Computation and comparison of value signals in simple perceptual and economic choices

Thesis by
Gabriela de Oliveira Penna Tavares

In Partial Fulfillment of the Requirements for the
degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2018
Defended December 11, 2017

Gabriela de Oliveira Penna Tavares
ORCID: 0000-0002-8624-1758

ACKNOWLEDGEMENTS

I feel so incredibly lucky to have been able to pursue a PhD at Caltech, and humbled by the support that I received along the way from so many wonderful people.

I am extremely grateful to my adviser, Antonio Rangel, for believing in my potential right from the start and for his unfailing encouragement throughout the past 4 years. Observing his excitement and passion for science has been an inspiration that I will carry with me for the rest of my life.

Pietro Perona was the first professor I met at Caltech. He was an essential collaborator in two of my projects (described in Chapters 2 and 3 of this dissertation), and brilliantly led the CNS program for most of my PhD. His wisdom, mentorship, and kindness guided me through many difficult times.

I would also like to express my gratitude to the other two members of my thesis committee, John O'Doherty and Ralph Adolphs, not only for great academic advice and scientific discussions, but also for inviting me to join many of their lab meetings, dinners, and trips.

My masters thesis supervisor at Imperial College London, Aldo Faisal, is the person who first sparked my academic interest in neuroscience. Working with him instilled in me much of the knowledge and confidence I needed in order to apply for and to pursue a PhD at Caltech.

I was also lucky to work with some incredibly talented people throughout my PhD. I am grateful for the friendships and for the many things I have learned from former and current members of the Rangel, O'Doherty, and Adolphs labs. In particular, Giovanni Gentile has been a truly great mentor and friend who taught me a lot about doing science the right way, as well as everything I know about fMRI research.

The past 4 years have been a very happy time in my life. I am thankful for having shared so many wonderful moments with my Caltech colleagues and friends. The sense of community and friendship I have gained from knowing them gave meaning to my time in grad school. In particular, it has been an honor to be alongside my CNS family, Janis Hesse, Mason McGill, Matteo Ronchi, Ryan Cho, and Vineet Augustine, through the many ups and downs of this journey.

I could not have asked for better people with whom to share my DNA than my amazing family, Rachel, Cláudio, Milla, Antonia, Cee, Arthur, Fernando, Fau,

Nanda, Myrian, Paulo, Martha, Artur, Paulinha, Feli, João, and Doka. Their faith in me and unconditional love have been essential in getting me where I am today.

I am forever grateful to my parents, Rachel and Cláudio, not only for the substantial financial investment they made in my education, but also for teaching me that with hard work and perseverance I could achieve anything I wanted.

My grandparents are an inspiration to me every day. I am thankful to my grandma, Myrian, from whom I have learned about patience, grace, and strength, and to my grandpa, Paulo, from whom I have learned about curiosity, determination, and self-motivation.

I am also grateful to my Albanian family, Moza, Halit, Rudi, Jessie, Linda, Edmond, Kevin, and Alba, and I feel so fortunate to be a part of their wonderful community.

My Brazilian friends, Aninha, May, Naldo, Nalu, Tine, Tito, and Tuka, have taught me what true friendship is, and even after 15+ years, many of which I have been away, being with them still feels like home.

Finally, my husband Juri has been my one constant amid the challenges and uncertainty of the past 4 years. His dedication to science has inspired me greatly; his love, patience, and kindness continue to amaze me every day. He is the greatest gift Caltech has given me.

ABSTRACT

How do we choose between different foods from a restaurant menu, or between a vacation overseas and more money in our savings account? Certain mechanisms in our brains allow us to make these and many other kinds of decisions effectively and efficiently. In this dissertation, I describe three projects which aim to advance our understanding of the systems and algorithms involved in the process of human decision making.

Chapter 2 investigates the application of the attentional Drift-Diffusion Model to a perceptual decision making task. Perceptual decisions requiring the comparison of spatially distributed stimuli that are fixated sequentially might be influenced by fluctuations in visual attention. We used two psychophysical tasks with human subjects to investigate the extent to which visual attention influences simple perceptual choices, and to test the extent to which the attentional Drift-Diffusion Model provides a good computational description of how attention affects the underlying decision processes. We found that this model provides a reasonable quantitative description of the relationship between fluctuations in visual attention, choices, and response times. We also found evidence for the sizable attentional choice biases predicted by the model, and that exogenous manipulations of attention induce choice biases consistent with these predictions.

Chapter 3 compares two methods for fitting the parameters of the Drift-Diffusion Model using experimental data. A large number of studies have proposed that sequential integrator models of decision making, such as the Drift-Diffusion Model and its variants, provide a simple computational description of the algorithms used to make a large number of simple decisions. This is based on the fact that this class of models has been able to produce reasonably accurate descriptions of how choices, response times, and fixations are related to each other and to exogenous trial parameters, in a wide range of tasks. A difficult step in those studies is the estimation of a small number of free parameters to find the ones that explain the observed data best. The estimation method used in most studies is computationally very expensive since it approximates the likelihood of the observed data by simulating the model thousands of times and then counting the frequency with which the outcomes match the observed data. This problem is exacerbated with more complex models, such as the attentional Drift-Diffusion Model, or models with collapsing bounds, which contain a larger number of free parameters. We propose an alternative method for

estimating the free parameters which relies on computing only the probability of the actual observed data, bypassing the need for the additional simulations. We present the results of simulation tests which show that our approach provides two key advantages over the alternative widely used method: a smaller number of experimental trials is needed in order to obtain comparable estimation accuracy, and the execution time of the estimation algorithms is substantially reduced.

Finally, Chapter 4 studies simple economic choices involving two distinct classes of valuation systems: an experiential system, which assigns value based on the history of previous reward experiences with similar options, and a descriptive system, which computes values using information about the options and environment available at the time of decision. Although these two systems often assign similar relative desirability to the different options, they do not always do so. When conflict arises with the experiential system favoring one option and the descriptive system favoring another, the brain needs to resolve the conflict to select a single option. We present the results of a psychometric study designed to characterize the basic interactions of these two valuation systems, with and without conflict.

PUBLISHED CONTENT AND CONTRIBUTIONS

Tavares, Gabriela, Pietro Perona, and Antonio Rangel (2017). "The Attentional Drift Diffusion Model of Simple Perceptual Decision-Making". In: *Frontiers in Neuroscience* 11, p. 468. ISSN: 1662-453X. DOI: 10.3389/fnins.2017.00468. URL: https://www.frontiersin.org/article/10.3389/fnins.2017.00468.

G.T. collected the data and participated in designing the experiment, analyzing the data and writing the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

NOMENCLATURE

aDDM. Attentional Drift-Diffusion Model.

DDM. Drift-Diffusion Model.

fMRI. Functional magnetic resonance imaging.

ITI. Inter-trial interval.

MLA. Maximum Likelihood Algorithm.

PTA. Probability Table Algorithm.

RT. Response time.

Chapter 1

INTRODUCTION

Imagine you are at your local diner, trying to decide on something to eat for dinner. You browse through the menu, reading the names of familiar food items, many of which you have tried in the past. It is possible that you know exactly what you want to eat even before you look at the menu; perhaps if you eat spaghetti every night, you decide you will eat spaghetti once again, without a second thought on the matter. However, you may be in the mood for something else, in which case you may remember times when you have had other dishes at that same diner, and how much you enjoyed each of them, and that should help you make a decision. Or you may want to try something new, in which case you may imagine how much you will enjoy each item you are considering ordering.

Now suppose you are traveling through a new country, where you do not speak the local language, and are unfamiliar with the local cuisine. Again you find yourself at a restaurant trying to decide what to eat for dinner, but this time you have very little information from previous experiences with the kind of food that is available. How do you decide in this case? It may be useful to look at what other people have ordered, and try to infer based on appearance and smell what you might like to eat.

The scenarios described above involve very simple every-day decisions, but they show us that even simple problems such as deciding what to eat for dinner involve deliberation, memory and construction of mental projections about the future. In neuroscience, this corresponds to the idea that decision making recruits several different cognitive processes taking place in the brain. The current dissertation summarizes the results of three projects that contribute to the goal of characterizing the neurocomputational basis of decision making. This first chapter provides some context about the current state of the neuroscience literature on this and directly related topics.

1.1 Frameworks for the Study of Decision Making

In the past few decades, researchers in neuroscience, psychology, and economics have dedicated considerable effort to the study of decision making, and have also developed computational frameworks that aim to facilitate the analysis of the related

experimental results (Kahneman and Tversky, 1979; Jerome R Busemeyer and Johnson, 2004; Gold and Shadlen, 2007; Rangel, C. Camerer, and Montague, 2008; Glimcher and Fehr, 2013). Despite this effort, much is still unknown about the exact nature of the algorithms and computations taking place in the brain that allow us to make even simple decisions.

Rangel and colleagues formalized a framework for the study of the neurobiology of decision making (Rangel, C. Camerer, and Montague, 2008). They summarized five main processes that presumably take place in the brain during the course of a decision: representation (for a set of states and actions related to the decision), action valuation (assigning a value to each possible action), action selection (comparing the action values and generating a choice), outcome valuation (assigning a value to the outcome of the selected action), and learning (updating the representation, valuation and action selection processes according to the latest outcome). Each of these processes has been extensively studied in the literature, both in humans and in animal models. However, no consensus exists to date about the exact implementation of these processes in the brain, about how much of the neural circuitry is shared between them, and about whether the sequential organization described by Rangel et al. is indeed an accurate representation (Kable and Glimcher, 2009).

Similarly, Kable and Glimcher collected evidence from several studies to support their hypothesis of a two-stage neural mechanism for choice in primates, composed of a valuation stage and a choice stage (Kable and Glimcher, 2009). According to the authors, the valuation stage, in which values are assigned to alternatives, involves the ventromedial sectors of the prefrontal cortex and parts of the striatum, while the choice stage, in which an highly value option is selected and implemented, involves lateral prefrontal and parietal areas which are also part of a sensory-motor hierarchy.

A central piece in understanding the mechanism of decision making is the process through which values are constructed and compared against one another. Sequential integrator models constitute one prominent class of models that propose a solution to this problem. They involve the use of decision variables that accumulate evidence in favor of each of the available options, until one variable converges to a hard threshold, causing a choice to be made. In Decision Field Theory (DFT), for instance, Busemeyer and Townsend proposed a theoretical framework that uses such a sequential sampling process to generate decisions (Jerome R Busemeyer and Townsend, 1993). Using a similar approach, Ratcliff proposed the Drift-Diffusion Model (DDM) of choice (Ratcliff, 1978; Ratcliff and Rouder, 1998; Gold and Shadlen, 2002; Ratcliff

and McKoon, 2008), which is a simplification of the more general DFT where the dynamics of the model are approximated by linear systems. Finally, Usher and McClelland proposed the leaky competing accumulator model, a modification of the DDM in which non-linear decision units are used, information accumulation is subject to leakage, and representations of the alternative outcomes compete with each other through a process of lateral inhibition (Usher and McClelland, 2001).

1.2 Perceptual versus Economic Choice

An important distinction made in the study of decision making is that of perceptual versus economic (or value-based) decisions. While the former are about making discriminations that are based on sensory properties of the options available (for instance, deciding which of two circles is the largest), the latter involve making choices based on preference with the aim of maximizing economic value (for instance, deciding which of two brands of candy bar to eat or which of two slot machines to play).

The study of perceptual decision making has been a major focus of the animal research in neuroscience (Newsome and Pare, 1988; Newsome, Britten, and Movshon, 1989; Roitman and Shadlen, 2002; Hanks, Ditterich, and Shadlen, 2006). One popular paradigm that has been used in several experiments with non-human primates is the random dot motion task, first described by Newsome and colleagues (Newsome and Pare, 1988; Newsome, Britten, and Movshon, 1989). In this task, the animal sees a screen with several moving dots. A certain percentage of the dots moves in a specific direction, while the remaining ones move randomly, and the animal must decide which is the net direction of movement (left or right) by shifting its gaze towards one of two locations. The percentage of cohesively-moving dots in a particular trial determines the level of difficulty of the decision. Using this paradigm, researchers have determined that neurons in the macaque extrastriate cortex carry critical information about motion direction which allows the animal to report its decision through eye movement (Newsome and Pare, 1988; Salzman, Britten, and Newsome, 1990; Salzman, Murasugi, et al., 1992; Britten et al., 1992; Celebrini and Newsome, 1994; Celebrini and Newsome, 1995). Using the same task, Roitman and Shadlen found evidence for a process of accumulation of visual information in the macaque lateral intraparietal cortex (LIP), where neuronal activity during stimulus viewing predicted both the decision and the time required to reach it (Roitman and Shadlen, 2002). Later, Hanks et al. established a causal link between LIP neuronal activity and motion discrimination choices by stimulating LIP neurons and

observing an increase in the proportion of choices toward the response field of the stimulated neurons (Hanks, Ditterich, and Shadlen, 2006).

In humans, researchers have used neuroimaging data to investigate the mechanisms of perceptual decision making. For instance, Heekeren et al. used a categorization task with fMRI and found evidence that perceptual decisions are computed by integrating evidence from sensory processing areas, and that this integration may occur in dorsolateral prefrontal cortex (Heekeren et al., 2004). This result was in line with a study in monkeys which suggested the same mechanism in a similar area of cortex (Kim and Shadlen, 1999).

Whereas in perceptual decisions there is usually a correct answer which is independent of the decision maker's preferences, in economic or value-based decisions this is not the case. Choices that are based on the preference of the subject making the decision have been the focus of behavioral economics and the emerging field of neuroeconomics (Kahneman and Tversky, 1979; C. Camerer and Hua Ho, 1999; Wu, Delgado, and Maloney, 2009). Many interesting results related to how humans make decisions have been obtained, such as choice irrationalities and decision biases (Kahneman and Tversky, 1979), which are not typically found in the perceptual domain. Among the effects commonly observed are loss aversion, in which subjects exhibit greater sensitivity to losses than to equivalent gains when making decisions (Kahneman and Tversky, 1979; Novemsky and Kahneman, 2005; Rabin, 2000; Abdellaoui, Bleichrodt, and Paraschiv, 2007); hot hand fallacy (Gilovich, Vallone, and Tversky, 1985; Ayton and I. Fischer, 2004; Croson and Sundali, 2005), which corresponds to a positive recency effect; gambler's fallacy (Tversky and Kahneman, 1971; Ayton and I. Fischer, 2004; Croson and Sundali, 2005), which is a negative recency effect; and preference reversals, in which the preference for one option over another changes depending on context (Grether and Plott, 1979; Green et al., 1981; Tversky, Slovic, and Kahneman, 1990). In addition, evidence from neuroimaging studies in humans indicate an important role of emotion regulation processes subserved by the insula and areas of prefrontal cortex during economic decisions (Sanfey et al., 2003; Bechara and A. R. Damasio, 2005; Koenigs and Tranel, 2007).

It remains an open question how much overlap there is between the algorithms used by the human brain to make perceptual and value-based decisions, although some evidence for a common mechanism has been recently described (Polania et al., 2014; Frydman and Nave, 2016). Polania et al. obtained electroencephalography

(EEG) recordings during a task using both perceptual and value-based decisions based on the same stimuli (Polania et al., 2014). The authors found parietal gamma-frequency oscillations in both types of choice, supporting an evidence accumulation process in that area, and a similar frontal signal present during value-based decisions only. They also found that fronto-parietal gamma coupling related to the accuracy of value-based decisions only, and thus concluded that while parietal regions may encode a common decision variable in perceptual and economic choices, frontal regions perform an additional evidence accumulation process that is unique to the economic case.

1.3 Value Representation

It is generally believed that, in any kind of economic decision, one requirement is that the brain must be able to evaluate the alternatives in order to compare them and make a choice. This hypothesis has led to the search for value and expected reward representation in the primate brain. A study in non-human primates has shown, for instance, that neurons in the orbitofrontal cortex encode the value of offered and chosen goods during economic decisions, independently of visuospatial factors and motor responses, meaning that these values do not modulate activity related to sensory or motor processes (Padoa-Schioppa and Assad, 2006).

In humans, the search for value representation in the brain has been performed particularly through the use of functional magnetic resonance imaging (fMRI) technology (J. P. O'Doherty, 2004; Gottfried, O'Doherty, and Raymond J Dolan, 2003). Several fMRI studies have established, for instance, that the subjective value of potential rewards is explicitly represented in the human brain. This has been observed for different sensory modalities, such as taste (O'Doherty et al., 2001; Small et al., 2003; J. O'doherty et al., 2004; Todd A. Hare et al., 2008), olfaction (Gottfried, Deichmann, et al., 2002; Rolls, Kringelbach, and De Araujo, 2003; Anderson et al., 2003), somatosensory (Rolls, J. O'Doherty, et al., 2003), auditory (Blood et al., 1999), and visual (Aharon et al., 2001; J. O'Doherty et al., 2003), as well as for financial rewards (Elliott et al., 2003; McClure et al., 2004; Knutson et al., 2005; N. D. Daw, J. P. O'doherty, et al., 2006; Kable and Glimcher, 2007).

Evidence for a common currency that allows for comparison between potential rewards of different types has also been found. Montague and Berns developed a computational model that predicted that neural activity in the orbitofrontal-striatal circuit may support the conversion of reward values into a common scale, allowing

for the comparison between future actions or stimuli. More recently, Chib et al. found that an area in the ventromedial prefrontal cortex correlates with valuations for different categories of goods (Chib et al., 2009), and McNamee et al. found a region in medial prefrontal cortex where such value representations were independent of stimulus category (McNamee, Rangel, and J. P. O'doherty, 2013), supporting the idea of a common currency that allows for comparison of goods across categories.

1.4 Drift-Diffusion Models

In light of the extensive experimental support for the idea that a process of evidence accumulation takes place in the brain during decisions, researchers in neuroscience have developed sequential integrator models of choice that aim to make predictions for choice and response time in several decision tasks. An important family of these models, known as the Drift-Diffusion Model (DDM) (Ratcliff, 1978; Ratcliff and Rouder, 1998; Gold and Shadlen, 2002; Ratcliff and McKoon, 2008), has been successfully applied in the analysis of data from a large variety of tasks with non-human primates (Shadlen and Newsome, 2001; Gold and Shadlen, 2001; Ratcliff, Cherian, and Segraves, 2003; Churchland, Kiani, and Shadlen, 2008; Kiani, Hanks, and Shadlen, 2008; Bennur and Gold, 2011; Shadlen and Kiani, 2013), humans (Heekeren et al., 2004; Philiastides, Ratcliff, and Sajda, 2006; Tosoni et al., 2008; Ho, S. Brown, and Serences, 2009; Krajbich, C. Armel, and Rangel, 2010; Krajbich and Rangel, 2011; Krajbich, Lu, et al., 2012; O'connell, Dockree, and Kelly, 2012), and rodents (Brunton, Botvinick, and Brody, 2013; Erlich et al., 2015).

In monkeys, electrophysiology recordings have indicated that neural activity in parts of the brain involved in the selection and preparation of eye movements reflects both the direction of an impending gaze shift and the quality of the sensory information that motivates that response. This has been shown in the lateral intraparietal area (Shadlen and Newsome, 1996; Shadlen and Newsome, 2001), superior colliculus (Horwitz and Newsome, 1999; Ratcliff, Cherian, and Segraves, 2003), frontal eye field (Bichot and Schall, 1999), and dorsolateral prefrontal cortex (Kim and Shadlen, 1999). The time course of the neural response recorded in these areas supports the idea that they accumulate sensory signals relevant to the selection of a target for an eye movement.

In humans, fMRI and EEG evidence provide some support for a similar process taking place in the human brain during perceptual decisions. Heekeren et al. used

a categorization task with fMRI and found that activity in an area of dorsolateral prefrontal cortex was greater during easy decisions than during difficult decisions, covaried with the difference signal between two category-selective regions in the ventral temporal cortex, and predicted behavioral performance in the task (Heekeren et al., 2004). Philiastides et al. used a cued paradigm with EEG and identified a component in the EEG data reflecting task difficulty that arose about 220 milliseconds after stimulus presentation, between two components predictive of decision accuracy, supporting the time course of an evidence accumulation process (Philiastides, Ratcliff, and Sajda, 2006).

More recently, studies have begun to uncover behavioral and neural evidence for accumulator processes during economic decisions in humans (Krajbich, C. Armel, and Rangel, 2010; Hunt et al., 2012; Philiastides and Ratcliff, 2013). Because traditional evidence accumulator models describe differences in activity between different populations of selective cells and ignore nonselective activity, these models cannot make predictions for imaging data. To get around this issue, Hunt et al. developed a biophysical implementation of a competition model and used it to make explicit predictions for magnetoencephalography data collected during a value-based decision task. The authors found that parietal and prefrontal signals matched closely with the predictions of this model, indicating a process of evidence accumulation taking place in those areas.

Because many decision tasks require visual fixations on spatially-distributed stimuli, the DDM has been extended to incorporate overt attention in the form of visual fixations, leading to the development of the attentional Drift-Diffusion Model (aDDM) (Krajbich, C. Armel, and Rangel, 2010). In the aDDM, the value of items gets discounted if they are not currently being fixated by the subject. This model has been previously shown to make better predictions for choices and response times than the traditional DDM, in the context of value-based decisions in humans (Krajbich, C. Armel, and Rangel, 2010; Krajbich and Rangel, 2011; Krajbich, Lu, et al., 2012). In Chapter 2 I present a study that using the aDDM to make these same predictions for perceptual decisions.

1.5 Learning

In many decision situations, we may want to base our choice on the outcome of similar past decisions in an attempt to maximize our utility from the current outcome. This allows us to use information we have learned in the past to make

better decisions in the present moment. For this reason, the study of decision making is closely linked to the study of learning.

The field of reinforcement learning (Sutton and Barto, 1998) provides us with a framework for modeling how systems (such as a machine, a person or an animal) can make predictions about future outcomes based on previously experienced rewards, therefore optimizing their choices. Neurophysiology recordings from dopaminergic neurons in the macaque brain show evidence for a prediction error signal that is compatible with reinforcement learning models. And many experiments in humans show that reinforcement learning models can accurately describe learning behavior (Frank, Seeberger, and O'reilly, 2004; Schönberg et al., 2007; Bogacz, McClure, et al., 2007; Gershman, Pesaran, and N. D. Daw, 2009; Gläscher, N. Daw, et al., 2010; N. D. Daw, Gershman, et al., 2011), indicating that learning algorithms implemented by the human brain are compatible with the reinforcement learning framework.

Another important framework in the study of learning is that of Bayesian inference. Many researchers argue that the brain represents uncertainty about its environment, and the success of Bayesian methods in modeling human behavior in perception and sensorimotor control supports the idea that the brain represents sensory information probabilistically, in the form of probability distributions (Knill and Pouget, 2004). Behavioral experiments have confirmed that human observers not only take uncertainty into account in a wide variety of tasks, but do so in a way that is nearly optimal (Knill and Richards, 1996; Beers, Sittig, and Der Gon, 1999; Ernst and Banks, 2002; Körding and Wolpert, 2004; Stocker and Simoncelli, 2006). In addition, a study by Ma et al. suggests that the variability in neuronal populations represents probability distributions over stimuli, which the authors call probabilistic population codes (Ma et al., 2006).

1.6 Multiple Decision Systems

Taking into account the many experiments in human learning and decision making, it becomes apparent that there is not a single system responsible for the computation of all kinds of decision. Rather, experimental evidence points to the existence of multiple systems, with different parts of the literature focusing on different aspects of the dissociation between them (Kahneman, 1973; Anthony Dickinson, 1985; Sloman, 1996; Greer and Levine, 2006). Furthermore, recent experiments have focused on understanding how the brain allocates control between these different

systems in order to make better decisions according to environmental context, previous experience, available cognitive resources, among many other factors (N. D. Daw, Niv, and Dayan, 2005; Gläscher, N. Daw, et al., 2010; N. D. Daw, Gershman, et al., 2011; Otto, Gershman, et al., 2013; Lee, Shimojo, and J. P. O'Doherty, 2014; Akam, R. Costa, and Dayan, 2015; Kool, F. A. Cushman, and Gershman, 2016; Kool, Gershman, and F. A. Cushman, 2017).

In the learning and decision literature, one largely studied dichotomy is that between a model-free system, which relies on previous experiences (rewards) to compute the current value of the available options, and a model-based system, which uses planning and a model of the environment to choose the best option (Ray J Dolan and Dayan, 2013; J. P. O'Doherty, 2015; Doll et al., 2015). A common approach for dissociating between these two systems with human subjects is the two-step Markov decision task (N. D. Daw, Gershman, et al., 2011). In this task, the subject must make two consecutive binary decisions. In the first stage, each choice has a high probability of transitioning to one of two second-stage states (common transition), and a low probability of transitioning to the other state (rare transition). In the second stage, each choice is associated with a probability of obtaining a reward, and these probabilities drift slowly and independently over time. The subject's consecutive choices allow researchers to identify whether the strategy being used is model-based or model-free. For instance, if the subject makes a choice on the first stage and a rare transition occurs, leading to a reward on the second stage, a model-free strategy will lead to a higher probability of making the same choice on the first stage, since a reward was obtained. On the other hand, a model-based strategy will lead to a higher probability of choosing the alternative option on the first stage, since the subject has access to a model of these transitions and knows that the alternative option is the one more likely to take them to the rewarded state. This task has been used in many studies (N. D. Daw, Gershman, et al., 2011; F. Cushman and Morris, 2015; Otto, Gershman, et al., 2013; Otto, Raio, et al., 2013; Gläscher, N. Daw, et al., 2010), although some have criticized its ability to reliably dissociate between the two learning strategies (Akam, R. Costa, and Dayan, 2015).

In the study of the model-based vs. model-free dichotomy, a common theme is that of an efficiency-cost trade-off. Whereas the model-free system is computationally cheaper and more unreliable, and is therefore able to generate faster decisions but with a higher risk for errors, the model-based system tends to provide more reliable choices but at the same time incurs a significantly higher computational

cost (Gershman and N. D. Daw, 2012), which can lead to longer decision times as well as higher energy expenditure. Several studies have made use of this trade-off to develop computational models that aim to describe how the brain switches between these two systems.

Another instance of a dual decision system is one stemming from the behavioral economics literature, and relates to an experiential vs. a descriptive mode of computing option values (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). In an experiential setting, the subject relies on values learned from experience which must be retrieved from memory at the time of choice, whereas in a descriptive setting the subject computes option values from information that is explicitly described at the time of choice. In behavioral economics, several studies have described a phenomenon called the description-experience gap, in which consistent choice differences are found depending on how the alternatives are presented to the subject (Barron and Erev, 2003; Hertwig, Barron, et al., 2004; Hertwig and Erev, 2009). Looking at risk, subjects choosing between descriptive options tend to be risk averse in the gain domain and risk seeking in the loss domains, as predicted by Prospect Theory (Kahneman and Tversky, 1979), whereas in experiential choices this effect is reversed. In terms of probabilities, another prediction of Prospect Theory is that subjects tend to overweight low probability events and underweight large probability events when those events are fully described, but the opposite is true when subjects learn about these events from experience.

The differences between descriptive and experiential decision making have also been studied in the neuroscience literature (Jessup, Bishara, and Jerome R Busemeyer, 2008; FitzGerald et al., 2010). Jessup and colleagues performed a behavioral study and found that feedback had an critical role in determining whether subject overweighted or underweighted probabilities, regardless of whether or not fully descriptive information was presented (Jessup, Bishara, and Jerome R Busemeyer, 2008). In an fMRI study with humans, FitzGerald et al. found differential sensitivity to learned and described values and risk in brain regions associated with reward processing (FitzGerald et al., 2010), concluding that the neural encoding of decision variables was strongly influenced by the manner in which value information was presented to the subjects.

An interesting question that arises from the study of multiple learning and decision systems is that of arbitration. How is control allocated to each decision system?

Do they compete or cooperate? Several studies have attempted to understand this process (Lee, Shimojo, and J. P. O'Doherty, 2014; Doya et al., 2002; N. D. Daw, Niv, and Dayan, 2005; Beierholm et al., 2011; Economides et al., 2015; Russek et al., 2017; Kool, F. A. Cushman, and Gershman, 2016; Kool, Gershman, and F. A. Cushman, 2017; Miller, Botvinick, and Brody, 2017). Notably, Lee and colleagues found evidence for an arbitration mechanism in the human brain that allocates control over behavior to the model-based and model-free systems as a function of their reliability (Lee, Shimojo, and J. P. O'Doherty, 2014). Kool et al. (Kool, Gershman, and F. A. Cushman, 2017) studied the same arbitration process under a cost-benefit framework, suggesting that humans perform on-line cost-benefit analysis to switch between learning systems.

