefficient obtained for the difference between Bayesian estimates is not significant.

Table 4.7: Choice logit mixed effects model: trials where $\pi = 0$.

Regressor	**Estimate**	**Std. Error**	**z value**	**p-value**
Intercept	-0.0009044	0.1854154	-0.005	0.996
$Q_L - Q_R$	6.5582566	1.0949574	5.990	2.1e-09
Bayesian estimate diff.	1.2407964	1.2383145	1.002	0.316

Finally, to carry out the main test of interest, we computed a mixed effects logistic regression of choice applied to all trials, using the uncertainty measure to modulate value differences. In this model the dependent variable was the probability of choosing left, and the independent variables were the weighted lottery expected value difference, the weighted fractal Q value difference, the weighted lottery expected value difference multiplied by the uncertainty measure, and the weighted fractal Q value difference multiplied by the uncertainty measure. The values of the uncertainty measure were scaled between 0 and 1 to facilitate model convergence. The resulting coefficients for this model are shown in **Table 4.8**. We found no effect of modulation of the uncertainty measure on the lottery expected value difference ($p = 1$) or on the fractal Q value difference ($p = 0.32$). To further confirm this result, we ran a similar analysis but separated the regressors into two mixed effects models. In the first model we used only the weighted lottery expected value difference and the weighted lottery expected value difference multiplied by uncertainty as regressors; in the second model, we used the weighted fractal Q value difference and the weighted fractal Q value difference multiplied by uncertainty. In both models, we again found no effect of the uncertainty measure on choice ($p = 0.37$ for the lottery expected value difference modulated by uncertainty, and $p = 0.39$ for the fractal Q value difference modulated by uncertainty).

Table 4.8: Choice logit mixed effects model with Bayesian uncertainty.

Regressor	**Estimate**	**Std. Error**	**z value**	**p-value**
Intercept	2.344	3.278	0.715	0.47465
$w\|EV_L - EV_R\|$	9.050	1.367	6.621	3.56e-11
$w\|Q_L - Q_R\|$	9.370	3.104	3.018	0.00254
uncertainty $\times w\|EV_L - EV_R\|$	-2.470	1.625	-1.520	0.12856
uncertainty $\times w\|Q_L - Q_R\|$	-2.001	5.571	-0.359	0.71953

It is important to note that the metric we used as the uncertainty is an indirect measure, as it was computed from the posteriors generated by an ideal Bayesian learner and therefore does not necessarily correspond to the uncertainty experienced by subjects during the task. Motivated by the reliability measure used by Lee and colleagues (Lee, Shimojo, and J. P. O'Doherty, 2014), we performed an additional test using an alternative measure of uncertainty which is more directly related to the Q-learning model, and which we defined as the sum of the absolute prediction errors for the left and right fractals. We then ran a similar mixed effects model to the one described above, but replacing the previous uncertainty measure with the alternative one based on prediction errors. The resulting coefficients for this model are shown in **Table 4.9**. As can be seen in the table, we again found no modulation effect of uncertainty on subjects' choices.

Our results from this subsection indicate that the Bayesian uncertainty measure, defined as the sum of the standard deviations from the posterior distributions for left and right fractal probabilities, exerted no influence on the relative weight given to the descriptive and the experiential decision systems. We similarly found no influence of an uncertainty measure based on the prediction errors from the Q-learning model.

Table 4.9: Choice logit mixed effects model with uncertainty based on prediction errors.

Regressor	Estimate	Std. Error	z value	p-value
Intercept	0.01309	0.10385	0.126	0.900
$w\|EV_L - EV_R\|$	6.76754	0.83850	8.071	6.97e-16
$w\|Q_L - Q_R\|$	6.68773	1.50362	4.448	8.68e-06
uncertainty $\times w\|EV_L - EV_R\|$	0.26221	0.75247	0.348	0.727
uncertainty $\times w\|Q_L - Q_R\|$	1.69346	1.56166	1.084	0.278

Role of Conflict Between the Two Systems

The value integration models we used to study subjects' behavior in our task suggest that the presence of conflict between the descriptive and the experiential systems should be irrelevant to the computation of choices. Here we define conflict as the condition in which one valuation system recommends one choice, while the other recommends a different choice. In our paradigm, this corresponds to trials where $(EV_{\text{left}} - EV_{\text{right}}) \times (Q_{\text{left}} - Q_{\text{right}}) < 0$. To investigate whether this conflict indeed had no impact on choices in our data, we performed a comparison between conflict and no-conflict trials.

