DECISION FIELD THEORY-PLANNING: A COGNITIVE MODEL OF PLANNING AND DYNAMIC DECISION MAKING

Jared M. Hotaling

Submitted to the faculty of University Graduate School in partial fulfillment of the requirements for the degree Doctor of Philosophy in the department of Psychological and Brain Sciences and the Cognitive Science Program

Indiana University

November 2013

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Doctoral Committee

Jerome R. Busemeyer, Ph.D. (Chair)

Richard M. Shiffrin, Ph.D.

Peter M. Todd, Ph.D.

Robert M. Nosofsky, Ph.D.

August 14, 2013

Acknowledgments

I would like to thank my advisor, Jerome Busemeyer, for giving me the freedom to pursue my academic interests and the guidance to help me along the way. Thanks also to Rich Shiffrin for all of your support throughout my time at Indiana University, and for teaching me that all models are wrong. Thank you Peter Todd for your insightful feedback and unique perspective, and thank you Rob Nosofsky for your attention to detail and alternative explanations. I would also like to thank the many people who contributed to Indiana's amazing cognitive science community. In particular, I would like to thank Dan Yurovsky and George Kachergis for our many conversations, scientific or otherwise, as well as Joe Houpt, Jennifer Trueblood, Greg Cox, Drew Henrickson, Carlos Zednik, and Paul Williams. You all played a critical part in how I think about the world and how I try to understand it. Thanks also to the members of Busemeyer and Shiffrin labs. To my parents, Mike and Roseann, without your unwavering support and encouragement I never would have been able to accomplish this. Finally, thanks to my wife, Kate, whose love and devotion helped me across the finish line and on to the next adventure.

Jared M. Hotaling

Decision Field Theory-Planning: A Cognitive Model of Planning and Dynamic Decision Making

The world is full of complex environments in which individuals must plan a series of choices to obtain some desired outcome. In these situations entire sequences of events, including one's future decisions, should be considered before taking an action. Backward induction provides a normative strategy for planning, in which one works backward, deterministically, from the end of a scenario. However, it often fails to account for human behavior. I propose an alternative account, *Decision Field Theory-Planning*, in which individuals plan future choices on the fly through repeated mental simulations. A key prediction of DFT-P is that payoff variability produces noisy simulations and reduces sensitivity to utility differences. In two multistage risky decision tree experiments I obtained this payoff variability effect, with choice proportions moving toward 0.50 as variability increased. I showed that DFT-P provides valuable insight into the strategies that people used to plan future choices and allocate cognitive resources across decision stages.

Contents

List of Figures

List of Tables

1. Introduction

Risky decisions present an individual with a choice between uncertain alternatives. Much of the research into how people make risky decisions has focused on isolated choices with each decision being made independently from one another. Unfortunately, many choices outside of the laboratory are not so easily separated from the other actions an individual takes, and instead involve planning across sequences of events and actions. Although there are many ways to describe such complex choice problems, I will use the terms *multistage decision making* and *dynamic decision making* interchangeably. To begin, imagine a scenario where Sam, an employee at a marketing firm, is contemplating asking for pay raise. This is not a decision that Sam should take lightly, without first considering the various ways in which his choice might interact with other events and actions. Figure 1 uses a decision tree to represent the complexities to this scenario. Decision trees like these are commonly used to represent multistage decision problems. The tree consists of three kinds of nodes: decision nodes, chance nodes, and outcome nodes. Decision nodes (DNs) represent points where the decision maker chooses an action. Chance nodes (CNs) denote random events that occur in the world, which the decision maker cannot control. Finally, each branch of the tree terminates with an outcome node (ON) representing a final consequence.

In this example, it is clearly important that Sam think ahead before making an initial choice because the decisions he will later make at DN2a or DN2b (denoted DN2a/b from here on) partially determine the outcome of asking for a raise. For instance, Sam should consider what he will do at DN2b if his boss refused his request. Perhaps such a rejection would sour his working relationship to the point where he would leave his job in search of another (CN2c). If so, choosing to ask for a raise would be quite risky indeed. On the other hand, if Sam would be content to continue working at the firm, despite being spurned, then he may have little to lose by taking a risk at DN1. The present research investigates how people solve such multistage risky decision problems. In particular, I focus on the degree to which individuals use information about future choices and events to inform their decision making at present. In Chapter 1 I introduce backward induction as a normative theory of dynamic decision making, and summarize previous research into its utility as a descriptive model of human decision making. In Chapter 2 I present an experiment testing whether, and to what degree, people spontaneously plan future choices in a laboratory setting. Chapter 3 presents a new cognitive model of dynamic decision making aimed at describing the mental computations that support planning. In Chapter 4 I test a fundamental prediction of the model using an experiment focused on the effects of payoff variability on multistage risky decisions, and demonstrate that a competing psychological model cannot account for the observed data. Chapter 5 presents a replication of the key findings from Chapter 4, along with a quantitative fit of my model to group data. Finally, with Chapter 6 I demonstrate how the model can be used to explain the actions of individual decision makers, as well as differences in behavior across individuals.

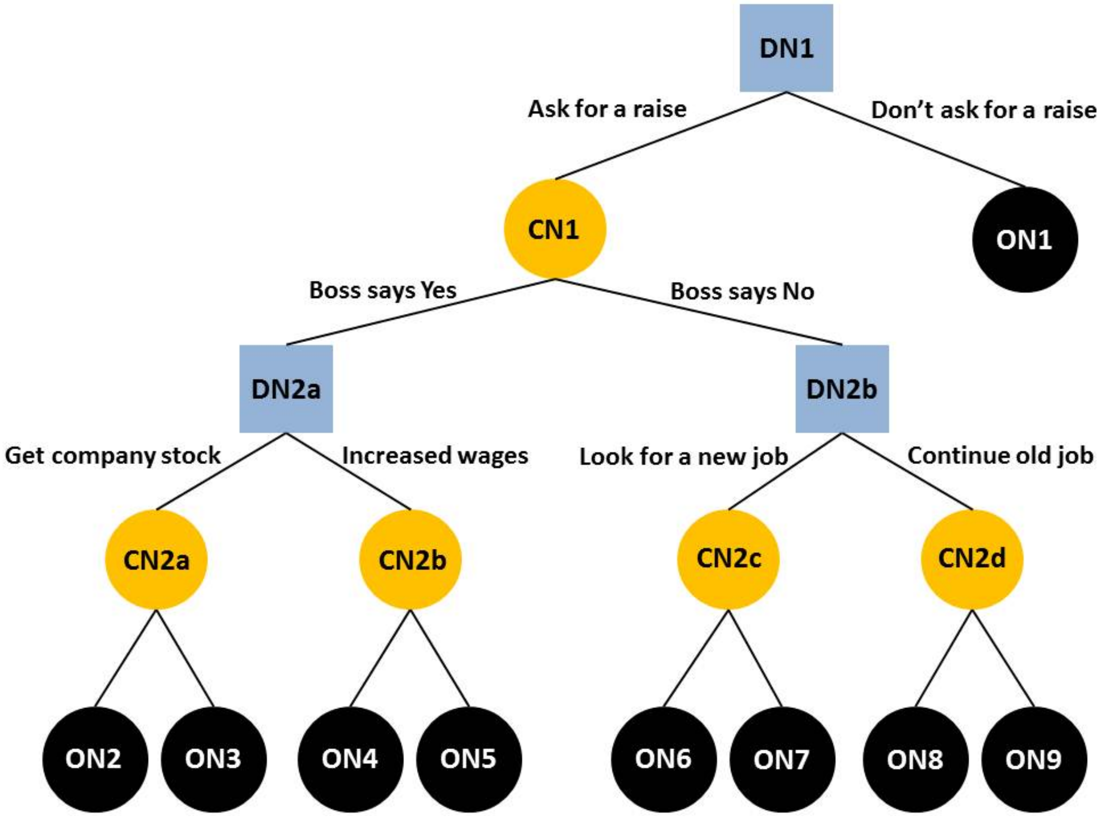

Figure 1.1. Example dynamic decision scenario, represented as a decision tree. Decision nodes (DNs) represent points where you make a choice. Chance nodes (CNs) represent points where an uncertain external event occurs. Outcome nodes (ONs) represent potential final outcomes.

Backward Induction

Within decision science, planning is normally treated as a problem of *backward induction* (see Bertsekas, 1976; DeGroot, 1970; Keeney & Raiffa, 1976; Raiffa, 1968; Von Winterfeldt & Edwards, 1986). This refers to a normative strategy involving thinking backward from the end of a decision tree. The basic idea is that the individual begins at the end of the tree and prunes unwanted branches while working backward to the beginning. The process begins by assigning a utility value to each ON, representing its worth to the decision maker. Using the tree in Figure 1 as an example, ON1–9 would first be assigned the values $\mu_1 - \mu_9$, respectively. Next, an *expected utility* is computed as the weighted average utility for each of CN2a-d. For instance, EU(CN2c), the expected utility for CN2c, is calculated by weighting μ_6 and μ_7 by their probabilities, p_6 and p_7, respectively and summing the results. Third, final DNs (DN2a/b in Figure 1) are each assigned an expected utility by taking the maximum value of the CNs branching out of them. All other branches are pruned and subsequently ignored. This process can be repeated indefinitely for multistage decision trees of an arbitrary length, until utilities are propagated to the first DN. At this point, the decision maker need only choose the alternative with the highest expected utility, and execute their planned decisions upon reaching later DNs. This theory is valuable because following the prescriptions of backward induction maximizes one's long-term expected utility.

Continuing the earlier example, assuming Sam already knows the utility of each outcome, backward induction dictates that he next calculate expected utilities for CN2a-d and use these to plan choices for DN2a/b. Imagine that choosing company stock (CN2a) offers Sam a greater expected value than choosing a wage increase (CN2b). According to the normative theory, Sam should prune the branch extending down to CN2b and carry EU(CN2a) up the tree to DN2a. This pruning reflects the assumption that, if Sam asks for a raise and his boss agrees, Sam will choose the action that is, on average, best for him. The same optimization is done at DN2b. Sam should next compute a weighting average of DN2a and DN2b, based on the likelihood of his boss saying Yes versus No, to get the expected utility of CN1. All that remains is a choice between ON1 and CN1. Sam is again assumed to take the option with greatest utility, and to continue down the tree exactly as planned.

Tests of Backward Induction as a Descriptive Model

The theory of backward induction has great value as a normative model, and has been used successfully to examine choices in contexts such as sequential games (Fudenberg & Tirole, 1991), investment (Dixit & Pindyck, 1994), business valuation (Anderson, 2009), and economic growth and taxation (Ljunggvist & Sargent, 2000). However, several researchers have questioned its value as a descriptive model of human decision making. Much of this work has focused on one of its foundational assumptions: *dynamic consistency* (Machina, 1989; Sarin & Wakker, 1998), which requires that decision makers follow through on their plans to the end. This is critically important to backward induction because planned decisions cause the decision maker to ignore information related to forgone prospects. Failing to follow through on one's plans thus jeopardizes the entire strategy by changing the relevant information. In a seminal study, Cubitt, Starmer, and Sugden (1998) tested this consistency principle with a between-subjects design, in which each individual made a single choice between the same two gambles presented in one of several forms. In the relevant conditions, participants were told that would begin with a lottery. With probability *p* they would receive nothing. With probability 1-*p*, they would be allowed to choose between the two gambles. In the *planned* choice condition, participants were asked which gamble they would choose if they won the lottery. In the *final* choice condition, participants were told that they had already won the lottery, and were asked to make a choice. Cubitt et al. found a significant difference between planned and final choices, suggesting that participants were not dynamically consistent. This result is especially noteworthy, given the relative simplicity of these decision problems when compared to the sort of multistage decision scenarios introduced at the beginning of this chapter.

Busemeyer and colleagues extended this work with several studies demonstrating further violations of dynamic consistency (Barkan & Busemeyer, 2003; Busemeyer, Weg, Barkan, Li, &

Ma, 2000; J. G. Johnson & Busemeyer, 2001). These employed within-subjects designs, in which participants repeatedly chose between similar gambles. Busemeyer at al. used a simple asymmetric tree structure offering a sequence of three DNs offering a choice between a sure payoff and a CN, which had a 50/50 chance of continuing to the next DN or exiting the tree with no payoff. The final DN, *D*, offered a choice between, a sure payoff, *S*, and a gamble, *G*, involving a 50/50 chance of receiving reward, *R*, or punishment, *P*. On some trials, at the beginning of the tree participants were asked to plan what they would choose if they found themselves at *D*. The computer would then automatically carry out the plan if *D* was reached. On other trials with the same tree, participants simply made a final choice upon reaching *D*. Busemeyer et al. found that when planning the choice they would make at *D*, participants tended to gamble, but when they actually reached *D*, they often opted for the sure payoff. They concluded that people are dynamically inconsistent, and cannot necessarily predict their future choices. Johnson and Busemeyer found that dynamic inconsistency increased with the length of the decision tree. That is, the further ahead a person is asked to plan, the more likely she is to deviate from that plan. These findings further challenge the idea that the normative model can account for people's behavior.

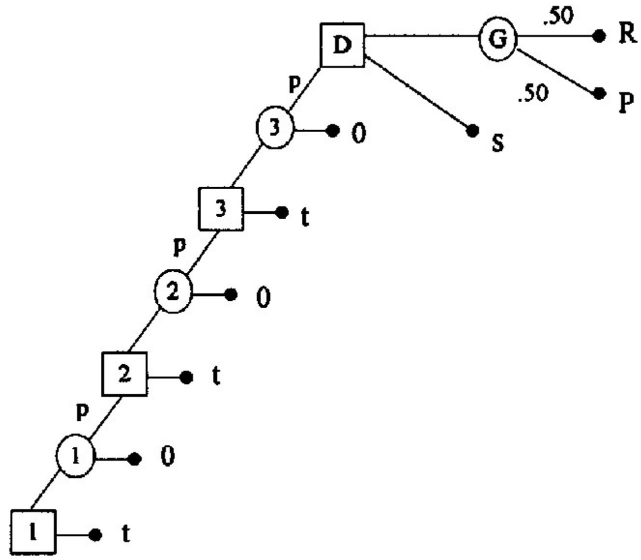

Figure 1.2. Example decision tree used in Busemeyer et al. (2000).

More recently, Hey and colleagues have taken a more indirect approach to testing people's planning in multistage decision making (J. Bone, Hey, & Suckling, 2009; J. D. Bone, Hey, & Suckling, 2003; Carbone & Hey, 2001; Hey & Knoll, 2007, 2011). Carbone and Hey (2001) conducted a study using large 3-stage decision trees comprised of interleaved binary DNs and CNs, with sixty-four terminal ONs. Participants were given unlimited time to explore and plan their way through the trees. Each of the ON values could be uncovered with a mouse click. At each CN or DN participants could open an electronic notebook and record a comment, plan a choice, or calculate an expected value. After exploring, participants worked their way through the tree, choosing a direction to move at each DN and reacting to events at each CN. Every action made by the participant during both phases of the experiment was recorded. Carbone & Hey replayed these actions to evaluate how people solved the problems, and found that many individuals did use some version of backward induction, but did not prune the tree according to their planned choices. Although they did look ahead to ONs to inform their strategies, they ignored the *principle of optimality,* in that they seemed unaware that they could make optimal choices in the future.

Using a simpler tree structure, Bone, Hey, and Suckling (2003) found that decision makers would often pass up a sure payoff at an early DN, only to settle for a smaller sure payoff at a later DN. This is puzzling since one would expect that any participant who gambles at the early DN would be seeking the risky alternative at the later DN. Hey and Knoll (2007) again used three-stage decision trees like those used by Carbone and Hey, but this time with a dominance structure such that, assuming participants' preferences were rational, those who planned one, two, or three steps ahead would each exhibit a unique pattern of choices. Their results showed that almost every participant could be classified as either planning all three stages completely or not planning at all. Hey and Knoll (2011) directly investigated individual differences in strategy use with yet another three-stage decision tree study. They examined the entire sequence of actions each decision maker made – every node they uncovered, every notebook entry they recorded, every choice they made – and tried to infer a strategy. Although, no specific classification scheme is reported, Hey and Knoll identified several distinct strategies. For example, some individuals were found to ignore significant amounts of data in order to minimize effort, while others approximated an optimal strategy. Forward workers, who employed simple, fast strategies, were distinguished from backward workers who attempted to backward induce, with varying levels of success. Dimperio (2009) conducted several multistage choice experiments and found similarly mixed results. After applying several competing models, two distinct groups emerged, one which carefully planned future choices and another that treated future decisions as if they were random CNs, outside of their control. He found that the dominant strategy was to make an initial decision based on one's tolerance for risk, rather than on the basis of future payoffs.

Summary and Overview of Current Work

The picture painted by the above research is rather bleak. Cubitt et al. (1998) and Busemeyer et al. (2000) found significant violations of dynamic consistency, implying that people cannot, or at least often fail to, accurately predict their future choices. The work of Hey and colleagues further supported this conclusion and demonstrated that, although decision makers often engage in some form of planning, they often lack the ability to properly implement normative strategies. Dimperio (2009) also found substantial individual differences in people's ability to plan ahead, and observed that most participants used risk itself, rather than expected payoff, to guide their decisions. These conclusions are hard to reconcile with the success that many, if not most, people have in coping with complex, dynamic real world environments. How then do people engage in dynamic decision making, if not through some form of backward induction? Before I can begin to answer this question, it would seem critical to test whether planning can be demonstrated in a laboratory setting. In the next chapter, I present a study designed to do just that by showing that individuals are sensitive to the distribution of payoffs across future DNs.

2. Experiment 1: Do people plan?

Although everyday observations attest to people's ability to navigate complex decision scenarios, much of the above research demonstrates the difficulty of studying dynamic decision making systematically in the laboratory. The first step in my investigation is therefore to develop a multistage decision tree paradigm in which individuals spontaneously plan ahead. In this section, I present an experiment designed to show that relatively small monetary incentives are sufficient to elicit planning. This finding provides an important foundation for further investigation into people's dynamic decision making processes.

The primary manipulation in Experiment 1 involves how payoffs are distributed across future DNs. The idea is that if individuals plan ahead they should be sensitive to this manipulation. This is because a rational decision maker can exploit any asymmetry in the distribution of payoffs by choosing the more attractive option. If individuals show sensitivity to future payoff distributions, they must be considering their future actions and planning to make optimizing choices. If individuals are not sensitive to the payoff distribution manipulation, this would suggest that they are not planning future choices for these DNs, but may instead treat them like CNs outside of their control.

A second goal of Experiment 1 is to assess the effect of semantic labels on individuals' choices. Some theorize that humans have evolved a number of expert systems, which they selectively employ, depending on their environment (for a review, see Barkow, Cosmides, & Tooby, 1992; Hirschfeld & Gelman, 1994). For instance, Cosmides and Tooby (1992) argue that when placed in a situation related to social exchanges, individuals will employ a reasoning system specially tailored to that domain. They use the example of the Wason card selection task (Wason, 1966), which was originally used to demonstrate the difficulty people often have with deductive reasoning. In the task, participants are shown four cards and are asked to test a rule of the form “if a card has a vowel on one side, then it has an even number of the other side”. They must test the rule by turning over the minimum number of cards possible. The classic effect is that individuals fail to use the most efficient strategy, and instead search for confirmatory evidence. This suggests that untrained individuals do not possess knowledge of logical rules and do not spontaneously reason according to them. However, Cosmides (1989) found that when the logical rule is expressed using a social context, most people correctly identify the correct two cards. It would appear that when individuals are able to apply their knowledge of social interactions they reason differently, and naturally adhere to the rules of logic. In Experiment 1, some decision trees were presented in abstract form, while others were given socially relevant cover stories. If domain-specific reasoning was present, I expected the amount of planning to be different, perhaps higher, when trees had semantic labels.