1.7 The Value of Behavioral Studies in Understanding the Brain

Despite not providing us with any direct measurements of neural activity, behavioral studies can aid our understanding of the brain mechanisms related to decision making in several ways. First, computational models that attempt to explain complex cognitive processes can make very precise predictions about the resulting behavioral data, therefore testing these predictions through behavioral metrics can help us to rule out certain models, compare across different models, and adjust the existing models when necessary. Second, we can establish a precise correspondence between behavioral phenomena observed in human and animal studies. When studying animal models, it is possible to perform sophisticated clinical and causal tests by relying on techniques such as genetic mutations, pharmacological interventions, and optogenetics, but in most cases these tests are not available when dealing with human subjects. Behaviors such as choices, response times, and visual fixations can be easily recorded from both humans and animals and used as a bridge across the phenomena observed in each species. Moreover, reflexive behaviors such as heart rate, skin conductance, blink rate, and changes in pupil size and reactivity can be used to provide further insight about the neuro-psychological state of subjects during different experimental conditions. Finally, in some cases the models being tested cannot make predictions for imaging data, so using behavioral data can provide validation for the effectiveness of the model in explaining choice mechanisms. This is the case for the DDM, which relies on the difference in neural activity between cell populations, and not on the kind of averaged response that is obtained from techniques typically used with humans, such as fMRI and EEG. Because single-unit recordings can be extremely difficult to obtain from human subjects, it is helpful to

test the predictions of this model on choice and response time data obtained from behavioral studies.

1.8 Developmental, Clinical and Comparative Aspects

The present dissertation describes two empirical studies for which data was collected at the California Institute of Technology in Pasadena, California. It is critical to note that this data was taken exclusively from a WEIRD (western, educated, industrialized, rich and democratic) population (Henrich, Heine, and Norenzayan, 2010). Furthermore, most of the subjects tested were in early adulthood, and were either undergraduate or graduate university students at a science and technology-focused institution. This means that, like most scientific results currently obtained at research universities, the results presented here are limited by the biases intrinsic to the sample population. Therefore, it is important to consider how the process of decision making in general, and our results in particular, may differ across different age groups, clinical conditions, and other species. In this section I review results from previous studies that provide insight into some of the potential differences that may occur in these cases.

It is well documented that significant changes in brain structure occur throughout lifespan. Research in neuroscience has begun to investigate how these changes may affect cognitive function such that behaviors during childhood, adolescence and elder years may deviate from what is typically encountered for an adult population (Giedd et al., 1999; Harbaugh, Krause, and Vesterlund, 2001; Mather et al., 2004; Galvan et al., 2006; Mata et al., 2011; Samanez-Larkin and Knutson, 2015). Neuroimaging studies have found, for instance, that stimulus value signals in the amygdala are down-modulated in younger adults (Mather et al., 2004), and that ventral striatum response to monetary rewards dramatically decreases from adolescence into adulthood (Galvan et al., 2006).

Choice biases and irrationalities have also been studied in different age groups. A study of risky decision making in older adults found no evidence of systematic age-related differences in risk taking and no differences in tendency to take risks when choices are framed as gains versus losses, but suggested that older adults tend to make more mistakes in choices overall, indicating cognitive limitations rather than different risk preferences in that population (Mata et al., 2011). Harbaugh et al. investigated the endowment effect (when the minimum compensation people are willing to accept in return for giving up a good they already possess exceeds the

amount that they are willing to pay to acquire the same good) in children but found no evidence that this effect decreases with age (Harbaugh, Krause, and Vesterlund, 2001). On the other hand, Galvan et al. used fMRI to study risk-taking behavior in adolescents and found evidence that subcortical systems become disproportionately activated relative to top–down control systems in this population, biasing their actions toward immediate over long-term gains (Galvan et al., 2006).

Several clinical conditions have also been found to correlate with significant differences in terms of decision making behavior when compared to neurotypical populations. Gillan et al. found a strong association between deficits in goal-directed control in a two-step Markov decision task and symptoms such as compulsive behavior and intrusive thought (Gillan et al., 2016). The researchers also found an association between these same task deficits and self-reported conditions that are typically characterized by a loss of control over behavior, such as eating disorders, impulsivity, obsessive-compulsive disorder, and alcohol addiction. Relatedly, Foerde et al. found that individuals with anorexia nervosa engaged the dorsal striatum, an area typically associated with action selection and control, more than healthy controls when making food choices, and that activity in their fronto-striatal circuits was correlated with actual food consumption the following day (Foerde, Steinglass, et al., 2015). Numerous other studies have related cognitive impairments and faulty decision making behavior to conditions such as addiction (Bechara and H. Damasio, 2002; Bechara, 2005; Schoenbaum, Roesch, and Stalnaker, 2006; Schoenbaum, Roesch, and Stalnaker, 2006), depression (Murphy et al., 2001; Apkarian et al., 2004), and anxiety (Raghunathan and Pham, 1999; Maner et al., 2007), as well as neurological disorders such as Parkinson's disease (Brand et al., 2004; Mimura, Oeda, and Kawamura, 2006; Foerde and Shohamy, 2011) and obsessive-compulsive disorder (Cavedini et al., 2002; Lawrence et al., 2006).

In previous sections I described a large body of animal research that provides insight about the decision mechanisms that exist in the mammalian brain, particularly in primates. Support for evidence accumulation processes, for instance, was first obtained from several cortical areas in the macaque brain (Shadlen and Newsome, 2001; Gold and Shadlen, 2001; Ratcliff, Cherian, and Segraves, 2003; Churchland, Kiani, and Shadlen, 2008; Kiani, Hanks, and Shadlen, 2008; Bennur and Gold, 2011; Shadlen and Kiani, 2013), and only more recently there have been attempts to generalize these results through behavioral and neuroimaging studies in humans (Heekeren et al., 2004; Philiastides, Ratcliff, and Sajda, 2006; Tosoni et al., 2008; Ho, S. Brown,

and Serences, 2009; Krajbich, C. Armel, and Rangel, 2010; Krajbich and Rangel, 2011; Krajbich, Lu, et al., 2012; O'connell, Dockree, and Kelly, 2012).

Regarding dual systems of decision making, one dichotomy that has been extensively explored in the animal literature, and which is strongly related to the model-based and model-free systems in human studies, is that of habitual vs. goal-directed decision making. Evidence from several studies supports the existence of a habitual system, which uses previous experience to generate a decision while disregarding the current state of the animal and of its environment, and a separate goal-directed system, which takes these states into account (Adams and Anthony Dickinson, 1981; Anthony Dickinson and B. Balleine, 1994; A Dickinson et al., 1995; Tricomi, B. W. Balleine, and J. P. O'Doherty, 2009; B. W. Balleine and J. P. O'doherty, 2010). One popular technique for dissociating between these systems which has been used with rodents is outcome devaluation (Adams and Anthony Dickinson, 1981). In this approach, the animal is taught to press a lever in order to receive a food reward. Once learning has been consolidated, the animal is fed to satiety and again presented with the lever. A persistence in lever-pressing behavior means the animal is acting out of habit, while an interruption of the behavior means the animal has taken into account the fact that it is now satiated and therefore does not need to press the lever for a food reward (i.e., the food reward has been devalued). In rodents, control between the habitual and goal-directed systems appears to involve interactions in dorsal striatum and in the cortico-basal ganglia network within which the striatum is embedded, although details about the arbitration process still need to be worked out (B. W. Balleine and J. P. O'doherty, 2010; Gremel and R. M. Costa, 2013).

Overall, it is clear that many differences exist in decision making behaviors depending on the population or species being studied. Despite these differences, certain phenomena, such as choice biases and irrationalities, appear to be pervasive across human populations, including a large array of age groups and clinical conditions. Furthermore, some of the neural mechanisms described here, such as evidence accumulation processes and dual decision systems, have been extensively studied in several mammalian species. The variations that do exist across populations and species are indeed quite useful as they allow us to further explore how a variety of structural and chemical differences present in the brain may impact the decision making process.

Chapter 2

THE ATTENTIONAL DRIFT-DIFFUSION MODEL OF SIMPLE PERCEPTUAL DECISION MAKING

2.1 Introduction

Over the last two decades, neuroscientists and psychologists have devoted considerable effort to understanding the neurocomputational basis of decision making. The goal has been to understand which are the variables encoded at the time of decision, what are the algorithms used to combine them into a decision, and how these processes are implemented and constrained by the underlying neurobiology. Considerable progress has been made in understanding simple perceptual decisions (e.g., determining the net direction of motion in a field of noisy moving dots) and simple value-based choices (e.g., choosing between two food snacks). Interestingly, a qualitatively similar class of algorithms has been shown to provide a good description for the accuracy and response time patterns in both perceptual (Ratcliff and Rouder, 1998; Gold and Shadlen, 2001; Gold and Shadlen, 2007; Smith and Ratcliff, 2004; Ditterich, 2006; Brunton, Botvinick, and Brody, 2013) and value-based choices (M. M. Mormann, Malmaud, et al., 2010; Hunt et al., 2012; Philiastides and Ratcliff, 2013; Hutcherson, Bushong, and Rangel, 2015), although many important details remain to be worked out (Bogacz, 2007; Summerfield and Tsetsos, 2012; Tsetsos, Gao, et al., 2012; Brunton, Botvinick, and Brody, 2013; Orquin and Loose, 2013; Shadlen and Kiani, 2013; Teodorescu and Usher, 2013). Despite important differences among the various models that have been proposed, all of the algorithms are built around the idea that decisions are made by accumulating noisy evidence in favor of the different alternatives, and that choices are made when the weight of accumulated evidence in favor of one of the options becomes sufficiently strong. For this reason, they are often described as sequential integration models. There is also a growing understanding of how the brain implements these processes in both perceptual (Shadlen and Newsome, 2001; Roitman and Shadlen, 2002; Heekeren et al., 2004; Philiastides, Ratcliff, and Sajda, 2006; Churchland, Kiani, and Shadlen, 2008; Kiani, Hanks, and Shadlen, 2008; Tosoni et al., 2008; Ho, S. Brown, and Serences, 2009; Bennur and Gold, 2011; O'connell, Dockree, and Kelly, 2012) and value-based choice (Basten et al., 2010; Philiastides, Biele, and Heekeren, 2010; Todd A Hare, Schultz, et al., 2011; Hunt et al., 2012; Polania et al., 2014; Rustichini

and Padoa-Schioppa, 2015).

Since many decision tasks require the comparison of spatially distributed stimuli, two important questions are whether decisions are affected by how visual attention is deployed during the process of choice, and if so, how do sequential integrator models need to be modified to incorporate the role of attention. For example, in the context of perceptual choice, if a subject is shown two lines of different length on the left and right sides of the screen and has to decide which one is longer, how does the pattern of fixations to the two stimuli affect their decision, if at all? Or in the context of value-based choice, if a subject is shown two food stimuli, how does their pattern of fixations affect which of the two foods they choose to eat?

This problem has been studied in the realm of value-based choice. Krajbich et al. (Krajbich, C. Armel, and Rangel, 2010; Krajbich, Lu, et al., 2012; Krajbich and Rangel, 2011; Towal, M. Mormann, and Koch, 2013) found that a modification of the popular Drift-Diffusion Model, which they call the attentional Drift-Diffusion Model (aDDM), provides a quantitatively accurate description of the relationship between visual attention, choices and response times in several value-based tasks. The aDDM builds on previous work by Busemeyer and collaborators, who proposed an alternative class of sequential integrator models in which attention plays a role (Roe, Jermone R Busemeyer, and Townsend, 2001). In the aDDM, attention influences choices by increasing the relative weight given to evidence related to the attended stimulus. As a result, the model predicts that exogenous shifts of attention can cause systematic choice biases, which is consistent with the results of several studies (Shimojo et al., 2003; K. C. Armel, Beaumel, and Rangel, 2008; Todd A Hare, Malmaud, and Rangel, 2011; Pärnamets et al., 2015; Kunar et al., 2017).

Given that a remarkably similar set of algorithms have been shown to be at work in perceptual and value-based choice tasks in which attention plays no role, it is natural to hypothesize that the aDDM might also provide a reasonable computational description of the role of visual attention in simple perceptual decisions, and that exogenous shifts in attention (i.e., unrelated to the perceptual properties of the stimuli) might causally bias choices as predicted by the aDDM. Here we present the results of two experiments designed to test these hypotheses.

Testing the extent to which the aDDM is able to provide a satisfactory quantitative description of the role of visual attention in simple perceptual choices is interesting for several reasons. First, previous experiments have shown that attention can affect perceptual choices using divided attention paradigms (Wyart, Myers, and

Summerfield, 2015), spatial pre-cuing paradigms (Posner, Snyder, and Davidson, 1980; Smith, Ratcliff, and Wolfgang, 2004; Carrasco, 2011), and serial dependence paradigms (J. Fischer and Whitney, 2014). However, the algorithmic or computational description of this effect remains an open question (Summerfield and Egner, 2013). Second, an important open question in cognitive neuroscience is whether the same algorithms are at work in different domains and systems whenever the problem they are trying to solve is sufficiently similar. This view is consistent with the fact that sequential integrator models are able to accurately describe two-alternative forced choices in domains ranging from memory, to perception, to economic choice (Gold and Shadlen, 2007; Ratcliff and McKoon, 2008; Starns, Ratcliff, and McKoon, 2012; Shadlen and Kiani, 2013). However, since perception and value-based choice are made on the basis of different evidence (i.e., perceptual inputs vs. reward predictions), attention might operate through very different channels in these two cases, and thus we cannot assume ex ante that it might have a computationally similar effect in both types of decisions.

2.2 Materials and Methods

Subjects

In Experiment 1 we tested 25 subjects (10 female, mean age 23), which included Caltech students and staff as well as members of the surrounding community. Subjects were advised to use glasses for eyesight correction as needed. Each subject completed 1,344 decision trials, split into 4 identical experimental sessions, spread across 4 different days. Subjects received a $15 show-up fee in each day and a $40 bonus for completing all sessions, as well as additional earnings based on performance, as described below. In Experiment 2 we tested 20 subjects (9 female, mean age 25). Each subject completed 336 trials in a single session, and received a $15 show-up fee, as well as additional earnings based on performance. The experiments were approved by Caltech's IRB and all subjects provided informed consent prior to participation.

Experiment 1

Experiment 1 consisted of four identical sessions, collected on four separate days, within a period of two weeks. Each experimental session was divided into 12 blocks of 28 decision trials. At the beginning of each block, subjects were shown for 5 seconds a line depicting a target orientation chosen from the set $\{20°, 35°, 55°, 70°\}$, as shown in **Figure 2.1A**. This excludes vertical and horizontal orientations, which

would have made subsequent choices too easy. Each orientation was chosen as the target three times per session, in random order.

In each decision trial subjects were shown two oriented lines, on the left and right sides, with an eccentricity of 16 degrees from the center of the screen (**Figure 2.1B**). The relative orientation between each of the two lines and the target, denoted by Δ, was chosen from the set $\{-15°, -10°, -5°, 0°, 5°, 10°, 15°\}$ (**Figure 2.1C**). The subjects' task was to decide which of the two lines had an orientation closest to the target. They were allowed to take as long as needed to make a choice, and indicated their choice with a button press ("A" for left and "L" for right). Let Δ_{left} and Δ_{right} denote the relative angular distance between the target and the left and right items, respectively. The two choice stimuli shown in the trial were not allowed to have the same Δ. Uniform sampling subject to these constraints led to 42 different trial conditions. Each was used eight times per session, in random order. Subjects saw a blue box around the chosen item in each trial, but they did not receive feedback about the correctness of their decisions during the task.

Stimuli were presented on a 1,280 × 1,024 screen, placed $\sim$ 50 cm from the subjects' eyes. Subjects were required to keep their hands on the response buttons for the entire task, so they could enter responses without looking at the keyboard. Subjects' fixation patterns were recorded at 500 Hz using an EyeLink 1000 Plus desktop-mounted eye-tracker with head support. Fixations and saccades were determined using the eye-tracker's accompanying software package. The eye-tracking system was calibrated at the beginning of each session, and again whenever the eye-tracker lost the subject's eye (which only occurred four times during all sessions of both experiments). Before each decision trial, subjects were required to maintain a continuous fixation on a central cross for 500 ms before the items would appear, which ensured that every trial began with a fixation on the same central location.

In order to familiarize subjects with the targets, they also completed a training task at the beginning of each block (**Figure 2.1D**). Here, subjects were shown a single oriented line in the center of the screen, and had to decide whether or not the line shown had the same orientation as the target. They were allowed to take as long as needed to make the decision, which they then indicated with a button press ("A" for no and "L" for yes). Subjects received immediate feedback for 1 s after every decision indicating its correctness. The training task ended after six correct decisions in a row.

The target stimulus for each block was shown once at the beginning of the block

(before the training trials), once immediately after the training trials, and again after every 5 decision trials. At the end of each experiment session, we selected 25 decision trials at random, and subjects received an additional payment of $1 for each correct response.

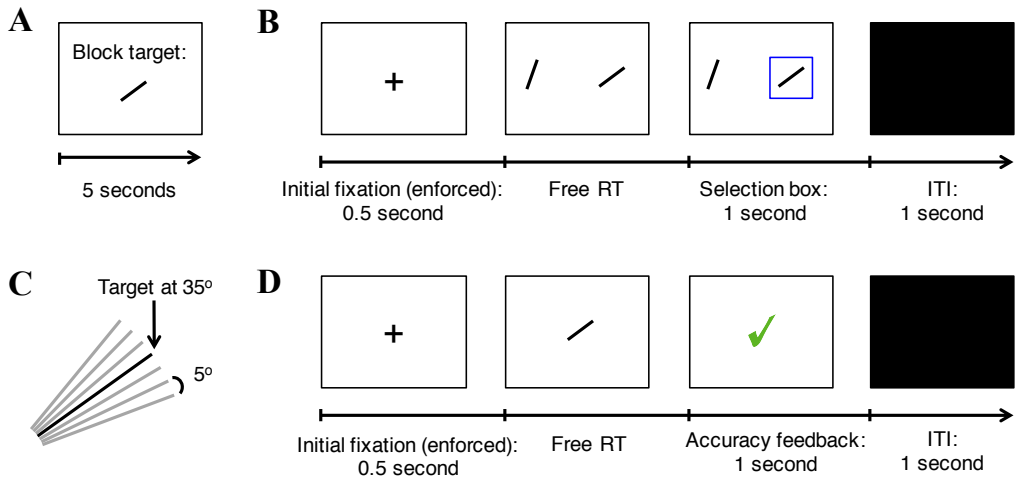

Figure 2.1: Summary of Experiment 1. (A) In the beginning of each block of trials a new target orientation was shown for 5 s. The target was shown again immediately after the training trials, and again after every 5 decision trials. (B) Trial structure for our simple perceptual decision task. In each trial subjects must choose the stimulus (left or right) with the orientation closest to the target. (C) Diagram showing all seven possible item orientations, in increments of 5°, given a target oriented at 35°. (D) Trial structure for training trials.

Experiment 2

The structure of a typical trial in Experiment 2 is depicted in **Figure 2.2**. Experiment 2 was similar to Experiment 1, except for the following differences. First, each subject completed a single session with 12 blocks of 28 trials each. Second, in each trial we randomly selected one of the two items on the screen to be the bias-target item. We used the following procedure to bias fixations toward that item. Unbeknownst to the subjects, we required a minimum amount of cumulative fixation time to each item: 800 ms for the bias-target item and 200 ms for the other one. In every trial we kept track of the cumulative fixation durations to each stimulus and, as soon as the minimum requirement for both was met, the items disappeared and the subject was prompted to make a choice. Third, subjects were told (without deception) that both the duration of decision trials and the item that appeared first would be chosen at random every trial, but were not told that the

procedure was designed to bias fixations. To minimize awareness of the nature of the experimental manipulation, which relies on giving subjects control over the duration of trials through their fixations, we set the maximum duration for each evaluation period to 3 seconds. If the minimum fixation requirements for both items were not met within that period of time, the subject was prompted to make a decision. Trials in which the 3-second boundary was binding were removed from additional analyses since they exhibited unusual fixation patterns (24.3% over the entire group; across subjects, min = 6.8% and max = 55.6%). This trial exclusion rule was chosen a priori to minimize the chance that subjects would become aware of the experimental manipulation. Furthermore, the exclusion of these trials did not qualitatively affect any of the reported results. Fourth, subjects were only allowed to enter their decisions after the decision prompt appeared, and could take as much time as needed to do so. Fifth, we refer to trials in which the bias-target item was fixated longer (and in which the stopping condition was reached before 3 seconds) as effective manipulation trials. Note that not all trials were effective since the contingencies described above allow for the possibility that subjects fixate more on the non-bias-target item. In order to increase the fraction of effective trials, which is the manipulation of interest, the bias-target item was always displayed on the screen first, and the other item was only added after a certain delay, which counted toward the total fixation time for the bias-target. The duration of this lag was between 100 and 500 ms, and was calibrated separately for each subject at the beginning of the experiment using the following staircase procedure. The lag started at a value of 300 ms, and was adjusted with a step of 30 ms. After every 3 consecutive effective trials (i.e., trials in which the bias-target item was fixated longer) the lag was decreased by 30 ms, and after a single ineffective trial (i.e., one where the bias-target item was not fixated longer), the lag was increased by the same amount. A total of 48 trials were used in the staircase procedure, and the value of the delay at the 48th trial was then used throughout the remainder of the task (duration: mean = 450 ms, SD = 35 ms). Sixth, at the end of the experiment we randomly selected 20 decision trials, and subjects received an additional payment of $1 for each correct response in this set. Seventh, at the end of the task, subjects completed a questionnaire in which we asked if they found anything strange about the timing of the items being displayed and the decision prompt. None of the subjects reported finding a connection between their fixations and the duration of the trials.

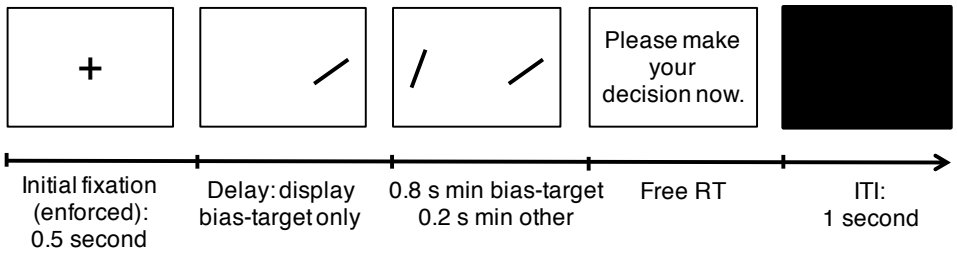

Figure 2.2: Summary of Experiment 2.

Fixations

All recorded fixations were classified as either item fixations (to either the left or right stimuli on the screen) or "blank" fixations. Trials in which "blank" fixations accounted for more than 50% of the response time were discarded from further analysis (mean percentage discarded trials across all subjects: 5.2%, min = 0.15%, max = 25.8%). Furthermore, if a "blank" fixation was recorded between two fixations to the same stimulus, the observation was converted into a fixation on that item. This is justified by the fact that this type of "blank" fixations tend to be very short, and are likely to be the result of blinking or eye-tracker noise (duration: mean = 41 ms, SD = 133 ms). If a "blank" fixation was recorded between fixations on different items, then it was grouped into that trial's inter-fixation transition time and used as such in the analyses below.

Group Model Fitting

We used maximum likelihood estimation (MLE) to fit the aDDM to the pooled group data. The model has three free parameters (d, θ, and σ) which we fitted using only the odd-numbered trials, so that the even-numbered trials could be used to test its out-of-sample predictions (see the Results section for a description of the aDDM).

The MLE procedure was carried out in multiple steps. In step 1 we defined a coarse grid of parameter combinations, denoted by Ω_1, which was given by the cross product of the sets $\{0.001, 0.005, 0.01\}$ for d, $\{0.1, 0.5, 0.9\}$ for θ, and $\{0.01, 0.05, 0.1\}$ for σ. Although this set only has nine points, it was selected because it spans a wide range of potential parameter combinations. We computed the likelihood of the choices and response times (RTs) observed in the odd-numbered trials, conditional on the observed pattern of fixations in each trial, for each vector of parameters in Ω_1. This was done by simulating the aDMM using the algorithm described in detail

in Chapter 3. All time data (including RTs, latencies, fixation durations, and inter-fixation transition times) were binned into 10 ms steps. We then selected the vector of parameters in Ω_1 with the highest log-likelihood as the first candidate solution, which for our data was given by $d = 0.005$, $\theta = 0.1$ and $\sigma = 0.05$. Let $(d_1, \theta_1, \sigma_1)$ and ML_1 denote, respectively, the best set of parameters and the likelihood that it explains the data that arises from the first step.

The algorithm then proceeded inductively until a stopping criterion was reached. Let Ω_t denote the search set used in step t, and $(d_t, \theta_t, \sigma_t)$ and ML_t denote the best candidate solution at this step. Step $t+1$ then proceeded as follows. A new grid of 9 potential vector parameters, denoted by Ω_{t+1}, was constructed as the cross product of the sets $\{d_t - \frac{\Delta d_t}{2}, d_t, d_t + \frac{\Delta d_t}{2}\}$, $\{\theta_t - \frac{\Delta \theta_t}{2}, \theta_t, \theta_t + \frac{\Delta \theta_t}{2}\}$, and $\{\sigma_t - \frac{\Delta \sigma_t}{2}, \sigma_t, \sigma_t + \frac{\Delta \sigma_t}{2}\}$, where Δd_t, $\Delta \theta_t$, and $\Delta \sigma_t$ correspond to the parameter step sizes used in Ω_t. Note that Ω_{t+1} included $(d_t, \theta_t, \sigma_t)$, as well as a finer grid around it.

The MLE step was then repeated again. The algorithm continued until the improvement in the MLE of the proposed parameter solution was <1%. For our data, the convergence process was accomplished in 7 steps, and resulted in an estimate of $d = 0.0041$, $\sigma = 0.063$, and $\theta = 0.36$.

Out-of-Sample Group Simulations

In order to test the ability of the model to predict out of sample, we used the aDDM with the best fitting parameters for the odd-numbered trials to predict data group patterns in the even-numbered trials.

Critically, the predictions were made conditional on the relative orientation of the stimuli, but not on the actual fixation patterns observed in the even trials. To understand why, note that due to the randomness in the aDDM algorithm, two trials with identical stimuli and fixations might lead to different choices and RTs. As a result, two runs of the same trial can result in different outcomes even if they initially exhibit identical fixations. In addition, if the aDDM is an approximately accurate description of the underlying processes, the pattern of fixations can vary widely over repeated decisions with an identical pair of stimuli, a fact that is observed in the data. For these reasons, our out-of-sample predictions condition on the relative orientation of the stimuli, including the effect that this has on the fixation process, as described below, but not on the actual realized fixations. This allows us to test the ability of the aDDM to account out-of-sample for key patterns in the data conditional only on independent variables like the relative orientation of the stimuli.

For each of the 42 trial conditions we simulated 400 trials of the model, while sampling fixations, latencies, and inter-fixation transitions from the empirical distributions, pooling the even-numbered trials from all subjects. Initial latency (i.e., the delay between stimulus appearance and the first item fixation) and subsequent inter-fixation transitions were sampled, each from its own distribution, without any further conditioning. To maximize the extent to which the simulated fixations matched the observed fixations, item fixations were sampled as follows. First, they were partitioned into 3 groups, corresponding to first, second, and other middle fixations. Additionally, item fixations were conditioned on the relative proximity difference between the fixated and the unfixated items, $r_{\text{fixated}} - r_{\text{unfixated}}$, since this matched the observed fixation patterns well. Note that the pool of fixations used to simulate the model excluded final fixations. According to the aDDM, a maximal fixation duration is drawn at the fixation outset and runs its course unless a barrier is cross beforehand. As a result, observed final fixations are truncated, and using them would bias the simulations (under the maintained hypothesis that the aDDM is correct).