First, we computed a logistic choice mixed effects model using conflict and no-conflict indicator variables to modulate the total value difference between the left and right options. The coefficients for this model are shown in **Table 4.10**. We found a significant difference between the coefficients obtained for conflict and no-conflict trials ($p = 1.18 \times 10^{-5}$), indicating an impact of conflict on choice.

Table 4.10: Choice logit mixed effects model with conflict/no-conflict indicators.

Regressor	Estimate	Std. Error	z value	p-value
Intercept	-0.03015	0.12393	-0.243	0.808
conflict	0.09732	0.08784	1.108	0.268
conflict $\times(V_L - V_R)$	6.33325	0.52512	12.061	<2e-16
no-conflict $\times(V_L - V_R)$	7.97761	0.62846	12.694	<2e-16

We then computed an equivalent mixed effects model for RT, but using the absolute value of the total value difference between left and right. The coefficients for this model are shown in **Table 4.11**. Here again we found a significant difference between the coefficients obtained for conflict and no-conflict trials ($p = 1.18 \times 10^{-5}$), indicating an impact of conflict on RTs.

Table 4.11: RT mixed effects model with conflict/no-conflict indicators.

Regressor	Estimate	Std. Error	t value
Intercept	1.41448	0.06062	23.332
conflict	-0.04961	0.02517	-1.971
conflict $\times \lvert V_L - V_R \rvert$	-0.33477	0.09648	-3.470
no-conflict $\times \lvert V_L - V_R \rvert$	-0.68870	0.05769	-11.938

Figure 4.17 shows the choice and RT curves, as a function of the left and right integrated value difference, separating conflict from no-conflict trials. Qualitatively, these curves show a small difference in both choices and RTs between these two types of trials.

Overall, our analysis of the conflict between the descriptive and experiential valuation systems indicated an influence of conflict on choices. Nevertheless, our simple computational model of choice based on a non-linear transformation of the variable π was able to explain much of the variance present in the choice data without taking conflict explicitly into account. Future work should further investigate how conflict impacts choices by explicitly incorporating conflict into the choice models.

4.5 Discussion

In this chapter we discussed the results from an experiment testing the interactions between a descriptive and an experiential valuation systems. Our task design required subjects to take into account both types of values simultaneously in order to make a choice, and included a variable π that provided an exogenous manipulation of the relative relevance of the two valuation systems. The psychometric evidence from this experiment suggested a very simple model of arbitration between systems: choices were made based on a total weighted value signal with a relative system weight which was responsive to the exogenous variable π and which we found to be non-linear in most subjects. Additionally, we found only a small influence of the relative strength of preference (i.e., to what extent one option was better than the other) and no influence of the relative uncertainty on the weight of the two systems, and obtained evidence for a small difference in the interaction between systems on conflict vs. no-conflict trials.

While our analysis focused on choices, it would be a straightforward extension to model RTs in conjunction with choices by applying a Drift-Diffusion Model to our behavioral data. It would be particularly interesting to investigate the role of attention in this task by using an attentional Drift-Diffusion Model taking into account subjects' visual fixations, as we described in Chapter 2 of this dissertation. One potential mechanism through which visual attention affected subjects' behavior in our task is that attention fluctuated in a way consistent with the weight curves, and differently in each subject. Therefore, one direction for future work is to test this hypothesis through a behavioral experiment including eye-tracking, where one could check if the individual biases toward the descriptive or the experiential system correlate with the fixation time given to the corresponding options.

In our tests of the effect of the relative strength of preference in the two valuation systems, as measured by the absolute difference between lottery expected values and fractal Q values, we found evidence for only a small impact on choices. This is interesting because this type of modulation is predicted by models in which the two valuation systems compete for control based on the relative strength of their signals, which is akin to confidence or strength of preference. Instead, the results presented here are consistent with a simpler value integration of a total value signal which is then used to carry out choices.