Method

Participants. Forty five students from Indiana University were recruited via flyers posted on campus and an advertisement on the university’s job listing forum for a “paid decision making study”. Payment was based on performance in the experiment, but participants received an average of $8 apiece.

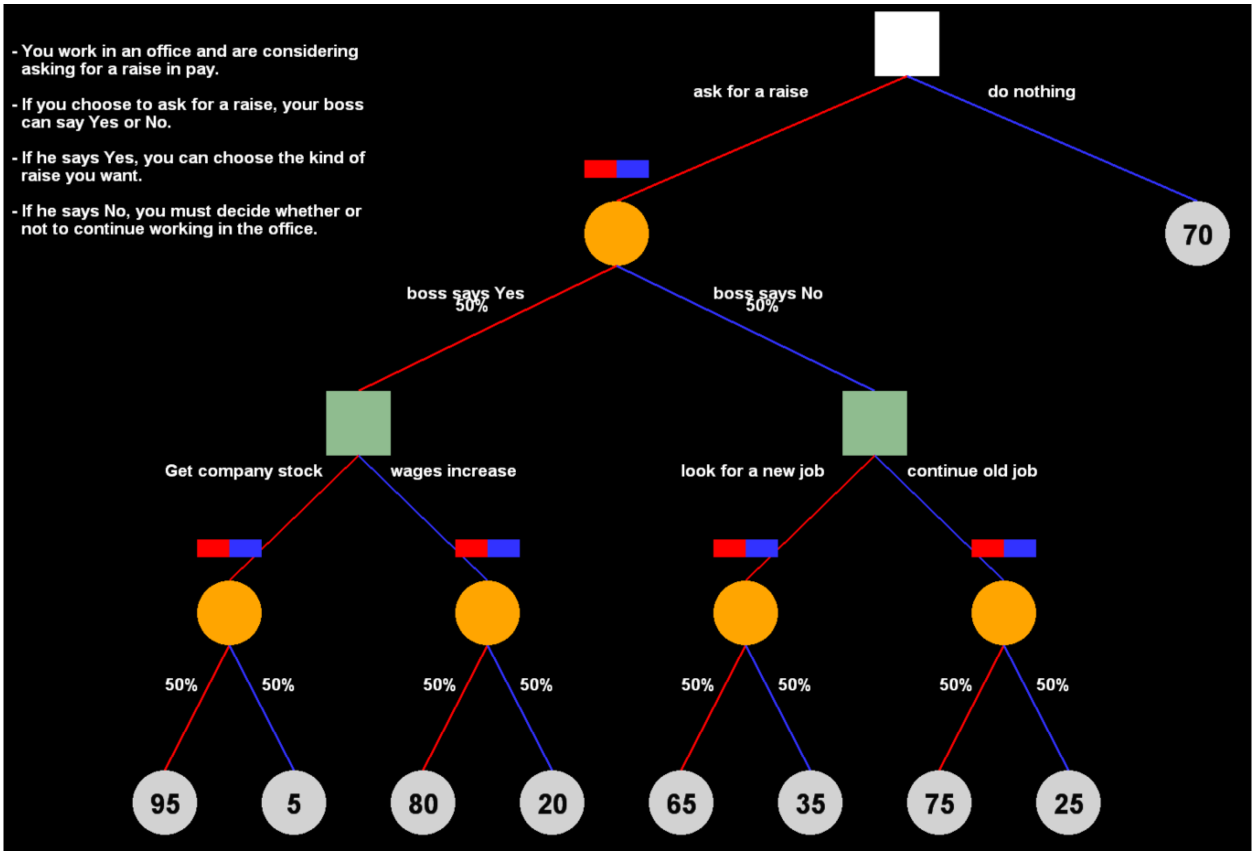

Figure 2.1. Screenshot from Experiment 1, showing a No Imbalance, Concrete trial, with the marker at DN1.

Stimuli and Design. The experiment was programed in Matlab, using the Psychophysics Toolbox. Figure 2.1 shows an example decision tree. All trees in the experiment had the same node structure introduced in Figure 1.1. DNs were shown as light green squares, CNs as yellow circles, and ONs as white circles. At all times, a white square indicated the participant's current position in the tree. Figure 2.2 shows three trees superimposed on one another to demonstrate the *Imbalance* variable. In the No Imbalance tree on top, each of CN2-CN5 have an expected value of 50. In the Low Imbalance tree, CN2 and CN4 have an expected value of 30, while CN3 and CN5 have an expected value of 70. In the High Imbalance tree, CN2 and CN4 have an expected value of 10, while CN3 and CN5 have an expected value of 90. If participants plan ahead, they should be more likely to gamble at DN1 on trials with high imbalance than on those with no imbalance. Note that this is true regardless of participants' risk preference because the minimum outcome for each of CN3 and CN5 increases with *imbalance*. However, if participants are not planning their future choices, they may only attend to the mean payoff for each future DN, and show no effect of Imbalance on choices at DN1.

The decision tree in Figure 2.1 is from the *Concrete* condition, in which a cover story was used to ground the problem in a real world scenario. Text describing the scenario was displayed in the top left corner of the screen on each Concrete trial. Each path branching out of CN1 and each DN was also labeled according to the cover story. Four different cover stories were used. For each cover story, four trees were created. Three corresponded to the three levels of Imbalance, while the fourth was an unrelated filler trial meant to obscure the experimental design. Slightly different ON values were use across cover stories to give the trials a more varied look, but the underlying structure and Imbalance manipulation was the same for each. Cover stories 1 and 3 used trees with transition probabilities of 0.50 at all CNs. Cover stories 2 and 4 used trees with uneven CN probabilities. This entire design was duplicated in the *Abstract* condition (see Figure 2.2), but with no cover story provided and branches being given only generic labels.

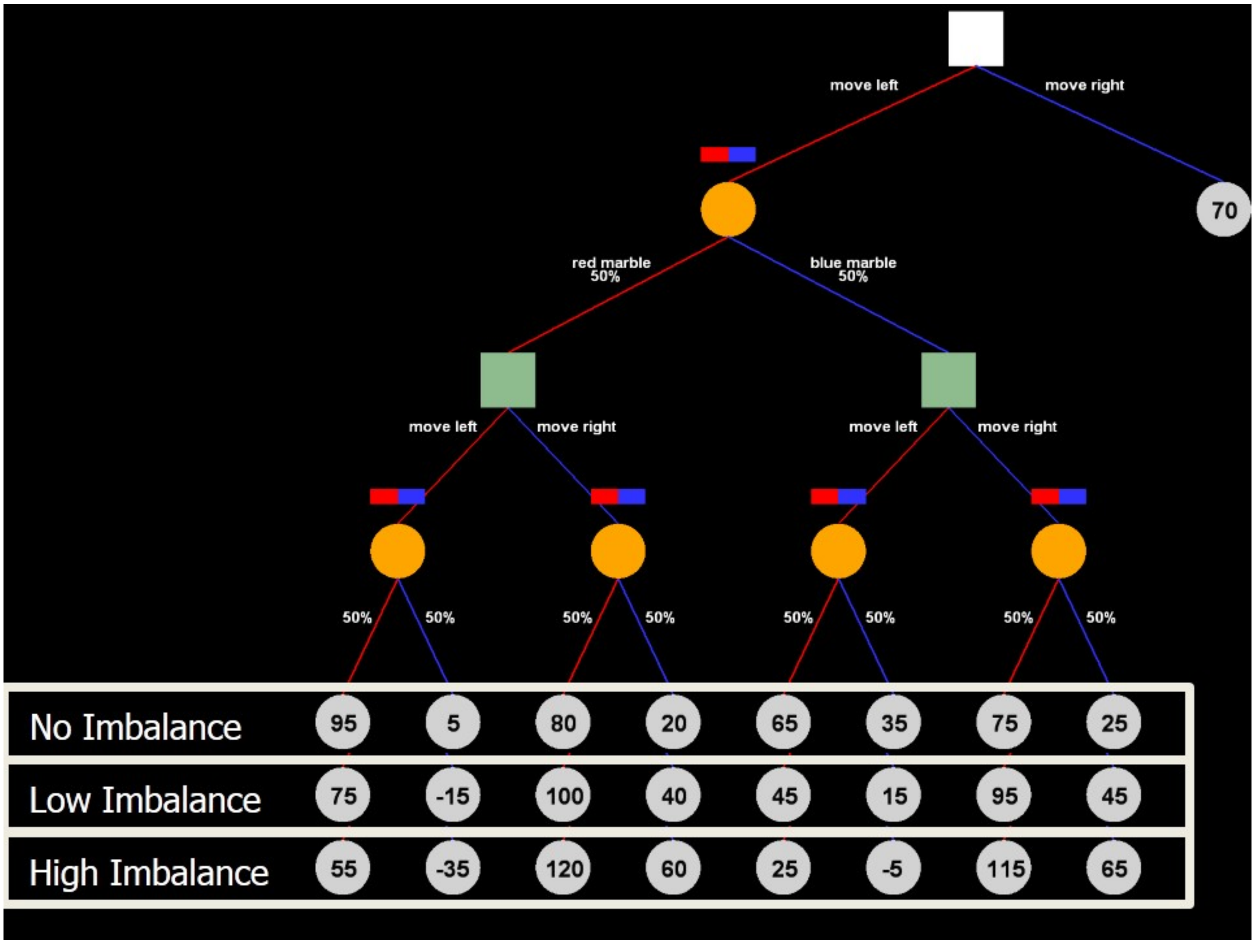

Figure 2.2. Three trees from Experiment 1, superimposed on one another to demonstrate the Imbalance manipulation. These trees are Abstract.

Procedure. Consenting participants were seated at a computer and began by watching a tutorial familiarizing them with the task. Each of the three node types was introduced and explained. Participants were told that they would earn points based on the decision outcomes from each trial, and that those would be converted to money to determine their payment. After completing the tutorial and practice trials, the main experiment began. Each trial presented a new decision tree, with a white marker at DN1. Participants made an initial choice by clicking the mouse on either CN1 or ON1. If they selected the certain payoff, the marker moved to ON1. If they chose to gamble, the marker moved to CN1, where an animation was shown of a marble being drawn from an urn. Drawing a red marble would send the marker down the path to the left (to DN2a), while a blue marble would send it down the path to the right (to DN2b). When the marker reached a second DN, participants made a second choice by click on one of the two remaining CNs. After making a final choice, the marker would move to the indicated CN, where a second marble and urn animation would determine the payoff. Once an ON was reached, either by choosing the certain payoff or by completing the gamble, its points were added to the total score displayed on screen and the trial ended.

Each decision tree appeared once during the experiment. Half of participants received the sixteen Concrete trees first, followed by the sixteen Abstract trees. The others received trees in the opposite order. Within each block of sixteen, trees with each cover story were presented in the same repeating order four times. That is, a random order was chosen for the stories (e.g. 2, 4, 1, 3) and one tree from each story was presented in that order, until all trees were presented. The four versions of each story (three level of Imbalance, plus filler) were then randomly assigned to the four cycles. This randomization method was aimed at separating similar trials, so as to avoid demand characteristics, while also counterbalancing any story order effects across individuals.

Results

My analyses focused on choices made at DN1, where planning was key to maximizing one's expected earnings. Figure 2.3 shows the effect of Imbalance on the probability of choosing the risky option at DN1 for concrete and abstract trees. As payoff imbalance across DN2a/b increased, the probability of gambling at DN1 increase from 0.52, for No Imbalance, to 0.88, for High Imbalance. This finding suggests that participants considered what choices they would make at DN2a/b before making an initial decision at DN1 and realized that they could exploit unbalanced DNs with future optimal choices. For each individual I calculated a planning score by subtracting the probability of gambling on No Imbalance trials from the probability of gambling on High Imbalance trials. Across individuals, the mean score was 0.36 (*SE* = 0.05), and this effect was highly consistent, with only one participant having a score below 0. Somewhat contrary to my hypothesis, the Imbalance effect was roughly the same for both concrete and abstract trees. Participants appeared to plan ahead equally, regardless of whether or not the decision problem was given a cover story intended to aid reasoning.

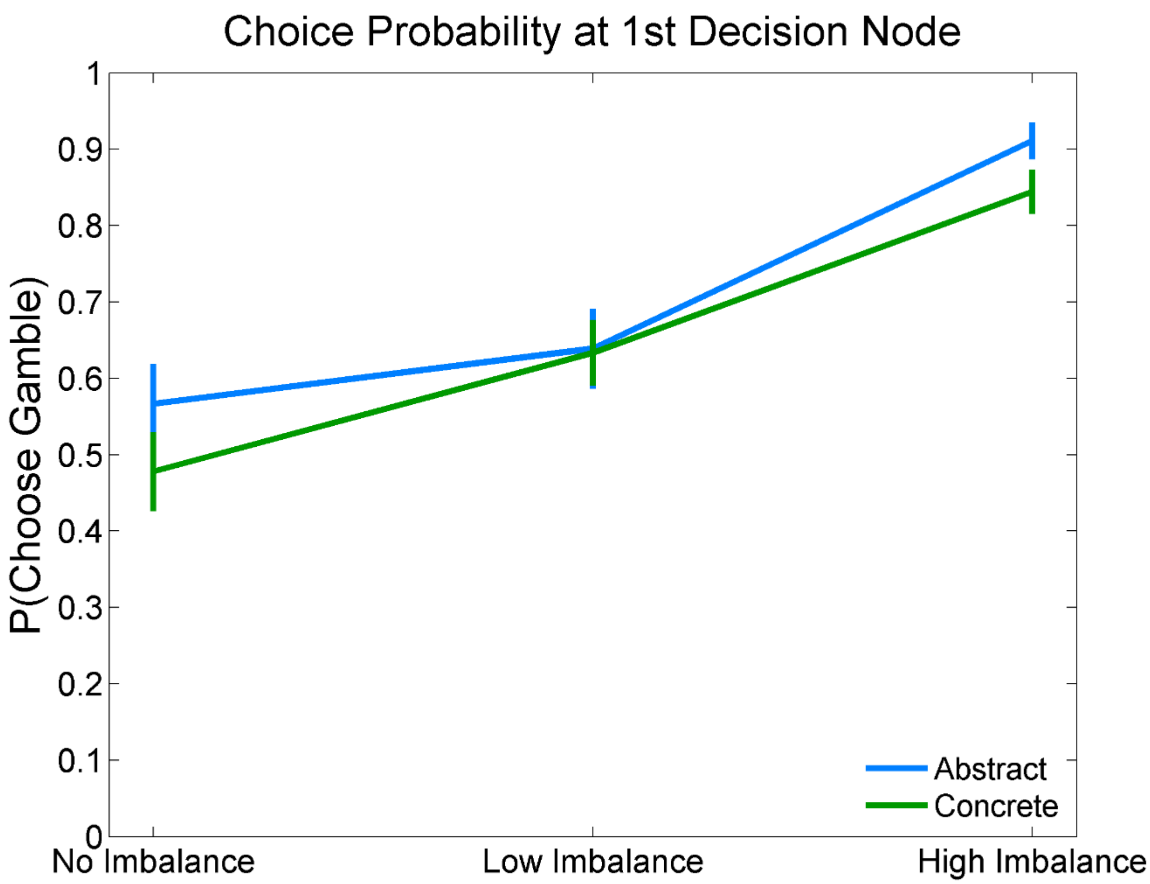

Figure 2.3. Mean probability of choosing to gamble at DN1 in Experiment 1, as a function of Imbalance and Abstractness

To confirm these effects I ran a 3 (Imbalance) x 3 (Abstractness) x 2 (Order) mixed models analysis of variance (ANOVA); however, after finding that the order in which concrete and abstract trees were present did not have an effect, this factor was collapsed for a 3 (Imbalance) x 3 (Abstractness) repeated measures ANOVA. A main effect of Imbalance was found, $F(2,44) = 36.31$, $MSE = 11.87$, $p < .001$, but there was no main effect of Abstractness, nor an Imbalance x Abstractness interaction.

Discussion

In light of the rather pessimistic findings presented in Chapter 1, the present results provide some reassurance that human planning can be studied in a controlled laboratory environment. My findings are inconsistent with the idea that people do not plan, but rather treat future DNs like random events. I also find little evidence of participants deterministically maximizing expected utility, as backward induction predicts. Instead, they appear to probabilistically increase their rate of gambling as Imbalance increases; a result difficult to reconcile with the standard normative model. In the next chapter I test a modified version of backward induction that allows for stochastic choice and find it also fails to account for critical data patterns.

The results of Experiment 1 also suggest that one need not provide a cover story for decision trees in order to elicit planning. That said, the design of this study confounds the maximum payoff of the gamble with Imbalance, so it remains a possibility that these effects could, in part, be due to risk seeking utility functions. In the future an Imbalance manipulation that uses CN probabilities to change the expected value of DN2a/b, while keeping ON values constant would help rule out this explanation. In all, the results of Experiment 1 point to a new planning mechanism. Participants do not appear to plan optimally, as backward induction predicts, nor do they seem insensitive to distribution of payoff across future DNs, as a model that treats these as random events would predict. In the next chapter I introduce a new cognitive model of dynamic decision making aimed at providing a psychologically plausible account of how people cope with complex choice scenarios.

3. DFT-P: A Cognitive Model of Dynamic Decision Making

The results of Experiment 1, coupled with the previous research in behavioral economics and cognitive psychology presented in Chapter 1, suggest the normative strategy described by backward induction is insufficient for explaining behavior in multistage decision tasks. A new theory is therefore needed. Judgment and decision making researchers have traditionally developed theories at the higher and more abstract levels, opting to understand and predict behavior using simple, intuitive concepts. Their approach is typically agnostic to the underlying mechanisms that give rise to behavioral tendencies. In this chapter I take a different approach, and introduce a new model of planning and dynamic decision making, *Decision Field Theory-Planning* (DFT-P). This model focuses on how preferences are constructed from internal cognitive processes. As its name implies, it is based on *Decision Field Theory* (Busemeyer & Townsend, 1993; Hotaling, Busemeyer, & Li, 2010; Roe, Busemeyer, & Townsend, 2001). DFT-P adopts the attention-driven information sampling and deliberation mechanism of Decision Field Theory, and extends it to situations where an individual must plan decisions in a multistage decision tree. Although the model can, in principle, be applied to situations where there are many alternatives, I will focus on the version of DFT-P for decision trees with binary DNs. I will continue to use the hypothetical decision tree scenario in Figure 1.1 for explanation, though the model can be applied to any acyclic multistage decision problem.

The central claim of DFT-P is that people sample information about complex decision scenarios through repeated rapid mental simulations. As they deliberate over which action to take at present, they imagine possible sequences of events, and simulated their own future decisions. Returning to the earlier example, when Sam imagines the consequences of asking for a raise (DN1), he does so by simulating a likely sequence of events. He might first imagine that his boss will say yes, and will offer a choice between the stocks and wages. He then envisions choosing the stock option, which he imagines will yield a large payoff because the economy will strengthen. He compares this simulated outcome to the known outcome of not asking for a raise. On the basis of this sample, Sam will take a step toward asking for a raise. However, at the next moment Sam's attention may focus on less rosy possibilities. Now he imagines that his boss rejects his request, leaving him to decide whether to quit his job or continue. He imagines quitting and then being unable to find work due to a bad job market. Comparing this to the certain outcome of not asking for raise produces subjective evidence pushing Sam toward the safe alternative. In this fashion, he repeatedly simulates sequences of events, based on what he thinks will happen in the future. Simulations differ from moment to moment as he attends to different aspects of the scenario. Sam's attentional biases may lead him to over- or under-sample certain outcomes, with significant consequences. Over time, he accumulates a sufficiently strong preference to make a choice. The reader should note that I have described DFT-P's mental simulations as proceeding via conscious thought primarily as means of developing the proper intuitions. In reality it is likely that decision makers are not explicitly aware of the fast, repeated simulations that drive their choices at a subconscious level. This is by no means a unique feature of the present model, but rather a property shared with cognitive theories of various behavioral phenomena.