Each trial was simulated by sampling latencies, fixations, and inter-fixation transitions as needed to carry out the simulation to its completion. In particular, each simulation began with a sampled latency, during which only white Gaussian noise was added to the relative decision value. Following this, fixations alternated between the left and right items such that, if the first fixated item was left, the second one would be right, and so on. The first fixation was chosen to be left with probability 0.65, which equals its empirical frequency. The maximum first fixation duration was sampled from the pool of first fixations, conditioned on $r_{\text{fixated}} - r_{\text{unfixated}}$.

The simulation for a trial was terminated if the aDDM crossed a decision barrier during the course of a fixation. After each item fixation, an inter-fixation transition duration was sampled, and if a simulation happened to terminate on a transition, it was discarded, since this was not commonly observed in the data (mean percentage of trials across all subjects: 14.4%, min = 8.2%, max = 25.9%). We also simulated the model without discarding simulations that ended on transitions, but did not find any significant differences from the results presented here.

Model Comparison

In order to explore the role of attention in explaining the data, we carried out an additional set of analyses designed to test the best fitting aDDM with the best fitting

standard DDM, which equals the special case of $\theta = 1$, where attention does not matter. To do this, we first re-estimated the model in the odd-number trials under the restriction that $\theta = 1$ to find the best fitting standard DDM. We then carried out three different out-of-sample prediction exercises.

First, we predicted choices and RTs in the even trials using the best fitting DDM. This was done by simulating the model 100,000 times for each potential combination of r_{left} and r_{right}, and then making predictions by sampling choices and RTs from the resulting simulations conditional on the stimulus orientations in each trial.

Second, we predicted choices and RTs in the even trials using the best fitting aDDM, conditional on net fixation time (i.e., total fixation time on left minus total fixation time on right). This was also done by simulating the model 100,000 times for each potential combination of r_{left}t and r_{right}, and then sampling choices and RTs from the resulting simulations, but this time conditional on both the stimulus orientations and the overall net fixation time observed in the even trial.

Third, we predicted choices and RTs in the even trials using the best fitting aDDM, and conditional on the observed fixations. To do this, we simulated the aDDM for each even trial assuming the same values of r_{left} and r_{right}, and that the fixation process was identical to the one seen in the trial up to its RT. If the simulation did not lead to a choice by the observed RT, additional fixations were sampled using the fixation process described above. The outcomes of the simulation were used as the choice and RT predicted for each even trial.

Goodness-of-Fit Measures

For binary variables, we report Efron's pseudo R-squared as a measure of goodness-of-fit, which corresponds to the squared correlation between the predicted values and the actual values. For non-binary variables, we report a number of goodness-of-fit measures, which are designed to test the similarity between the predicted and the observed data patterns. Each pattern involves a relationship between an independent (e.g., differences in relative proximity) and a dependent variable (e.g., RTs). Similar to previous work (Krajbich, C. Armel, and Rangel, 2010), these measures were computed as the p-values on the coefficients of a weighted least squares regression, in which the dependent variables were given by the difference between each subject's mean and the average value predicted by the model, and the weights were given by the inverse of the variance.

Data and Code

The data and code used in the analyses are available at the Rangel Neuroeconomics Lab website (`www.rnl.caltech.edu`).

2.3 Results

In order to investigate the role of visual attention in perceptual decision making, we carried out two different experiments. Experiment 1 was designed to test the extent to which the aDDM provides a reasonable quantitative description of the relationship between visual attention (as measured by fixations), choices, and response times (RTs) in simple perceptual decisions. Experiment 2 was designed to test a key prediction of the aDDM, namely, that exogenous shifts in attention can bias perceptual decisions in favor of the attended item.

The first experiment, depicted in **Figure 2.1**, required subjects to make simple perceptual decisions about line orientations (see Materials and Methods for details). At the beginning of each block of trials, subjects were shown an oriented bar for 5 seconds, which served as the target for the entire block (**Figure 2.1A**). The orientation of the target was chosen from the set $\{20°, 35°, 55°, 70°\}$. In each decision trial subjects were shown two oriented bars, one on the left and one on the right, and had to decide which of them had an orientation closest to the target orientation by pressing a button (**Figure 2.1B**). The angular distance between each of the lines and the target, denoted by Δ, was chosen randomly from the set $\{-15°, -10°, -5°, 0°, 5°, 10°, 15°\}$ (**Figure 2.1C**), with the constraint that the two stimuli could not have an equal orientation. Note that the correct response depends only on the angular distance, which is a relative orientation measure. For example, if $\Delta_{\text{left}} = -10°$ and $\Delta_{\text{right}} = 15°$, the correct response is left, and if the two stimuli are equidistant to the target (e.g., if $\Delta_{\text{left}} = -10°$ and $\Delta_{\text{right}} = 10°$, then either choice is considered correct. In order to motivate subjects to perform the task, a subset of the trials was selected at random at the end of the experiment and subjects earned \$1 for each correct choice.

In order to familiarize the subjects with the stimuli, they also participated in a training task at the beginning of each block in which they were shown one oriented bar at a time and had to judge if it had the same orientation as the target (**Figure 2.1D**). Training was administered until a pre-specified performance criterion was reached on each block. See Materials and Methods for details.

Perceptual aDDM

The aDDM provides an algorithmic description of how information is integrated over time in order to make a binary perceptual choice, and of the role that fixations play in this process. As illustrated in **Figure 2.3**, the model assumes that choices are made by dynamically computing a relative decision value (RDV) signal, which at any instant provides an estimate of the relative attractiveness of the two options. The RDV begins at zero and a choice is made the first time it crosses one of two pre-established decision barriers: one at +1, indicating a choice for left, and one at -1, indicating a choice for right.

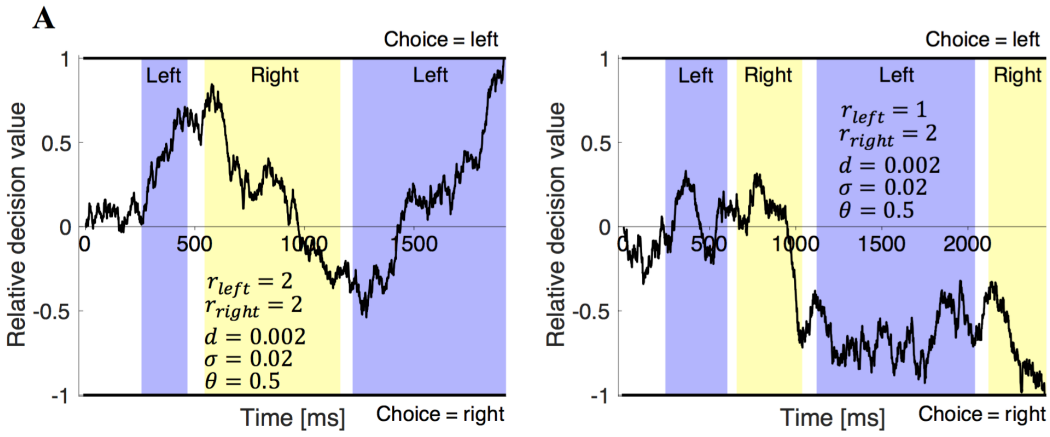

Figure 2.3: Two sample runs of the aDDM. (A) The two items have the same relative proximity, and the left item is chosen after 1,909 ms. (B) The right item has greater relative proximity, and is chosen after 2,444 ms.

The predictions of the model depend heavily on the dynamics of the RDV. Let RDV_t denote its value at time t within the course of a single decision. At every time step Δt, its change is given by $\mu\Delta t + \varepsilon_t$, where ε_t is i.i.d. zero mean white Gaussian noise with standard deviation σ, and μ is the deterministic change in the RDV over the time step, often called the slope of the process. A critical assumption of the aDDM is that the slope of the RDV signal depends on the location of the fixations at each time step. In particular, $\mu = 0$ until the first fixation to one of the two stimuli occurs, as well as during non-stimuli fixations and inter-fixation transitions, while $\mu = d(r_{\text{left}} - \theta r_{\text{right}})$ during fixations to the left option, and $\mu = d(\theta r_{\text{left}} - r_{\text{right}})$ during fixations to the right option. Here, d is a positive constant that controls the speed of integration, θ is a parameter between 0 and 1 that measures the size of the attentional bias, and r_{left} and r_{right} are the relative proximities of the left and right items shown in the trial.

The relative proximity of an option is a measure of its attractiveness, which in this task is given by the negative of the absolute value of Δ, and can only take four values: $\{-15°, -10°, -5°, 0°\}$. For ease of interpretation, and ease of comparison with related studies (Krajbich, C. Armel, and Rangel, 2010; Krajbich, Lu, et al., 2012; Krajbich and Rangel, 2011), we normalized the relative proximities to the scale $\{0, 1, 2, 3\}$, with 3 denoting the best possible proximity (i.e., an orientation equal to the target), and 0 denoting the worst possible proximity (i.e., an angular distance of either $-15°$ or $+15°$). We chose the range to be from 0 to 3 because there were 4 possible values for the angular distance between an item and the target. **Figure S3** illustrates the transformation from the angular distance scale to the relative proximity scale, see Supplementary Materials from Tavares, Perona, and Rangel, 2017.

The aDDM also makes a critical assumption about the fixation process. It allows fixations to depend on properties of the stimuli (as described below), but it assumes that the fixation process is otherwise independent of the state of the RDV signal. In other words, it assumes that there is no feedback from the path of the decision process to the propensity to fixate on stimuli. We return to this important assumption in the Discussion section.

No other major restrictions are placed on the fixation process, except for those that are reflected in the empirically observed properties of the fixations. First, the model assumes that the location of the first fixation and its latency are independent of the relative proximity of the two stimuli. To be precise, it assumes that the first fixation is to the left item with a constant probability p, and that the latency of this first fixation is drawn from a fixed distribution. Second, subsequent fixations alternate between the left and right items. Third, a maximal fixation duration is drawn from a distribution at the beginning of each fixation, and the fixations run their course unless a choice is made by crossing a barrier before the end of the fixation. In this case, the process terminates and the duration of the last fixation is truncated. The distribution of maximal fixation durations is allowed to depend on the fixation number, and on the difference in relative orientation between the fixated and the unfixated stimuli (see Materials and Methods for details). Importantly, we only sample from non-last fixations, i.e., fixations that were not terminated when the subject makes a choice. Fourth, fixations are separated by inter-fixation transitions that are drawn from another fixed distribution. As with the fixations, a maximal inter-fixation transition duration is drawn from this distribution at the beginning of

each transition, and runs its course unless a barrier is crossed before it terminates.

Several aspects of the model are worth highlighting. First, the model has three free parameters: d, θ, and σ (the time step used for binning the data was fixed at 10 ms). This follows from the fact that, in this class of models, multiplying the size of the barriers, d and σ by a common positive constant does not change the predictions of the model. As a result, we can fix the size of the barriers to +1 and -1 without any loss of generality. Second, if $\theta = 1$, the model reduces to the standard Drift-Diffusion Model (DDM) (Ratcliff, 1978; Gold and Shadlen, 2002; Gold and Shadlen, 2007; Ratcliff, Cherian, and Segraves, 2003; Ratcliff and Smith, 2004; Smith and Ratcliff, 2004; Bogacz, 2007; Ratcliff and McKoon, 2008) and therefore item fixations become irrelevant. Thus, the model includes as a special case the possibility that attention plays no role in choices. Third, if $\theta < 1$, the model predicts that changes in fixations can affect choices. The intuition for why this is the case is illustrated in **Figure 2.3A**, which depicts a sample run in which $r_{\text{left}} = r_{\text{right}} = 2$. In the absence of an attentional bias (i.e., when $\theta = 1$), the mean slope of the RDV signal is zero and the choice and RT are determined solely by the noise in the process. In contrast, when $\theta = 0.5$, as shown in the figure, the mean slope of the RDV signal is positive during left fixations and negative during right fixations (i.e., the integrator moves toward the fixated item on average). Fourth, when $\theta < 1$, exogenous shifts (i.e., unrelated to the perceptual properties of the stimuli) in fixations toward an item can bias choices toward that item, and the magnitude of the bias increases as θ decreases. For instance, as shown in **Figure 2.3B**, when $r_{\text{left}} < r_{\text{right}}$ and $\theta = 0.5$, the attentional bias strengthens the negative slope toward the right item barrier during right fixations. Fifth, the assumption that the fixation process does not depend on the state of the RDV signal implies that one can think of the aDDM as a model of the decision process that takes as given the empirical relationship between fixations and various non-aDDM variables (e.g., fixation number or relative proximity), presumably because the choices and fixations are controlled by distinct systems.

Basic Psychometrics

We began the analysis by characterizing the basic psychometrics of the task, which resembled the patterns commonly found in previous perceptual (Ratcliff, Cherian, and Segraves, 2003; Ratcliff, Philiastides, and Sajda, 2009; Churchland, Kiani, and Shadlen, 2008; Deco, Rolls, and Romo, 2010; Bode et al., 2012; Bowman, Kording, and Gottfried, 2012; Vugt et al., 2012; White, Mumford, and Poldrack,

2012; Brunton, Botvinick, and Brody, 2013; Ossmy et al., 2013) and value-based decision making studies (Gold and Shadlen, 2007; Krajbich, C. Armel, and Rangel, 2010; Krajbich, Lu, et al., 2012; M. M. Mormann, Malmaud, et al., 2010; Krajbich and Rangel, 2011; Hunt et al., 2012; Tsetsos, Chater, and Usher, 2012; Philiastides and Ratcliff, 2013). Choices were well described by a logistic function of the relative attractiveness of the two items with a significant but negligible bias (mixed effects logistic regression: constant = 0.08246, $p = 0.0115$, slope = 1.15047, $p < 10^{-16}$; **Figure 2.4A**). The mean frequency of correct trials across subjects was 86.3% (SD = 5.1%). Response times decreased as choice ease increased (mixed effects linear regression: slope = -277.77 ms, $p = 10^{-11}$; **Figure 2.4B**). We measured choice ease using the relative proximity difference between the items with the closest and farthest orientations to the orientation of the target. The mean response time was 1,849 ms (SD = 613 ms). Also consistent with previous studies (Krajbich, C. Armel, and Rangel, 2010; Krajbich and Rangel, 2011), we found that the number of fixations per trial decreased as choice ease increased (mixed effects linear regression: slope = -0.28 fixations, $p = 10^{-20}$; **Figure 2.4C**). The mean number of fixations was 2.83 (SD = 0.39). Together, these analyses showed that our perceptual task exhibits psychometric properties common in 2-alternative forced choice tasks, which are predicted by a wide class of sequential integrator models, including the aDDM.

Properties of Fixations

We recorded fixations using an eye-tracker, which allowed us to characterize their properties during the choice process. For this purpose, we classified each item fixation as "first," "middle" or "last," according to when it occurred within the trial. "Middle" fixations are those that are neither the first nor the last ones.

We found that the probability that the first fixation is to the item with closest orientation to the target was not significantly different from chance, and was independent of the relative proximity difference between the two stimuli (mixed effects linear regression: slope = -0.009, $p = 0.095$). Moreover, we found that the duration of first fixations increased with the relative proximity of the fixated item (mixed effects linear regression: slope = 30.55 ms, $p = 10^{-12}$; **Figure 2.5A**), decreased with the relative proximity of the unfixated item (mixed effects liner regression: slope = -11.36 ms, $p = 10^{-7}$), and increased with the relative proximity difference between fixated and unfixated items (mixed effects linear regression: slope = 17.67 ms, $p = 10^{-14}$; **Figure 2.5B**). We did not find a significant correlation between first fixation durations and choice ease (mixed effects linear regression: slope = -1.36

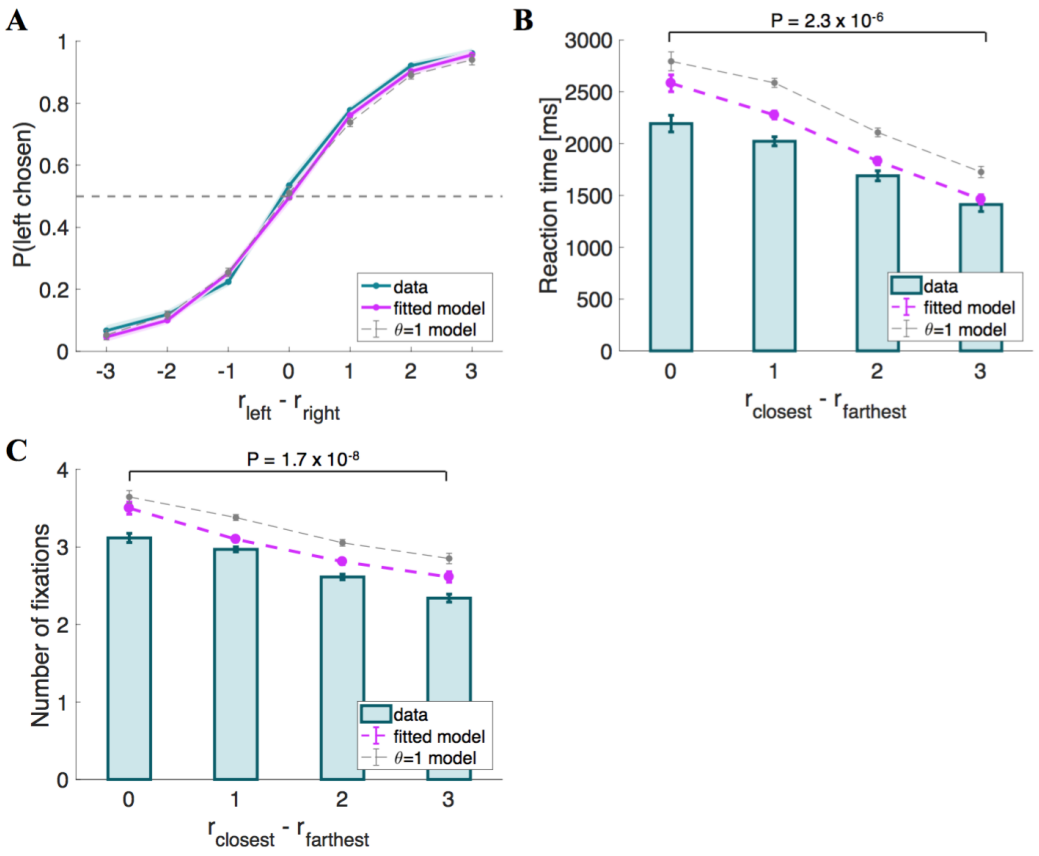

Figure 2.4: Basic psychometrics for Experiment 1. (A) Psychometric choice curve. (B) RT curve depicting mean response times vs. trial ease, as measured by the difference in absolute proximity between the correct and incorrect options. (C) Mean number of fixations vs. trial ease. Subject data includes only even-numbered trials. Fitted model is the best fitting aDDM with free θ, and $\theta = 1$ corresponds to the DDM. Error bars show 95% confidence intervals for the data pooled across all subjects, and across all simulated trials in the case of the data predicted by the models. Tests are based on a paired two-sided t-test.

ms, $p = 0.55$; **Figure 2.5C**).

When looking at middle fixations, we found that their duration increased with the relative proximity of the fixated item (mixed effects linear regression: slope = 59.9 ms, $p = 10^{-17}$; **Figure 2.5D**), decreased with the relative proximity of the unfixated item (mixed effects linear regression: slope = -40.187 ms, $p = 10^{-5}$), and increased with the relative proximity difference between the fixated and the unfixated items (mixed effects linear regression: slope = 40.77 ms, $p = 10^{-21}$; **Figure 2.5E**). Finally, we found that middle fixation durations decreased as choices became easier (mixed effects linear regression: slope = -24.4 ms, $p = 0.00075$; **Figure 2.5F**).

These findings show that the observed fixations patterns are consistent with the assumptions of the aDDM described above. Importantly, note that these analyses did not include last fixations because their duration is endogenous in the aDDM, even under the maintained hypothesis that fixation durations and locations are not affected by the state of the choice process. The endogeneity of the last fixations follows from the simple fact that they are terminated whenever the choice is made upon crossing a barrier.

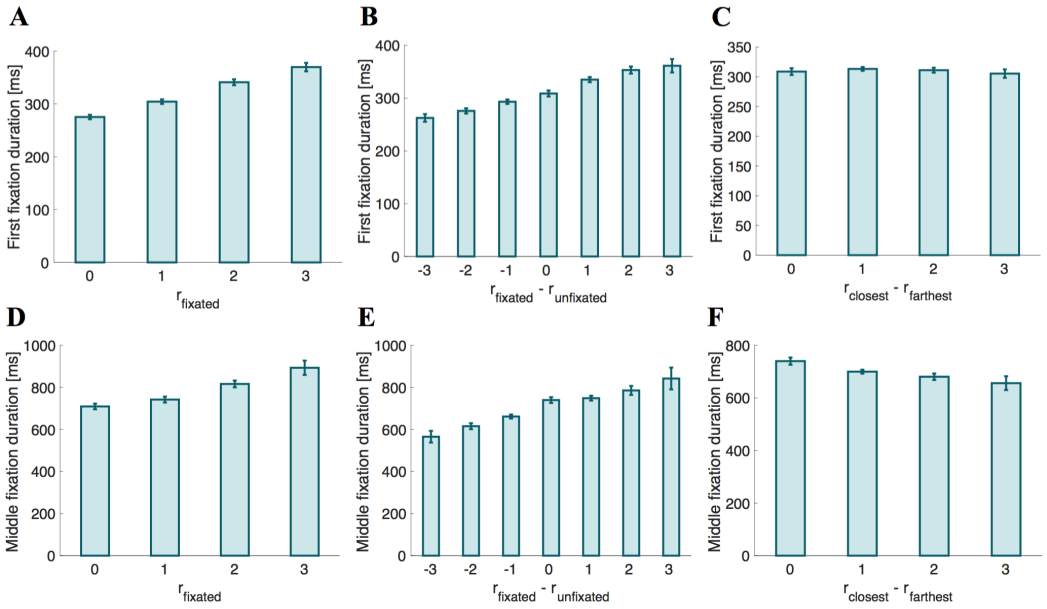

Figure 2.5: Fixation properties. (A) First fixation duration as a function of the relative proximity of the fixated item. (B) First fixation duration as a function of the relative proximity difference between the fixated and the unfixated items. (C) First fixation duration as a function of choice ease. (D) Middle fixation duration as a function of the relative proximity of the fixated item. (E) Middle fixation duration as a function of the relative proximity difference between the fixated and the unfixated items. (F) Middle fixation duration as a function of choice ease. Error bars show 95% confidence intervals for the data pooled across all subjects.

Model Fitting

We divided the data into even- and odd-numbered trials, used the odd trials to fit the free parameters of the aDDM using maximum likelihood estimation (see Materials and Methods for details), and then tested the predictions of the model out-of-sample using the even trials. The best fitting parameters resulting from the group-level MLE were $d = 0.0041$, $\sigma = 0.063$, and $\theta = 0.36$ (using a time step of size 10 ms). Since θ is much smaller than 1, this suggests a sizable attentional bias.

We then used the best fitting parameters to simulate behavior in the even-trials, and compared it to the actual observed data (see Materials and Methods for details). The simulated data provided a reasonably good qualitative and quantitative match to the observed out-of-sample behavior. The psychometric choice curve (Efron's pseudo $R^2 = 0.086$; **Figure 2.4A**) predicted that choices are a logistic function of relative orientation differences, and that response times (goodness-of-fit: $p = 0.10$, **Figure 2.4B**) and number of fixations (goodness-of-fit: $p = 0.22$, **Figure 2.4C**) decrease as choice ease increases.

In the Supplementary Materials (Tavares, Perona, and Rangel, 2017) we present three additional sets of results that might be of interest to the reader. First, we provide individual subject fits. Second, we estimate a non-linear version of the aDDM and find that the best fitting model is approximately linear (as in the basic aDDM). Third, we fit the aDDM separately for trials occurring immediately after target display, where the memory of the target orientation is fresh, and trials occurring 4 trials after target display, where the memory of the target orientation might have dissipated. The best fitting parameters in both cases are very similar, which suggests that this was not an issue affecting performance in the task.

Model Predictions

We next tested for several basic predictions of the aDDM.

First, the model predicts that final fixations should be shorter than middle fixations. This prediction follows from the fact that, according to the model, last fixations are interrupted when the RDV reaches one of the barriers, cutting the last fixation short. We found this to be the case in our data, as both second fixations as well as other middle fixations (middle fixations excluding second fixations) are significantly longer than last fixations ($p = 10^{-9}$ and $p = 10^{-15}$, respectively; **Figure 2.6A**).

Second, the model predicts that subjects should exhibit a bias toward choosing the last fixated item, even in trials where they have fixated on both of them. This prediction follows from the fact that when $\theta < 1$, the RDV moves toward the decision barrier of the fixated item unless it is significantly less desirable than the other item. For example, when $\theta = 0.5$, the RDV moves toward the left barrier when fixating left as long as $r_{\text{left}} > 0.5 r_{\text{right}}$. This pattern was observed both in the data and the simulations (Efron's pseudo $R^2 = 0.11$; **Figure 2.6B**).

Third, the model predicts a very specific relationship between the duration of the last fixation and the pattern of previous fixations. At any point in time within the trial, we

can compute the relative fixation time of the fixated item, given by the total fixation time on that item thus far minus the total fixation time on the other item thus far. The model predicts that in trials where the last fixated item is chosen, the duration of the last fixation decreases with its relative fixation time computed at the beginning of the last fixation. This effect is due to the nature of the RDV signal: when $\theta < 1$, the longer an item is fixated in the trial, the more the signal will move toward its barrier, so the last fixation to the other item (which will eventually be chosen) will have to be longer so that the signal can move back toward its barrier. As shown in **Figure 2.6C**, this effect was present in the data (mixed effects linear regression: slope = -0.28, $p = 10^{-11}$) and the simulations (goodness-of-fit: $p = 0.48$).

Choice Biases

The aDDM also predicts several choice biases when $\theta < 1$, which we tested next.

First, the model correctly predicts a last-fixation bias: subjects are more likely to choose the last item fixated in the trial (Efron's pseudo $R^2 = 0.11$ for left last fixated, and 0.084 for right last fixated; **Figure 2.7A**), due to the fact that the relative proximity of the unfixated item is underweighted in the RDV integration process. Note that a sizable bias effect can be seen both in our data as well as in the simulations (for instance, when $r_{\text{left}} - r_{\text{right}} = 0$, the difference in the probability of choosing left when left was last fixated vs. when right was last fixated is 0.51 for the data, and 0.26 for the simulations).

Second, the model predicts that the probability of choosing an item increases with its overall relative fixation time. This follows from the fact that, because the RDV moves toward the barrier of the fixated item (unless it is significantly worse than the other one), the RDV is more likely to move in the direction of an item's barrier when that item is being fixated than when it is not. Consistent with this, we found a strong association between overall relative fixation times and choices (Efron's pseudo $R^2 = 0.14$; **Figure 2.7B**). However, a concern with this test is that overall relative fixation times and relative proximity are correlated. To correct for this, we computed a corrected choice probability curve by subtracting from each trial's choice (1 for left and 0 for right) the average probability of choosing left for that particular relative proximity difference. This curve provides an uncontaminated measure of the effect of relative fixation at the time of choice, under the assumptions of the aDDM. As shown in **Figure 2.7C**, the observed choice bias was sizable, and matched well the simulated data (goodness-of-fit: $p = 0.038$).