We also investigated the impact of uncertainty within the valuation systems by using the standard deviation of the Bayesian posteriors for the fractal probabilities as a measure of uncertainty in the experiential system, and found no effect on choice. The lack of an effect observed here may have been due to having an incorrect measure of the uncertainty driving the competition, since the metric was computed from the posteriors generated by an ideal Bayesian learner, which does not necessarily correspond to the uncertainty experienced by subjects during the task. However, using the prediction errors from the Q-learning model as a measure of uncertainty similarly did not reveal any effects. A more comprehensive comparison is needed to better understand this interaction.

When looking at conflict between the two valuation systems (i.e., trials where the two systems disagreed on what was the best option), we found a small influence of conflict on subjects' choices. However, we did not attempt to model this interaction explicitly. Future work should further investigate how conflict affects choices and response times.

An interesting effect present in our data is that the negative coefficient for the absolute fractal value difference in $\pi = 0$ trials was twice as large as the negative coefficient for the absolute lottery value difference in $\pi = 1$ trials. We hypothesize this is due to the different ways in which these two types of values were computed. Since the individual fractal values were retrieved from memory and required no explicit computation, the preference for one fractal over the other could be decided during the probability screen, before the subject even saw the choice options. This means that, when $\pi = 0$, the subject could make a choice between left and right before they saw the choice screen, since the lotteries presented were irrelevant, leading to faster RTs. On the other hand, the lottery expected values always had to be computed explicitly once the choice screen appeared, and because in $\pi = 1$ trials they were especially relevant, RTs in those trials tended to be longer.

Two additional features of our experimental paradigm are worth noting. First, both rewards drawn from the fractals were shown in every trial. Therefore, in trials where the actual reward received was one drawn from a fractal, the reward in the opposite (non-chosen) side was counter-factual, i.e., it was not actually experienced by the subject. In trials where the reward was drawn from a lottery, both rewards shown for the fractals were counter-factual. In future work, it may be important to model the learning rates related to these rewards more carefully, for instance, by using distinct learning rates based on whether the reward was received or not, as well as based on whether the reward was drawn from the side selected by the subject or not. Second, it may be useful to understand potential interactions between model-free and model-based learning within the experiential system as applied to our task: model-free learning may occur when a reward is received and used to update Q value estimates, whereas model-based learning may occur when a reward is not received, but merely observed and used to update a model of the reward structures in the task. These additional investigations may reveal interesting differences in the learning process across subjects, and may further elucidate the interactions between the two valuation systems.

Finally, while this chapter focused on the analysis of behavioral data, we also collected neuroimaging data from all participants. In future work, we will analyze the fMRI data by testing predictions based on the behavioral modeling results presented here. In particular, it will be important to check whether the descriptive and experiential systems have corresponding distinct value representations in the brain. Such a result would either replicate the findings by FitzGerald et al. (FitzGerald et al., 2010), or potentially lead to different regions of activation than those obtained by the authors. Based on previous findings (Bechara, H. Damasio, et al., 1999; N. D. Daw, J. P. O'doherty, et al., 2006; Gläscher, Hampton, and J. P. O'doherty, 2008; Todd A Hare, C. F. Camerer, and Rangel, 2009; Chib et al., 2009; McNamee, Rangel, and J. P. O'doherty, 2013), we also expect to see a combined value signal, incorporating the values from both systems, in ventromedial prefrontal cortex. A related interesting direction is to look for an activation correlated with the weight w from our non-linear model, above and beyond the value of π, which is the linear setting, and of a simple step function, which corresponds to a winner-takes-all setting.