DFT-P belongs to the broad class of sequential sampling models, which posit that information is accumulated in real-time during deliberation, until preference for one option reaches some critical threshold and a decision is made (see Ratcliff & Smith, 2004 for a review). When applied to risky decision making, these models propose that individuals collect information by thinking about the possible outcomes of each available action. From moment to moment, different outcomes come to mind and are compared, producing fluctuations in preference. Figure 3.1 illustrates an example deliberation process where the decision maker considers two alternatives. The horizontal axis represents deliberation time and the vertical axis represents the preference state. As attention switches stochastically, the strengths and weaknesses of each alternative cause the preference state to evolve in a noisy manner. Eventually, sufficient evidence accumulates in favor of Option A that the corresponding decision threshold is reached. Since they describe a process that extends over time, sequential sampling models can be used to simultaneously account for both choices and response times, allowing for better testing cognitive theories of decision making.

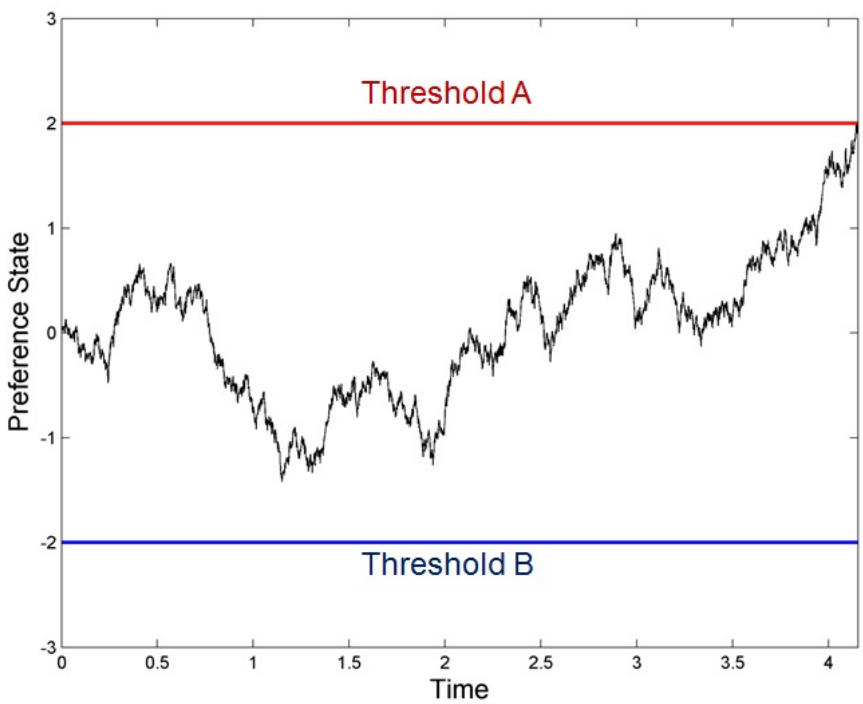

Figure 3.1. A noisy sequential sampling process for a decision between two alternatives.

Model Specification

More formally, DFT-P posits that at each time step a mental simulation is run for each alternative, and that the resulting outcomes are compared to produce a momentary valence, *V*. The preference state at time *t* is defined as the sum of the previous preference state and valence:

$$P(t) = P(t-1) + V(t-1). \tag{1}$$

P(0) is the initial preference state, and can be used to explain initial bias and carry-over effects from past experience or previous decisions. Across time, the theoretical mean input to the sampling process is therefore the expectation of the valence:

$$\mu = E[V(t)]. \tag{2}$$

For a specific decision tree, like that shown in Figure 2.1, this value can be calculated by taking the difference of mean inputs across alternatives. For each choice alternative, at any level of the tree, this is simply the sum of all possible outcomes of choosing that alternative, weighted by the likelihood of mentally simulating each outcome. At a final DN (DN2a or DN2b), for each alternative, *i*, branching to *k* possible outcomes, o_{ik}, the mean input to the deliberation process is:

$$v_i = \sum_k S(o_{ik}). \tag{3}$$

S is the *simulated utility function*, and represents the influence of selective attention on the mental simulations that drive deliberation (see the following section for more details). For decisions that involve planning a future choice (DN1), valences are produced in much the same way, with one exception. Here, Sam must mentally simulate an entire path through the tree, from DN1 to a final ON. This begins with him first imagining what will happen at CN1. Depending on the simulated event at CN1, Sam will then imagine himself at either DN2a or DN2b. He must now simulate what he will do in the future if he reaches that DN. This proceeds via the same choice mechanism describe above for 1-step decisions. After simulating this second choice, one last simulation is done for the final CN, allowing Sam to sample an ON. The value of this node is compared to the certain payoff at ON1 to produce a momentary valence. From moment to moment, Sam attends to different aspects of the scenario, and probabilistically simulates different paths through the tree. Valences accumulate according to Equation 1 until $P(t)$ exceeds some threshold, θ. The model assumes that Sam chooses a θ value at the beginning of his deliberation, based on how cautious he intends to be.

Thus, according to DFT-P, as Sam deliberates over what to do at DN1, he does not explicitly plan to make a decision at DN2a/b. Instead, moment to moment, he imagines making one choice or the other, without committing to a particular action[1]. As random fluctuations in attention produce different simulations, $V(t)$ also fluctuates across time. The variance in valences is therefore a theoretically important concept, defined as:

$$\sigma^2 = E[V(t) - \mu]^2. \tag{4}$$

This key feature of the DFT-P is used to empirically test the model against human behavior in Experiments 2 and 3.

[1] Under the present formulation of the model, Sam's imagined future choices do not constrain the decision he will make when he actually reaches DN2a/b; however, memory-based dependencies could be incorporated into DFT-P's assumptions.

Mental Simulation

For decision makers interested in maximizing expected value, it is important to sample ONs according to its likelihood. However, humans often show a preference for choices that maximize their own subject utility. The simulated utility function, *S*, introduces a similar subjectivity to DFT-P. It allows for the possibility that mental simulations are biased to over- or under-sample particular ONs. For instance, Sam might optimistically imagine that random events at CN2a-d will yield positive outcomes, leading him to over-sample gains (relative to their true likelihoods) and under-sample losses (relative to their true likelihoods). Such a bias would produce decisions based predominantly on gains, rather than losses. In its current form, *S* is defined as a function of outcome value alone (see Equation 3). This is primarily a simplifying assumption of the model, though, given the saliency of ON values in decision tree stimuli like those shown in Figure 2.1, it seems reasonable. In other scenarios, however, it is likely that additional factors, such as CN probabilities, node labels, branch lengths, etc. could also play a role in guiding mental simulations toward or away from certain outcomes. As we learn more about the factors influencing planning and multistage choice, the formulation of *S* may change to reflect other influences on likelihood of mentally simulating an outcome. This remains an important goal for future research.

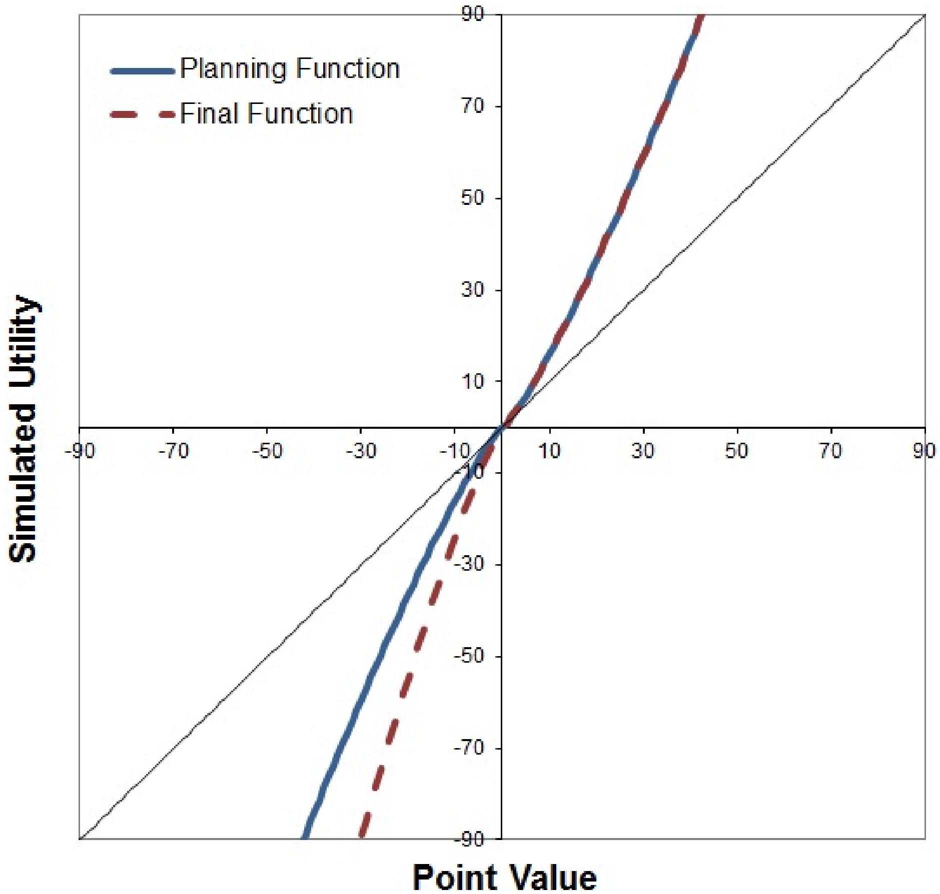

Figure 3.2. Example simulated utility functions for planned and final decisions. Planning and Final curves are identical for gains, but differ in the loss domain.

Figure 3.2 presents two example simulated utility functions. The horizontal axis represents the objective value of an outcome, while the vertical axis represents the subjective utility produced by simulating an outcome. If an individual were to simulate outcomes without bias according to their true likelihoods, then these functions would fall on the diagonal. However, DFT-P posits that individuals may distribute attention unevenly across ONs, causing mental simulations to favor particular outcomes. Any divergence of *S* from the diagonal indicates that outcomes in that range were over- or under-simulated in participants' deliberation. DFT-P allows attention to shift across stages. The solid curve corresponds to simulations done while deliberating at DN1 and imagining future choices at DN2a and DN2b. It indicates that during planning attention disproportionately favors large gains and large losses, compared to moderate ones. Thus, when simulating future random CN events an individual would tend to imagine the more extreme of the two outcomes. The degree to which *S* deviates from the diagonal denotes the amount of attentional distortion of this kind. The dashed curve in Figure 3.2 corresponds to final decisions made at DN2a/b. Its shape indicates that simulations at this stage are also biased toward extreme outcomes, however the greater deviation below the diagonal for negative outcomes means that losses tend to loom larger for final decisions. These example *S* functions could be used to produce behavior much like that observed by Busemeyer et al. (2000), where participants were initially risk seeking and planned to choose the gamble, but later became more cautious of potential losses and chose the safe alternative. As will become evident in Chapters 5 and 6, after fitting DFT-P to human choice data, the resulting simulated utility functions can provide valuable insight into how individuals sample information across planned and final decision stages[2].

Analytic Solution for Choice Proportions

Although the model posits that individuals engage in many mental simulations, obtaining predictions from DFT-P does not require computer simulations. Instead we can use the mathematical derivation:

[2] Under the current implementation of the DFT-P the probability of simulating the left versus right path descending from a CN is assumed to be the probability given by the experimenter, and the *S* function is used to account for attentional bias in sampling. I chose to represent this bias as subjective utility at ONs, rather than subjective probability weighting at CNs, for the sake of computational simplicity. Although these two methods for representing sampling bias can mimic each other, and can therefore be thought of as roughly equivalent, the probability weighting approach would be more true to the underlying theory.

$$P(choose\ risky\ option) = \frac{\exp[4 \cdot \theta \cdot \left(\frac{\mu}{\sigma}\right)] - \exp[2 \cdot (\theta - z) \cdot \left(\frac{\mu}{\sigma}\right)]}{\exp[4 \cdot \theta \cdot \left(\frac{\mu}{\sigma}\right)] - 1} \qquad (5)$$

where θ is the threshold for making a choice, and z is the initial bias in favor of the risky alternative (see Equation A1 in Busemeyer & Townsend, 1993 for the derivation). This equation highlights several key features of the theory. First, choices will generally favor the higher utility option. When μ is near 0, choices will be quite random, but as μ becomes more extreme, individuals will move toward deterministically choosing the higher utility alternative. Two factors control sensitivity to mean utility differences. The first is the decision threshold, θ. High thresholds produce slower, more deliberate choices, based on many samples, whereas low thresholds yield more random decisions based on only a few simulations. DFT-P allows for the possibility that individuals may conserve cognitive resources by setting a low threshold and conducting relatively few simulations at some DNs, but raise their threshold for others.

Payoff variability is the second factor affecting sensitivity to utility differences, and will be the focus of the next chapter. This refers to the variability in outcome values across ONs in a decision tree[3]. When ON values are highly variable, simulations will sometimes produce valences that strongly favor one option and other times strongly favor the alternative. This produces large fluctuations in preference and more random decisions. With low ON variability, each simulation will reach a similar payoff and produce a similar valence. Here, preference will accumulate more smoothly toward the higher utility option.

[3] The Imbalance variable in Experiment 1 can be seen as a special case of payoff variability, where variability across DNs is specifically manipulated. Experiments 2 and 3 involved a more general payoff variability manipulation aimed at testing DFT-P's basic prediction.

Accounting for Violations of Dynamic Consistency

In Chapter 1, I showed how several studies found systematic violations of key rational axioms underlying the normative backward induction model. It is therefore reasonable to expect the new competing model, DFT-P, to address this shortcoming. Fortunately, DFT-P has several features that allow for dynamic inconsistency. The first, and most trivial, is the stochastic nature of the deliberation process. According to the model, from moment to moment, mental simulations feed noisy input to the system. This makes preferences inherently variable, and allows for some of the inconsistencies seen in the literature. Second, individuals may set different decision thresholds for planned versus final choices, with the former being more impulsive than the latter. Third, as individuals move from planning to executing their decisions, DFT-P allows for shifts in attentional focus. Returning to the earlier example, Sam may initially give equal attention to the potential risks and rewards of asking for a raise (DN1), but may later focus primarily on potential gains when choosing between company stock and increased wages (DN2a). With his simulations at DN2a now based on a different distribution of information, preferences may change in a dynamically inconsistent way.

Relating DFT-P to Other Research

The concept of decision making through simulation is not new. In the field of naturalistic decision making (Nemeth & Klein, 2010), Klein and colleagues have proposed the *Recognition-Primed Decision* (RPD) model (Klein, 1993, 1998) to describe how people make quick decisions in complex real-world environments, such as hospitals, forest fires, and stock exchanges. According to the RPD model, people first use situational cues to identify a potential action. Next, they run a mental simulation of the likely consequence of that action. If the simulation produces an acceptable outcome the action is taken, otherwise the next most promising action is simulated and the process repeats until a satisfactory option is found. Unlike DFT-P, RPD is noncompensatory, in that it does not compare alternatives, but rather evaluates them in isolation and chooses the first to meet some criterion. This difference is understandable, considering the origins of each model. DFT-P was developed for situations with relatively low time pressure and a small number of well-defined alternatives. RPD, on the other hand, was designed to operate in highly chaotic, ill-defined scenarios where individuals are under significant pressure to make decisions quickly. Careful deliberation in such environments may not be possible.

In the field of neuroscience, Suzuki et al. (2012) conducted a study in which individuals were shown to use “direct recruitment” of their own mental processes to predict the choices of others. In one condition, participants repeatedly chose between two stimuli with unknown constant reward probabilities, p and 1-p, respectively. In another condition, participants predicted which stimulus another individual would choose. Suzuki et al. found that the same reinforcement learning model accounted for choices participants made in both conditions. That is, the same process by which people learned p, when choosing for themselves, also governed their mental simulations of others’ learning. An fMRI analysis showed the same neural activation pattern previously associated with reward prediction error was present in both conditions. Suzuki et al. concluded that the ventromedial prefrontal cortex was key to the shared representation of one’s own learned decision strategies as well as one’s simulation of others’. With DFT-P, I propose that individuals use mental simulation not only when imagining the decisions of others, but also when imagining their own future decisions for the purpose of planning.

The hypothesis that planning proceeds through forward-looking mental simulations, rather than backward induction, is also supported by research into so-called hippocampal place cells. These neurons show activation patterns in which individual cells selectively fire at specific locations along an animal's path, thus forming a kind of neural map (O'Keefe & Nadel, 1978; Wilson & McNaughton, 1993). Using experiments where rats learned to navigate branching mazes in search of food rewards, A. Johnson and Redish (2007) and Dragoi and Tonegawa (2010) found evidence that rats were simulating potential future outcomes while deliberating over which paths to take. Upon reaching a decision point in a familiar maze, animals would pause before choosing a path to continue along. During this time hippocampal place cells showed a forward-sweeping pattern of activation consistent with mental simulation. Cells would fire in a sequence corresponding to imagining walking down one branch of the maze, followed by the other. This is strikingly similar to the planning mechanism described by DFT-P, and considering that the rats were essentially learning to solve multistage decision trees, can be taken as rather strong support for the model's mental simulation framework. Pezzulo and colleagues (Chersi & Pezzulo, 2012; Pezzulo, Rigoli, & Chersi, 2013) built on this work with a computational model for how rats learned to navigate. They describe a neural circuit in which the hippocampus is primarily involved in mental simulation of possible navigation paths, while the ventral striatum is responsible for evaluating the associated reward expectancies.

Summary

DFT-P represents a new approach to understanding how individuals cope with the complexities of dynamic decision making. It posits that individuals do not explicitly commit to planned decisions. Instead, it describes a cognitive process whereby future choices are planned on the fly, through repeated mental simulation. As an individual imagines different possible sequences of events, simulated outcomes are compared and preferences evolve in a noisy fashion until a threshold is reached. Unlike the normative backward induction model, DFT-P allows for probabilistic choice, dynamic inconsistency, uncertainty about future decisions, and sensitivity to outcome variance. In the next section, I use another decision tree experiment to test this final prediction and demonstrate DFT-P's value as an explanatory tool.

4. Experiment 2: A Test of DFT-P

As noted earlier, a key prediction of DFT-P is that individuals should be sensitive to payoff variability during planning. This is because, in scenarios with many extreme outcomes, simulations will sample very different outcomes moment to moment. With valences sometimes strongly supporting one action and other times supporting the alternative, preference may swing back and forth between options, increasing the likelihood that a decision boundary is hit due to random chance. In this section I present an experiment which tests this fundamental prediction. The results show that people's choices become less sensitive to expected utility as payoff variability increases. This finding supports the mental simulation framework of DFT-P. Finally, I demonstrate how key parameters of the model control the predicted size of this effect.