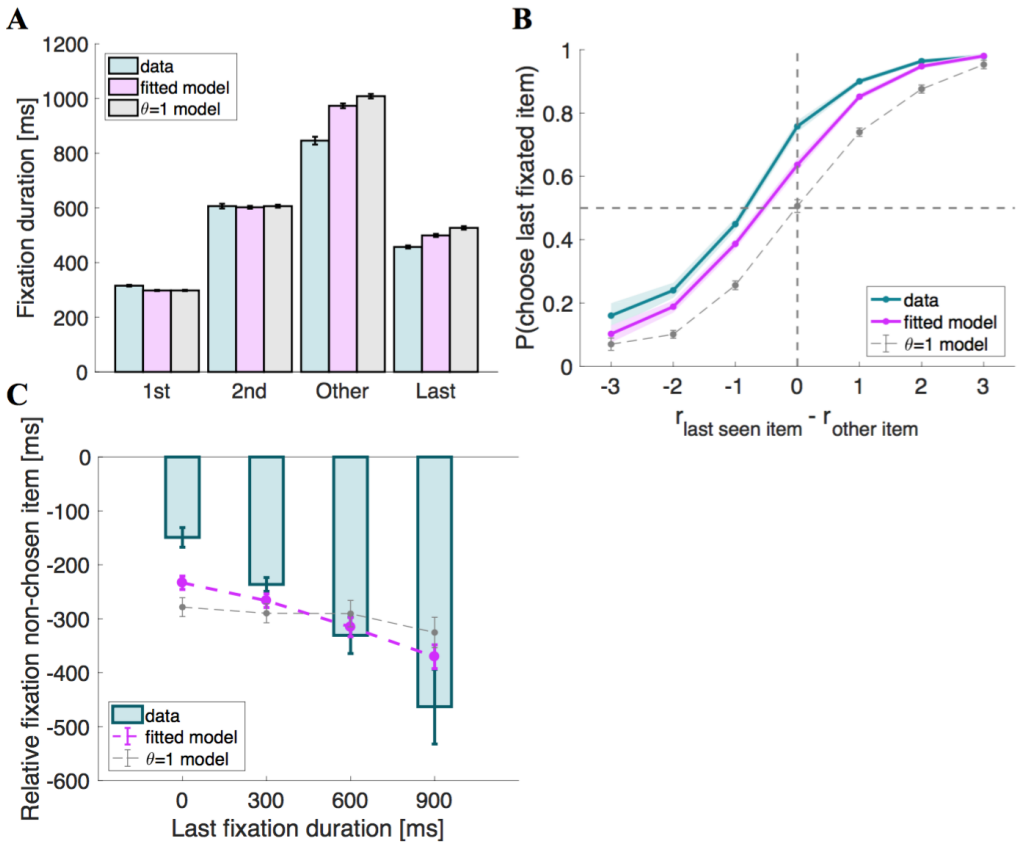

Figure 2.6: Model predictions. (A) Fixation duration by type. As predicted by the model, last fixations are shorter than middle fixations. Note that, except for last fixations, the match between the data and the model is a direct consequence of our fixation sampling process. (B) Probability that the last fixation is to the chosen item as a function of the relative proximity difference between the last fixated item and the other item. In the absence of a bias effect, the probability at 0 should be around 0.5; due to the bias, the observed probability is significantly larger than 0.5. (C) Amount of time spent looking more at item A before the last fixation (to item B), as a function of the duration of that last fixation. Subject data includes only even-numbered trials. The fitted model is the best fitting aDDM with free θ, and $\theta = 1$ corresponds to the DDM. Error bars show 95% confidence intervals for the data pooled across all subjects, and across all simulated trials in the case of the models.

Finally, the model also predicts that the likelihood of choosing the first seen item increases with the duration of the first fixation. This was observed in the data (mixed effects linear regression: slope = 0.00075, $p = 10{-}35$; **Figure 2.7D**) and in the simulations (Efron's pseudo $R^2 = 0.088$). The effect was still present after correcting for relative proximity differences, by subtracting from each trial's choice the average probability of choosing the first-seen item for that particular relative

proximity difference (goodness-of-fit: $p = 0.021$; **Figure 2.7E**).

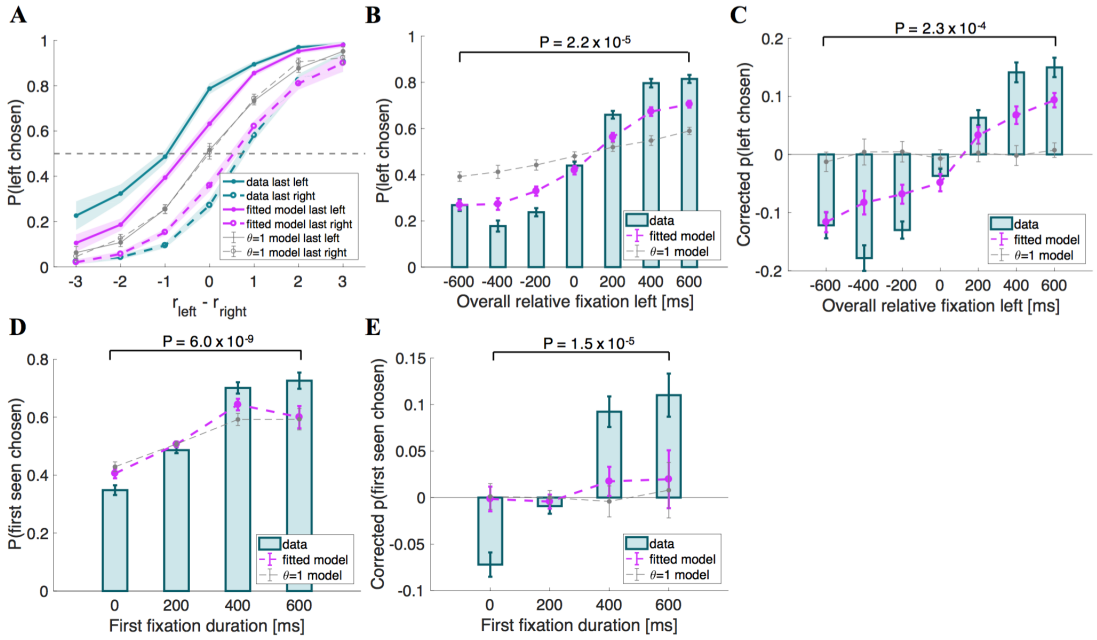

Figure 2.7: Choice biases. (A) Psychometric choice curves conditioned on the location of the last fixation. (B) Probability that the left item is chosen as a function of the excess amount of time for which the left item was fixated during the trial. (C) Analogous to (B), except subtracting the average probability of choosing left for each relative proximity difference. (D) Probability that the first-seen item is chosen as a function of the duration of that first fixation. (E) Analogous to (D), except subtracting the average probability of choosing the first-seen item for each relative proximity difference. Subject data includes only even-numbered trials. Fitted model is the best fitting aDDM with free θ, and $\theta = 1$ corresponds to the DDM. Error bars show 95% confidence intervals for the data pooled across all subjects, and across all simulated trials in the case of the models. Tests are based on a paired two-sided t-test.

Model Comparison

As discussed above, the aDDM reduces to the standard DDM, in which fixations do not affect choices or RTs, when $\theta = 1$. This provides an additional way of exploring the role of visual attention, by comparing the ability of the best fitting aDDM and the best fitting DDM to explain the data out-of-sample.

This was done in several steps. First, we re-estimated the model in the odd-numbered trials under the restriction that $\theta = 1$, which amounts to finding the best fitting DDM. We found that the best fitting parameters in this case were $d = 0.0024$ and $\sigma = 0.062$.

In contrast, the best fitting parameters for the aDDM were $d = 0.0041$, $\sigma = 0.063$, and $\theta = 0.36$.

Second, we used these best fitting parameters to carry out the same out-of-sample predictions described above, but for the best fitting DDM. As shown throughout the figures, the results show that when $\theta = 1$ the model cannot account for key aspects of the data patterns and choice biases. For example, consider **Figure 2.6B**, which shows that there is a sizable choice bias in favor of the last fixation. As the figure shows, the best fitting aDDM can account for this pattern, which follows from the overweighting of r_{fixated} relative to $r_{\text{unfixated}}$. In contrast, the best fitting standard DDM cannot explain this pattern since attention does not matter in that model.

Third, we compared the ability of the two models to predict choices and RTs out-of-sample, which provides a test of the value of fixation data in predicting choices and RTs (see Materials and Methods for details). We compared the accuracy of three types of predictions. For the standard DDM we predicted choices and RTs in the even-numbered trials, conditional only on r_{left} and r_{right}, since fixations do not matter in this case. For the aDDM we carried out two different sets of predictions. In one of them we predicted choices and RTs, conditional on r_{left}, r_{right} and on the net fixation time on the left item (i.e., total fixation time on left minus total fixation time on right) in each even trial. In the other, we made predictions based on r_{left}, r_{right}, and, as much as possible, on the actual path of fixations observed on each trial. The most accurate predictions were made by the best fitting aDDM conditional on net fixation time (choice prediction accuracy = 72.3%, average RT absolute error = 1.92 s). The second most accurate predictions were made by the other aDDM exercise (choice prediction accuracy = 70.1%, average RT absolute error = 2.49 s). The least accurate predictions were made by the best fitting DDM (choice prediction accuracy = 68.4%, average RT absolute error = 2.69 s). To put these numbers in perspective, note that the standard deviation of RTs is 1.87s. Together, these results suggest that incorporating fixation information in the way specified by the aDDM improves the out-of-sample predictions made by this class of models.

It might seem counter intuitive that predictions that condition on the observed fixations as much as possible perform worse than those that condition only on observed net fixation time. However, this makes sense once the randomness in the aDDM model is taken into account. Since the model entails significant randomness, repeated runs of the model with the same stimuli will result on different choices, RTs, and net fixation times, even if they follow the same fixation pattern as much

as possible. As a result, repeated runs of the exact same trials can result in net fixations times that are significantly different from those observed in the trial that we are trying to predict, even if the repeated runs require that fixations follow the same process as long as possible (see Section Materials and Methods for details). In contrast, the other set of aDDM predictions are made using runs of the aDDM that result in nearly identical net fixation times. This leads to better predictions because net fixation time on the two items affects the average slope of integration in the diffusion model, and thus choices and RTs. Consistent with this, the standard deviation in the simulated net fixation times was 276 ms in the first aDDM prediction model, 886 ms in the second aDDM prediction model, and 2,603 ms in the DDM prediction model.

Experiment 2: Causal Test of the Attentional Effect

So far we have found that the aDDM provides a reasonable quantitative and qualitative account of the relationship between fixations, choices and RTs in our perceptual task. Importantly, while the aDDM predicts a causal impact of attention on perceptual choice, the evidence presented so far is only correlational. We addressed this issue using an experimental paradigm that manipulates item fixation times with the aid of an eye-tracker, and which has been previously shown to causally affect subjects' choices on a moral decision task (Pärnamets et al., 2015).

The task, depicted in **Figure 2.2**, consisted of a modification of the previously described perceptual choice task. The key difference is that in each trial we randomly selected one of the two items on the screen to be the bias-target for that trial, and implemented the following procedure to bias fixation toward that item (see Materials and Methods for details). Unbeknownst to the subjects, we defined a minimum period of time required for them to fixate on each item before a decision could be made: 800 ms for the bias-target and 200 ms for the non-bias-target. We then used the eye-tracker to record the duration of each fixation and, as soon as the minimum requirement for both items was met, the items disappeared from the screen, and the subject was asked to make a choice. Note that this requirement does not guarantee that the bias-target will be fixated longer, since it only establishes a minimum amount of time for each item to be fixated, but not a maximum amount. To ensure that subjects would not become aware of our manipulation, we set the maximum duration for each decision trial at 3 seconds. If the minimum fixation requirements for both items were not met within that period, the subject was prompted to make a decision, and the trial would be discarded from all future analyses. Overall, 24.3%

of the trials (1,633 trials) were discarded in this manner. We refer to trials in which the manipulation was effective (i.e., in which the bias-target was fixated longer) as effective trials. In order to increase the number of effective trials, we also attempted to guide subjects' first item fixations toward the bias-target (Hikosaka, Miyauchi, and Shimojo, 1993), so that it would have a better chance of being fixated longer. In particular, the bias-target stimulus was always displayed first, and the other stimulus appeared on the screen after a short delay (duration: mean = 450 ms, SD = 35 ms). Using this manipulation, 74.3% of the non-discarded trials were effective (i.e., a total of 3780 trials were effective).

To check the success of the experimental manipulation, we compared the overall relative fixation time advantage of the left item in all trials from Experiment 1 vs. all effective trials in Experiment 2. As shown in **Figure 2.8A**, in Experiment 1 the left time advantage increased with the difference in relative proximity, while in Experiment 2 the bias-target was fixated longer regardless of the items' relative proximity. In effective trials, the mean total fixation time on the bias-target was 814.2 ± 85.8 ms, while for the other item it was 509.2 ± 46.1 ms.

As predicted by the aDDM, we found that the probability of choosing the left item was higher on trials where the left item was the bias-target than on those where the right item was the bias-target (**Figure 2.8B**). To appreciate the magnitude of the bias, note that when $r_{\text{left}} - r_{\text{right}} = 0$, the probability of choosing left increases by 14% across the two conditions (χ^2 statistic = 16.51, $p = 10^{-5}$).

Importantly, the bias obtained here is comparable in magnitude to what we found in Experiment 1. In particular, the slope of the corrected choice curve in **Figure 2.7C** is 0.02, which implies that a shift in relative fixation of 300 ms (which is similar in size to the one induced by the experimental manipulation in Experiment 2) should induce about a 6% increase in the probability of choosing the item. **Figure 2.8C** shows a comparison of this effect between Experiments 1 and 2, illustrating that the quantitative effects of the causal manipulation in Experiment 2 are consistent with the measurements from Experiment 1, providing additional support for the validity of the causal manipulation.

A natural concern with these results is that the observed effect might have been due, in part, to priming: since the first fixation was manipulated to be to the bias-target item, this could have primed the subjects to bias their choices in this direction (Meyer and Schvaneveldt, 1971; Nedungadi, 1990). Another concern is that, by discarding trials in which the manipulation was not effective, we are biasing the results toward

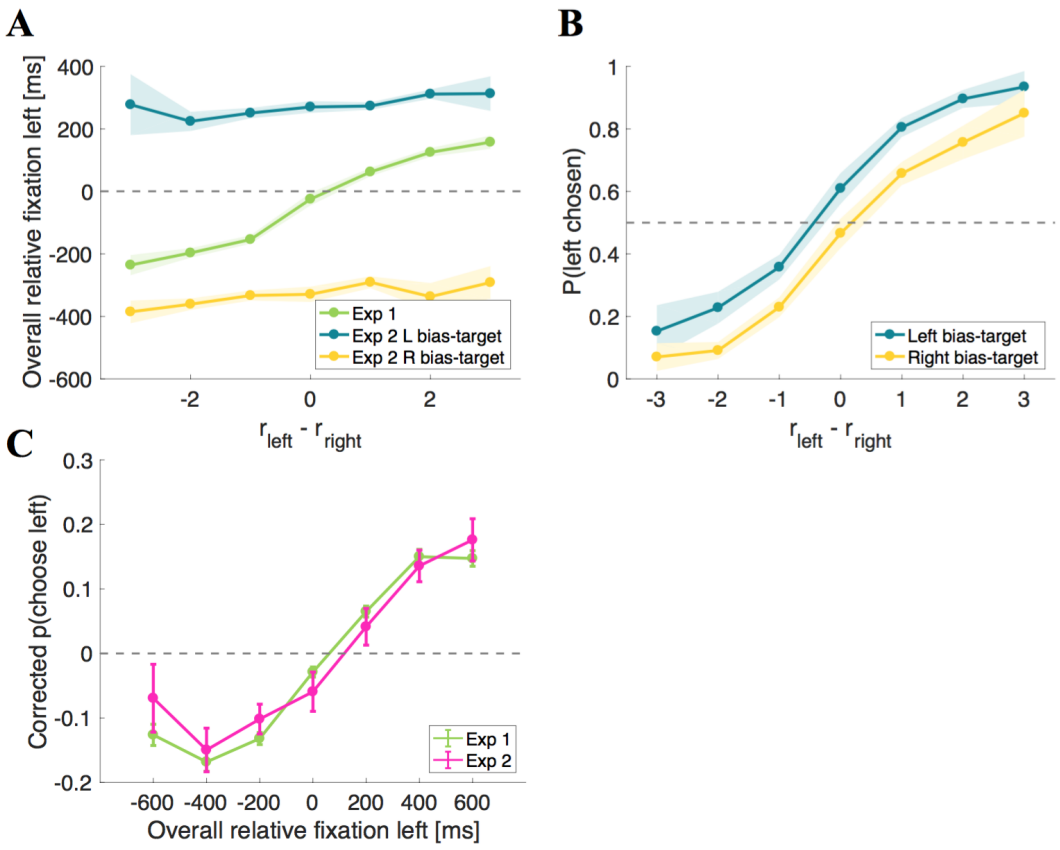

Figure 2.8: Causal test of the attentional effect. (A) Time advantage of the left item over the right item comparing effective trials from Experiment 2 to all trials from Experiment 1. (B) Psychometric choice curves conditioned on the bias-target item. Subjects choose the bias-target with higher probability. (C) Corrected probability that the left item is chosen as a function of the excess amount of time for which the left item was fixated during the trial, comparing all trials from Experiment 1 to all trials from Experiment 2. Corrected probabilities are obtained by subtracting from each trial's choice (1 for left and 0 for right) the average probability of choosing left for the relative proximity difference from that trial. Error bars and shaded error bars show 95% confidence intervals for the data pooled across all subjects.

the hypothesis that longer fixations increase the probability of choosing the fixated option. To address these issues, we split the trials into two sets based on whether or not the bias-target was the longest fixated item in the trial (i.e., effective vs. ineffective trials), regardless of which item was fixated first, and compared the size of the bias in these two types of trials (**Figure 2.9**). If the observed effect were exclusively due to priming, one would expect a similar choice bias in both groups of trials. In contrast, the aDDM predicts a stronger choice bias in the trials in which the bias-target was fixated longer. Consistent with this, we found that the choice-bias

was larger in trials where the bias-target was fixated longer than in those when it was fixated less (comparison of the individual biases in logit regressions: mean constant difference 1.45 vs. 0.57, paired t-test $t = 7.13$ and $p = 10^{-8}$ vs. $t = 0.27$ and $p = 0.79$; when the bias-target was longest fixated, **Figure 2.9A**, mean total fixation time was 814 ms for the bias-target and 509 ms for the non-bias-target, and when the bias-target was least fixated, **Figure 2.9B**, mean total fixation time was 801 ms for the bias-target and 1,132 ms for the non-bias-target).

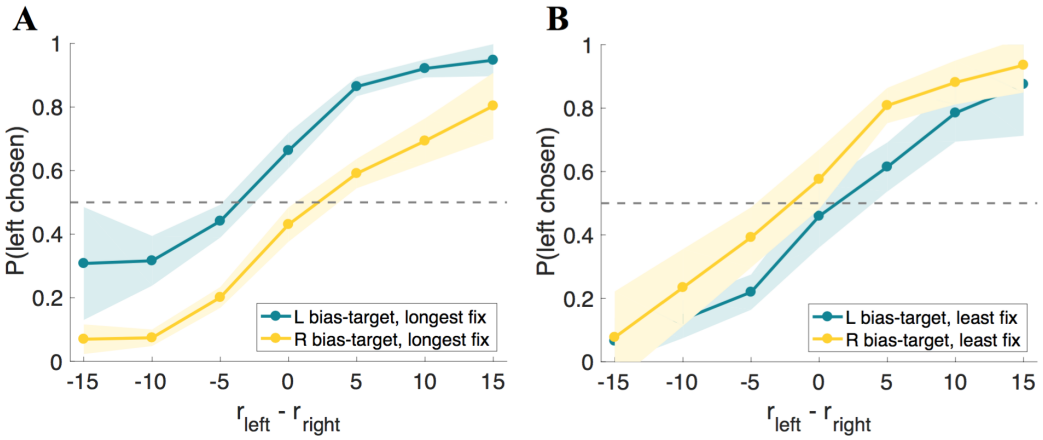

Figure 2.9: Experiment 2 choice curves, effective vs. ineffective trials. Experiment 2 psychometric choice curves conditioned on the bias-target item and on whether the bias-target was the longest (A; effective trials) or least (B; ineffective trials) fixated item in the trial (as measured by the total fixation time on each item).

2.4 Discussion

The results described above provide evidence consistent with the hypothesis that the aDDM gives a plausible algorithmic description of the impact of attention in simple perceptual decision making. Experiment 1 shows that the model is able to provide a reasonably good (although not perfect) quantitative description of the relationship between fluctuations in visual attention, choices and response times. Experiment 2 shows that the impact of attention in choice predicted by the aDDM is causal, and of a qualitatively similar size as that predicted by the best fitting model in Experiment 1.

The imperfect match between our data and the model simulations could be due to at least two important factors. First, the model we used to simulate the fixation process is a simplistic approximation, which, under the assumptions of the aDDM, adds noise to the simulations. Second, the results presented for the model simulations are

averaged across trials, so the fact that they are not conditioned on the same fixations present in the data also adds noise.

The version of the aDDM tested here is virtually identical to the one used in previous value-based choice studies (Krajbich, C. Armel, and Rangel, 2010; Krajbich, Lu, et al., 2012; Krajbich and Rangel, 2011; Towal, M. Mormann, and Koch, 2013). The only difference is that in value-based choice the evidence that is integrated is composed of noisy measurements of preference for the stimuli, whereas here it is noisy perceptual signals about line orientation. This suggests that a similar simple class of algorithms with only three free parameters is able to provide a quantitative characterization of several complex behavioral patterns in the data, such as the impact of relative fixation durations, or the impact of first fixations. This provides further support for the view that the brain utilizes similar algorithms, and perhaps similar neural architectures, for sufficiently similar classes of problems, even if they operate in domains as different as perception and value-based choice.

Suggestively, the attentional bias we found in this study ($\theta = 0.36$) is substantial and of similar size to the bias found in previous value-based studies ($\theta = 0.3$) (Krajbich, C. Armel, and Rangel, 2010). This result leads us to speculate that attentional biases might be sizable in any simple decision task (perceptual or value-based) in which fixations facilitate the evidence gathering process.

An important feature of the aDDM is that it posits a causal impact of attention on choice. In particular, it assumes that the evidence related to fixated items is weighted more heavily during the decision process, and as a result choices can be biased toward a stimulus by increasing its share of fixations. Furthermore, when the attentional bias parameter is much smaller than 1, the predicted biases can be sizable. Previous studies of value-based choice with exogenously manipulated fixations have found causal effects in the predicted direction, but of smaller magnitude than predicted by the model (Shimojo et al., 2003; K. C. Armel, Beaumel, and Rangel, 2008). One potential explanation for the small effect sizes is that the experimental manipulations had limited success in shifting attention. Experiment 2 provides evidence consistent with this interpretation. Here we utilized a different experimental manipulation of attention that was able to shift fixations and found a causal effect of a similar magnitude as the one predicted by the model. In fact, our attentional manipulation was inspired by a recent study of attention and moral decisions, which also found sizable effects (Pärnamets et al., 2015).

A critical assumption of the aDDM is that the fixation process is independent from

the sequential integration process. More precisely, random fluctuations in the RDV variable that guides the choices do not affect the fixation process. We emphasize that this does not rule out the possibility that stimulus properties orthogonal to the decision process might affect fixations. Indeed, a recent study in the domain of value-based choice showed that the fixation process was affected by low level visual features (Towal, M. Mormann, and Koch, 2013), and several others have provided evidence that the value integration process is modulated by the saliency of the stimuli (Tsetsos, Chater, and Usher, 2012; Tsetsos, Moran, et al., 2016). Instead, the key assumption of the model is that the actual integration process does not affect the fixation process. Thus, the aDDM can be thought of as a model of the decision process, taking as given the exogenous and potentially stochastic fixation process. In this study, as well as previous ones (Krajbich, C. Armel, and Rangel, 2010; Krajbich and Rangel, 2011; Towal, M. Mormann, and Koch, 2013), this is implemented by taking the fixation process to be the one that best describes the one observed in the data.

We do not view our results as providing evidence for the hypothesis that the attentional process is not influenced by the state of the decision process variables. In fact, studies of value-based choice with large numbers of items have found that fixations are shifted toward the best items several seconds into the decision process (Dawling et al., 2011). Instead, our results suggest that these additional influences on attention have a limited impact on the choice process, since most of the effects are already accounted for by the simpler aDDM. However, we also conjecture that "top-down" modulations of attention are more likely to occur in more complex decisions with longer response times. A full characterization of how the decision process affects attention, and how this feeds back to the choices, is a critical open question for ongoing research. For example, a recent study has shown that the choice bias toward the last fixated item shown here can arise in multiple types of integrator models, even when attention is entirely random and independent of the choice process (Dawling et al., 2011; Mullett and Stewart, 2016).

The aDDM provides a simple way of introducing attention in sequential integrator models of choice, by adding an extra parameter to the most basic version of the Drift-Diffusion Model. However, similar modifications could be introduced to a number of other sequential integrator models of choice, including leaky-accumulator models (Usher and McClelland, 2001), neural network models of the choice process (Wong and Wang, 2006; Hunt et al., 2012), or more complex versions of the

Drift-Diffusion Model (Churchland, Kiani, and Shadlen, 2008; Ratcliff and McKoon, 2008; M. M. Mormann, Koch, and Rangel, 2011; Hawkins et al., 2015), among others. Such modifications would have qualitatively similar effects, provided that the assumption that fixations are orthogonal to the state of the decision process is maintained. Given the active debate in the literature about which sequential integrator model provides the best description of the underlying processes, an important direction for future research is to carry out a systematic comparison of the attentional versions of all of these models.

The aDDM models the effects of attention at a high level of abstraction. Another important question for future research is to characterize the neural mechanisms behind the attentional effects captured by the model. For example, does attention operate at the perceptual stage, prior to the integration of the information by the decision process, or does it operate in the decision process itself? Does attention operate through similar channels in perceptual and value-based choice, and if so, why does it have a similar effect on choice?

Chapter 3

A FORWARD LIKELIHOOD METHOD FOR PARAMETER ESTIMATION IN DRIFT-DIFFUSION MODELS

3.1 Introduction

Making decisions is an essential part of our daily lives, and much effort has been put into developing computational models that can explain the principles and algorithms underlying simple decisions (Jerome R Busemeyer and Johnson, 2004; Rangel, C. Camerer, and Montague, 2008; Glimcher and Fehr, 2013; Kahneman and Tversky, 1979). An important class of models is that of Drift-Diffusion Models (DDMs), where the decision-maker accumulates evidence for each option over time, until a relative decision value signal converges to one of two barriers and a decision is made (Stone, 1960; Ratcliff, 1978; Ratcliff, Cherian, and Segraves, 2003; Ratcliff and Smith, 2004; Laming, 1979; Usher and McClelland, 2001; Smith, 1995; Smith and Ratcliff, 2004; Ditterich, 2006; Bogacz, 2007; Gold and Shadlen, 2001; Gold and Shadlen, 2002). Previous work has shown that DDMs implement an optimal decision making process that is equivalent to a sequential likelihood ratio test (Gold and Shadlen, 2001; Gold and Shadlen, 2002; Gold and Shadlen, 2007; Bogacz, 2007; Bogacz, E. Brown, et al., 2006; Reddi and Carpenter, 2000). Furthermore, the applicability of DDMs has been consistently verified in a variety of contexts, including perceptual decision tasks, where choice is based on physical properties of the options presented (Ratcliff, Cherian, and Segraves, 2003; Smith and Ratcliff, 2004; Ratcliff and Rouder, 1998; Ditterich, 2006; Gold and Shadlen, 2001; Brunton, Botvinick, and Brody, 2013; Gold and Shadlen, 2007), as well as value-based (economic) decisions tasks, where choice is based on an individual's subjective preferences (M. M. Mormann, Malmaud, et al., 2010; Tsetsos, Usher, and McClelland, 2011; Hutcherson, Bushong, and Rangel, 2015; Philiastides and Ratcliff, 2013; Hunt et al., 2012).

A critical step in behavioral studies of decision making is the estimation of the model parameters that best explain the observed psychometric data. Given this, there is considerable interest in developing fast, reliable methods that can be used to fit the behavioral models to experimental data. When estimating the parameters of the DDM in its simplest version, it is possible to obtain a closed-form solution for

the likelihood of the model (Lehmann, 1983). However, the inclusion of additional parameters, which are often needed to account for key aspects of the data, requires likelihoods to be computed through a process of numerical approximation, which can be slow and unreliable.

Several different methods can be used to approximate the likelihood of a DDM. However, in general these methods require slow computations and a large amount of experimental data in order to generate good approximations. In this chapter we present an algorithm for fitting the DDM which aims to generate reliable parameter estimates while requiring significantly less time and data than the more widely used approaches.

We begin this section by providing a brief overview of the related literature. We then provide a mathematical description of the DDM and briefly discuss the traditional maximum likelihood approach used to fit the parameters of the model.