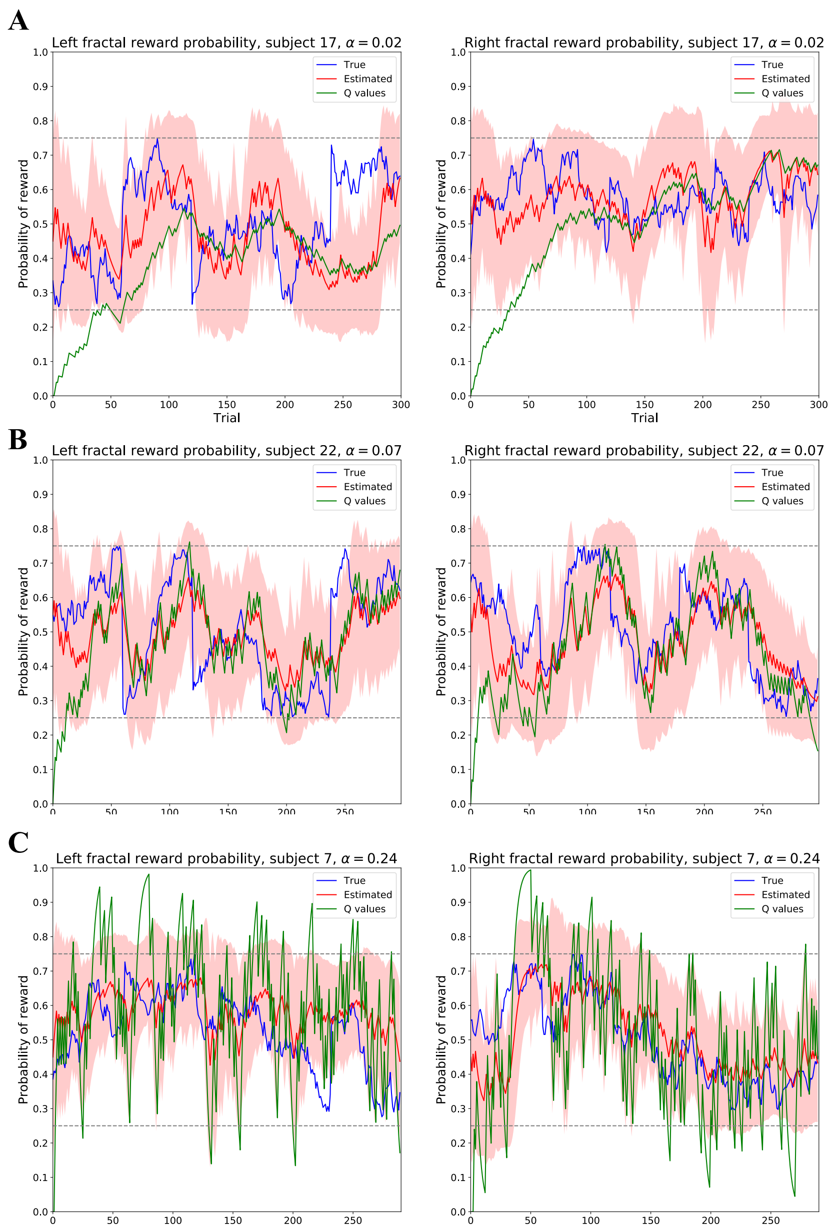

Figure 4.15: Bayesian update model estimates. The mean and 95% confidence interval for the Bayesian posteriors (in red) are shown for 3 different subjects. Also displayed are the true values of the fractal probabilities (in blue) and the Q values obtained from the Q-learning model (in green). (A) Example trial sequence with low learning rate, $\alpha = 0.02$. (B) Example trial sequence with medium learning rate, $\alpha = 0.07$. (C) Example trial sequence with high learning rate, $\alpha = 0.24$.