The Payoff Variability Effect

The so-called *payoff variability effect* is well established for traditional single-stage decisions (Barron & Erev, 2003; Busemeyer, 1985; Erev, Ert, & Yechiam, 2008; Katz, 1962; Myers & Katz, 1962; Myers & Sadler, 1960; Myers, Suydam, & Gambino, 1965; Suydam & Myers, 1962). To demonstrate the effect, imagine a task where participants choose between pairs of the following options: 1) a low risk gamble (LR) to gain or lose 5¢, with equal probability, 2) a high risk gamble (HR) to gain or lose 33¢, with equal probability, 3) a certain loss (CL) of 1¢, and 4) a certain gain (CG) of 8¢. The critical finding first involves a comparison of the LR vs. CL condition and the HR vs. CL condition. Participants typically choose LR from the set {LR,CL} more than they choose HR from the set {HR,CL}. This result indicates risk aversion, and suggests that LR has greater utility than HR. However, participants also choose LR from the set {LR,CG} less than they choose HR from the set {HR,CG}. In contradiction to the previous effect, this constitutes risk seeking and implies that HR has greater utility than LR.

The payoff variability effect is important because it cannot be easily explained by strong utility models or ratio of strength models because these traditional approaches represent subjective utilities as static and context independent. It suggests that static models of decision making neglect important aspects of the psychological processes that produce choice. Encouragingly, Busemeyer and Townsend (1993) showed that DFT naturally predicts the payoff variability effect. This is because its noisy sampling process produces diminishing sensitivity to utility differences with greater outcome variance. This diminishing sensitivity is also a property of DFT-P, however now the prediction is that the payoff variability effect may occur even when individuals plan future decisions involving distant outcomes

The payoff variability effect has been demonstrated across many contexts from one-shot decisions under uncertainty (see Busemeyer & Townsend, 1993) to repeated decisions under ambiguity (see Erev & Barron, 2005). Busemeyer (1985) showed that the payoff variability effect occurred within individuals. The experiment presented below tests for a similar effect when future decisions are being planned in a multistage decision problem. DFT-P predicts that greater payoff variability produces noisier sampling during mental simulations, and moves choice proportions toward 0.50. Thus, for trees with a certain loss—where we expect participants to prefer the gamble—increased variability should decrease preference for the gamble. For trees with a certain gain—where we expect participants to prefer the safe alternative—increased variability should increase preference for the gamble. Will preferences show a payoff variability effect, even when outcomes are distant, as DFT-P predicts, or will people simply attend to expected utilities and ignore variability, as prescribed by backward induction?

Method

Participants. Forty three students from Indiana University were recruited via flyers posted on campus and an advertisement on the university's job listing forum for a "paid decision making study". Payment was based on performance in the experiment, but participants received an average of $8 apiece. Due to the extremeness of some ONs, actual payments were calculated using a logistic function centered on $8 with asymptotes of $5 and $11.

Stimuli and Design. Experiment 2 used the same paradigm and decision tree structures introduced in Experiment 1. Again, participants began by choosing between a certain payoff at ON1 and a multistage gamble at CN1. Since concrete cover stories did not have a substantial effect on decisions in Experiment 1, only abstract trees were used in Experiment 2. One unique aspect of the present task is that the gambles involved a second decision at DN2a/b between a *disadvantageous CN* (with a negative expected value (EV)) and a *neutral CN*, which offered a gain or loss of *X* points, with equal probability (i.e. EV = 0). In this sense, DN2a and DN2b were equivalent, though ON values were slightly different across the two DNs, so as to make the experimental design less obvious to participants.

Three critical factors were manipulated factorially across trials: *certain payoff* amount, *variability* of neutral CNs, and *negativity* of disadvantageous CNs. The certain payoff amount at ON1 had three levels: 8, 1, or -1 points. These were selected to assess the effect of payoff variability in circumstances where participants would typically prefer to gamble (i.e. ON1 = -1) as well as in those where participants would typically prefer the safe option (i.e. ON1 = 8). Variability of the neutral CNs was high (e.g. gain or lose 33), medium (e.g. gain or lose 19), or low (e.g. gain or lose 5). Figure 4.1 demonstrates this variability manipulation with three trees superimposed on one another. Since, while deliberating at DN1, participants should easily recognize which future decisions they must make to avoid the disadvantageous CNs, their valuation of the gamble was predicted to derive primarily from the neutral CNs. By independently manipulating the value of the certain payoffs (ON1) and the variability of the gambles (i.e. the size *X*) I could assess the impact of variability across different dynamic decision contexts. DFT-P predicts that, across all levels of certain payoff, individuals' preferences for gambling at DN1 should move toward indifference as variability increases.

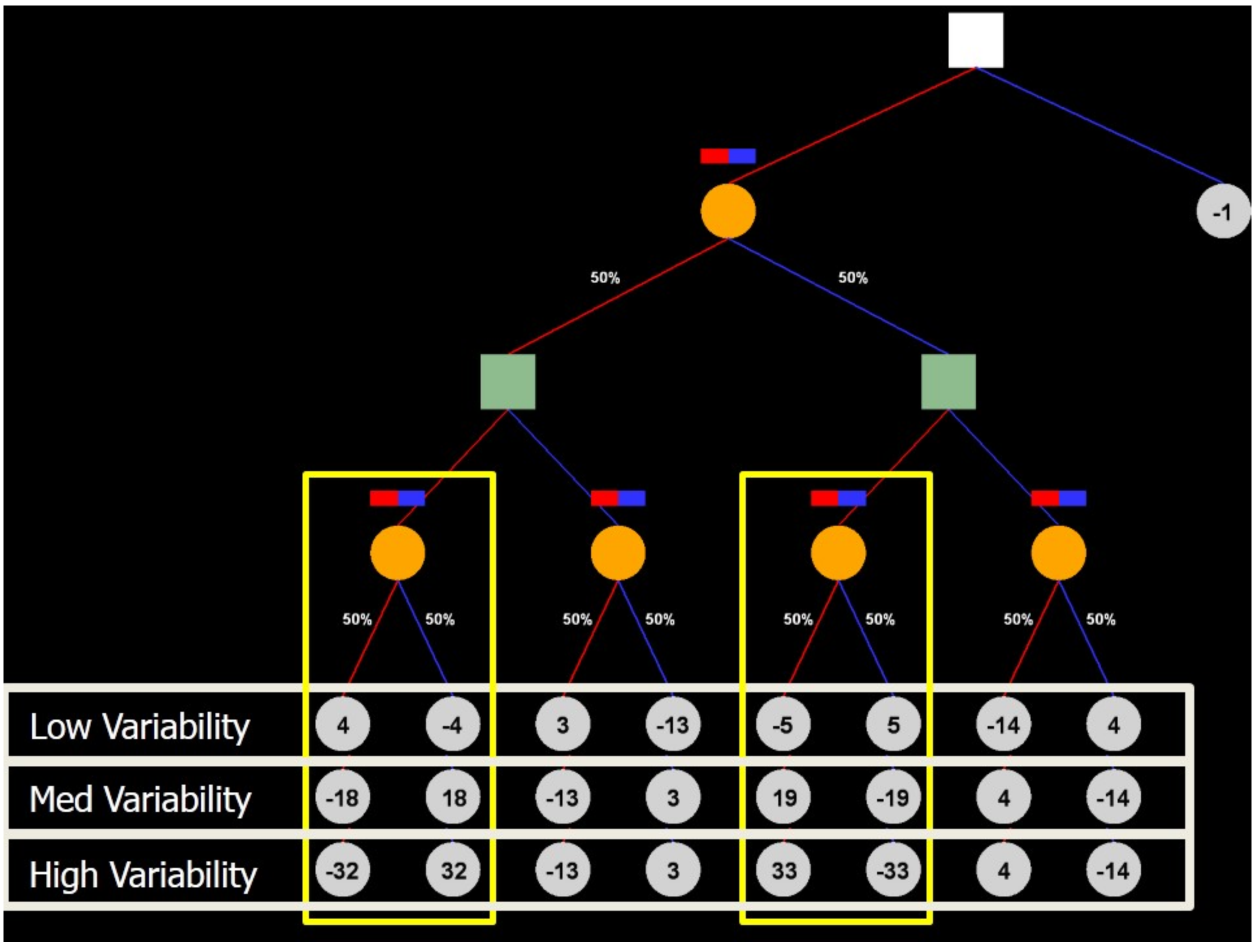

Figure 4.1. Three trees from Experiment 2, superimposed on one another to demonstrate the Payoff Variability manipulation. These trees have low Negativity and Certain Payoff of -1.

The model also makes a potentially interesting prediction regarding the negativity of disadvantageous CNs. According to DFT-P, people do not form a single static plan regarding what action to take at a future DN. Instead, each time they simulate a path through the tree, they imagine making a decision at each DN they meet. As a result, the proportion of mental simulations that involve moving left versus right at a given DN determine the relative impact of each payoff on earlier decisions. Returning to the example in Figure 1.1, if while deciding whether to ask for a raise Sam always imagines choosing company stock over increased wages, the outcomes related to better wages will never be sampled. In some rare cases this could lead to a counterintuitive effect, whereby decreasing some of the payoffs in a multistage gamble may increase Sam's willingness to gamble. For instance, if ON2 and ON3 are slightly more attractive than ON4 and ON5, while deliberating at DN1 Sam will sometimes simulate moving left at DN2a and other times simulate moving right, and each of ON2-5 will contribute to his valuation of CN1. However, if ON4 and ON5 are made less attractive (i.e. more negative) Sam's simulations at DN2a will increasingly result in choices for CN2a, meaning that ON4 and ON5, which have lower values, will be sampled less and less. Thus, it is possible that making parts of the gamble (ON4 and ON5 in this example) more negative may actually increase people's willingness to gamble because fewer of their mental simulations result in negative outcomes. In the present experiment, negativity of the disadvantageous CNs was high (EV = -15), medium (EV = -10), or low (EV = -5). I was interested in testing whether high negativity would actually increase gambling at DN1 due to mental simulations almost always avoiding the obviously disadvantageous CNs. Perhaps when negativity was low participants would be less certain about what future choice to make at DN2a/b and would therefore end up sampling more from disadvantageous CNs, making the gamble look less attractive.

A factorial combination of the three independent variables produced 27 unique tree types. Each appeared once per block of trials. I created equivalent versions of each tree type to use across blocks by rearranging ONs such that the underlying structure was preserved but the appearance of the tree was unique.

Procedure. The same procedure as in Experiment 1 was used. Tree types were shown in a random order, once per block, with each individual completing four blocks, for a total of 108 trials.

Results & Discussion

To more directly compare my findings to those in the payoff variability effect literature, I focused my analyses on individuals' willingness to choose the risky gamble versus the certain payoff at DN1. Figure 4.2 shows the effect of payoff variability and the certain payoff amount on the probability of choosing the risky option at DN1. As expected, for each level of certain payoff choice proportions moved toward 0.50 as variability increased. When the certain payoff was -1, participants appeared risk-averse, and the probability of gambling decreases from 0.76 to 0.63 as variability increased. However, when the certain payoff was 8, participants appeared risk-seeking, and the probability of gambling increases from 0.17 to 0.27 with variability. This confirms the fundamental prediction of DFT-P that increased payoff variability leads to more variable mental simulations and therefore more random choices. To check that this payoff variability effect occurred at the individual level, a contrast score was calculated for each participant by comparing the difference between certain payoff -1 and 8 for low versus high variability. Across individuals, a mean contrast value of 0.23 was found (*SE* = 0.05), indicating that the effect of certain payoff value on the probability of gambling was substantially less when payoff variability was high. This was highly consistent across individuals, with thirty of forty three participant having a positive contrast score.

The probability of making the risky choice at DN1 decreased slightly (from 0.52 to 0.46) as the negativity of the disadvantageous CNs increased. This effect, though not what I had hypothesized, is nonetheless consistent with DFT-P, which predicts that increased negativity will increase gambling for only a narrow range of decision scenarios and model parameters. One explanation for the obtained result is that the negativity manipulation may have been too strong. That is, as negativity increased, the decrease in the value of disadvantageous ONs was greater than the decrease in frequency of sampling those ONs. This would lead to more negative valences, on average, associated with the gamble, and a decrease in choices for the gamble. In support of the notion that the Negativity manipulation swamped any decrease in the likelihood of sampling disadvantageous ONs, the vast majority of choices at DN2a/b were for a neutral CN, with only 3.78% of responses in favor of the disadvantageous CNs. To investigate this matter in the future would therefore require a more subtle approach where participants are less certain of their choices at DN2a/b.

A 3 (variability) x 3 (negativity) x 3 (certain payoff) repeated measures ANOVA confirmed these observations. Main effects were found for certain payoff, $F(2,42) = 76.13$, $MSE = 83.67$, $p < .001$, and negativity, $F(2,42) = 12.39$, $MSE = 2.28$, $p < .001$, along with a negativity x certain payoff interaction, $F(4,84) = 3.47$, $MSE = 0.35$, $p < .01$. Importantly, a significant interaction between variability and certain payoff, $F(4,84) = 10.60$, $MSE = 1.88$, $p < .001$, was obtained, confirming that the effect of payoff variability differed significantly across certain payoff amounts.

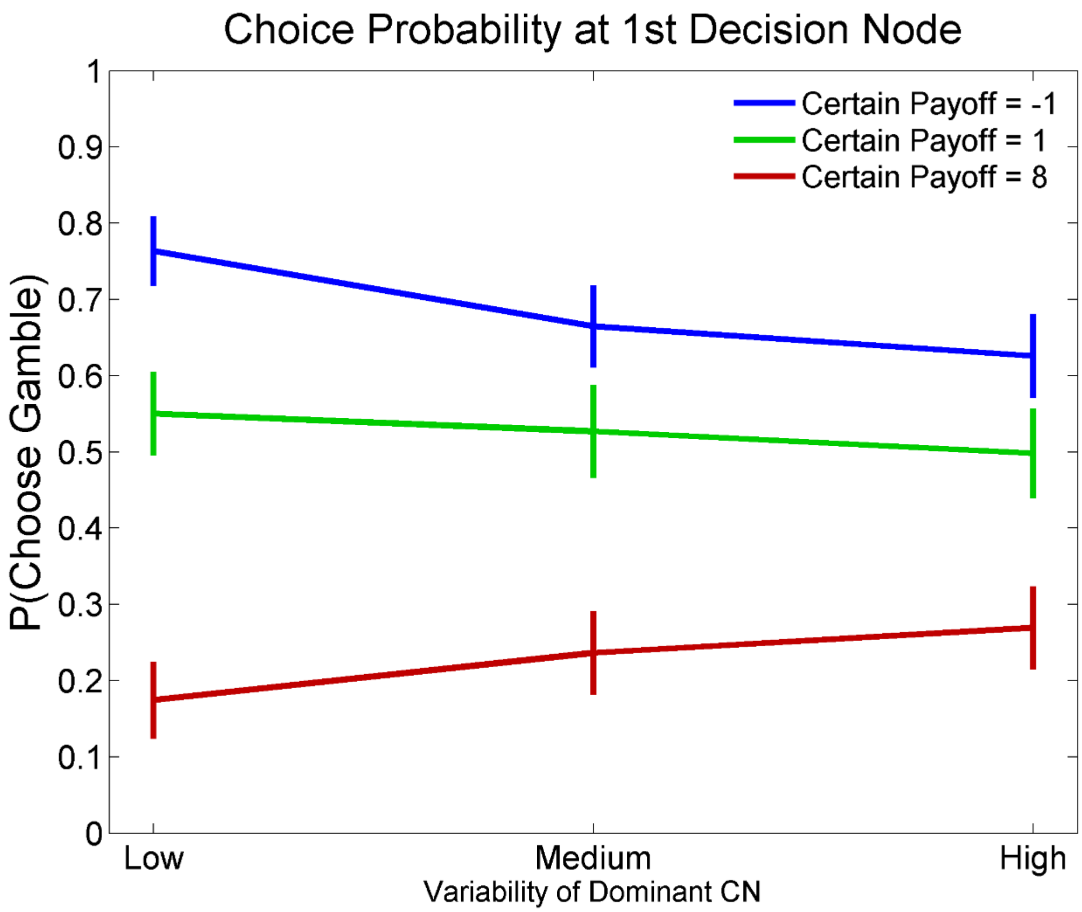

Figure 4.2. Mean probability of choosing to gamble at DN1 of Experiment 2, as a function of Variability and Certain Payoff Amount. Error bars are between-subject standard errors.

<u>Modeling Results</u>. A *static model* was built to test whether it is possible to account for the observed payoff variability effect without the dynamic sequential simulation process proposed by DFT-P. This model assumes that choice probability is a logistic function of the subjective utility difference across two alternatives. Although it assumes that individuals work backward from the end of the tree, unlike in a pure backward induction model planning is probabilistic. This means that planned decisions yield a choice probability that is used to compute an expected utility for future DNs. The model relaxes the assumption that individuals prune tree branches that they plan to avoid. Instead, it posits that individuals use their uncertainty about future decisions

to weight the outcomes on each side of a DN. Critically, the static model is backward-looking and insensitive to payoff variability. To test if it might nonetheless account for the payoff variability effect, a parameterized version of model was fit to the response data from Experiment 2. To convert ON values to subjective utility, the value of ON1 was first subtracted from each of ON2-ON9. This results in a utility function centered on the certain payoff, where outcomes are defined as deviations from ON1. Next, a standard nonlinear utility function, of the form U(x) = x^{α} for gains and U(x) = $-\lambda|x|^{\beta}$ for losses, was used. Choice proportions for DN2a/b were calculated using the function $\frac{1}{1+e^{-sd}}$, where s controls sensitivity to utility difference d, and values were propagated up the tree according to the process described above. The best fitting values for α, β, λ, and s were found using the Nelder-Mead method (Nelder & Mead, 1965) for maximizing the likelihood of observed data. Figure 4.3 show its predictions for choices at DN1. Clearly the static model does not yield the observed convergence toward 0.50. This demonstrates that even with a nonlinear utility function and probabilistic backward induction, which lend the normative model considerable flexibility, a static model cannot produce the payoff variability effect. Instead, regardless of the utility parameters selected, these parameters must produce a qualitatively similar effect on preference (either risk-seeking or risk-averse) for each level of Certain Payoff.

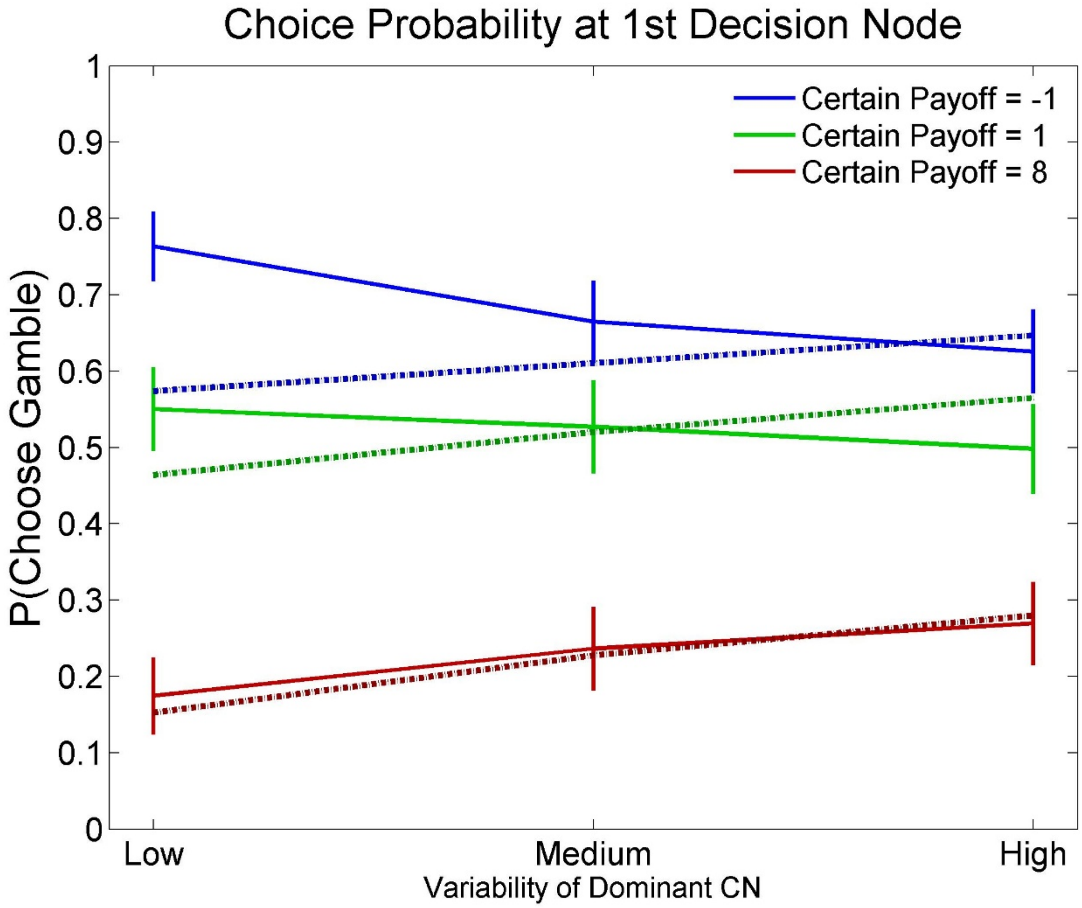

Figure 4.3. Predictions of static model (dashed lines) and mean data for DN1 of Experiment 2.