Review of Existing Literature

Several previous studies have addressed the comparison between different methods for fitting DDMs. Most notably, Ratcliff and Tuerlinckx compared three different methods for DDM parameter estimation, namely, maximum likelihood, chi-square, and weighted least squares, and made recommendations for when to use each one of them depending on characteristics of the data (Ratcliff and Tuerlinckx, 2002). The maximum likelihood approach involves generating RT histograms from the data and from model simulations, then multiplying the corresponding bins in order to generate an approximation for the likelihood (more details about this method are given later in this section). The chi-square fitting method also uses RT histograms, but in this case the histograms are used to compute a chi-square statistic, which can then be minimized by parameter adjustment. Finally, the weighted least squares method relies on minimizing the sum of squared differences between observed and predicted accuracies plus the sum of squared differences between observed and predicted quantile response times.

In their study, Ratcliff and Tuerlinckx found that the maximum likelihood approach was superior in terms of the bias and standard deviations of the estimated parameters relative to the parameter values used to generate the data. On the other hand, they also found that the same method was sensitive to the presence of outliers and variability in non-decision processes, which led to poor fitting in several cases. The authors discussed the issues that arise when contaminant response times or variability in

non-decision processes are present in the data, and how to adapt each of the three methods in order to deal with these issues.

Diederich and Busemeyer reviewed matrix methods that can be applied to different sequential sampling models, including the DDM (Diederich and Jerome R Busemeyer, 2003). The Markov chain approximation approach they used is very similar to the one we present in this chapter, but while they focused on a general mathematical description of the method to be used with any type of diffusion process, here we focused on its application to the particular problem of fitting DDMs to empirical data. Another key difference is that we also apply the algorithm to a modified version of the model that includes an additional experimental parameter and makes use of visual fixations. This makes for an interesting application since visual fixations vary from trial to trial and are independent from the decision process.

Finally, Brunton et al. applied a method similar to the one we describe here to fit a DDM to data from a decision making experiment in rats (Brunton, Botvinick, and Brody, 2013). Their model fitting procedure was based on transition probabilities and included neural data as well as behavioral data.

Drift-Diffusion Model

In the simplest version of the DDM, the value of the decision variable starts at zero and evolves at a constant rate over time until it reaches one of two barriers, $+B$ and $-B$, corresponding to the two choices (or items) presented in the task (for convenience, the notation to be used throughout the remainder of this chapter is provided in **Table 3.1**). Once one of the barriers is reached, a decision is made and the chosen option is the one that corresponds to the barrier reached by the signal. The relative decision value (RDV) at each time step, V_t, is computed as

$$V_t = V_{t-1} + \delta. \tag{3.1}$$

The step δ is drawn from a normal distribution $N(\mu, \sigma)$, whose mean μ is computed as

$$\mu = d(r_{\mathrm{A}} - r_{\mathrm{B}}), \tag{3.2}$$

where d is a parameter that controls the speed of integration of the signal, and r_{A} and r_{B} are the values of options A and B, respectively. Note that we use the word

"value" here to mean the attractiveness of an option in the context of the choice task, which can represent either a physical attribute in the case of perceptual decisions, or a measure of utility in the case of economic decisions.

Table 3.1: Notation for Chapter 3.

Symbol	Description
Θ_m	Set of model parameters.
Θ_c	Set of experimental condition parameters.
V_t	Relative decision value at time t.
δ	Step in the relative decision value.
r_A	Value of item A.
r_B	Value of item B.
B	Fixed DDM parameter; magnitude of the decision barriers.
d	Free DDM parameter; controls speed of signal integration.
σ	Free DDM parameter; standard deviation of normal distribution.
θ	Free aDDM parameter; controls attentional bias.
μ	Mean of normal distribution.

The model comprises two free parameters, d and σ, as well as an additional parameter, B, which determines the size of the barriers and which can be set to 1 for simplicity, as long as σ is kept free. For the remainder of this chapter, we will use as a convention the fact that the upper barrier, +1, corresponds to a decision for item A, while the lower barrier, -1, corresponds to a decision for item B.

We can differentiate between the model's set of parameters, Θ_m, and the set of parameters that depend on experimental conditions, Θ_c. The set of model parameters Θ_m includes, in the simplest case, the free parameters d and σ, which can be fitted through a grid search maximum likelihood estimation procedure. More sophisticated versions of the DDM can incorporate additional parameters to control the rate at which the size of the barriers changes over time, to account for visual and motor delays, as well as other improvements that are beyond the scope of this chapter. The set of experimental condition parameters Θ_c is determined at the time the experiment is designed. The parameters in Θ_c can vary from trial to trial and include, at least, the values of the two items, r_{A} and r_{B}. Other examples include the trial's level of difficulty, the trial's visual features (such as saliency, colors, and brightness), the subject's visual fixations, as well as other behavioral measurements.

During data collection, the key experimental measurements one obtains from sub-

jects are two: a response time (the time elapsed from the beginning of the trial until the subject reports the decision) and a choice (e.g., A vs. B) for each trial. Any method for estimating model parameters makes use of these measurements over a large number of trials, possibly in different experimental conditions.

Maximum Likelihood Algorithm

The widely used maximum likelihood approach for fitting the parameters of the DDM, which we refer to here as Maximum Likelihood Algorithm (MLA), can be described as follows. First, for each different experimental condition c (i.e., for each possible combination of values of the items A and B presented), we generate a histogram of the response times (RTs) conditional on the final choice. We simply count the number of data trials in each of the RT bins, separating the trials where item A was chosen from those where item B was chosen. After this step, we have two empirical histograms (for options A and B) for each of the K experimental conditions, resulting in $2K$ empirical histograms total. Second, for each model m and each experimental condition c, we generate S simulations of the DDM using the parameter values given by m and the item values determined by c. Using these simulations, we again generate RT histograms conditional on choice, resulting in $2K$ simulation histograms per model. Finally, we compute an approximation of the likelihood of the data given the model m by first multiplying the number of data trials in each bin of the empirical histograms by the number of simulated trials in the equivalent bin of the simulation histograms obtained for model m, then summing up the resulting products. This procedure is repeated for each of the N different models (i.e., samples of the parameter set Θ_m), and the final model selected is the one that maximizes the likelihood over all trials.

From the description above, we can see that the MLA has a few inherent drawbacks. In terms of computation time, the execution of the algorithm depends on the number of distinct experimental conditions available in the empirical data set, K, and on the number of simulations generated for each model m and each experimental condition c, S. The number of experimental conditions K can be very large depending on the task. While there is no exact bound on how many simulations S are necessary in order to generate a good fit, previous studies have used on the order of 1000 simulations per model and experimental condition. This leads to potentially very large execution times, as well as some uncertainty about the quality of the fit, since a larger number of simulations can potentially lead to better results. Nevertheless, this approach has been successfully applied in many previous studies using DDMs

to model experimental data (Ratcliff and Rouder, 1998; Ratcliff and Smith, 2004; Shadlen and Newsome, 2001; Bennur and Gold, 2011; Bowman, Kording, and Gottfried, 2012; Todd A Hare, Schultz, et al., 2011; Hawkins et al., 2015; Ho, S. Brown, and Serences, 2009; Hutcherson, Bushong, and Rangel, 2015; Krajbich, C. Armel, and Rangel, 2010; Krajbich and Rangel, 2011; Krajbich, Lu, et al., 2012; M. M. Mormann, Malmaud, et al., 2010; Polania et al., 2014; Tsetsos, Chater, and Usher, 2012; Tsetsos, Gao, et al., 2012; Tsetsos, Moran, et al., 2016; Ossmy et al., 2013; Ditterich, Mazurek, and Shadlen, 2003).

3.2 Computing Likelihoods from a Probability Table

We now describe an alternative algorithm for estimating the parameters of DDMs, which we refer to as Probability Table Algorithm (PTA). Importantly, the execution time for this algorithm depends only on the number of models being tested, N, and on the number of distinct experimental conditions available in the empirical data set, K, as the method does not require any simulations to be produced during the fitting process. This makes for a much more efficient likelihood computation when compared to the traditional MLA described above.

Given a trial's experimental condition c (i.e., the values of the two options r_{A} and r_{B}) and a set of parameter values $\langle \hat{d}, \hat{\sigma} \rangle$, the algorithm computes, at each time step of the total duration of the trial, the probability of a choice being made for each of the options A and B. In this algorithm, the RDV space is discretized into states corresponding to small equally sized bins. Then, at each time step, the probability that the RDV signal is in the range corresponding to each of these bins is computed, as well as the probability that the signal has crossed each of the two decision barriers.

In the next section we will describe in more detail how to obtain the probability table. For now, we assume that such a table can be obtained, and show how this table can be used to compute the likelihood of a trial's experimental data given a model m. **Algorithm 1** can be used to estimate the likelihood of a single trial from a two-alternative forced-choice task, given a set of DDM parameters. In the DDM likelihood computation, the model to be tested contains fixed values for parameters d and σ, the experimental condition consists of the values of the two options, and the behavioral data consists of the subject's response time and final choice.

The algorithm works as follows. **Line 1:** load trial experimental condition: value of item A, *r_A*, and value of item B, *r_B*. **Line 2:** load trial data: response time, *RT*, and *choice* (we assume 1 is used to represent item A and -1 to represent item B).

Line 3: initialize the parameters of the model with the values to be tested: *d*, *sigma*, and the magnitude of the signal barriers, *B*. **Line 4:** initialize the size of the step for the RDV dimension, *state_step*, and the size of the step for the time dimension, *time_step*, to be used in the probability table computation. **Line 5:** the algorithm will run for a number of time steps, *num_time_steps*, which is defined as the duration of the trial, *RT*, divided by the size of the time step, *time_step*. **Line 6:** set the mean of the normal distribution of transitions between RDV states, *mu*, according to the model (**Equation 3.2**). **Line 7:** call the procedure defined in **Algorithm 2** to obtain vectors *prob_up_crossing* and *prob_low_crossing*, which correspond to the probabilities of the RDV signal crossing the upper and lower barriers over time, respectively. **Line 8-12:** if item A was chosen (i.e., *choice* is equal to 1), the likelihood returned is equal to the probability of crossing the upper barrier at the end of the trial duration; on the other hand, if item B was chosen (*choice* is equal to -1), the likelihood is equal to the probability of crossing the lower barrier.

Algorithm 1 DDM likelihood estimation for a single trial.

```
1:  Load trial experimental condition: r_A, r_B
2:  Load trial data: RT, choice
3:  Initialize model parameters: d, sigma, B
4:  Initialize algorithm hyper-parameters: state_step, time_step
5:  num_time_steps = RT / time_step
6:  mu = d * (r_A - r_B)
7:  prob_up_crossing, prob_low_crossing = compute_probability_table(mu,
    sigma, B, num_time_steps, state_step)
8:  if choice == 1 then
9:      return prob_up_crossing[end]
10: else
11:     if choice == -1 then
12:         return prob_low_crossing[end]
```

3.3 Computing the Probability Table

We now describe the concept of the algorithm in more detail. To calculate the likelihood for a particular experimental condition *c*, both the RDV and the time dimensions are discretized. As illustrated in **Figure 3.1**, this transforms the state-space of the model into a two-dimensional table. Columns denote time, increasing from left to right, while rows denote the state of the RDV. At any discretized time point, the algorithm assigns a probability mass to each RDV bin, which measures the likelihood that the value of the RDV signal falls within that bin, at that time,

conditional on a decision not having made yet (because a decision barrier has not been crossed). Importantly, note that the sum of these likelihoods in a particular column need not add to 1 if there is a positive probability that a decision can be made before that time.

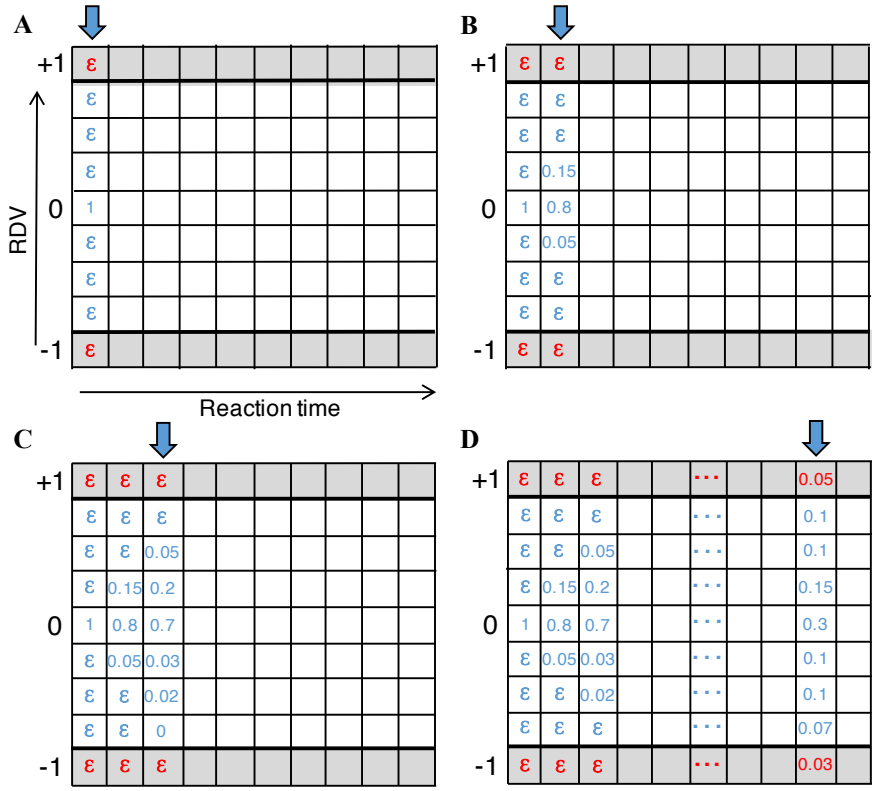

Figure 3.1: Toy example of likelihood estimation using the Probability Table Algorithm. At each time step, the probability of the RDV signal being located at each bin inside the barriers (white cells) is computed. The shaded cells indicate the probability of the signal crossing each of the two barriers. (A) A two-dimensional table is used to compute probabilities for the position of the RDV signal at each time step. Each column in the table corresponds to a time step in the trial. At time zero, the signal is in the zero-bin with probability 1. We use ε instead of zero for all other bins, to account for the small probability of a spurious key press in the experiment (i.e., a subject might incorrectly press a key, submitting their response before a decision is actually made). (B-C) As time progresses, the probability mass of the signal spreads throughout the RDV bins. (D) After some amount of time, the probability of the signal crossing each of the two barriers (shaded cells) becomes larger than ε. The likelihood for a particular trial is given by the probability of crossing the barrier corresponding to the trial's choice at the time step corresponding to the trial's RT.

The computation starts at time zero (first column of the table) with the entire probability of the RDV signal placed in the zero-bin with probability 1 (**Figure**

3.1A). This reflects the assumption that in the DDM the value of the RDV is zero at the beginning of every trial. We use ε (a very small number) in all bins other than the zero-bin to account for the small probability of a spurious key press early in the trial. The columns of the table are then filled out from left to right. Let P_t^i denote the probability of bin i at time step t. For any RDV bin j between the barriers, the likelihood of the signal being in RDV bin j at time $t + 1$ is given by

$$P_{t+1}^j = \sum_i P_t^i \times P_t^{i \to j}, \tag{3.3}$$

where $P_t^{i \to j}$ denotes the probability of transition from i to j. This transition probability can be computed from **Equations 3.1** and **3.2** as follows. The change in the RDV in a single time step is given by $N(\mu, \hat{\sigma})$. The mean of the normal distribution is fixed throughout the trial and can be computed as $\mu = \hat{d}(r_A - r_B)$. The probability of transition $P_t^{i \to j}$ can therefore be approximated as the value of the probability density function $N(\mu, \hat{\sigma})$ for δ, where δ is the difference in mean RDV values between bins j and i.

At every time step, the probability of the RDV reaching each of the two barriers is also calculated. In particular, the likelihood of the model crossing the upper barrier at time $t + 1$ is given by

$$P_{t+1}^{\text{up}} = \sum_i P_t^i \times P_t^{i \to \text{up}}, \tag{3.4}$$

where $P_t^{i \to \text{up}}$ denotes the probability of going from an RDV within bin i to crossing the upper barrier at time step t, and is given by the probability of a draw from $N(\mu, \hat{\sigma})$ that exceeds $1 - i$. P_{t+1}^{down} is defined analogously.

For every unique experimental condition c, the model is simulated by filling out the table for a certain period of time, which should be at least as long as the longest trial in the dataset with that experimental condition. The likelihood for a particular trial is then given by $P_{\text{RT}}^{\text{up}}$, if the individual chose A, and by $P_{\text{RT}}^{\text{down}}$, if the individual chose B (according to the convention for the barriers established in section 3.1), at the time that the observed RT in the trial is reached. The process is illustrated in **Fig. 3.1**, in which the barrier crossing probabilities are depicted as additional top and bottom rows shaded in gray. Once the likelihood for all trials has been computed using this method, the overall likelihood (or, equivalently, the negative log-likelihood) for

each model *m* can be obtained, with the largest likelihood being associated with the optimal parameter values among the models tested, for that particular set of trials.

A layout of the method to compute the probability table for the RDV signal is given in **Algorithm 2**. This algorithm is implemented as a callable function which takes as arguments the fixed model parameters, *mu*, *sigma* and *B*, the number of time steps that defines the number of columns in the probability table, *num_time_steps*, and the size of the RDV state bins to be used in the probability table, *state_step*. In order to obtain the probability of the RDV signal being in each bin at each time step, the function performs a convolution between the probabilities in the previous time step and a Gaussian kernel. After filling out the whole probability table, the function returns two vectors of length *num_time_steps*, which contain the probabilities of the RDV signal crossing the upper and lower barriers over time.

The algorithm works as follows. **Lines 2-3:** the magnitude of the two barriers is determined by *B*. **Line 4:** save the value provided by the caller for *state_step*. **Line 5:** save the number corresponding to half the number of state bins. **Line 6:** correct the size of the *state_step* to make sure all bins are the same size. **Line 7:** the state bins range from the negative barrier to the positive barrier, and the size of each bin is determined by the corrected *state_step*; we make sure each bin is represented by the value at the center of its range. **Line 8:** sample the space of all possible value differences between RDV state bins. **Line 9:** save the number corresponding to half the size of the kernel space; this value will be needed to access the correct portions of the convolution output later. **Line 10:** obtain the kernel, which is a Gaussian probability density function over the kernel sample space. **Line 11:** normalize the kernel to make sure all probabilities add up to 1. **Lines 12-13:** create the probability table. At time zero, the probability of the signal being in any of the state bins is epsilon (to account for spurious key presses), except for the state bin of value zero, which contains the signal with probability one (this is because at the beginning of each trial the RDV is equal to zero). **Lines 14-15:** initialize the probabilities of crossing each barrier throughout the trial with ε. **Line 16:** iterate over the duration of the trial, *num_time_steps*. **Line 17:** convolve the probabilities of the RDV bins from the previous time step with the kernel. **Line 18:** the probabilities for the state bins inside the barriers at the current time step are given by the middle portion of the convolution output. **Line 19:** the probability of crossing the lower barrier at the current time step is given by the sum of the probabilities in the lower portion of the convolution output. **Line 20:** the probability of crossing the upper

barrier at the current time step is given by the sum of the probabilities in the upper portion of the convolution output. **Line 21:** Return arrays *prob_up_crossing* and *prob_low_crossing*.

Algorithm 2 Computation of the probability table.

```
1:  procedure COMPUTE_PROBABILITY_TABLE(mu, sigma, B, num_time_steps, state_step)
2:      up_barrier = B
3:      low_barrier = -B
4:      approx_state_step = state_step
5:      half_num_state_bins = ceiling(B / approx_state_step)
6:      state_step = B / (half_num_state_bins + 0.5)
7:      state_bins = (low_barrier + state_step/2) : state_step : (up_barrier - state_step/2)
8:      kernel_samples = state_step * (low_barrier - (8 * sigma / state_step) : 1 : up_barrier + (8 * sigma / state_step))
9:      half_kernel_size = (length(kernel_samples) - 1) / 2
10:     kernel = normal_pdf(kernel_samples, mu, sigma)
11:     kernel = kernel / sum(kernel)
12:     prob_table = epsilon * ones(length(state_bins), num_time_steps)
13:     prob_table[state_bins == 0, 0] = 1
14:     prob_up_crossing = epsilon * ones(num_time_steps)
15:     prob_low_crossing = epsilon * ones(num_time_steps)
16:     for t = 1 to num_time_steps do
17:         new_prob_states = convolve(prob_table[:, t-1], kernel)
18:         prob_table[:,t] = new_prob_states[half_kernel_size + (1 : length(state_bins))]
19:         prob_low_crossing[t] = sum(new_prob_states[1 : half_kernel_size])
20:         prob_up_crossing[t] = sum(new_prob_states[length(new_prob_states) - half_kernel_size : end])
21:     return prob_up_crossing, prob_low_crossing
```

3.4 Computational Complexity

In this section we present a brief discussion comparing the computational complexity of the PTA versus the traditional MLA for computing the likelihood of the experimental data given a particular set of DDM parameters.

In the MLA, we first need to generate the empirical RT histograms, which corresponds to counting the available data trials, and is therefore $O(M)$, where M is the number of trials in the dataset. Generating the simulations is $O(K \times S \times T)$, where

K is the number of experimental conditions, S is the number of simulations to be generated per experimental condition, and T is the expected number of time steps in a trial. Since there are $K \times S$ simulated trials, generating the simulation RT histograms is $O(K \times S)$. Finally, multiplying the bins in the empirical histograms by the corresponding ones in the simulation histograms is $O(2 \times K \times H)$, where H is the number of bins used in creating the histograms. The computational complexity of this method is therefore $O(M + KTS + KH)$.

In the PTA, we need to fill out one probability table for each experimental condition, and each table contains $T \times C$ cells, where T is the expected number of time steps in a trial, and C is the number of RDV bins used in the table. Therefore, the algorithm's overall computational complexity is $O(K \times T \times C)$, where K is the number of experimental conditions.

The main factor that makes the PTA computationally advantageous when compared to the MLA is that the execution time for the former depends only on the number of experimental conditions present in the dataset (K), whereas the execution time for the latter depends both on the number of data trials available (M), as well as on the number of simulations generated to create the RT histograms ($K \times S$).

3.5 Experiments

Simple DDM with 2 Free Parameters

In order to compare the two algorithms for computing DDM likelihoods, we carried out the following experiments. In Experiment 1, using a simple DDM with 2 free parameters (d and σ), we first generated artificial data by simulating the model with fixed parameter values, $\langle d, \sigma \rangle = \langle 0.01, 0.1 \rangle$. We then attempted to recover these parameters by applying the PTA and the MLA, performing a grid search over the 2-dimensional parameter space. Each point in the grid corresponds to a set of parameters $\langle d, \sigma \rangle$. The heat maps in **Figure 3.2** show the likelihood values computed for different grid sizes and different dataset sizes, comparing our approach to the traditional method. As shown in the figure, our approach can estimate the correct model with a small number of trials, while the traditional method converges on the wrong model when only a small dataset is available for a large grid. Additionally, we can see that the PTA provides smoother likelihood curves over the search space, which makes the algorithm more robust in the optimization process.

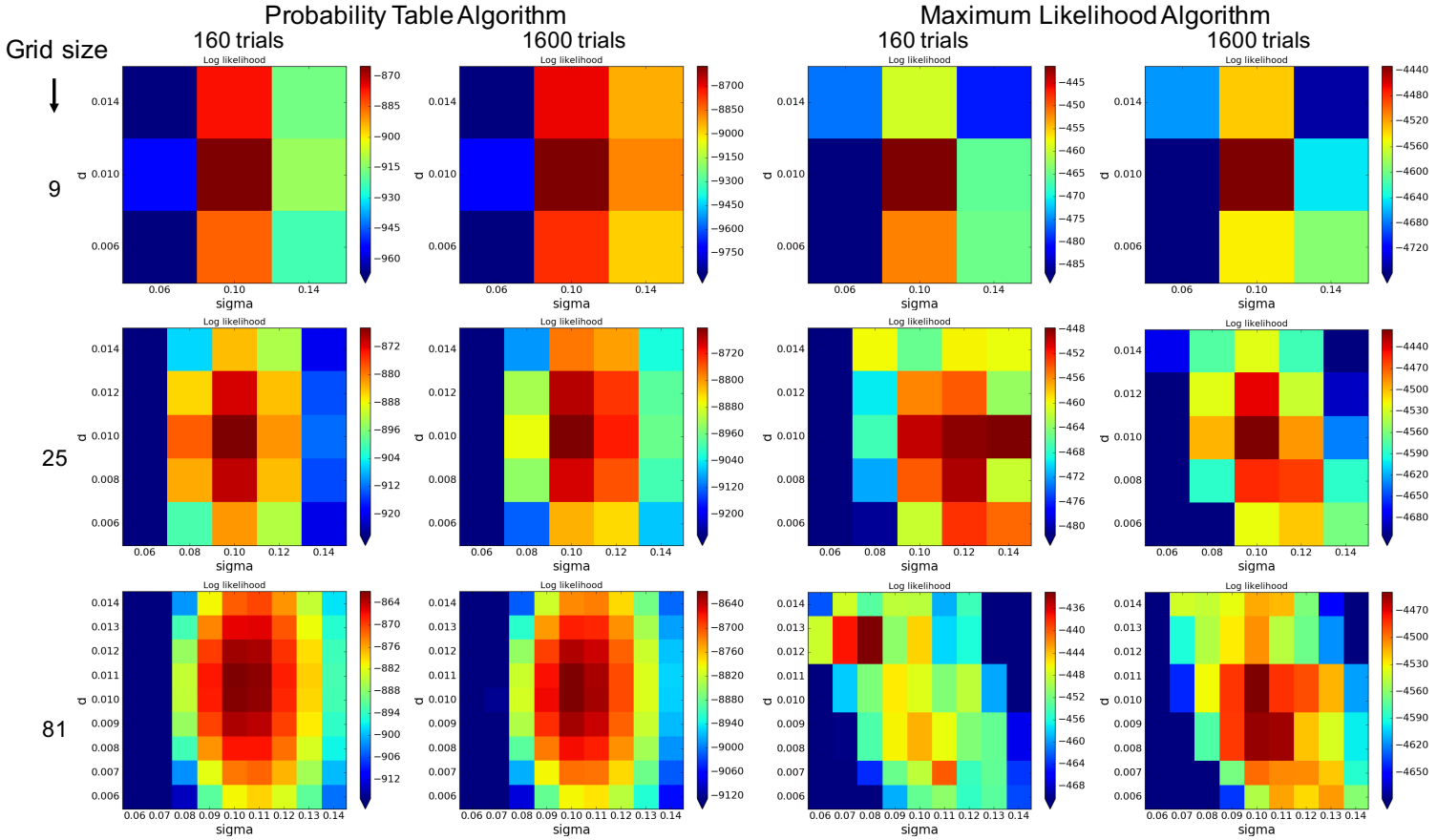

Figure 3.2: Likelihood heat maps. Left column: Probability Table Algorithm; right column: Maximum Likelihood Algorithm. The model used to generate the artificial data was $\langle d, \sigma \rangle = \langle 0.01, 0.1 \rangle$, so when estimated correctly, the maximum likelihood value should be at the center of the grid. Results are shown for three different parameter search grid sizes and for two dataset sizes.

Next, we compared the two algorithms in terms of the error rate obtained over several runs. We used the same artificial data obtained in Experiment 1 using fixed parameter values, $\langle \hat{d}, \hat{\sigma} \rangle$. Then, using each algorithm, we searched for the correct model by running the algorithm 100 times and at each run selecting the model in the grid with the maximum likelihood. The error rate is the fraction of runs that produced an incorrect model as a result. We varied the number of data trials available for estimation, as well as the size of the model search grid. For the MLA, the number of simulations used in the algorithm was kept fixed at 10,000. The results of this experiment are shown in **Figure 3.3A**. For all three grid sizes tested, the PTA consistently required less data trials to produce similar error rates.