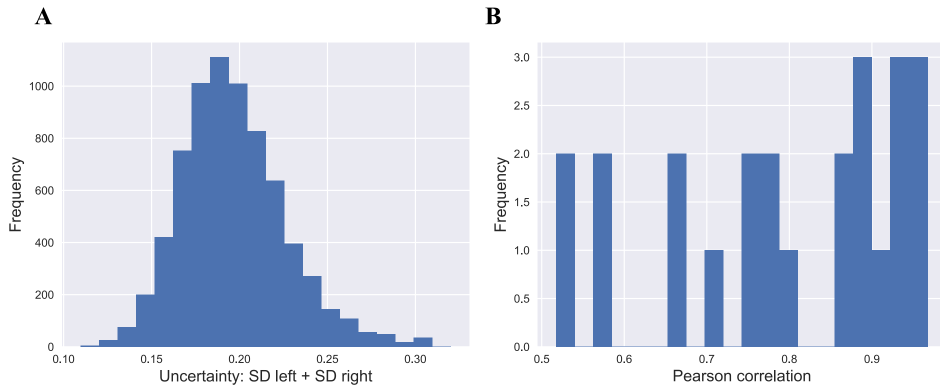

Figure 4.16: Validation of the Bayesian measure of uncertainty. (A) Histogram of the uncertainty measure, which corresponds to the sum of standard deviations from the posterior distributions obtained for left and right fractal probabilities, across all trials from all subjects. (B) Histogram of the Pearson correlation coefficients between the difference between the mean of the Bayesian posteriors and the difference between the Q values, for all trials from all subjects.

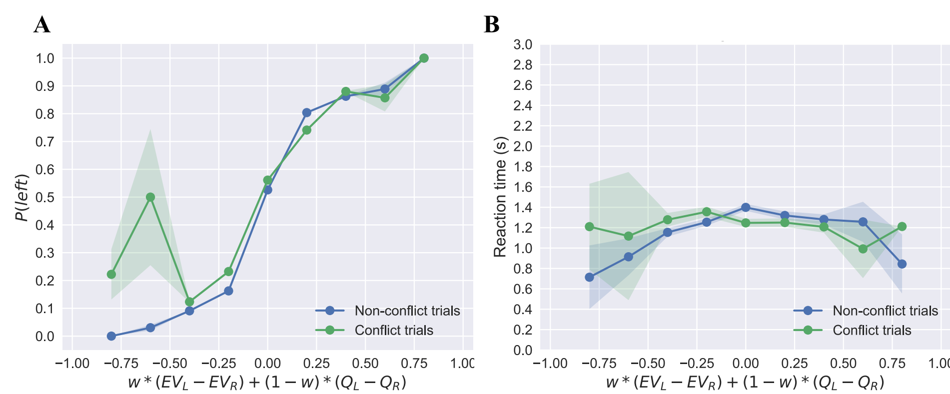

Figure 4.17: Basic psychometrics, conflict vs. no-conflict trials. (A) Choice curve as a function of total value difference, with trials grouped by whether or not there was conflict between the two systems. (B) Choice curve as a function of total value difference, with trials grouped by whether or not there was conflict between the two systems. Data was aggregated from all subjects. Shaded error bars show 95% confidence intervals for the data pooled across all subjects.

Chapter 5

DISCUSSION

5.1 Summary of Results

In this dissertation I have summarized the results of three projects that approach the subject of human decision making from a behavioral, cognitive and computational neuroscience perspective. These projects contribute to the study of the neurocomputational basis of decision making by proposing models that make precise predictions about simple decisions in various types of settings, then testing these predictions in the behavioral data acquired from human subjects.

One such model which has been extensively used in the computational and cognitive neuroscience literature is the Drift-Diffusion Model (DDM). The DDM is a type of sequential integrator model in which evidence for the available options accumulates over time until it reaches a threshold, leading to a decision being made. In Chapter 2, I discussed the application of a variation of the DDM, called the attentional DDM (aDDM), in which visual fixations bias choices towards the fixated item, to a perceptual decision making experiment with human subjects. Our results from this experiment showed that the aDDM can make reasonably accurate qualitative and quantitative predictions for choice and response time data from a perceptual decision task. Importantly, the aDDM was able to explain choice biases observed in our data that could not be accounted for by the version of the model without an attentional component. Moreover, our causal manipulation showed that artificially increasing fixation time on one of the options effectively increases the probability of choosing that option, which is in line with the model predictions. These results extend those obtained by Krajbich and colleagues in the context of economic decision making (Krajbich, C. Armel, and Rangel, 2010; Krajbich and Rangel, 2011; Krajbich, Lu, et al., 2012), providing some generalization of the ability of the aDDM to account for behavioral data stemming from different choice domains and to explain the role of attention in the decision process.

Chapter 3 also dealt with the DDM: we compared two different methods which can be used to fit this type of model to experimental data. The first method, called the Maximum Likelihood Algorithm (MLA), has been widely used in the literature, and involves generating simulations of the model to be compared against the data, thus