I turn next to DFT-P. An extremely simple version of the model was created to test if, as expected, the model's forward-looking, simulation-based online planning produced a qualitatively similar convergence of choice proportions toward indifferences. For simplicity's sake a parameter-free linear simulated utility function was assumed (see Figure 3.2). This represents the idea that mental simulations sampled ON values according to their true likelihoods, without any attention-driven bias to over- or under-sample outcomes based on their magnitudes. The model also assumes no bias for or against risky alternatives. To test the bounds of DFT-P's ability to produce the payoff variability effect, several threshold parameters were selected. Since choices at DN2a/b almost always favored the obviously superior neutral CN, a high threshold (θ_2 = 10) was used to predict choices at DN2a/b. This parameter was used both when simulating future choices for DN2a/b, during deliberations at DN1, and when making final choices at DN2a/b. It denotes the amount of subjective evidence collected each time a future decision is simulated, as well as before a final decision is made to end the trial at DN2a/b. Note that the full version of DFT-P allows for the possibility that individuals may set different thresholds for planned versus final decisions. A different threshold parameter, θ_1, controlled the amount of preference required to make a choice at DN1. Figure 4.4 shows the model's predictions for three different values of θ_1. Across the entire range, DFT-P predicts a payoff variability effect similar to that observed in Experiment 2. For high values of θ_1 the effect is large because a high threshold allows for many simulations, and therefore high sensitivity to mean payoff differences when Variability is low. When θ_1 is low, the effect is still present, but is much less striking.

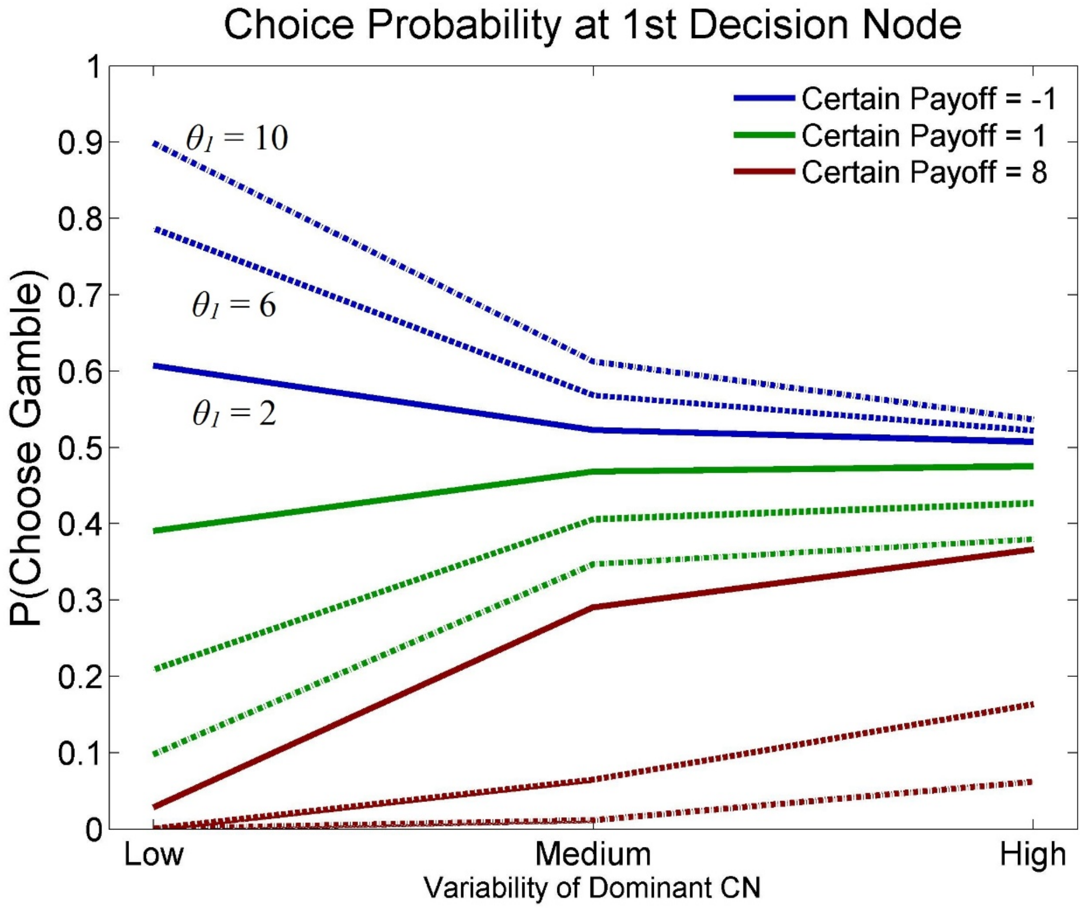

Figure 4.4. Predictions of DFT-P at DN1 for three values of θ_1. Solid lines are for θ_1 = 2. Small dotted line are for θ_1 = 6. Solid lines are for θ_1 = 10

Admittedly, this simplified form of the model yields an imperfect fit[4], however the goal here is not to find the optimal parameters for a quantitative fit to data. Rather, I aim to demonstrate that even a simple version of DFT-P naturally predicts the sort of interaction between certain payoff amount and outcome variability observed in Experiment 2. In the next section, I test a fully parameterized version of the model and evaluate its quantitative fit to response data.

---

[4] For instance, because the model assumes no response bias or nonlinear utility functions it predicts that when the certain payoff is 1 choice probabilities will be below 0.50 (since the gamble has an expected value of 0) and will increase with variability. In the next section I use a response bias parameter to account for participants' overall tendency to prefer gambling at DN1.

5. Experiment 3: Replication and Quantitative Fit

In Experiment 2 I found that participants were sensitive to ON variability in a way consistent with the payoff variability effect. I presented two models, the Static Model and a simplified version of DFT-P. The Static model failed to produce the interaction between payoff variability and Certain Payoff Amount observed in participants' choices. DFT-P, however, was shown to produce a qualitatively correct pattern of results, across a range of decision threshold values. In this chapter I present a study replicating the payoff variability effect using a slightly improved experimental design. In Experiment 2, the decision at DN2a/b was quite easy, and participants had little trouble identifying the neutral CNs. This is not ideal, from the perspective of model fitting, as many models would correctly predict that participants almost always choose the higher expected value CNs at DN2a/b. In the present experiment the expected values of CN2-CN5 were made much more similar in an effort to increase the difficulty of decision at DN2a/b. This necessitates a removal of the Negativity manipulation, though, as explained below, the design of Experiment 3 still allows for a test of DFT-P's counterintuitive negativity prediction. After presenting the results of Experiment 3, I use a fuller version of DFT-P to produce a quantitative fit to group data.

Method

Participants. Forty students from Indiana University were recruited and paid using the same procedure as in Experiment 2.

Design. Experiment 3 used a design similar to Experiment 2. Again, DN1 involved choosing between a certain payoff at ON1 and a multistage gamble at CN1 and DN2a/b consisted of a decision between two CNs. However, in Experiment 3 the difference in expected value across CN2-CN5 was small. On all trials both DN2a and DN2b presented a choice between an *advantageous* and a *disadvantageous* CN. On one third of trials (*Version 1*) the disadvantageous CN had an expected value of -0.5 and advantageous CN had an expected value of 0. On another third of trials (*Version 2*) the disadvantageous CN had an expected value of 0 and advantageous CN had an expected value of 0.5. On the final third of trials (*Version 3*) the disadvantageous CN had an expected value of -0.5 and advantageous CN had an expected value of 0.5. The version variable served to obscure the experimental design, while also increasing the number of ON values that appeared in the experiment. This reduced the likelihood that individuals would use simple decision strategies based on associations between ON values and the advantageousness of the CNs. Furthermore, comparing Version 3 to Version 2 allows for a second test of DFT-P's counterintuitive prediction that reducing the expected value of disadvantageous CN may increase people's willingness to gamble by increasing the likelihood that mental simulations sample ONs from advantageous CNs.

Certain payoff amount and payoff variability were manipulated similar to as in Experiment 2. The certain payoff at ON1 was 8, 3, or -1 points. Unlike in Experiment 2, payoff variability for both advantageous and disadvantageous CNs was manipulated together. Figure 5.1 demonstrates high, medium, and low variance trees superimposed on one another. Again, DN2a and DN2b were equivalent, in terms of the expected value of the CNs below each. A 3 x 3 x 3 factorial combination of independent variables produced 27 unique tree types. Each appeared once per block of trials. I again created equivalent versions of each tree type to use across blocks by rearranging ONs to create superficially unique trees.

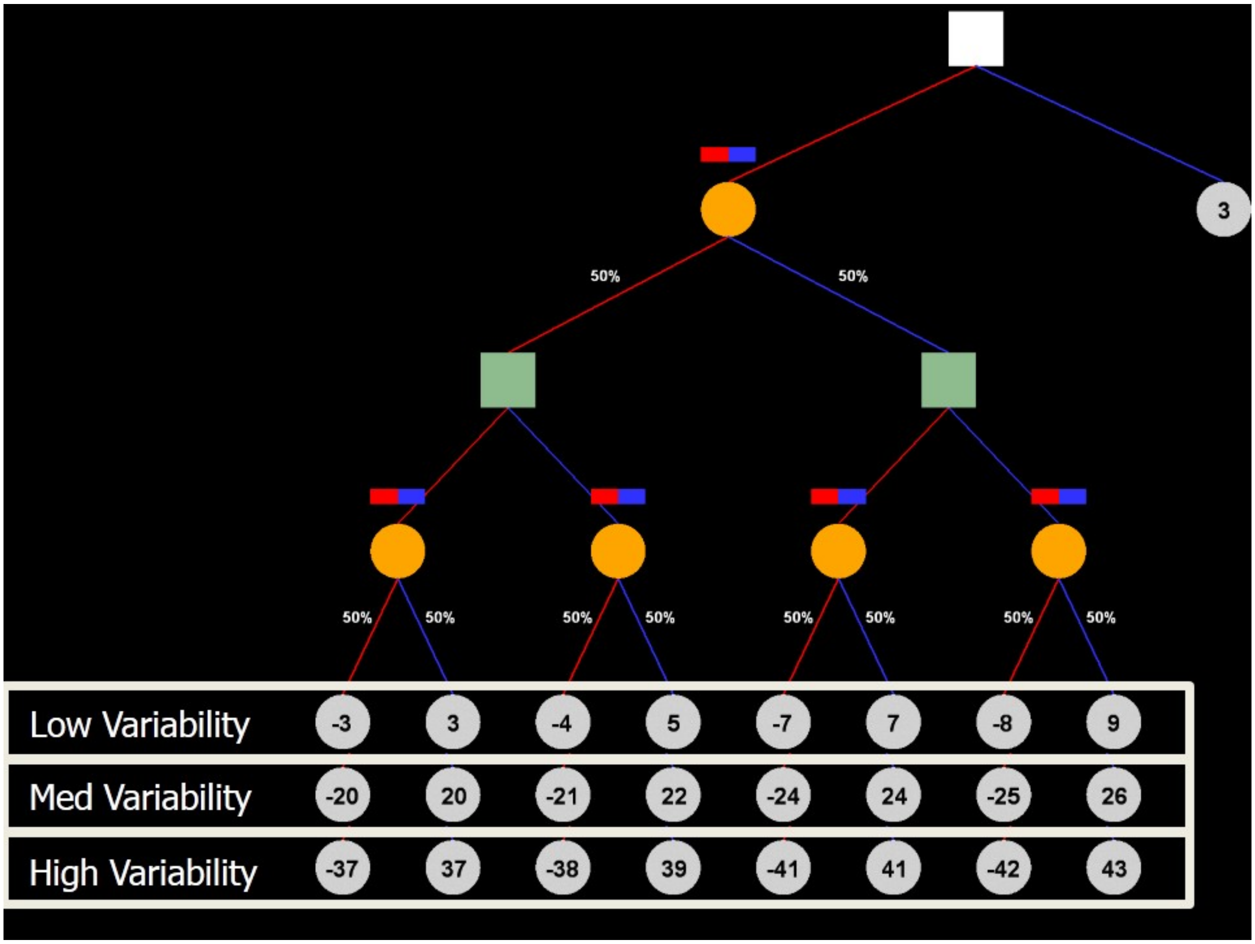

Figure 5.1. Three trees from Experiment 3, superimposed on one another to demonstrate the Payoff Variability manipulation. These trees are Version 2 and have a Certain Payoff of 3.

Procedure. Experiment 3 used a procedure identical to that of Experiment 2, except that each individual completing five blocks of decision tree trials, for a total of 135.

Results

Figure 5.2 shows that a familiar pattern of results was obtained. As in the previous experiment, as payoff variability increased choice proportions at DN1 moved toward 0.50 for each level of certain payoff. When the certain payoff was -1, participants again appeared risk-averse, and the probability of gambling decreases from 0.86 to 0.72 as variance increased. When the certain payoff was 8, participants appeared risk-seeking, and the probability of gambling increases from 0.19 to 0.38 with variance. Unsurprisingly, when the certain payoff was 3, participants preferred the safe alternative, mirroring their preference of the certain payoff of 1 in Experiment 2. Contrast scores were calculated for each individual as in Experiment 2, yielding a mean of 0.32 (SE = 0.05). Twenty eight out of forty participant had a positive contrast score. Thus, the payoff variability effect appeared again, though now with a different set of decision tree stimuli and a second set of participants. No difference was found in the probability of gambling at DN1 between Version 2 and Version 3 trees. As hoped, participants were more variable in their choices at DN2a/b, with 70.64% of responses in favor of the advantageous CNs. A 3 (variability) x 3 (Version) x 3 (certain payoff) repeated measures ANOVA confirmed these results. A main effect of certain payoff, $F(2,39) = 76.51$, $MSE = 102.75$, $p < .001$ was found, along with a variability x certain payoff interaction, $F(4,78) = 19.68$, $MSE = 4.84$, $p < .001$. The latter confirms that payoff variability moved preferences in opposite direction, depending on the level of certain payoff amount.

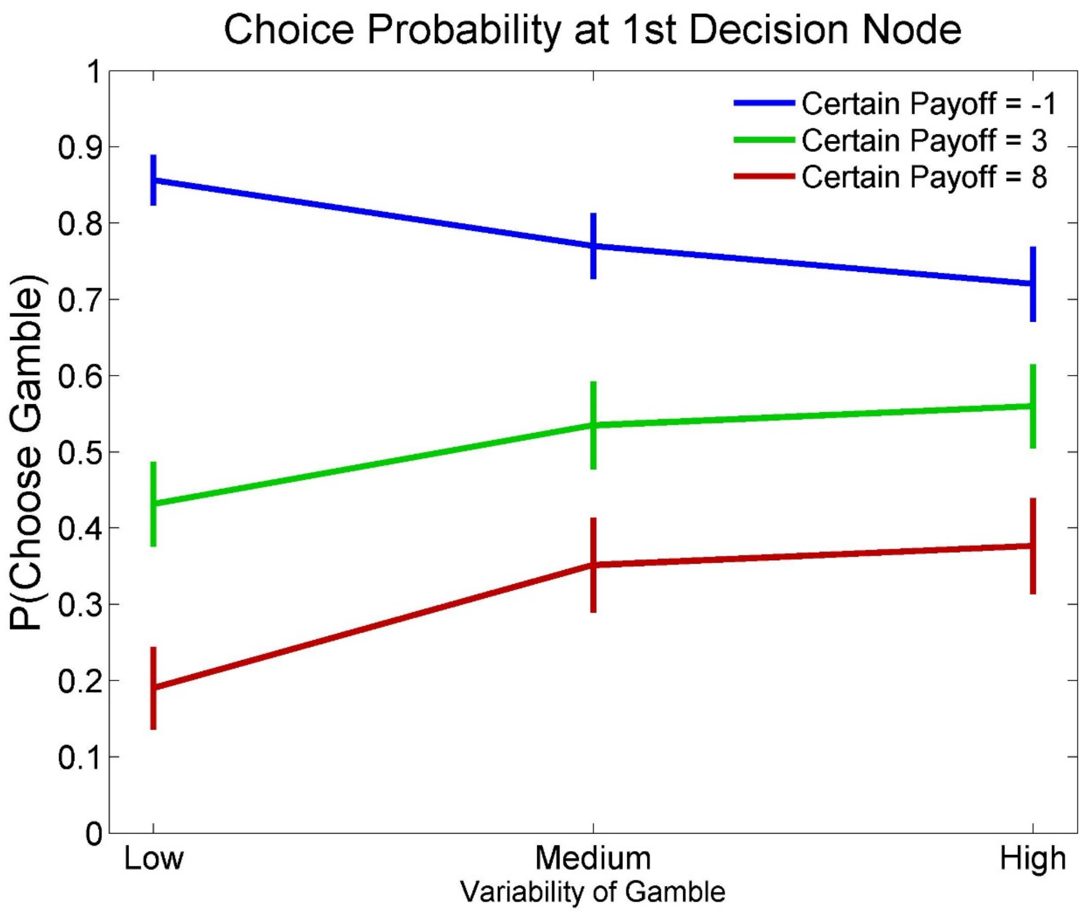

Figure 5.2. Mean probability of choosing to gamble at DN1 of Experiment 3, as a function of Variability and Certain Payoff Amount. Error bars are between-subject standard errors.

Modeling Results. Having previously demonstrated that a simplified version of DFT-P predicted a patterns of results consistent with the payoff variability effect, I now turn to a more rigorous test of the model. For this I employ a fully parameterized version that uses seven free parameters to predict the mean choice probabilities across the eighty one DNs in the experiment. As seen in Figure 5.3, DFT-P accounts very well for participants' choices across all 81 data points. Moreover, Figure 5.4 shows that DFT-P provides a good fit to the observed payoff variability effect.