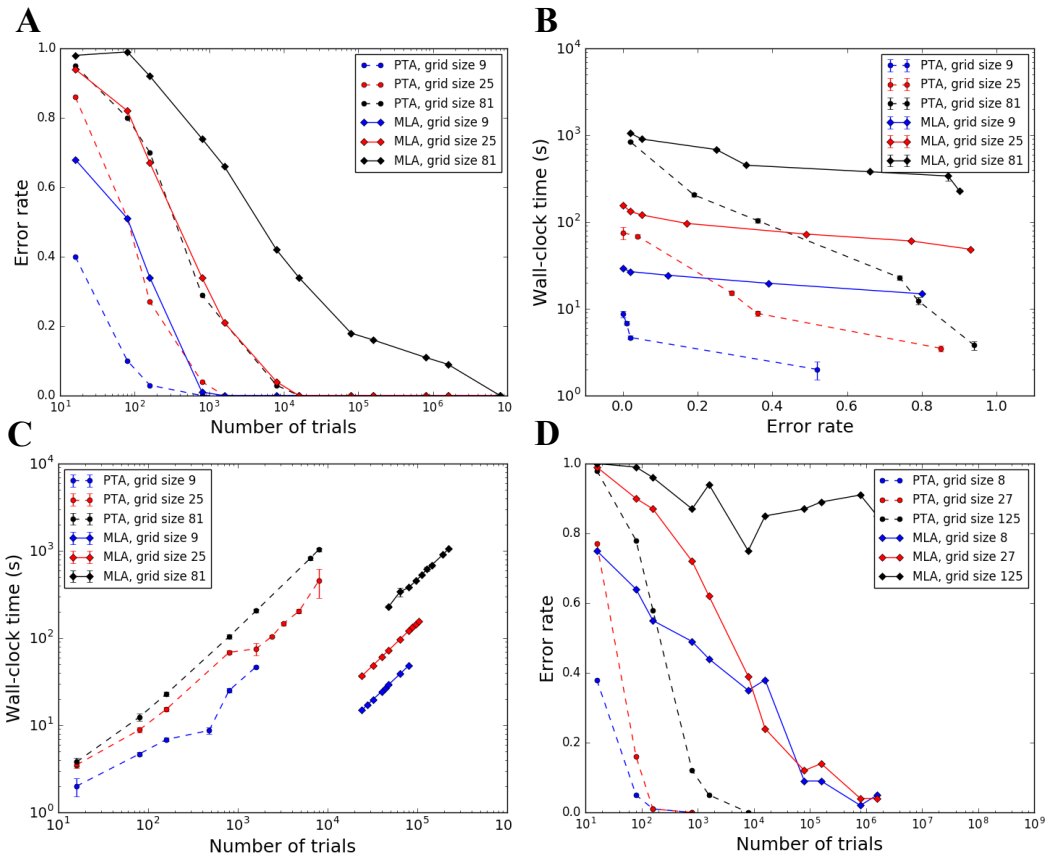

Figure 3.3: Experiments comparing Probability Table and Maximum Likelihood Algorithms. (A) Error rate vs. number of data trials used in the estimation procedure for the DDM. (B) Execution time vs. error rate (DDM). (C) Execution time vs. number of data trials (DDM). (D) Error rate vs. number of data trials used in the estimation procedure for the aDDM. PTA, Probability Table Algorithm; MLA, Maximum Likelihood Algorithm.

Finally, we compared the two algorithms in terms of their mean execution time. We

used the same artificial data from the two previous experiments. For each algorithm and each grid size, we obtained the error rate over 100 runs, as described above, and measured the execution time of each run. **Figure 3.3B** shows the mean and standard deviation of the wall-clock time versus the error rate for each algorithm and each grid size. For all three grid sizes tested, the PTA consistently required less time to produce similar error rates. We also show the increase in wall-clock time as the number of trials used in the estimation procedure increases (**Figure 3.3C**). Note that for a given amount of time the MLA could process a much larger number of data trials than the PTA, but this did not translate into smaller error rates, as indicated in **Figures 3.3A** and **3.3B**.

Attentional DDM with 3 Free Parameters

The classic DDM discussed above can be fitted using a closed-form solution for the likelihood, so an improved method for model fitting is not advantageous in practice. However, for more sophisticated versions of the model, an approximation for the likelihood may be necessary and can make the fitting process much faster. In order to demonstrate the advantages of our method when additional parameters are included in the model, we ran an experiment comparing the two methods (PTA and MLA) in terms of error rate vs. number of trials using a variant of the DDM which contains an additional free parameter. The attentional Drift-Diffusion Model (aDDM) (Krajbich, C. Armel, and Rangel, 2010; Krajbich and Rangel, 2011) includes a parameter θ, ranging between 0 and 1, which controls the attentional bias, discounting the value of the non-fixated item at each time step. Note that the aDDM takes into account subjects' visual fixations during the trials, and because each trial's set of fixations is unique, the number of experimental conditions corresponds to the number of trials in the dataset.

Using the same approach described above, we generated artificial data by simulating the aDDM with fixed parameter values, $\langle \hat{d}, \hat{\sigma}, \hat{\theta} \rangle$, and attempted to recover them using the two different methods. As before, each algorithm was run 100 times and the error rate was calculated over these runs. We varied the number of data trials and the size of the model search grid, while the number of simulations used in the MLA was kept fixed at 10,000. The results of this experiment are shown in **Figure 3.3D**. For all three grid sizes tested, our approach consistently required less data trials to produce similar error rates. In addition, note that the discrepancy in performance between the two methods is even more pronounced than in the case of two-parameter model (**Figure 3.3A**), suggesting that the PTA becomes increasingly

advantageous as the complexity of the model increases.

3.6 Discussion

In this chapter we have presented the Probability Table Algorithm, which can be used for the computation of DDM likelihoods when no closed-form solution is available for the model being fitted. This method provides an exact measure of the probability of the decision barrier being crossed at the end of a trial, and unlike the state-of-the-art Maximum Likelihood Algorithm, it does not require the generation of model simulations. We have shown through several tests that the PTA requires less data and achieves faster execution times when compared to the MLA, and that it becomes increasingly advantageous as the complexity of the model being tested increases. Furthermore, while we have focused our comparison on the MLA, since it is the most widely used method, we predict that similar advantages would be obtained over other simulation-based algorithms, such as chi-square and weighted least squares, which depend on producing RT histograms in the same way as the MLA.

One potential concern with using the PTA described here is when dealing with models that include variables that are not directly involved in the decision process, such as non-decision time. One solution to this problem would be to include the non-decision variable as an extra free parameter in the model, to be fitted in the same way as the other free parameters, as previously described. However, it may be desirable to allow this variable to vary across trials in order to better fit the experimental data. In this case, an alternative is to use a probability distribution, parameterized by a small number of hyper-parameters, to describe the non-decision variable, and to include these hyper-parameters as additional free parameters in the fitting process. The disadvantage in this case is that the PTA will become more computationally expensive as the number of free parameters in the model increases.

The approach we have described does not make any assumptions about the nature of the two-alternative force-choice task used to collect the data, and can therefore be used with a variety of experimental setups. It would also be a straightforward modification to adapt the algorithm for more than two choices presented in the task. Finally, while we have only provided the algorithmic implementation for a simple two-parameter version of the DDM, the implementation of the algorithm can be easily extended to account for additional free parameters, including collapsing barriers, initial bias, as well parameters aiming to capture specific experimental

conditions, such as the attentional bias in the aDDM (Krajbich, C. Armel, and Rangel, 2010; Krajbich and Rangel, 2011; Tavares, Perona, and Rangel, 2017).

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.

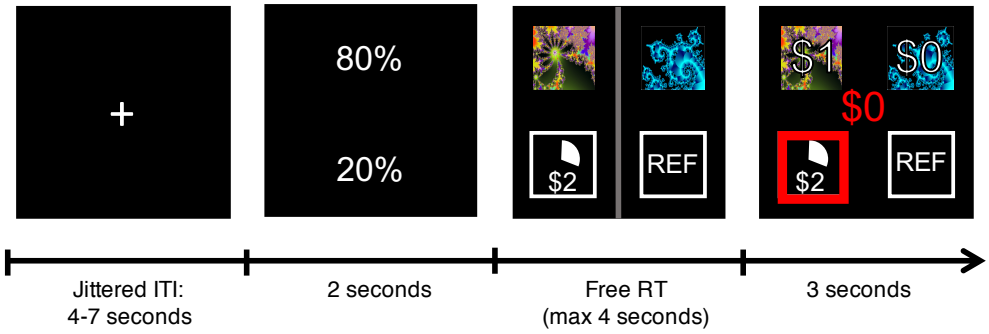

Figure 4.1: Task structure.

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),

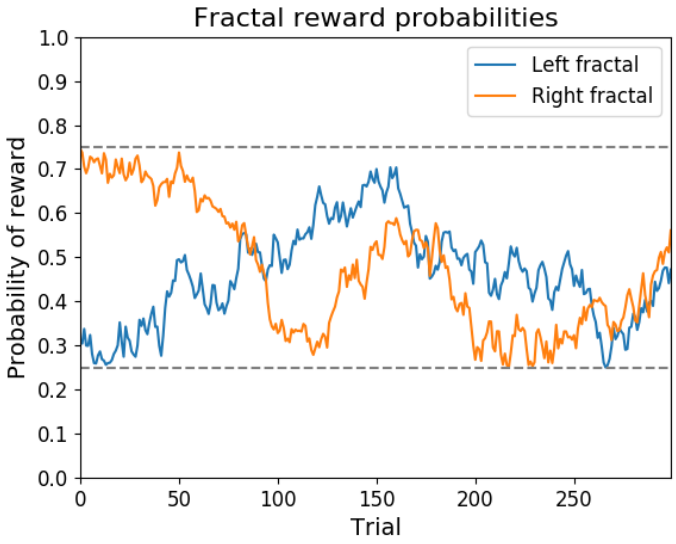

Figure 4.2: Example of fractal reward probabilities. The probabilities of receiving a reward associated with the fractals drifted slowly and independently over time between 0.25 and 0.75, according to a Gaussian random walk with $\sigma = 0.025$.

(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**.

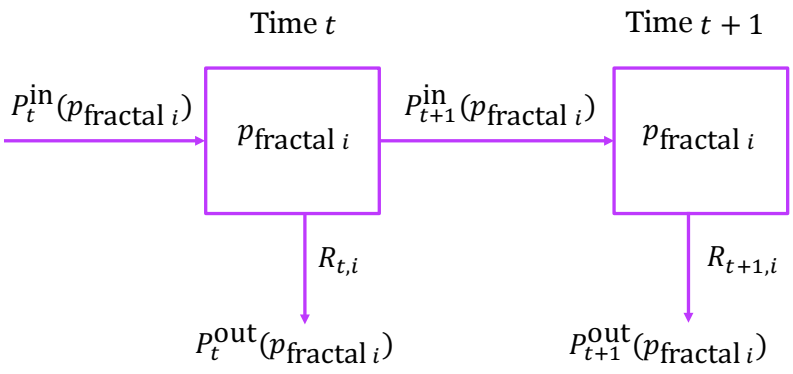

Figure 4.3: Diagram of the Bayesian update model for the fractal probabilities. $P_t^{\text{out}}(p_{\text{fractal }i})$ is the posterior distribution at time t for the probability associated with fractal i, where i is either left or right. $P_t^{\text{in}}(p_{\text{fractal }i})$ is the prior distribution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

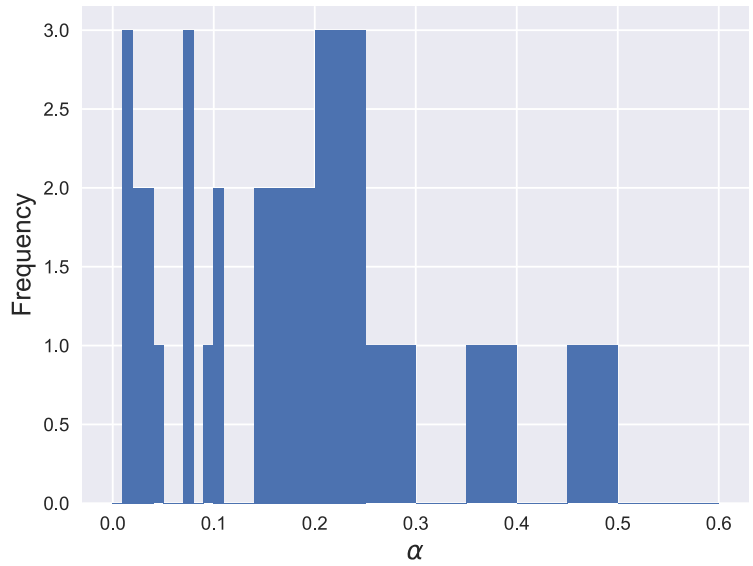

Figure 4.4: Histogram of the learning rate α used in the Q-learning model, fitted per subject.

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

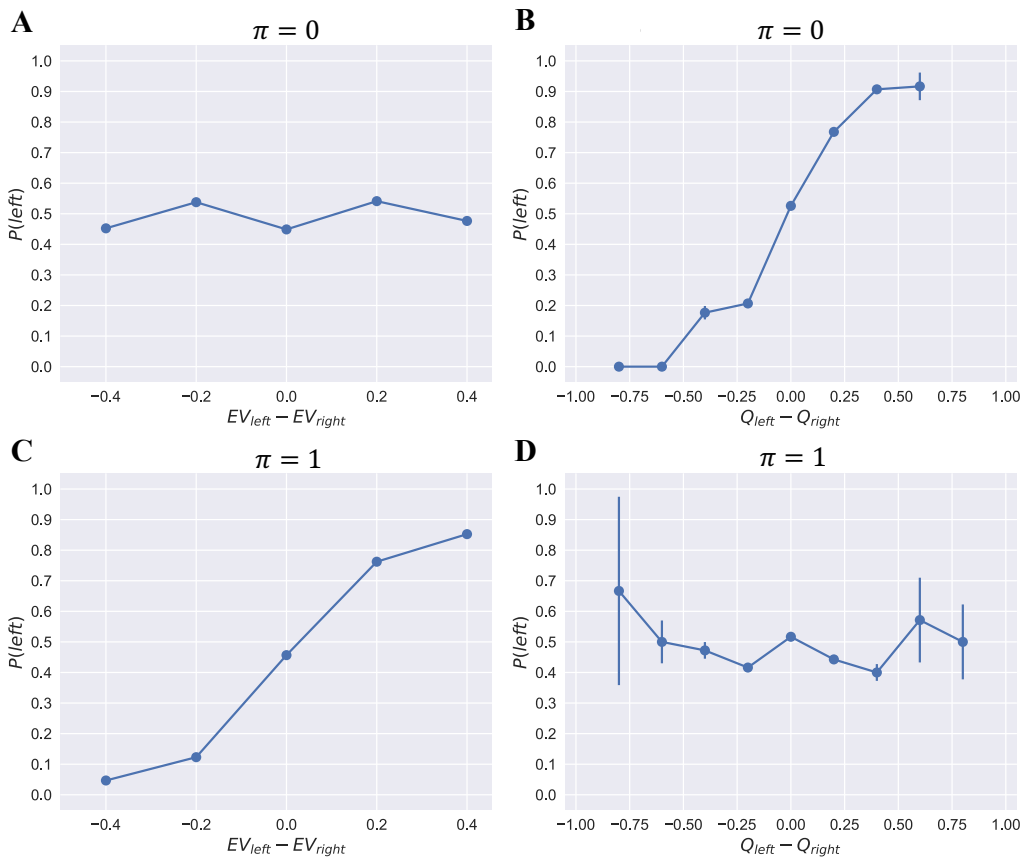

Figure 4.5: Psychometric choice curves as a function of the lottery expected value difference (A and C) and the fractal Q value difference (B and D). The top row shows only trials where $\pi = 0$ (A and B), whereas the bottom row shows only trials where $\pi = 1$ (C and D). Data was aggregated from all 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 $\times\lvert EV_L - EV_R\rvert$	-0.40933	0.14183	-2.886
$\pi = 1$ indicator $\times\lvert Q_L - Q_R\rvert$	-0.05352	0.14869	-0.360
$\pi = 0$ indicator $\times\lvert EV_L - EV_R\rvert$	0.11285	0.14630	0.771
$\pi = 0$ indicator $\times\lvert Q_L - Q_R\rvert$	-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

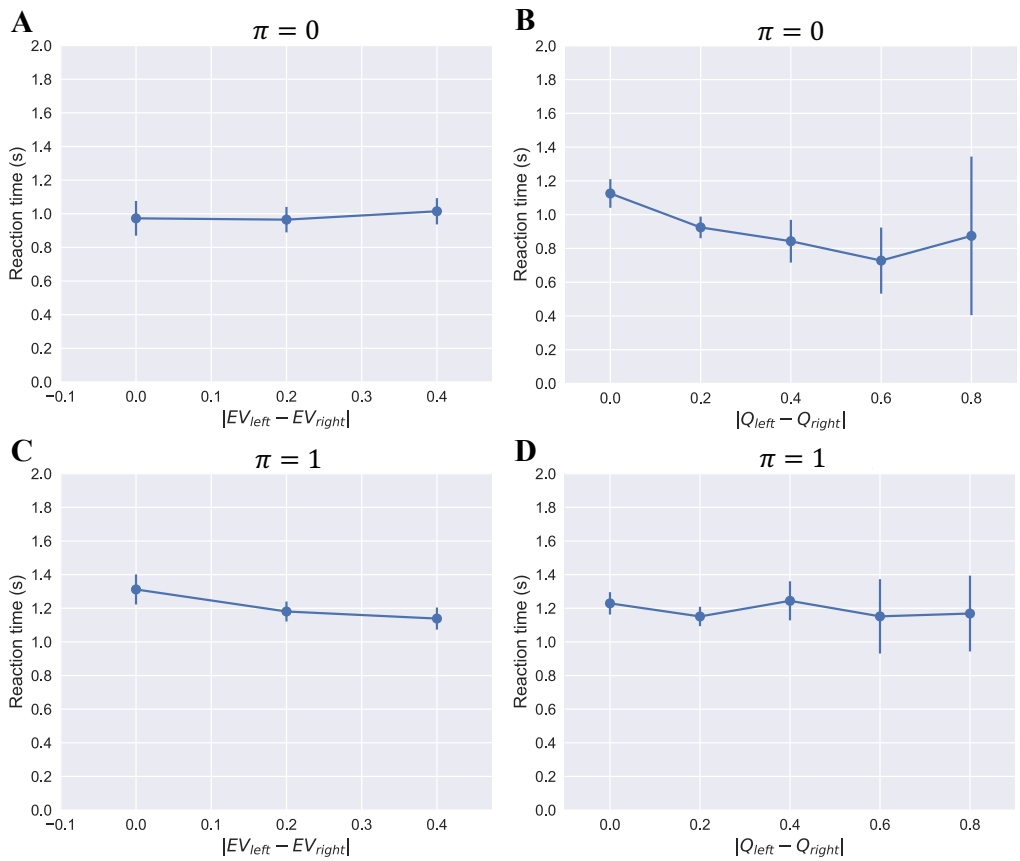

Figure 4.6: Response times as a function of the absolute lottery expected value difference (A and C) and the absolute fractal Q value difference (B and D). The top row shows only trials where $\pi = 0$ (A and B), whereas the bottom row shows only trials where $\pi = 1$ (C and D). Data was aggregated 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.

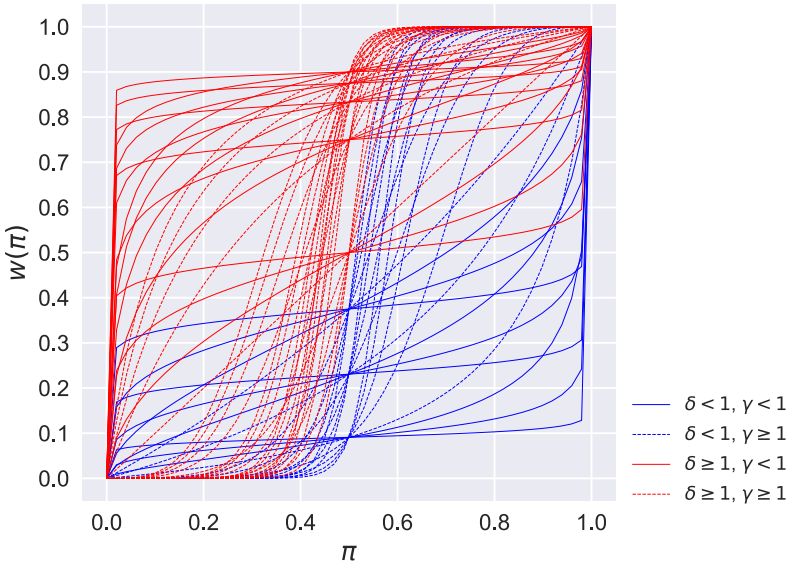

Figure 4.7: Different curves of the two-parameter weighting function. Curves were obtained with varying values of γ and δ, which were taken from the set $\{0.1, 0.3, 0.6, 1, 3, 5, 7, 9\}$. Curves are displayed in four groups according to whether or not γ is less than 1, and to whether or not δ is less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

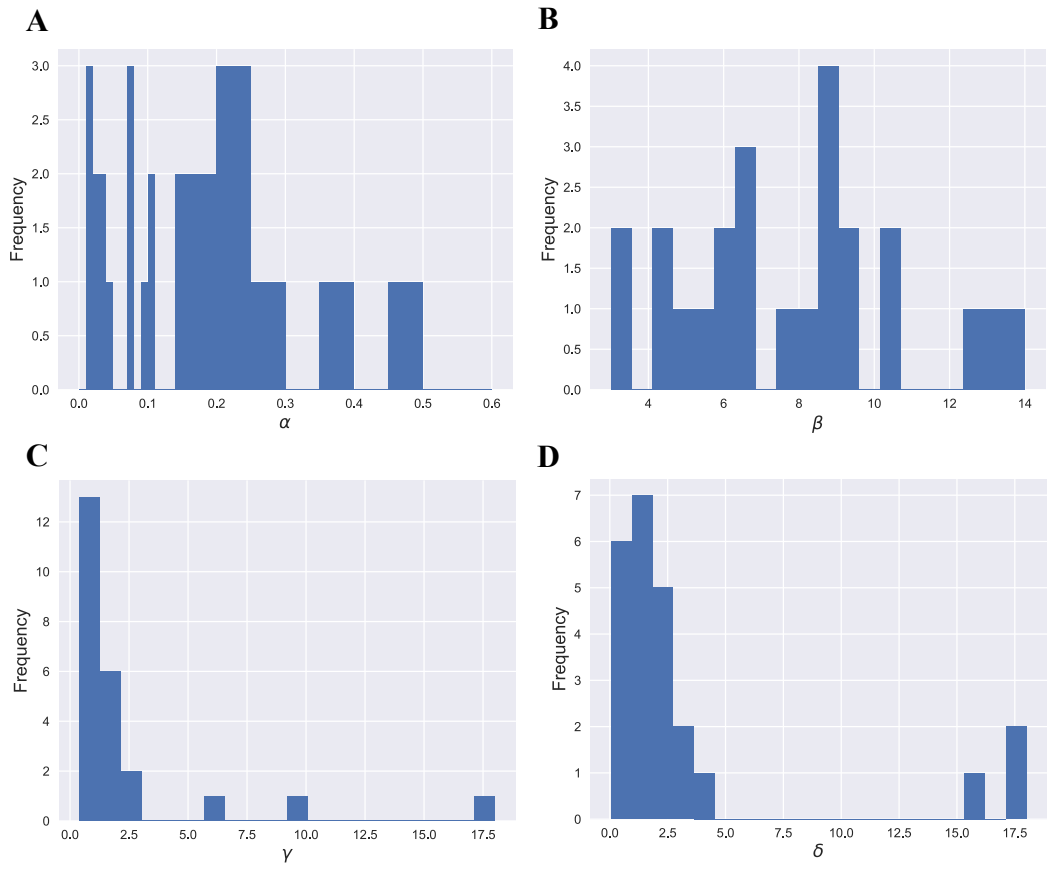

Figure 4.8: Histograms for the non-linear model parameters, fitted per subject. (A) α. (B) β. (C) γ. (D) δ.

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.

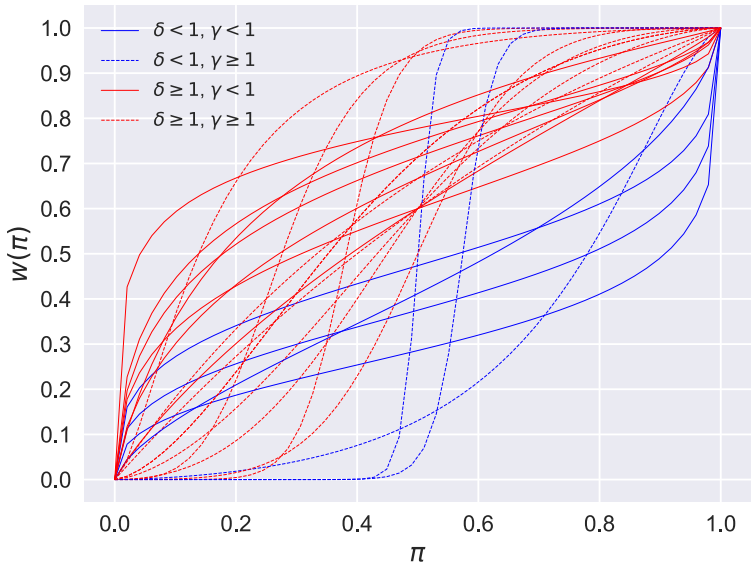

Figure 4.9: Curves obtained for the two-parameter weighting function for each subject. Curves are grouped into 4 groups according to whether or not γ is less than 1, and to whether or not δ is less than 1.

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.

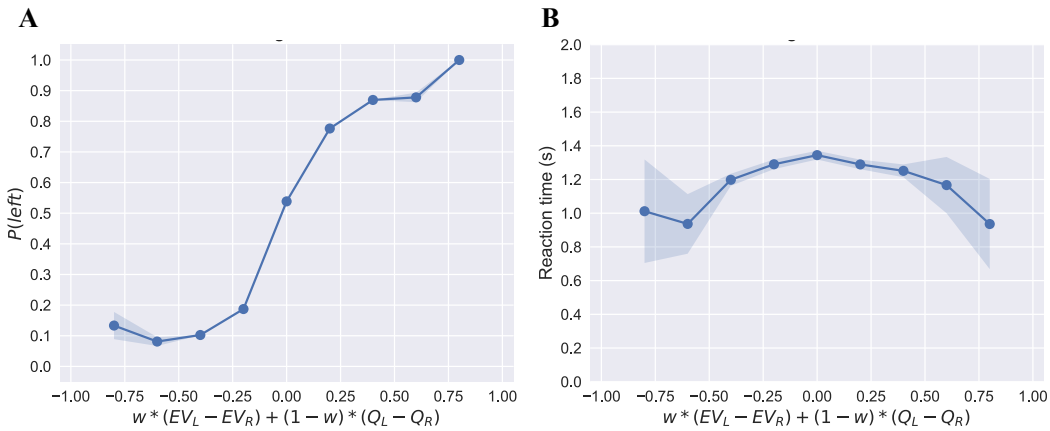

Figure 4.10: Basic psychometrics. (A) Choice curve as a function of value difference. (B) Response time as a function of value difference. Data was aggregated from all subjects. Shaded error bars show 95% confidence intervals for the data pooled across all subjects.

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.