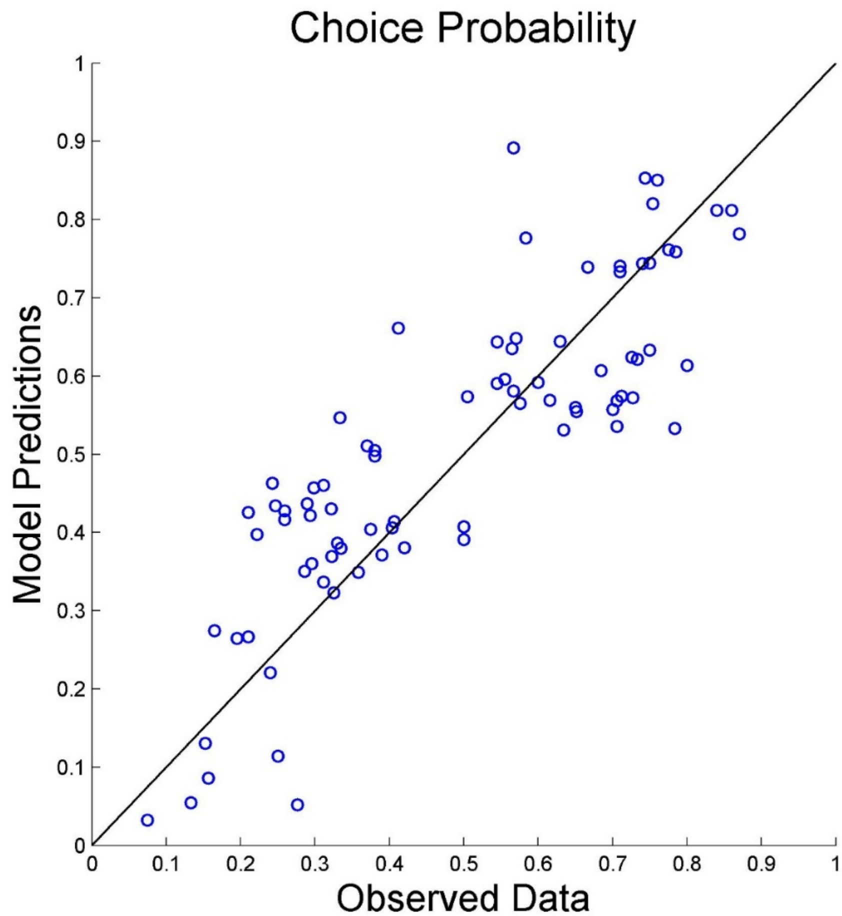

Figure 5.3. Scatter plot of mean choice probabilities from each DN in Experiment 3, compared to the predictions of DFT-P. Probabilities for DN2b are inverted to reduce overlapping of points.

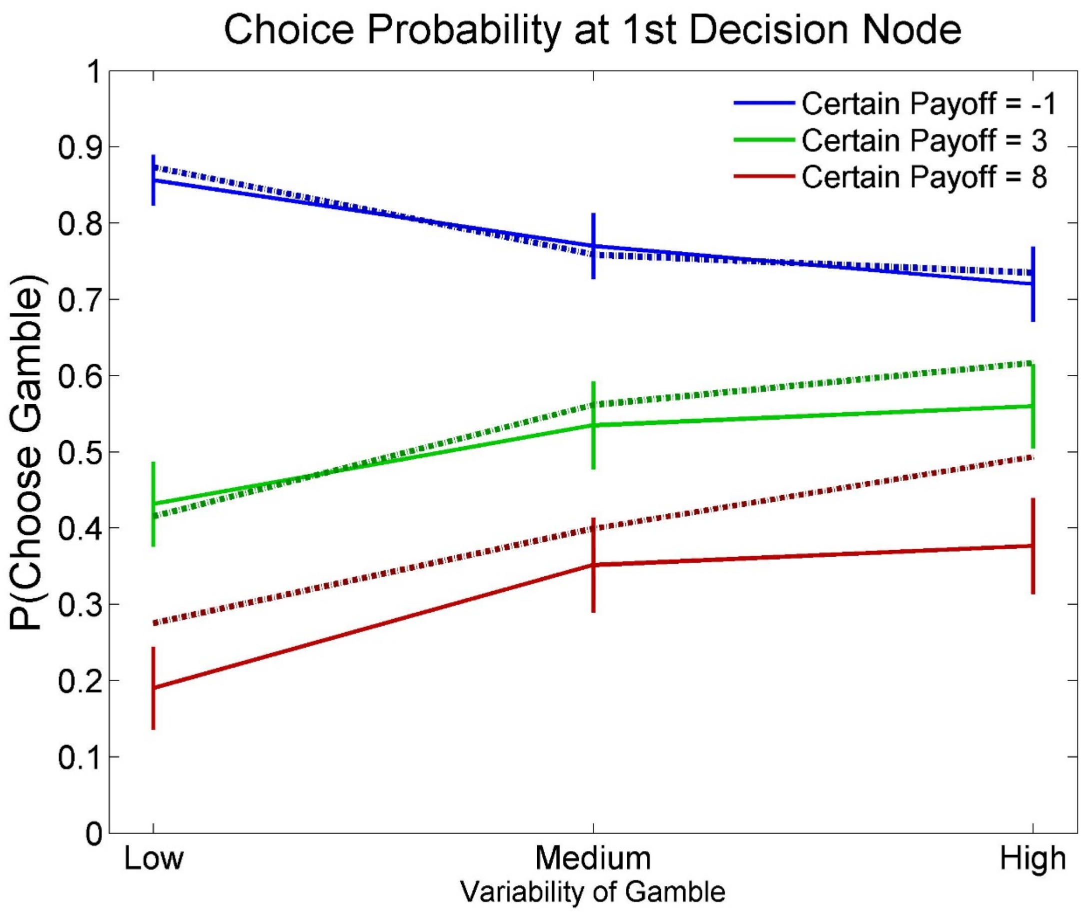

Figure 5.4. Predictions of DFT-P (dashed lines) and mean data for DN1 of Experiment 3.

The model's parameter values also provide insight into people's decision strategies. As in the previous chapter, I used two decision threshold parameters to represent the different amounts of deliberation participants engaged in at different points in the decision tree. Again, θ_1 was used for choices at DN1, while θ_2 was used for both planned and final decisions at DN2a/b. As shown in Table 5.1, a higher decision threshold (θ_2 = 4.378) was called for at DN2a/b than at DN1 (θ_1 = 0.868). Thus, when simulating future choices at DN1 participants based their initial actions on relatively few simulations. This make sense because mental simulation requires substantial cognitive resources and each simulation from DN1 requires a simulated decision at DN2a/b.

When later faced with a simpler final choice between immediate options at DN2a/b, participants were more careful, set a higher decision threshold, and accumulated more evidence before making a choice. A response bias in favor of the risky alternative was found, as indicated by a *z* value above 0, signifying that, all things being equal, participants preferred to gamble rather than take the safe option.

Table 5.1. Best fitting parameters of DFT-P for Experiment 3.

Best Fitting Parameters of DFT-P

Parameter		Description
z	0.220	Response bias in favor of risky alternative
θ_1	0.868	Decision threshold at DN1.
θ_2	4.378	Decision threshold at DN2a/b.
α_1	2.685	Over-simulating bias for extreme outcomes at DN1.
λ_1	0.623	Under-simulating bias for losses at DN1.
α_2	1.199	Over-simulating bias for extreme outcomes at DN2a/b.
λ_2	0.311	Under-simulating bias for losses at DN2a/b.

The other parameters of the model determine the simulated utility functions from which μ is calculated. It is common practice to posit that outcomes are judged relative to some reference. Model comparisons and participants' reports led me to define gains and losses on each trial as deviations from the certain payoff (ON1). I use a standard utility formula where the simulated utility of x points is x^{α} for gains and $-\lambda|x|^{\alpha}$ for losses. α represents how much selective attention causes participants to over-simulate (i.e. simulate more often than the true probability dictates) outcome values as they become more extreme, and λ represents the overall bias in the likelihood of simulating gains vs. losses. As with the decision thresholds, these simulated utility parameters change across decision stages to represent differences in sampling biases at DN1 versus DN2a/b. Figure 5.5 provides a comparison of the simulated utility functions for planned and final choices. Recall that if participants sampled outcomes according to their true probabilities, then these functions would run through the diagonal, indicating that there was no attentional distortion. The solid curve in Figure 5.5 shows the utility function derived for DN1 (using from α_1, and λ_1) as participants simulated the potential outcomes from gambling versus taking the certain payoff. Its divergence from the diagonal indicates that attention disproportionately favored large gains and losses, compared to moderate ones, during planning. λ_1 makes this most pronounced for negative outcomes, with small losses being largely ignored. The dashed line in Figure 5.5 is derived from α_2 and λ_2, and shows how mental simulations changes for final choices made at DN2a/b. The bias for simulating extreme outcomes diminished, while attention shifted away from potential losses. Hence, in this context of immediate payoffs, participants sampled each ON value at a rate closer to its true probability, but tended to imagine chance events as yielding gains, rather than losses. This means that participants made final decisions primarily by comparing the gains offered by each alternative.

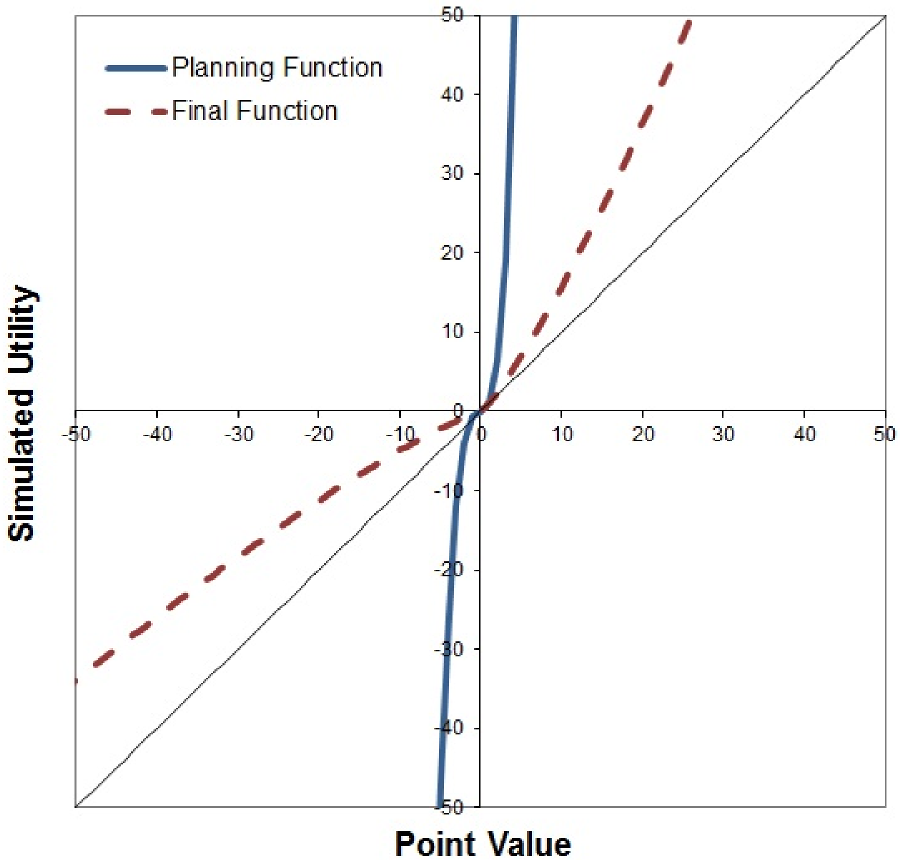

Figure 5.5. Simulated utility functions for planned and final decisions, derived from parameters of DFT-P in Experiment 3.

Discussion

Experiment 3 replicated the central finding from Experiment 2. When a complex multistage gamble was pitted against a certain gain, participants preferred the safe option, but this preference decreases as the riskiness of the gamble increased. In contrast, when the safe option was a certain loss, the gamble was preferred, but participants gambled less as riskiness increased. Again, these seemingly contradictory results portray individuals as both risk-seeking and risk-averse, depending on the context, and cannot be accounted for by backward induction or static utility models. This pattern naturally falls out of the noisy sequential simulation process driving DFT-P.

By fitting a fuller version of the model to the mean group data I demonstrated its ability to predict observed data precisely. The best fitting parameter values also provided insight into how participants collected evidence, and how this changed across planned and final decision stages. Participants were found to make fast choices, based on relatively few simulations, at DN1, but to collect more evidence and deliberate more extensively at the final decision stage. Overall, people tended to focus on extreme outcomes, good or bad, while thinking ahead about future decisions. For final choices between more immediate prospects, however, participants largely ignored potential losses and focused primarily on the large gains each alternative offered. In all, the experimental results and modeling provide strong support for the predictions of DFT-P, and shed new light on the topics of planning and dynamic decision making.

6. Fitting DFT-P to Individual Data

In the previous chapter I fit DFT-P to group data from Experiment 3, and showed that it accounted well for the central trends in participant's behavior. In this chapter I demonstrate the model's ability to fit data at the individual level, and therefore provide insight into differences in decision making across individuals. I again use the results of Experiment 3, selecting a small number of participants who exhibited distinct preferences. The goal of this chapter is to showcase how DFT-P can be used to discover strategy differences across individual decision makers. Although Experiment 3 yielded a substantial amount of data from each participant, its single day design is not ideally suited to the task at hand. That said, it is more than adequate for exploratory fits aimed at demonstrating DFT-P's capabilities. I have selected three participants from Experiment 3, who each display a unique pattern of response data. To each I fit a version of DFT-P similar to that used in Chapter 5, with one exception. In addition to the parameters used in Chapter 5, I have relaxed that assumption that individuals use the same decision threshold at DN2a/b for both simulated future choices (while deliberating at DN1) and making final choices. Thus, the present models use three separate decision thresholds, rather than two, for a total of eight free parameters.

Participant 1

Participant 1 shows a pattern similar to that of the group. Figures 6.1 and 6.2 give mean choice probabilities for DN1 and DN2a/b, respectively, along with the predictions of DFT-P. Although the model displays a larger effect of Variance when Certain Payoff if 8, it otherwise gives a good account of the rather standard payoff variability effect observed for this individual. Its fit to choices at DN2a/b is less impressive, though the data here are noisy. Figure 6.3 shows a scatter plot of all eighty one mean data points from the experiment. For most points in the middle of the scale DFT-P's prediction are quite accurate. However, Participant 1 is clearly acting more deterministic than the model in several situations. The frequency with which she produces choice probabilities at the extremes suggests that the model may overestimate her uncertainty at some DNs. Nonetheless, for almost all of these extreme points DFT-P correctly predicts which alternative will be chosen more often. Figure 6.4 shows the simulated utility functions for Participant 1 (see Table A.1 in the Appendix for the full set of parameter values), and indicates that she under-sampled extreme ONs during planning, but later focused almost exclusively on these outcomes when making a final choice, with gains being simulated more often than losses. Her decision threshold parameters indicate careful simulation of future choices at the beginning of each trial ($\theta_2 = 20$), but faster deliberation when making decisions at DN1 ($\theta_1 = 2.156$) and DN2a/b ($\theta_3 = 2.224$).

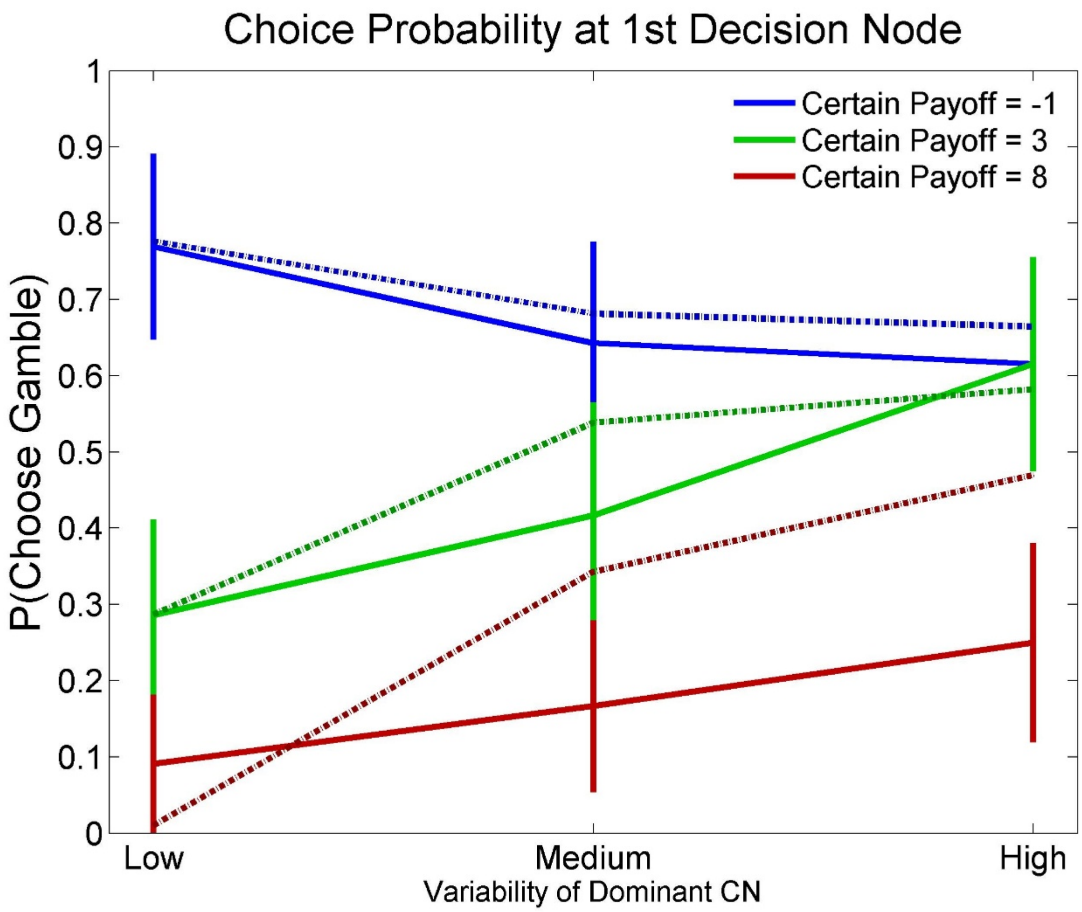

Figure 6.1. Predictions of DFT-P (dashed lines) and mean data at DN1 for Participant 1 of Experiment 3.

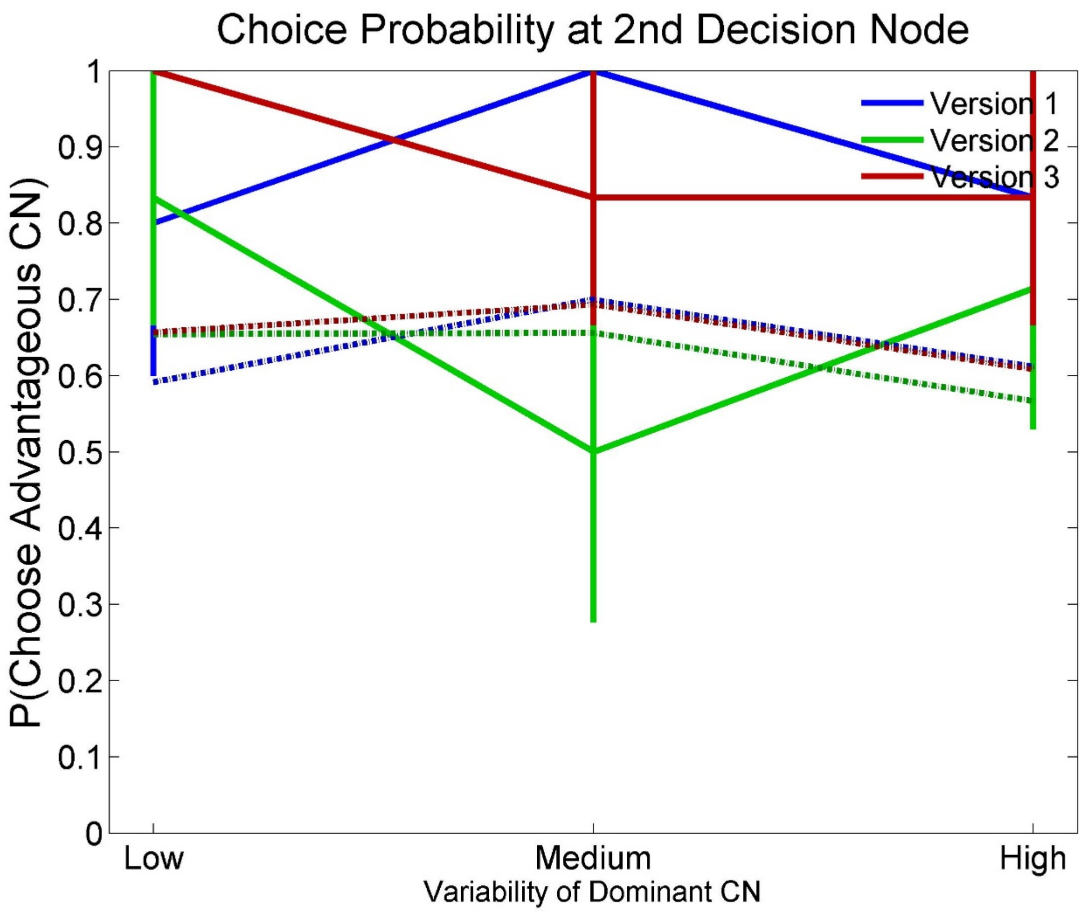

Figure 6.2. Predictions of DFT-P (dashed lines) and mean data at DN2a/b for Participant 1 of Experiment 3.

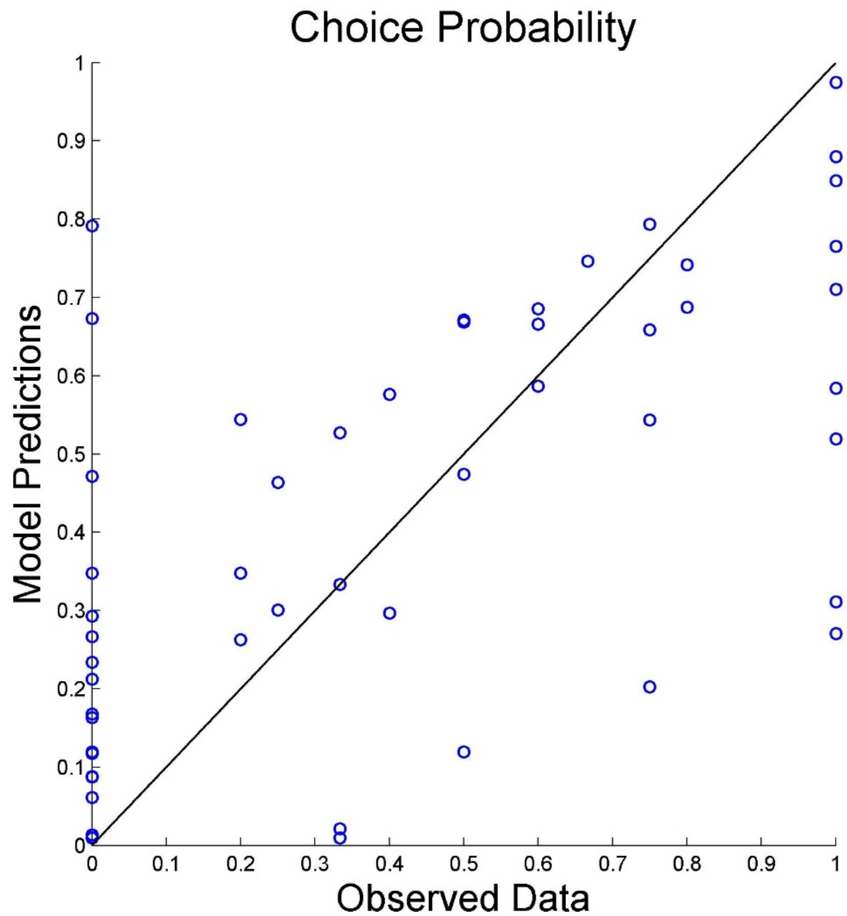

Figure 6.3. Scatter plot of mean choice probabilities for Participant 1 from each DN in Experiment 3, compared to the predictions of DFT-P. Probabilities for DN2b are inverted to reduce overlapping of points.

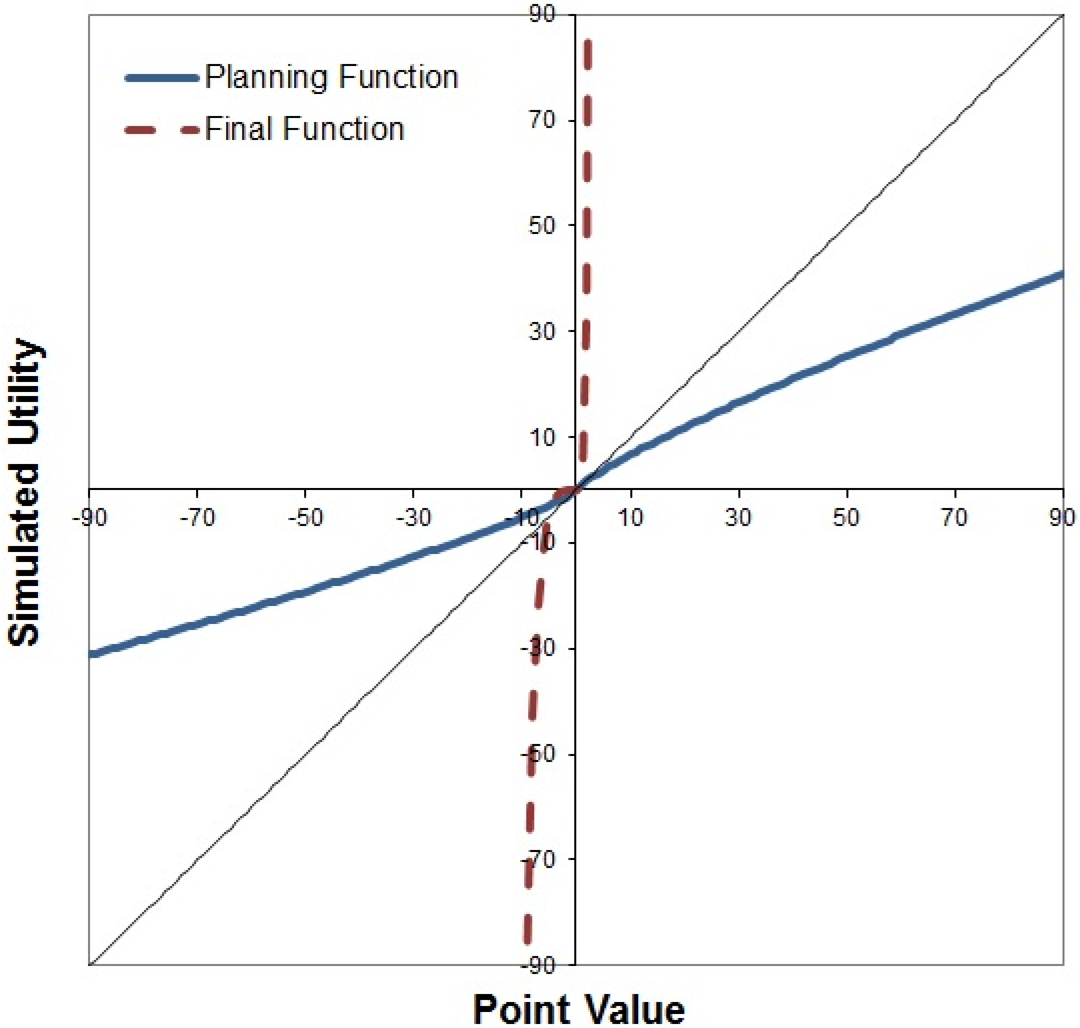

Figure 6.4. Simulated utility functions for planned and final decisions of Participant 1, derived from parameters of DFT-P in Experiment 3.

Participant 2

In contrast to Participant 1, Participant 2 displays a very difference pattern of responses. As shown in Figure 6.5, this individual has a very strong preference for risky alternatives and predominantly chose to gamble at DN1. Curiously, he tended to choose disadvantageous CNs more often than not at DN2a/b (see Figure 6.6). An examination of his simulated utility functions (Figure 6.8) sheds some light on this. For final choices, he devoted more attention to losses than gains, and since advantageous CNs happened to have slightly larger losses, he chose to avoid these nodes. Again, the scatter plot in Figure 6.7 shows greater determinism in the participant's behavior than predicted, but that DFT-P correctly identifies the preferred alternative.

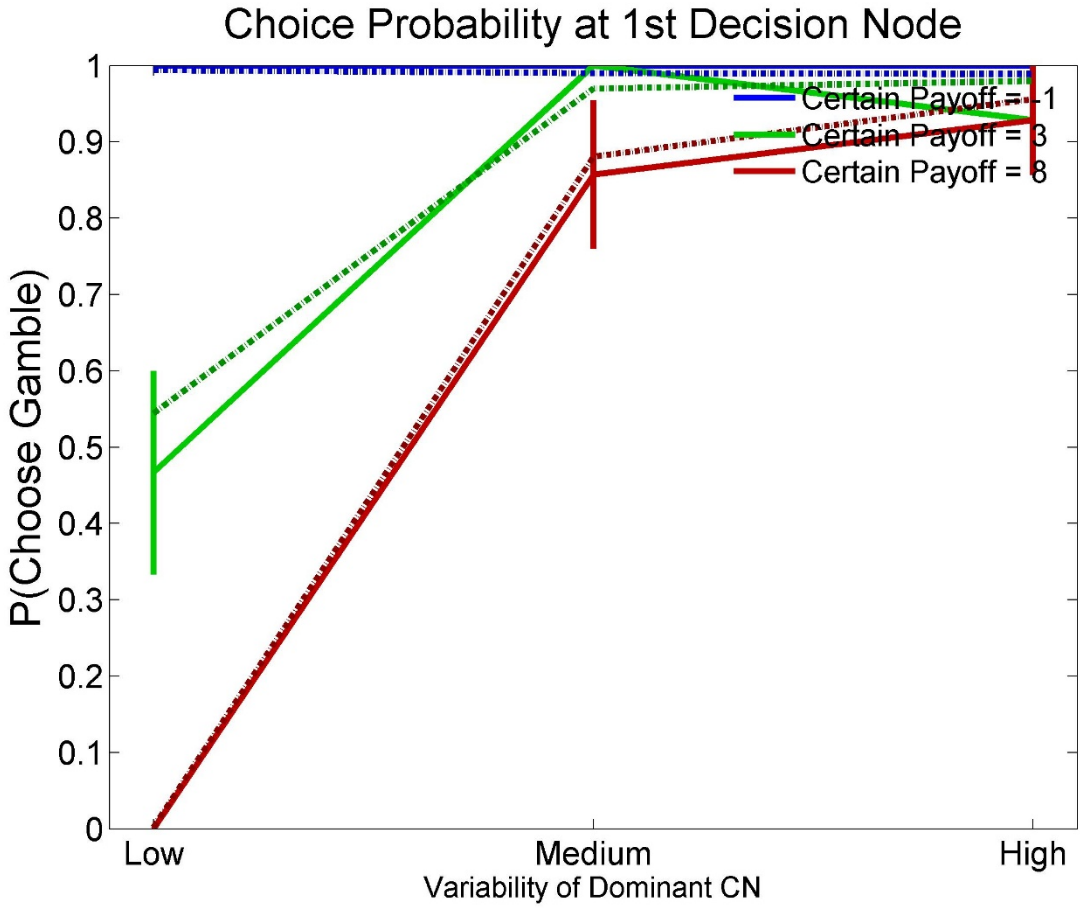

Figure 6.5. Predictions of DFT-P (dashed lines) and mean data at DN1 for Participant 2 of Experiment 3.

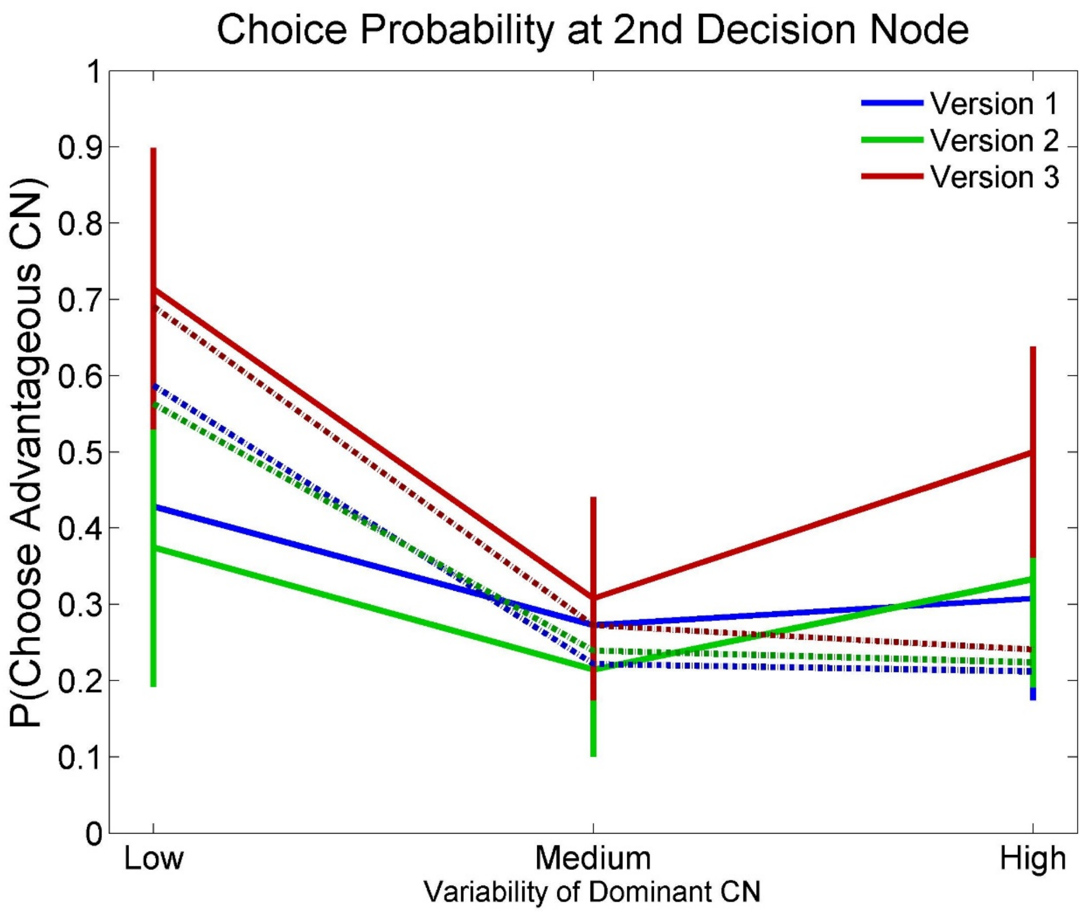

Figure 6.6. Predictions of DFT-P (dashed lines) and mean data at DN2a/b for Participant 2 of Experiment 3.

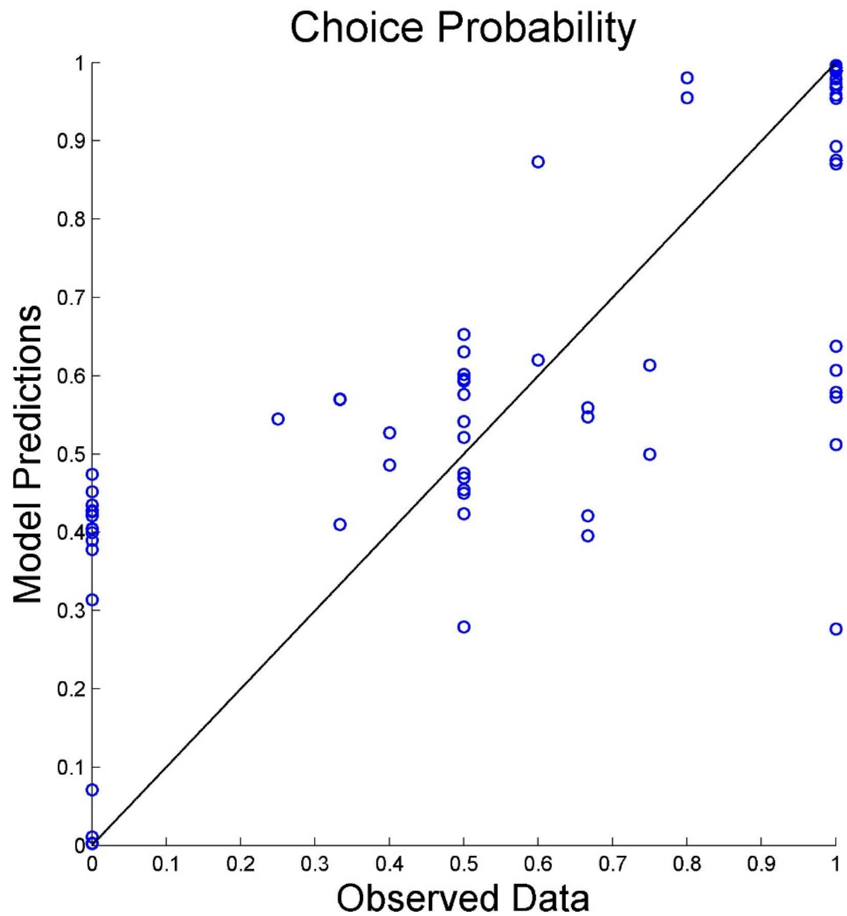

Figure 6.7. Scatter plot of mean choice probabilities for Participant 2 from each DN in Experiment 3, compared to the predictions of DFT-P. Probabilities for DN2b are inverted to reduce overlapping of points.

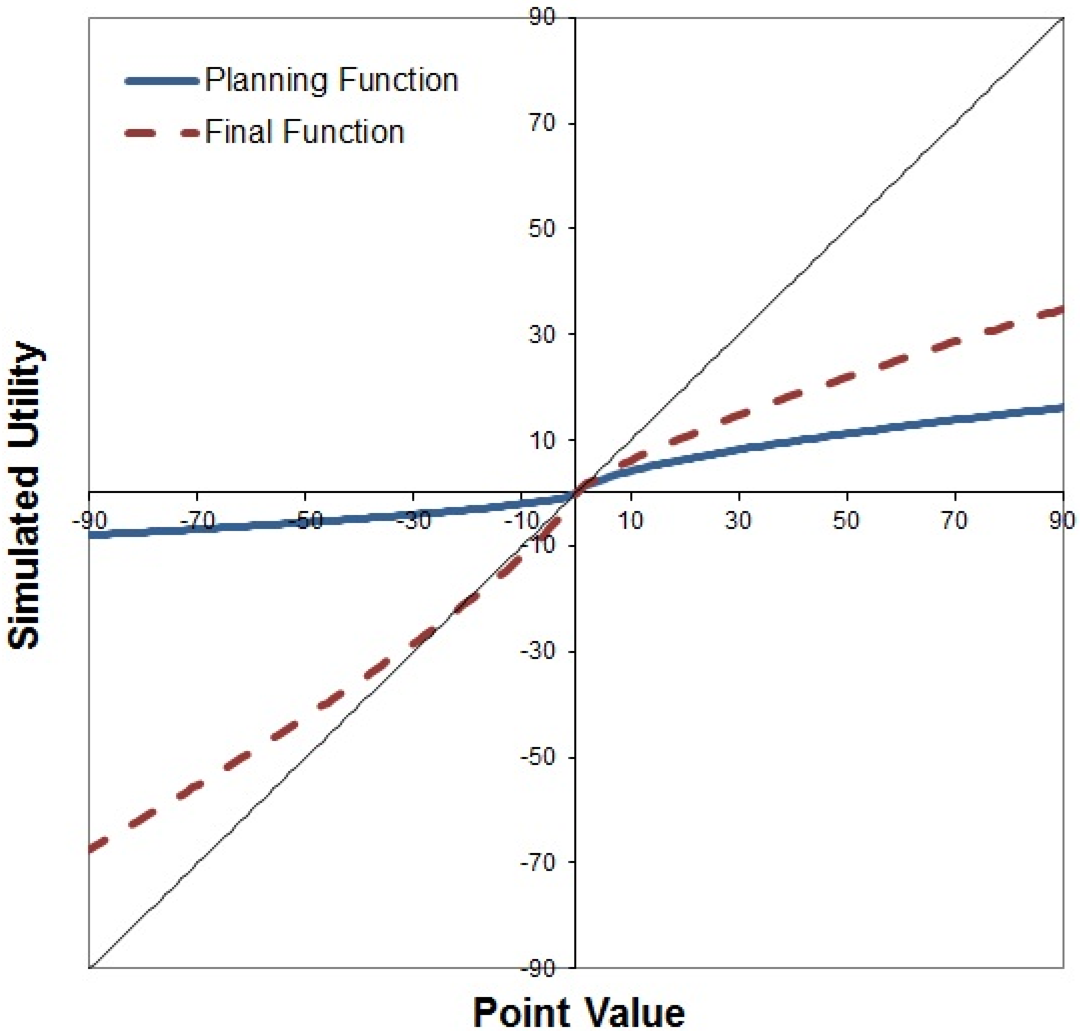

Figure 6.8. Simulated utility functions for planned and final decisions of Participant 2, derived from parameters of DFT-P in Experiment 3.