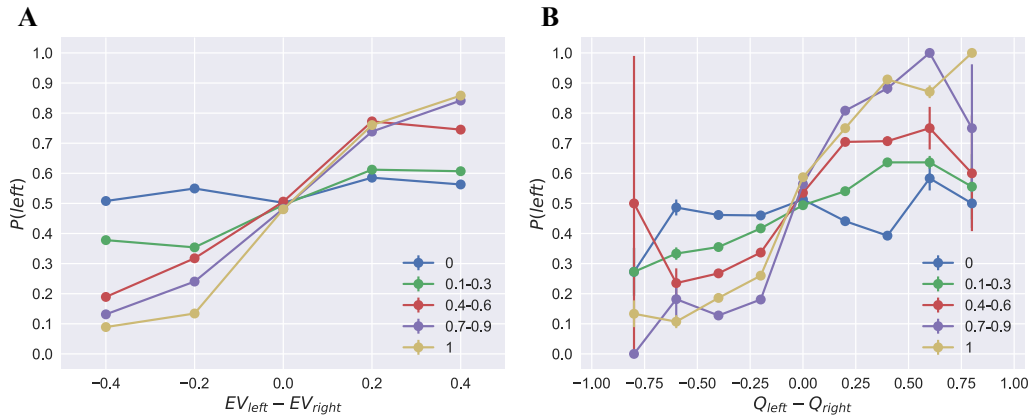

Figure 4.11: Choice curves grouped by weights. (A) Choice curves as a function of lottery expected value difference, grouped by w. (B) Choice curves as a function of fractal Q value difference, grouped by $1 - w$. Data was aggregated from all subjects. Error bars show 95% confidence intervals for the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

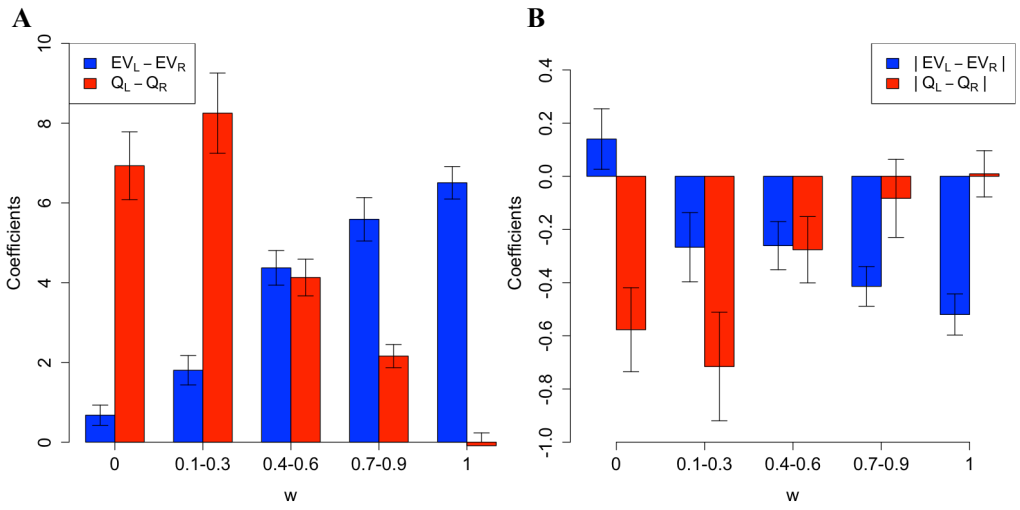

Figure 4.12: Coefficients of the mixed effects models, with trials grouped by w. (A) Coefficients for choice logit mixed effects model. (B) Coefficients for RT mixed effects model.

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 $|EV_{\text{left}} - EV_{\text{right}}| < |Q_{\text{left}} - Q_{\text{right}}|$).

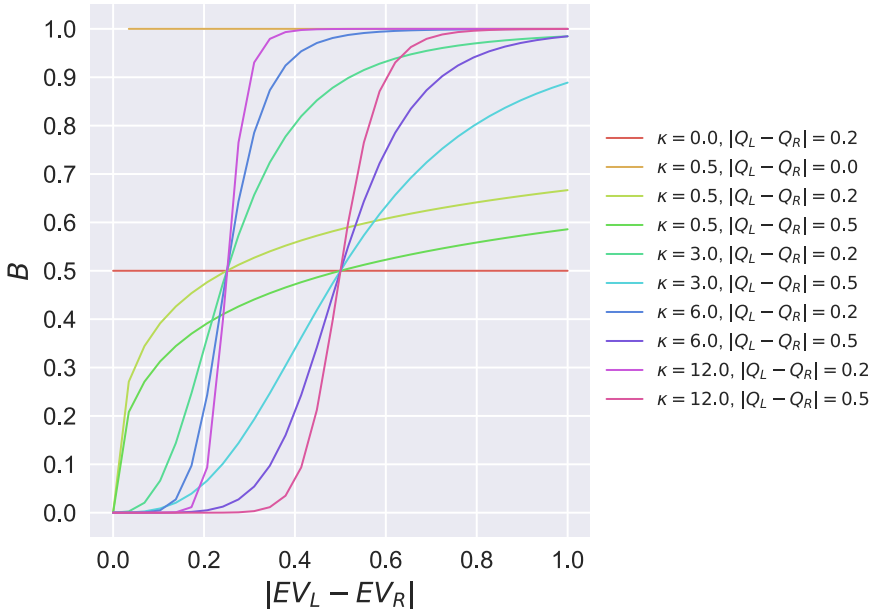

Figure 4.13: Examples of the weight adjustment curve B. Curves were obtained as a function of the absolute lottery expected value difference, for different values of κ and different values of the absolute fractal Q value difference.

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**.

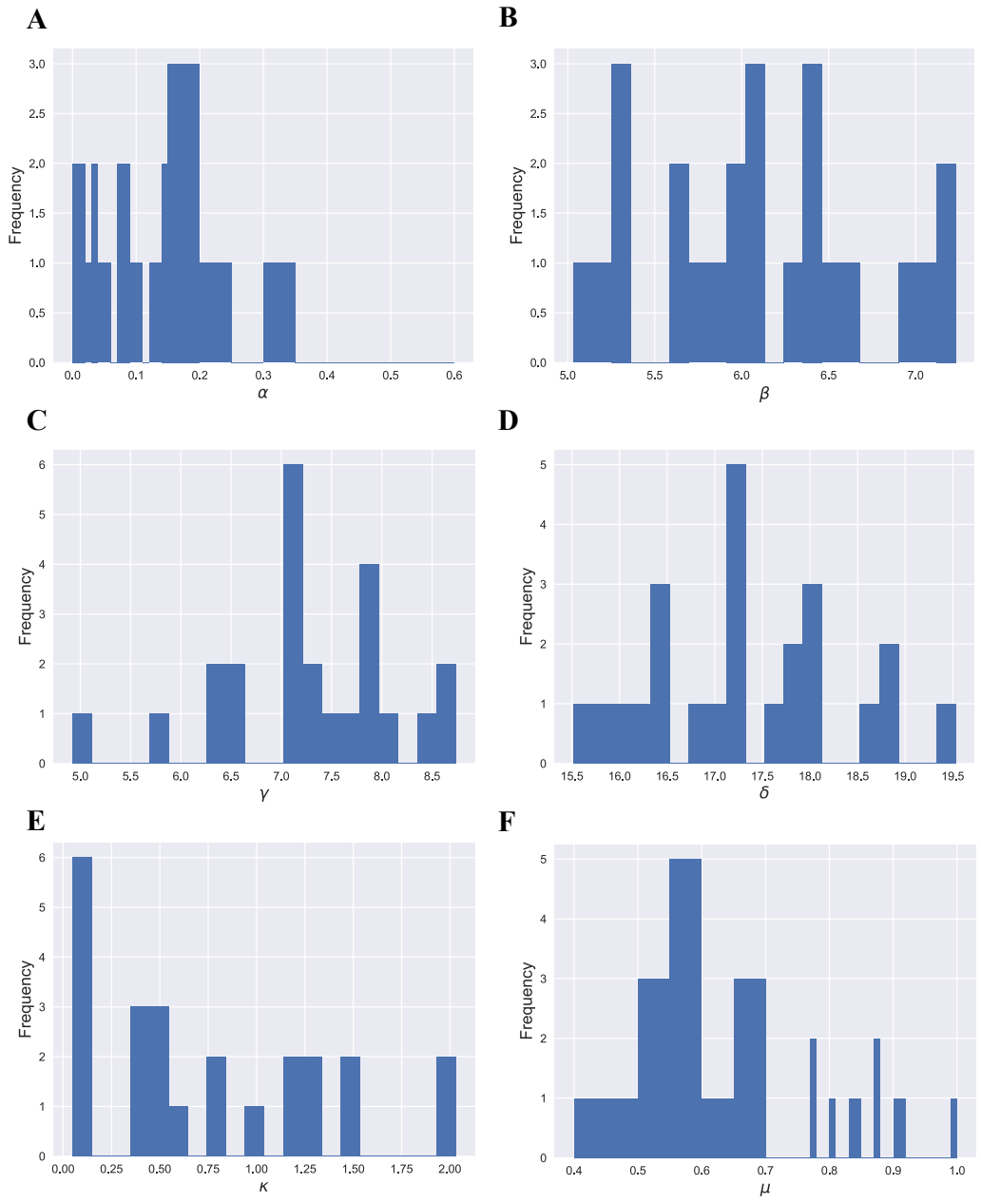

Figure 4.14: Histograms for the nested model parameters, fitted per subject. (A) α. (B) β. (C) γ. (D) δ. (E) κ. (F) μ.

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 ($p = 1.18\times10^{-5}$), 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.

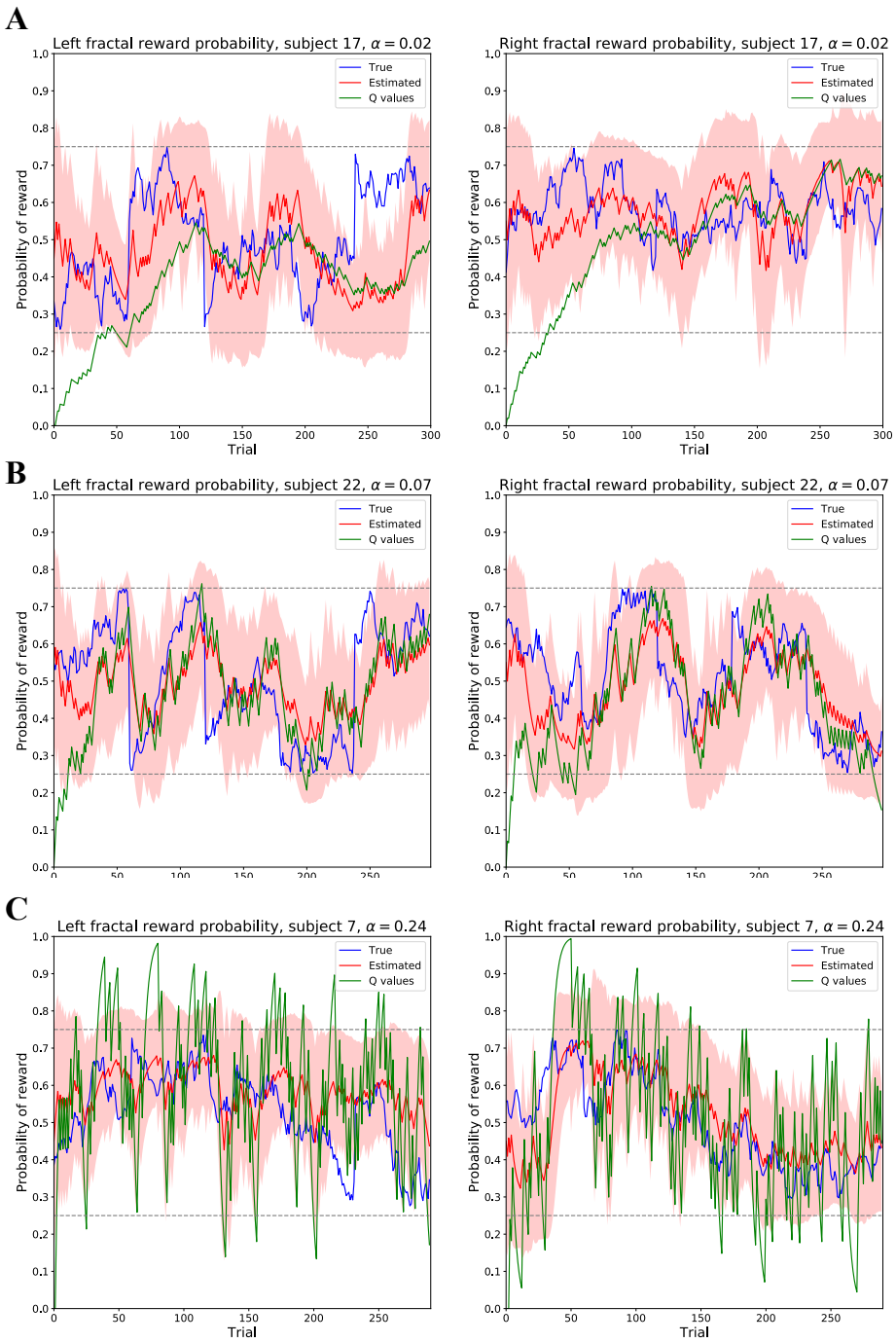

Figure 4.15: Bayesian update model estimates. The mean and 95% confidence interval for the Bayesian posteriors (in red) are shown for 3 different subjects. Also displayed are the true values of the fractal probabilities (in blue) and the Q values obtained from the Q-learning model (in green). (A) Example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

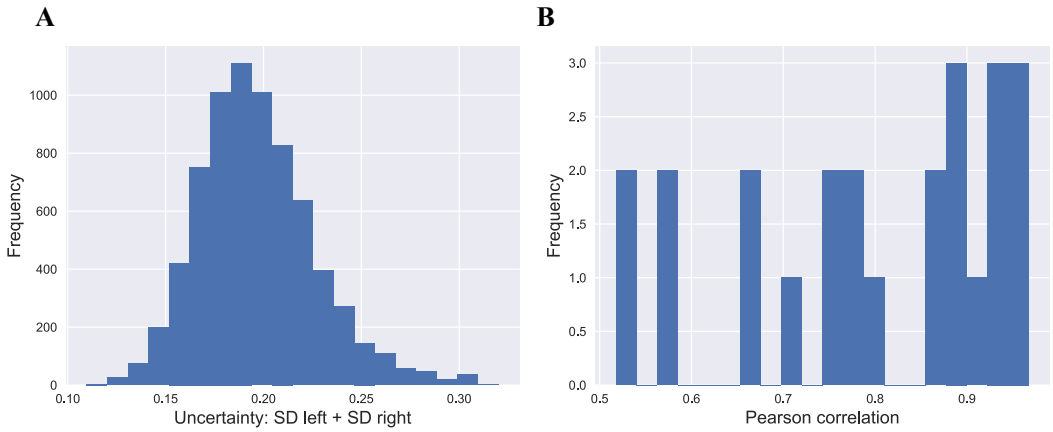

Figure 4.16: Validation of the Bayesian measure of uncertainty. (A) Histogram of the uncertainty measure, which corresponds to the sum of standard deviations from the posterior distributions obtained for left and right fractal probabilities, across all trials from all subjects. (B) Histogram of the Pearson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

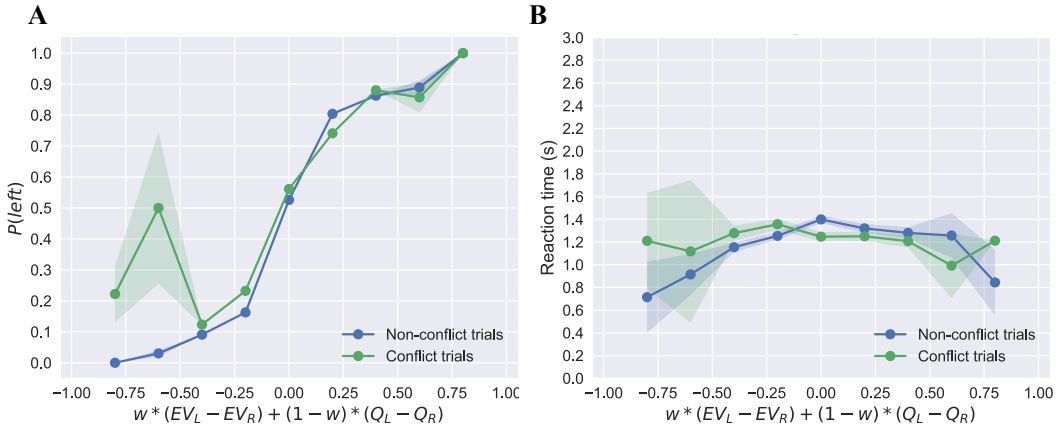

Figure 4.17: Basic psychometrics, conflict vs. no-conflict trials. (A) Choice curve as a function of total value difference, with trials grouped by whether or not there was conflict between the two systems. (B) Choice curve as a function of total value difference, with trials grouped by whether or not 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

providing an approximation of the model likelihood. The second, which we call the Probability Table Algorithm (PTA), requires Markov-chain type computations over the duration of each trial, leading to a precise likelihood value without the need for model simulations. We performed several experiments to test the ability of the two algorithms to obtain good likelihood estimates for artificial datasets, and compared them in terms of error rates, number of data trials used in the fitting procedure and total execution time. We found that using the PTA offered several advantages over the MLA, including faster computation times and a smaller number of data trials required in order to produce similar error rates. These advantages were especially pronounced when the number of free parameters in the model grew larger and across-trial variability, in the form of visual fixations, were present in the data.

In Chapter 4, I described an experiment aimed at addressing an important open question in decision neuroscience: when multiple sources of information, possibly acquired through different processes, must be integrated to generate a single decision, how does the brain modulate and combine their contributions in order to produce a single choice? Using a paradigm which required subjects to simultaneously consider values acquired through previous experience (experiential) as well as values calculated using information fully described at the time of choice (descriptive), we investigated the interactions between the mechanisms that compute descriptive and experiential values. We compared different behavioral models that combined the two types of values, and showed that a non-linear weighting model of the exogenous relevance of each system can successfully make predictions about subjects' choice behavior, as well as capture different levels of biases across subjects towards one system or another. We found an effect of the conflict between the two systems on choice, a small influence of the relative strength of preference within each system, and no effect of a measure of the uncertainty in the experiential system.

5.2 Future Directions

Collectively, the results presented in this dissertation expose a number of open questions in computational neuroscience and point to several directions for future research. Regarding the use of sequential integrator models in human decision making, for instance, further tests of the validity of the model are necessary. Our task focused on simple perceptual choices between two items, but many modifications could be tested to further validate our results. One potential direction would be to use the aDDM to model the effect of attention on other types of choices, such as tasks involving inter-temporal or moral decisions, or to use a larger number of

options in each trial. More generally, it is important to obtain further evidence, particularly from neural data, that this kind of accumulator model provides a good characterization of the underlying neural computations that take place in the brain during the course of a decision. One way to do this would be to test the aDDM with single-unit recordings data from a perceptual task, such as the random dot motion task, in non-human primates.

Our eye-tracking study used the aDDM to investigate the role of overt attention in perceptual decision making, as measured by visual fixations. An interesting modification of this experiment would be to use covert attention (for instance, by asking subjects to maintain a central fixation while attending to the choice options peripherally), which would give us an opportunity to check if foveation is a necessary condition for the attentional bias to take place, further elucidating the attentional process. Another useful modification would be to use a larger number of items to choose from in each trial. This would likely lead to very different patterns of fixations, and it would be interesting to investigate how these patterns may affect choices, and whether the aDDM could be adapted to account for these different fixation dynamics.

While our study focused on the impact of visual fixations on perceptual decisions, we did not address the opposite direction of this interaction, i.e., the way in which the decision process may guide a subject's fixations. The impact of the decision variable on attention and and on the fixation patterns that emerge is an essential open question in the study of decisions involving visual fixations. Understanding this interaction may lead to models that can predict fixation patterns and even potentially generate artificial sequences of fixations mimicking that of human subjects.

Our conclusion from Chapter 2 that the same computational model that explains the role of attention in economic choice can also be used to model perceptual choices leads to the question of how much overlap there is between the mechanisms involved in economic and perceptual decisions, and whether attention has a similar role in these two types of decision. Defining the precise neural implementations of these processes and establishing how much of the neural circuitry is shared between them through the use of neural data are very important topics for future study.

Another key open question in computational neuroscience is that of allocation of control between decision systems. A large body of research has provided evidence for multiple learning and decision mechanisms in the human brain. In Chapter 4 we investigated the interaction between descriptive and experiential valuation systems.

Our behavioral results require further validation from the associated neuroimaging data, such as evidence supporting the existence of the non-linear weighting of the two types of values, and evidence for individual tendencies to overweight one system over the other. Moreover, the collected fMRI data will be useful in further defining a more precise role for conflict, uncertainty and relative strength of preference on the arbitration between the experiential and descriptive valuation systems.

It would be interesting to test the DDM on the data from our descriptive-experiential paradigm, and to evaluate whether this model can be used to make predictions about response times as well as choices in our task. This would provide further validation of the non-linear weighting valuation model we described. Additionally, applying the same paradigm in an eye-tracking experiment and using an aDDM to incorporate fixations into the model would allow us to test the impact of attention on the arbitration between valuation systems. From the results presented in this dissertation, we hypothesize that a correlation may exist between stronger overt attention in the form of fixations towards one system and a larger choice bias for that same system, which would also be reflected in the parameters of an aDDM fitted to the data.

The study of dual systems of valuation and decision making has been a popular topic in recent decision making literature, with researchers defining dichotomies such as model-based versus model-free, and goal-directed versus habitual decision systems. Further investigation, particularly through causal experiments, is required to pin down the exact nature of the different decision mechanisms, how they interact and what kind of factors may impact them. Advances in the study of these questions have the potential to affect our current understanding and treatment strategies of several clinical conditions related to poor judgment and faulty decision making, such as addiction, eating disorders, and obsessive-compulsive disorder.

BIBLIOGRAPHY

Abdellaoui, Mohammed, Han Bleichrodt, and Corina Paraschiv (2007). "Loss aversion under prospect theory: A parameter-free measurement". In: *Management Science* 53.10, pp. 1659–1674.

Adams, Christopher D and Anthony Dickinson (1981). "Instrumental responding following reinforcer devaluation". In: *The Quarterly journal of experimental psychology* 33.2, pp. 109–121.

Aharon, Itzhak et al. (2001). "Beautiful faces have variable reward value: fMRI and behavioral evidence". In: *Neuron* 32.3, pp. 537–551.

Akam, Thomas, Rui Costa, and Peter Dayan (2015). "Simple plans or sophisticated habits? State, transition and learning interactions in the two-step task". In: *PLoS computational biology* 11.12, e1004648.

Anderson, Adam K et al. (2003). "Dissociated neural representations of intensity and valence in human olfaction". In: *Nature neuroscience* 6.2, pp. 196–202.

Apkarian, A Vania et al. (2004). "Chronic pain patients are impaired on an emotional decision-making task". In: *Pain* 108.1, pp. 129–136.

Armel, K Carrie, Aurelie Beaumel, and Antonio Rangel (2008). "Biasing simple choices by manipulating relative visual attention". In: *Judgment and Decision Making* 3.5, p. 396.

Ayton, Peter and Ilan Fischer (2004). "The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness?" In: *Memory & cognition* 32.8, pp. 1369–1378.

Balleine, Bernard W and John P O'doherty (2010). "Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action". In: *Neuropsychopharmacology* 35.1, pp. 48–69.

Barron, Greg and Ido Erev (2003). "Small feedback-based decisions and their limited correspondence to description-based decisions". In: *Journal of Behavioral Decision Making* 16.3, pp. 215–233.

Basten, Ulrike et al. (2010). "How the brain integrates costs and benefits during decision making". In: *Proceedings of the National Academy of Sciences* 107.50, pp. 21767–21772.

Bechara, Antoine (2005). "Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective". In: *Nature neuroscience* 8.11, pp. 1458–1463.

Bechara, Antoine and Antonio R Damasio (2005). "The somatic marker hypothesis: A neural theory of economic decision". In: *Games and economic behavior* 52.2, pp. 336–372.

Bechara, Antoine and Hanna Damasio (2002). "Decision-making and addiction (part I): impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences". In: *Neuropsychologia* 40.10, pp. 1675–1689.

Bechara, Antoine, Hanna Damasio, et al. (1999). "Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making". In: *Journal of Neuroscience* 19.13, pp. 5473–5481.

Beers, Robert J van, Anne C Sittig, and Jan J Denier van Der Gon (1999). "Integration of proprioceptive and visual position-information: An experimentally supported model". In: *Journal of neurophysiology* 81.3, pp. 1355–1364.

Beierholm, Ulrik R et al. (2011). "Separate encoding of model-based and model-free valuations in the human brain". In: *Neuroimage* 58.3, pp. 955–962.

Bennur, Sharath and Joshua I Gold (2011). "Distinct representations of a perceptual decision and the associated oculomotor plan in the monkey lateral intraparietal area". In: *Journal of Neuroscience* 31.3, pp. 913–921.

Bichot, Narcisse P and Jeffrey D Schall (1999). "Effects of similarity and history on neural mechanisms of visual selection." In: *Nature neuroscience* 2.6.

Blood, Anne J et al. (1999). "Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions". In: *Nature neuroscience* 2.4, pp. 382–387.

Bode, Stefan et al. (2012). "Predicting perceptual decision biases from early brain activity". In: *Journal of Neuroscience* 32.36, pp. 12488–12498.

Bogacz, Rafal (2007). "Optimal decision-making theories: linking neurobiology with behaviour". In: *Trends in cognitive sciences* 11.3, pp. 118–125.

Bogacz, Rafal, Eric Brown, et al. (2006). "The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks." In: *Psychological review* 113.4, p. 700.

Bogacz, Rafal, Samuel M McClure, et al. (2007). "Short-term memory traces for action bias in human reinforcement learning". In: *Brain research* 1153, pp. 111–121.

Bowman, Nicholas E, Konrad P Kording, and Jay A Gottfried (2012). "Temporal integration of olfactory perceptual evidence in human orbitofrontal cortex". In: *Neuron* 75.5, pp. 916–927.

Brand, Matthias et al. (2004). "Decision-making impairments in patients with Parkinson's disease". In: *Behavioural neurology* 15.3-4, pp. 77–85.

Britten, Kenneth H et al. (1992). "The analysis of visual motion: a comparison of neuronal and psychophysical performance". In: *Journal of Neuroscience* 12.12, pp. 4745–4765.

Brunton, Bingni W, Matthew M Botvinick, and Carlos D Brody (2013). "Rats and humans can optimally accumulate evidence for decision-making". In: *Science* 340.6128, pp. 95–98.

Busemeyer, Jerome R and Joseph G Johnson (2004). "Computational models of decision making". In: *Blackwell handbook of judgment and decision making*, pp. 133–154.

Busemeyer, Jerome R and James T Townsend (1993). "Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment." In: *Psychological review* 100.3, p. 432.

Camerer, Colin and Teck Hua Ho (1999). "Experience-weighted attraction learning in normal form games". In: *Econometrica* 67.4, pp. 827–874.

Carrasco, Marisa (2011). "Visual attention: The past 25 years". In: *Vision research* 51.13, pp. 1484–1525.

Cavedini, Paolo et al. (2002). "Decision-making heterogeneity in obsessive-compulsive disorder: ventromedial prefrontal cortex function predicts different treatment outcomes". In: *Neuropsychologia* 40.2, pp. 205–211.

Celebrini, Simona and William T Newsome (1994). "Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey". In: *Journal of Neuroscience* 14.7, pp. 4109–4124.

– (1995). "Microstimulation of extrastriate area MST influences performance on a direction discrimination task". In: *Journal of Neurophysiology* 73.2, pp. 437–448.

Chib, Vikram S et al. (2009). "Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex". In: *Journal of Neuroscience* 29.39, pp. 12315–12320.

Churchland, Anne K, Roozbeh Kiani, and Michael N Shadlen (2008). "Decision-making with multiple alternatives". In: *Nature neuroscience* 11.6, pp. 693–702.

Croson, Rachel and James Sundali (2005). "The gambler's fallacy and the hot hand: Empirical data from casinos". In: *Journal of risk and uncertainty* 30.3, pp. 195–209.

Cushman, Fiery and Adam Morris (2015). "Habitual control of goal selection in humans". In: *Proceedings of the National Academy of Sciences* 112.45, pp. 13817–13822.

Daw, Nathaniel D, Samuel J Gershman, et al. (2011). "Model-based influences on humans' choices and striatal prediction errors". In: *Neuron* 69.6, pp. 1204–1215.

Daw, Nathaniel D, Yael Niv, and Peter Dayan (2005). "Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control". In: *Nature neuroscience* 8.12, pp. 1704–1711.

Daw, Nathaniel D, John P O'doherty, et al. (2006). "Cortical substrates for exploratory decisions in humans". In: *Nature* 441.7095, pp. 876–879.

Dawling, Pam et al. (2011). “Search dynamics in consumer choice under time pressure: An eye-tracking study”. In: *The American Economic Review* 101.2, pp. 900–926.

Deco, Gustavo, Edmund T Rolls, and Ranulfo Romo (2010). “Synaptic dynamics and decision making”. In: *Proceedings of the National Academy of Sciences* 107.16, pp. 7545–7549.

Dickinson, Anthony (1985). “Actions and habits: the development of behavioural autonomy”. In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 308.1135, pp. 67–78.

Dickinson, Anthony and Bernard Balleine (1994). “Motivational control of goal-directed action”. In: *Animal Learning & Behavior* 22.1, pp. 1–18.

Dickinson, A et al. (1995). “Overtraining and the motivational control of instrumental action”. In: *Anim Learn Behav* 22, pp. 197–206.

Diederich, Adele and Jerome R Busemeyer (2003). “Simple matrix methods for analyzing diffusion models of choice probability, choice response time, and simple response time”. In: *Journal of Mathematical Psychology* 47.3, pp. 304–322.

Ditterich, Jochen (2006). “Stochastic models of decisions about motion direction: behavior and physiology”. In: *Neural Networks* 19.8, pp. 981–1012.

Ditterich, Jochen, Mark E Mazurek, and Michael N Shadlen (2003). “Microstimulation of visual cortex affects the speed of perceptual decisions”. In: *Nature neuroscience* 6.8, pp. 891–898.

Dolan, Ray J and Peter Dayan (2013). “Goals and habits in the brain”. In: *Neuron* 80.2, pp. 312–325.

Doll, Bradley B et al. (2015). “Model-based choices involve prospective neural activity”. In: *Nature neuroscience* 18.5, pp. 767–772.

Doya, Kenji et al. (2002). “Multiple model-based reinforcement learning”. In: *Neural computation* 14.6, pp. 1347–1369.

Economides, Marcos et al. (2015). “Arbitration between controlled and impulsive choices”. In: *NeuroImage* 109, pp. 206–216.

Elliott, Rebecca et al. (2003). “Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study”. In: *Journal of Neuroscience* 23.1, pp. 303–307.

Erlich, Jeffrey C et al. (2015). “Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat”. In: *Elife* 4, e05457.

Ernst, Marc O and Martin S Banks (2002). “Humans integrate visual and haptic information in a statistically optimal fashion”. In: *Nature* 415.6870, pp. 429–433.

Fischer, Jason and David Whitney (2014). “Serial dependence in visual perception”. In: *Nature neuroscience* 17.5, pp. 738–743.

FitzGerald, Thomas HB et al. (2010). “Differentiable neural substrates for learned and described value and risk”. In: *Current Biology* 20.20, pp. 1823–1829.

Foerde, Karin and Daphna Shohamy (2011). “The role of the basal ganglia in learning and memory: insight from Parkinson’s disease”. In: *Neurobiology of learning and memory* 96.4, pp. 624–636.

Foerde, Karin, Joanna E Steinglass, et al. (2015). “Neural mechanisms supporting maladaptive food choices in anorexia nervosa”. In: *Nature neuroscience* 18.11, pp. 1571–1573.

Frank, Michael J, Lauren C Seeberger, and Randall C O’reilly (2004). “By carrot or by stick: cognitive reinforcement learning in parkinsonism”. In: *Science* 306.5703, pp. 1940–1943.

Frydman, Cary and Gideon Nave (2016). “Extrapolative beliefs in perceptual and economic decisions: evidence of a common mechanism”. In: *Management Science*.

Galvan, Adriana et al. (2006). “Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents”. In: *Journal of Neuroscience* 26.25, pp. 6885–6892.

Gershman, Samuel J and Nathaniel D Daw (2012). “Perception, action and utility: The tangled skein”. In: *Principles of brain dynamics: Global state interactions*, pp. 293–312.

Gershman, Samuel J, Bijan Pesaran, and Nathaniel D Daw (2009). “Human reinforcement learning subdivides structured action spaces by learning effector-specific values”. In: *Journal of Neuroscience* 29.43, pp. 13524–13531.

Giedd, Jay N et al. (1999). “Brain development during childhood and adolescence: a longitudinal MRI study”. In: *Nature neuroscience* 2.10, pp. 861–863.

Gillan, Claire M et al. (2016). “Characterizing a psychiatric symptom dimension related to deficits in goal-directed control”. In: *Elife* 5, e11305.

Gilovich, Thomas, Robert Vallone, and Amos Tversky (1985). “The hot hand in basketball: On the misperception of random sequences”. In: *Cognitive psychology* 17.3, pp. 295–314.

Gläscher, Jan, Nathaniel Daw, et al. (2010). “States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning”. In: *Neuron* 66.4, pp. 585–595.

Gläscher, Jan, Alan N Hampton, and John P O’doherty (2008). “Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making”. In: *Cerebral cortex* 19.2, pp. 483–495.

Glimcher, Paul W and Ernst Fehr (2013). *Neuroeconomics: Decision making and the brain*. Academic Press.

Glöckner, Andreas et al. (2012). “Processing differences between descriptions and experience: A comparative analysis using eye-tracking and physiological measures”. In: *Frontiers in psychology* 3.

Gold, Joshua I and Michael N Shadlen (2001). “Neural computations that underlie decisions about sensory stimuli”. In: *Trends in cognitive sciences* 5.1, pp. 10–16.

– (2002). “Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward”. In: *Neuron* 36.2, pp. 299–308.

– (2007). “The neural basis of decision making”. In: *Annu. Rev. Neurosci.* 30, pp. 535–574.

Gottfried, Jay A, Ralf Deichmann, et al. (2002). “Functional heterogeneity in human olfactory cortex: an event-related functional magnetic resonance imaging study”. In: *Journal of Neuroscience* 22.24, pp. 10819–10828.

Gottfried, Jay A, John O’Doherty, and Raymond J Dolan (2003). “Encoding predictive reward value in human amygdala and orbitofrontal cortex”. In: *Science* 301.5636, pp. 1104–1107.

Green, Leonard et al. (1981). “Preference reversal and self control: Choice as a function of reward amount and delay.” In: *Behaviour Analysis Letters*.

Greer, John Michael and David K Levine (2006). “A dual-self model of impulse control”. In: *The American Economic Review* 96.5, pp. 1449–1476.

Gremel, Christina M and Rui M Costa (2013). “Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions”. In: *Nature communications* 4, p. 2264.

Grether, David M and Charles R Plott (1979). “Economic theory of choice and the preference reversal phenomenon”. In: *The American Economic Review* 69.4, pp. 623–638.

Hanks, Timothy D, Jochen Ditterich, and Michael N Shadlen (2006). “Microstimulation of macaque area LIP affects decision-making in a motion discrimination task”. In: *Nature neuroscience* 9.5, pp. 682–689.

Harbaugh, William T, Kate Krause, and Lise Vesterlund (2001). “Are adults better behaved than children? Age, experience, and the endowment effect”. In: *Economics Letters* 70.2, pp. 175–181.

Hare, Todd A, Colin F Camerer, and Antonio Rangel (2009). “Self-control in decision-making involves modulation of the vmPFC valuation system”. In: *Science* 324.5927, pp. 646–648.

Hare, Todd A, Jonathan Malmaud, and Antonio Rangel (2011). “Focusing attention on the health aspects of foods changes value signals in vmPFC and improves dietary choice”. In: *Journal of Neuroscience* 31.30, pp. 11077–11087.

Hare, Todd A, Wolfram Schultz, et al. (2011). "Transformation of stimulus value signals into motor commands during simple choice". In: *Proceedings of the National Academy of Sciences* 108.44, pp. 18120–18125.

Hare, Todd A. et al. (2008). "Dissociating the Role of the Orbitofrontal Cortex and the Striatum in the Computation of Goal Values and Prediction Errors". In: *Journal of Neuroscience* 28.22, pp. 5623–5630.

Hawkins, Guy E et al. (2015). "Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making". In: *Journal of Neuroscience* 35.6, pp. 2476–2484.

Heekeren, Hauke R et al. (2004). "A general mechanism for perceptual decision-making in the human brain". In: *Nature* 431.7010, pp. 859–862.

Henrich, Joseph, Steven J Heine, and Ara Norenzayan (2010). "Most people are not WEIRD". In: *Nature* 466.7302, pp. 29–29.

Hertwig, Ralph, Greg Barron, et al. (2004). "Decisions from experience and the effect of rare events in risky choice". In: *Psychological science* 15.8, pp. 534–539.

Hertwig, Ralph and Ido Erev (2009). "The description–experience gap in risky choice". In: *Trends in cognitive sciences* 13.12, pp. 517–523.

Hikosaka, Okihide, Satoru Miyauchi, and Shinsuke Shimojo (1993). "Focal visual attention produces illusory temporal order and motion sensation". In: *Vision research* 33.9, pp. 1219–1240.

Ho, Tiffany C, Scott Brown, and John T Serences (2009). "Domain general mechanisms of perceptual decision making in human cortex". In: *Journal of Neuroscience* 29.27, pp. 8675–8687.

Horwitz, Gregory D and William T Newsome (1999). "Separate signals for target selection and movement specification in the superior colliculus". In: *Science* 284.5417, pp. 1158–1161.

Hunt, Laurence T et al. (2012). "Mechanisms underlying cortical activity during value-guided choice". In: *Nature neuroscience* 15.3, pp. 470–476.

Hutcherson, Cendri A, Benjamin Bushong, and Antonio Rangel (2015). "A neurocomputational model of altruistic choice and its implications". In: *Neuron* 87.2, pp. 451–462.

Jessup, Ryan K, Anthony J Bishara, and Jerome R Busemeyer (2008). "Feedback produces divergence from prospect theory in descriptive choice". In: *Psychological Science* 19.10, pp. 1015–1022.

Kable, Joseph W and Paul W Glimcher (2007). "The neural correlates of subjective value during intertemporal choice". In: *Nature neuroscience* 10.12, pp. 1625–1633.

Kable, Joseph W and Paul W Glimcher (2009). "The neurobiology of decision: consensus and controversy". In: *Neuron* 63.6, pp. 733–745.

Kahneman, Daniel (1973). *Attention and effort*. Vol. 1063. Prentice-Hall Englewood Cliffs, NJ.

Kahneman, Daniel and Amos Tversky (1979). "Prospect theory: An analysis of decision under risk". In: *Econometrica: Journal of the econometric society*, pp. 263–291.

Kiani, Roozbeh, Timothy D Hanks, and Michael N Shadlen (2008). "Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment". In: *Journal of Neuroscience* 28.12, pp. 3017–3029.

Kim, Jong-Nam and Michael N Shadlen (1999). "Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque". In: *Nature neuroscience* 2.2, pp. 176–185.

Knill, David C and Alexandre Pouget (2004). "The Bayesian brain: the role of uncertainty in neural coding and computation". In: *TRENDS in Neurosciences* 27.12, pp. 712–719.

Knill, David C and Whitman Richards (1996). *Perception as Bayesian inference*. Cambridge University Press.

Knutson, Brian et al. (2005). "Distributed neural representation of expected value". In: *Journal of Neuroscience* 25.19, pp. 4806–4812.

Koenigs, Michael and Daniel Tranel (2007). "Irrational economic decision-making after ventromedial prefrontal damage: evidence from the Ultimatum Game". In: *Journal of Neuroscience* 27.4, pp. 951–956.

Kool, Wouter, Fiery A Cushman, and Samuel J Gershman (2016). "When does model-based control pay off?" In: *PLoS computational biology* 12.8, e1005090.

Kool, Wouter, Samuel J Gershman, and Fiery A Cushman (2017). "Cost-benefit arbitration between multiple reinforcement-learning systems". In: *Psychological Science* 28.9, pp. 1321–1333.

Körding, Konrad P and Daniel M Wolpert (2004). "Bayesian integration in sensorimotor learning". In: *Nature* 427.6971, pp. 244–247.

Krajbich, Ian, Carrie Armel, and Antonio Rangel (2010). "Visual fixations and the computation and comparison of value in simple choice". In: *Nature neuroscience* 13.10, pp. 1292–1298.

Krajbich, Ian, Dingchao Lu, et al. (2012). "The attentional drift-diffusion model extends to simple purchasing decisions". In: *Frontiers in psychology* 3.

Krajbich, Ian and Antonio Rangel (2011). "Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions". In: *Proceedings of the National Academy of Sciences* 108.33, pp. 13852–13857.

Kunar, Melina A et al. (2017). “The influence of attention on value integration”. In: *Attention, Perception, & Psychophysics*, pp. 1–13.

Laming, Donald (1979). “A critical comparison of two random-walk models for two-choice reaction time”. In: *Acta Psychologica* 43.6, pp. 431–453.

Lawrence, Natalia S et al. (2006). “Decision making and set shifting impairments are associated with distinct symptom dimensions in obsessive-compulsive disorder.” In: *Neuropsychology* 20.4, p. 409.

Lee, Sang Wan, Shinsuke Shimojo, and John P O’Doherty (2014). “Neural computations underlying arbitration between model-based and model-free learning”. In: *Neuron* 81.3, pp. 687–699.

Lehmann, EL (1983). “Theory of Point Estimation, New York: JohnWiley”. In: *LehmannTheory of Point Estimation1983*.

Ma, Wei Ji et al. (2006). “Bayesian inference with probabilistic population codes”. In: *Nature neuroscience* 9.11, pp. 1432–1438.

Maner, Jon K et al. (2007). “Dispositional anxiety and risk-avoidant decision-making”. In: *Personality and Individual Differences* 42.4, pp. 665–675.

Mata, Rui et al. (2011). “Age differences in risky choice: A meta-analysis”. In: *Annals of the New York Academy of Sciences* 1235.1, pp. 18–29.

Mather, Mara et al. (2004). “Amygdala responses to emotionally valenced stimuli in older and younger adults”. In: *Psychological Science* 15.4, pp. 259–263.

McClure, Samuel M et al. (2004). “Separate neural systems value immediate and delayed monetary rewards”. In: *Science* 306.5695, pp. 503–507.

McNamee, Daniel, Antonio Rangel, and John P O’doherty (2013). “Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex”. In: *Nature neuroscience* 16.4, pp. 479–485.

Meyer, David E and Roger W Schvaneveldt (1971). “Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations.” In: *Journal of experimental psychology* 90.2, p. 227.

Miller, Kevin J, Matthew M Botvinick, and Carlos D Brody (2017). “Dorsal hippocampus contributes to model-based planning”. In: *Nature Neuroscience* 20.9, pp. 1269–1276.

Mimura, Masaru, Reiko Oeda, and Mitsuru Kawamura (2006). “Impaired decision-making in Parkinson’s disease”. In: *Parkinsonism & related disorders* 12.3, pp. 169–175.

Mormann, Milica Milosavljevic, Christof Koch, and Antonio Rangel (2011). “Consumers can make decisions in as little as a third of a second”. In:

Mormann, Milica Milosavljevic, Jonathan Malmaud, et al. (2010). “The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure”. In:

Mullett, Timothy L and Neil Stewart (2016). "Implications of visual attention phenomena for models of preferential choice." In: *Decision* 3.4, p. 231.

Murphy, Fionnuala C et al. (2001). "Decision-making cognition in mania and depression". In: *Psychological medicine* 31.4, pp. 679–693.

Nedungadi, Prakash (1990). "Recall and consumer consideration sets: Influencing choice without altering brand evaluations". In: *Journal of consumer research* 17.3, pp. 263–276.

Newsome, William T, Kenneth H Britten, and J Anthony Movshon (1989). "Neuronal correlates of a perceptual decision". In: *Nature* 341.6237, pp. 52–54.

Newsome, William T and Edmond B Pare (1988). "A selective impairment of motion perception following lesions of the middle temporal visual area (MT)". In: *Journal of Neuroscience* 8.6, pp. 2201–2211.

Novemsky, Nathan and Daniel Kahneman (2005). "The boundaries of loss aversion". In: *Journal of Marketing research* 42.2, pp. 119–128.

O'connell, Redmond G, Paul M Dockree, and Simon P Kelly (2012). "A supramodal accumulation-to-bound signal that determines perceptual decisions in humans". In: *Nature neuroscience* 15.12, pp. 1729–1735.

O'Doherty, John P (2004). "Reward representations and reward-related learning in the human brain: insights from neuroimaging". In: *Current opinion in neurobiology* 14.6, pp. 769–776.

– (2015). "Multiple systems for the motivational control of behavior and associated neural substrates in humans". In: *Behavioral Neuroscience of Motivation*. Springer, pp. 291–312.

O'Doherty, John et al. (2001). "Representation of pleasant and aversive taste in the human brain". In: *Journal of neurophysiology* 85.3, pp. 1315–1321.

O'Doherty, John et al. (2003). "Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness". In: *Neuropsychologia* 41.2, pp. 147–155.

O'doherty, John et al. (2004). "Dissociable roles of ventral and dorsal striatum in instrumental conditioning". In: *science* 304.5669, pp. 452–454.

Orquin, Jacob L and Simone Mueller Loose (2013). "Attention and choice: A review on eye movements in decision making". In: *Acta psychologica* 144.1, pp. 190–206.

Ossmy, Ori et al. (2013). "The timescale of perceptual evidence integration can be adapted to the environment". In: *Current Biology* 23.11, pp. 981–986.

Otto, A Ross, Samuel J Gershman, et al. (2013). "The curse of planning: dissecting multiple reinforcement-learning systems by taxing the central executive". In: *Psychological science* 24.5, pp. 751–761.

Otto, A Ross, Candace M Raio, et al. (2013). "Working-memory capacity protects model-based learning from stress". In: *Proceedings of the National Academy of Sciences* 110.52, pp. 20941–20946.

Padoa-Schioppa, Camillo and John A Assad (2006). "Neurons in the orbitofrontal cortex encode economic value". In: *Nature* 441.7090, pp. 223–226.

Pärnamets, Philip et al. (2015). "Biasing moral decisions by exploiting the dynamics of eye gaze". In: *Proceedings of the National Academy of Sciences* 112.13, pp. 4170–4175.

Philiastides, Marios G, Guido Biele, and Hauke R Heekeren (2010). "A mechanistic account of value computation in the human brain". In: *Proceedings of the National Academy of Sciences* 107.20, pp. 9430–9435.

Philiastides, Marios G and Roger Ratcliff (2013). "Influence of branding on preference-based decision making". In: *Psychological science* 24.7, pp. 1208–1215.

Philiastides, Marios G, Roger Ratcliff, and Paul Sajda (2006). "Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram". In: *Journal of Neuroscience* 26.35, pp. 8965–8975.

Polania, Rafael et al. (2014). "Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making". In: *Neuron* 82.3, pp. 709–720.

Posner, Michael I, Charles R Snyder, and Brian J Davidson (1980). "Attention and the detection of signals." In: *Journal of experimental psychology: General* 109.2, p. 160.

Rabin, Matthew (2000). "Diminishing marginal utility of wealth cannot explain risk aversion". In: *Department of Economics, UCB*.

Raghunathan, Rajagopal and Michel Tuan Pham (1999). "All negative moods are not equal: Motivational influences of anxiety and sadness on decision making". In: *Organizational behavior and human decision processes* 79.1, pp. 56–77.

Rangel, Antonio, Colin Camerer, and P Read Montague (2008). "A framework for studying the neurobiology of value-based decision making". In: *Nature Reviews Neuroscience* 9.7, pp. 545–556.

Ratcliff, Roger (1978). "A theory of memory retrieval." In: *Psychological review* 85.2, p. 59.

Ratcliff, Roger, Anil Cherian, and Mark Segraves (2003). "A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions". In: *Journal of neurophysiology* 90.3, pp. 1392–1407.

Ratcliff, Roger and Gail McKoon (2008). "The diffusion decision model: theory and data for two-choice decision tasks". In: *Neural computation* 20.4, pp. 873–922.

Ratcliff, Roger, Marios G Philiastides, and Paul Sajda (2009). "Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG". In: *Proceedings of the National Academy of Sciences* 106.16, pp. 6539–6544.

Ratcliff, Roger and Jeffrey N Rouder (1998). "Modeling response times for two-choice decisions". In: *Psychological Science* 9.5, pp. 347–356.

Ratcliff, Roger and Philip L Smith (2004). "A comparison of sequential sampling models for two-choice reaction time." In: *Psychological review* 111.2, p. 333.

Ratcliff, Roger and Francis Tuerlinckx (2002). "Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability". In: *Psychonomic bulletin & review* 9.3, pp. 438–481.

Reddi, BAJ and RHS Carpenter (2000). "The influence of urgency on decision time". In: *Nature neuroscience* 3.8, pp. 827–830.

Roe, Robert M, Jermone R Busemeyer, and James T Townsend (2001). "Multialternative decision field theory: A dynamic connectionst model of decision making." In: *Psychological review* 108.2, p. 370.

Roitman, Jamie D and Michael N Shadlen (2002). "Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task". In: *Journal of neuroscience* 22.21, pp. 9475–9489.

Rolls, Edmund T, Morten L Kringelbach, and Ivan ET De Araujo (2003). "Different representations of pleasant and unpleasant odours in the human brain". In: *European Journal of Neuroscience* 18.3, pp. 695–703.

Rolls, Edmund T, John O'Doherty, et al. (2003). "Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices". In: *Cerebral cortex* 13.3, pp. 308–317.

Russek, Evan M et al. (2017). "Predictive representations can link model-based reinforcement learning to model-free mechanisms". In: *bioRxiv*, p. 083857.

Rustichini, Aldo and Camillo Padoa-Schioppa (2015). "A neuro-computational model of economic decisions". In: *Journal of Neurophysiology* 114.3, pp. 1382–1398.

Salzman, C Daniel, Kenneth H Britten, and William T Newsome (1990). "Cortical microstimulation influences perceptual judgements of motion direction". In: *Nature* 346.6280, pp. 174–177.

Salzman, C Daniel, Chieko M Murasugi, et al. (1992). "Microstimulation in visual area MT: effects on direction discrimination performance". In: *Journal of Neuroscience* 12.6, pp. 2331–2355.

Samanez-Larkin, Gregory R and Brian Knutson (2015). "Decision making in the ageing brain: changes in affective and motivational circuits". In: *Nature Reviews. Neuroscience* 16.5, p. 278.

Sanfey, Alan G et al. (2003). “The neural basis of economic decision-making in the ultimatum game”. In: *Science* 300.5626, pp. 1755–1758.

Schoenbaum, Geoffrey, Matthew R Roesch, and Thomas A Stalnaker (2006). “Orbitofrontal cortex, decision-making and drug addiction”. In: *Trends in neurosciences* 29.2, pp. 116–124.

Schönberg, Tom et al. (2007). “Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making”. In: *Journal of Neuroscience* 27.47, pp. 12860–12867.

Shadlen, Michael N and Roozbeh Kiani (2013). “Decision making as a window on cognition”. In: *Neuron* 80.3, pp. 791–806.

Shadlen, Michael N and William T Newsome (1996). “Motion perception: seeing and deciding”. In: *Proceedings of the national academy of sciences* 93.2, pp. 628–633.

– (2001). “Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey”. In: *Journal of neurophysiology* 86.4, pp. 1916–1936.

Shimojo, Shinsuke et al. (2003). “Gaze bias both reflects and influences preference”. In: *Nature neuroscience* 6.12, pp. 1317–1322.

Sloman, Steven A (1996). “The empirical case for two systems of reasoning.” In: *Psychological bulletin* 119.1, p. 3.

Small, Dana M et al. (2003). “Dissociation of neural representation of intensity and affective valuation in human gustation”. In: *Neuron* 39.4, pp. 701–711.

Smith, Philip L (1995). “Psychophysically principled models of visual simple reaction time.” In: *Psychological review* 102.3, p. 567.

Smith, Philip L and Roger Ratcliff (2004). “Psychology and neurobiology of simple decisions”. In: *Trends in neurosciences* 27.3, pp. 161–168.

Smith, Philip L, Roger Ratcliff, and Bradley J Wolfgang (2004). “Attention orienting and the time course of perceptual decisions: Response time distributions with masked and unmasked displays”. In: *Vision research* 44.12, pp. 1297–1320.

Starns, Jeffrey J, Roger Ratcliff, and Gail McKoon (2012). “Modeling single versus multiple systems in implicit and explicit memory”. In: *Trends in cognitive sciences* 16.4, pp. 195–196.

Stocker, Alan A and Eero P Simoncelli (2006). “Noise characteristics and prior expectations in human visual speed perception”. In: *Nature neuroscience* 9.4, pp. 578–585.

Stone, Mervyn (1960). “Models for choice-reaction time”. In: *Psychometrika* 25.3, pp. 251–260.

Summerfield, Christopher and Tobias Egner (2013). “Attention and decision-making”. In: *The oxford handbook of attention*.

Summerfield, Christopher and Konstantinos Tsetsos (2012). "Building bridges between perceptual and economic decision-making: neural and computational mechanisms". In: *Frontiers in neuroscience* 6.

Sutton, Richard S and Andrew G Barto (1998). *Reinforcement learning: An introduction*. Vol. 1. 1. MIT press Cambridge.

Tavares, Gabriela, Pietro Perona, and Antonio Rangel (2017). "The Attentional Drift Diffusion Model of Simple Perceptual Decision-Making". In: *Frontiers in Neuroscience* 11, p. 468. ISSN: 1662-453X. DOI: 10.3389/fnins.2017.00468. URL: https://www.frontiersin.org/article/10.3389/fnins.2017.00468.

Teodorescu, Andrei R and Marius Usher (2013). "Disentangling decision models: From independence to competition." In: *Psychological review* 120.1, p. 1.

Tosoni, Annalisa et al. (2008). "Sensory-motor mechanisms in human parietal cortex underlie arbitrary visual decisions". In: *Nature neuroscience* 11.12, pp. 1446–1453.

Towal, R Blythe, Milica Mormann, and Christof Koch (2013). "Simultaneous modeling of visual saliency and value computation improves predictions of economic choice". In: *Proceedings of the National Academy of Sciences* 110.40, E3858–E3867.

Tricomi, Elizabeth, Bernard W Balleine, and John P O'Doherty (2009). "A specific role for posterior dorsolateral striatum in human habit learning". In: *European Journal of Neuroscience* 29.11, pp. 2225–2232.

Tsetsos, Konstantinos, Nick Chater, and Marius Usher (2012). "Salience driven value integration explains decision biases and preference reversal". In: *Proceedings of the National Academy of Sciences* 109.24, pp. 9659–9664.

Tsetsos, Konstantinos, Juan Gao, et al. (2012). "Using time-varying evidence to test models of decision dynamics: bounded diffusion vs. the leaky competing accumulator model". In: *Frontiers in neuroscience* 6.

Tsetsos, Konstantinos, Rani Moran, et al. (2016). "Economic irrationality is optimal during noisy decision making". In: *Proceedings of the National Academy of Sciences* 113.11, pp. 3102–3107.

Tsetsos, Konstantinos, Marius Usher, and James L McClelland (2011). "Testing multi-alternative decision models with non-stationary evidence". In: *Frontiers in neuroscience* 5.

Tversky, Amos and Daniel Kahneman (1971). "Belief in the law of small numbers." In: *Psychological bulletin* 76.2, p. 105.

Tversky, Amos, Paul Slovic, and Daniel Kahneman (1990). "The causes of preference reversal". In: *The American Economic Review*, pp. 204–217.

Usher, Marius and James L McClelland (2001). “The time course of perceptual choice: the leaky, competing accumulator model.” In: *Psychological review* 108.3, p. 550.

Vugt, Marieke Karlijn van et al. (2012). “EEG oscillations reveal neural correlates of evidence accumulation”. In: *Frontiers in neuroscience* 6.

White, Corey N, Jeanette A Mumford, and Russell A Poldrack (2012). “Perceptual criteria in the human brain”. In: *Journal of Neuroscience* 32.47, pp. 16716–16724.

Wong, Kong-Fatt and Xiao-Jing Wang (2006). “A recurrent network mechanism of time integration in perceptual decisions”. In: *Journal of Neuroscience* 26.4, pp. 1314–1328.

Wu, Shih-Wei, Mauricio R Delgado, and Laurence T Maloney (2009). “Economic decision-making compared with an equivalent motor task”. In: *Proceedings of the National Academy of Sciences* 106.15, pp. 6088–6093.

Wunderlich, Klaus, Peter Dayan, and Raymond J Dolan (2012). “Mapping value based planning and extensively trained choice in the human brain”. In: *Nature neuroscience* 15.5, pp. 786–791.

Wyart, Valentin, Nicholas E Myers, and Christopher Summerfield (2015). “Neural mechanisms of human perceptual choice under focused and divided attention”. In: *Journal of neuroscience* 35.8, pp. 3485–3498.

INDEX