Participant 2

Participant 3 exhibits a pattern of responses somewhat similar to that of Participant 1, but with a much greater sensitivity expected payoff. Figures 6.9 and 6.10 show that at all three DNs she almost always chose the alternative with the greater expected value. Though, DFT-P fits choices at DN1 quite well, in underestimates how well the individual was able to identify the advantageous CNs at DN2a/b. Figure 6.11 shows the extent to which Participant 3's choice proportions cluster at the extremes of the scale, but also how well DFT-P accounts for this data. The simulated utility functions in Figure 6.12 indicate that Participant 3 attended more to gains than losses, which, along with high decision thresholds (see Table A.2 in the Appendix), yields preferences consistent with maximizing expected value.

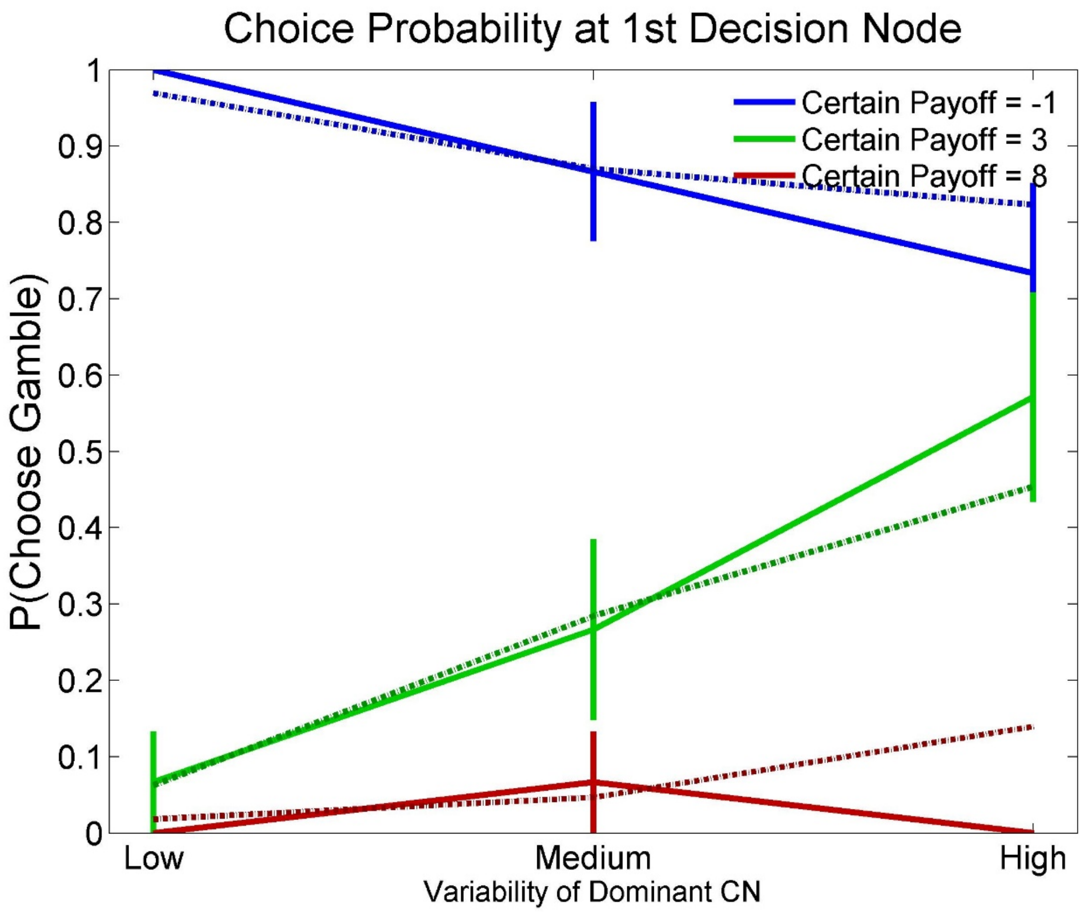

Figure 6.9. Predictions of DFT-P (dashed lines) and mean data at DN1 for Participant 3 of Experiment 3.

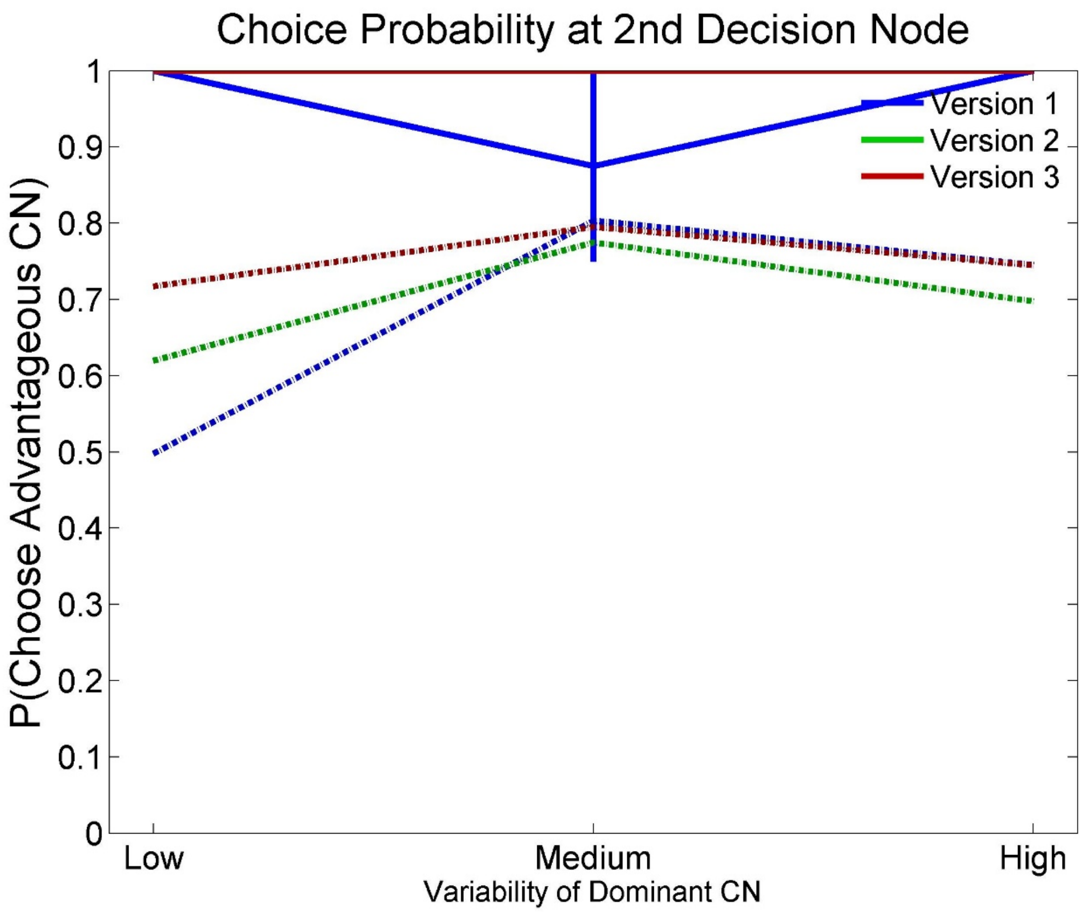

Figure 6.10. Predictions of DFT-P (dashed lines) and mean data at DN2a/b for Participant 3 of Experiment 3.

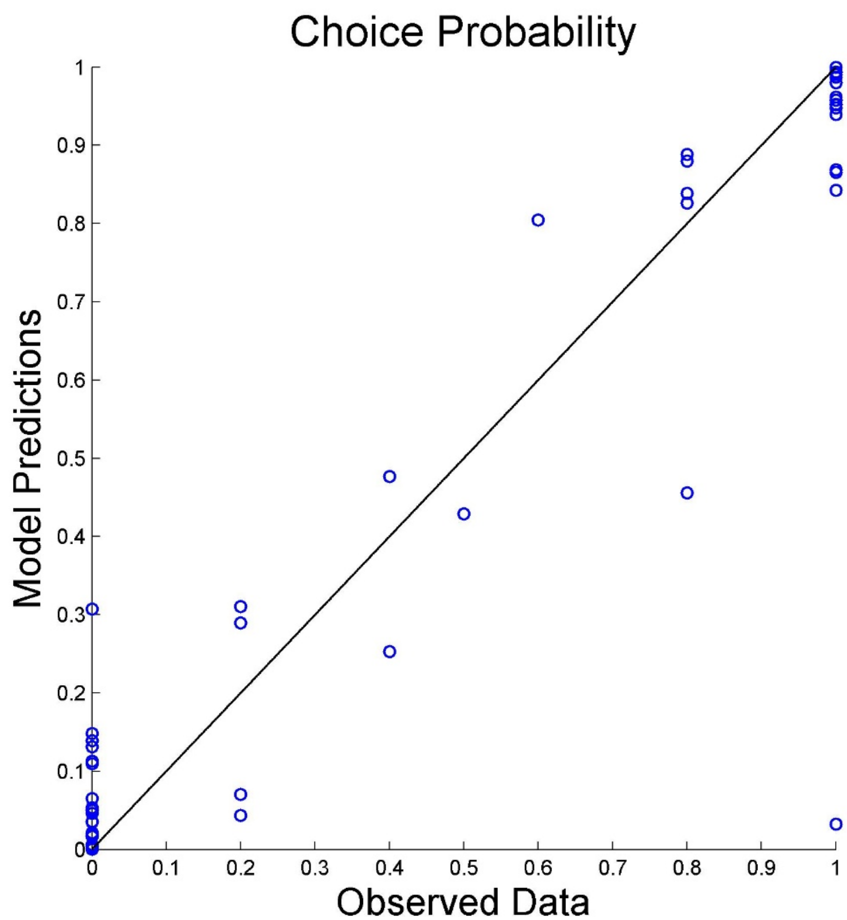

Figure 6.11. Scatter plot of mean choice probabilities for Participant 3 from each DN in Experiment 3, compared to the predictions of DFT-P. Probabilities for DN2b are inverted to reduce overlapping of points.

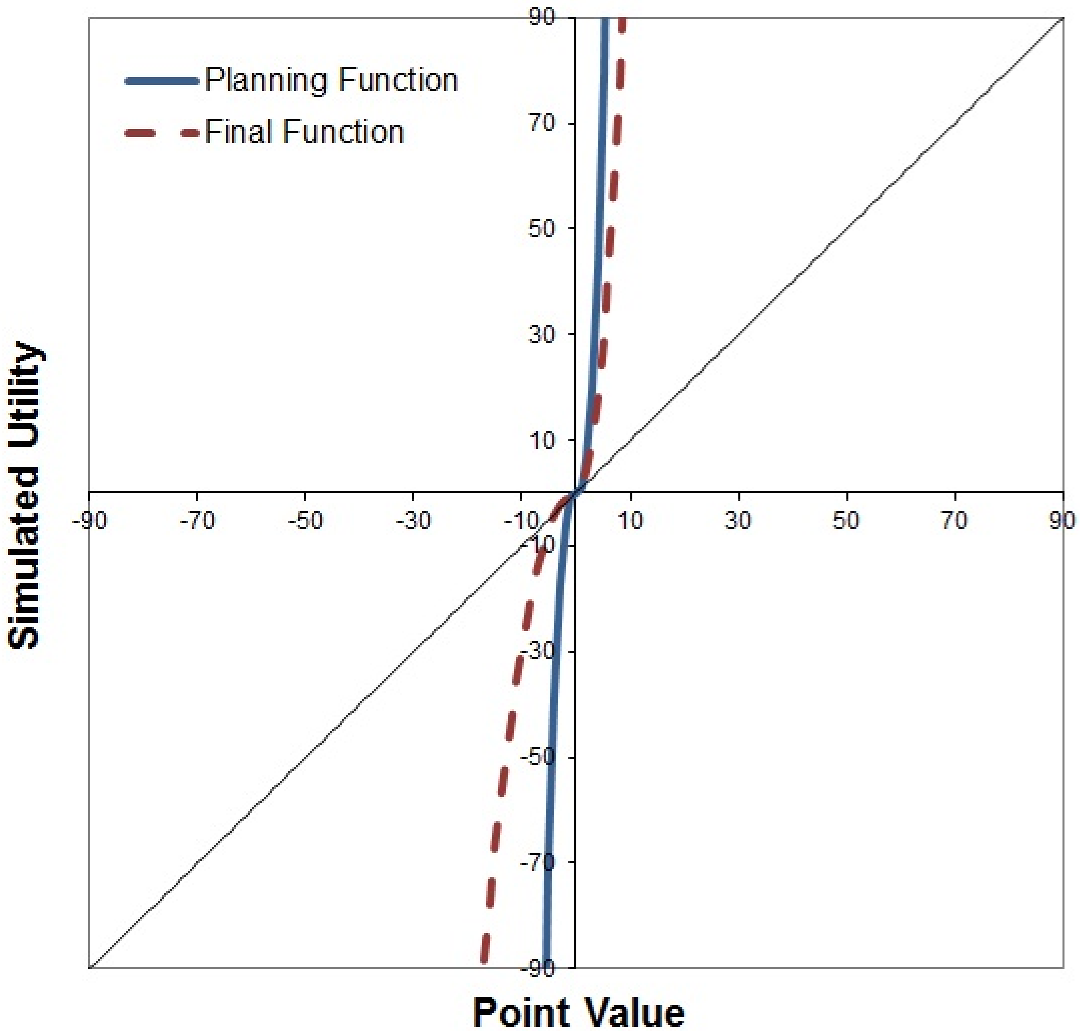

Figure 6.12. Simulated utility functions for planned and final decisions of Participant 3, derived from parameters of DFT-P in Experiment 3.

Discussion

In this chapter I showed how DFT-P can be used to model dynamic decision making at the level of individuals. Each of the three participants examined showed a unique set of preferences. These large differences highlight the difficulty of understanding human decision making in complex multistage choice environments. Compared to those in more traditional single-stage choice studies, these individuals are free pursue a broad range of strategies. They are potentially affected by myriad factors, both internal (e.g. attentional bias, risk preference, memory capacity, motivation) and external (e.g. expected value, payoff variability, computational complexity). However, by fitting DFT-P to each participant one can begin to explain these significant individual differences within its attention-driven mental simulation framework. The severely distorted simulated utility functions derived for some individuals suggest highly selective sampling that strongly favors some payoff information over others. That the model fits better when hugely exaggerating the value of extreme ONs indicates that these payoffs, when present, overshadowed all others. Despite the present successes, it is likely that a more sophisticated understanding of how individuals probe and perhaps simplify decision trees is required before much more progress can be made.

DFT-P's predictions were most accurate when participants were uncertain and variable in their decisions, often as a result of low simulated utility difference across alternatives and high variability in payoffs. When participants were relatively deterministic, and reliably chose the same alternative for each repetition of a tree, the model often produced insufficiently extreme choice probabilities. However, the present analysis does not involve a large amount of individual data, with each decision tree appearing only five times in Experiment 3. This is particularly problematic for estimating choice proportions at DN2a/b, where an observation is only possible on trials where the participant first chooses to gamble at DN1 and then experiences the appropriate event at CN1. With a greater number of trial repetitions, participants may have shown themselves to be more stochastic, in accordance with DFT-P's predictions. In all, the present demonstration shows DFT-P's promise, and suggests many potentially fruitful avenues of future study.

7. Summary

How people form plans and make decisions in dynamic environments has received relatively little attention from decision scientists, despite the ubiquity and importance of such decisions in everyday life. Much of the previous research has been rather pessimistic of people's ability to plan ahead in complex choice tasks. In he present article I first sought to establish that individuals spontaneously plan future decisions in a laboratory setting. In Chapter 1 I introduced backward induction as a normative theory of dynamic decision making, and summarized the previous research demonstrating that it failed to account for important aspects of human behavior. In Chapter 2 I presented a study showing participant were indeed sensitive to the distribution of payoffs across future DN nodes in a way that suggested they were planning ahead.

In Chapter 3 I introduced a new cognitive model of dynamic decision making, Decision Field Theory-Planning. The model represents a new approach to understanding how individuals cope with the complexities of dynamic decision making. It posits a set of cognitive processes whereby individuals plan future choices on the fly, through repeated forward-looking mental simulations. As they imagine different possible sequences of events, simulated outcomes are compared and preferences evolve in a noisy fashion until a threshold is reached. In Chapter 4 I tested a central prediction of the model by manipulating payoff variability in a decision tree task. Because, according to DFT-P, greater payoff variability introduces more noise to the mental simulation process, preference for a risky gamble should decrease for trees where the alternative is a certain loss, and increase for trees with a certain gain. Upon obtaining this result, I demonstrated the model's ability to produce the observed effect across a range of parameter values. In Chapter 5 I replication the payoff variability effect, and presented a quantitative fit of DFT-P to group data. An analysis of the best fitting parameter values also shed light on the types and amounts of information that produced people's choices. Finally, with Chapter 6 I demonstrated how fitting the model separately to individual decision makers allows one to investigate the sometimes large individual differences observed in dynamic decision tasks. In all, I found significant support for the ideas proposed by DFT-P and showcased it value as a tool for further investigation.

References

Anderson, P. L. (2009). The value of private businesses in the United States. *Business Economics, 44*(2), 87-108.

Barkan, R., & Busemeyer, J. R. (2003). Modeling dynamic inconsistency with a changing reference point. *Journal of Behavioral Decision Making, 16*(4), 235-255.

Barkow, J. H., Cosmides, L., & Tooby, J. (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making, 16*(3), 215-233.

Bertsekas, D. P. (1976). Dynamic programming and stochastic control.

Bone, J., Hey, J. D., & Suckling, J. (2009). Do people plan? *Experimental Economics, 12*(1), 12-25.

Bone, J. D., Hey, J. D., & Suckling, J. R. (2003). Do people plan ahead? *Applied Economics Letters, 10*, 277-280.

Busemeyer, J. R. (1985). Decision making under uncertainty. *Journal of Experimental Psycology: Learning, Memory, & Cognition, 11*, 538-564.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100*(3), 432-459.

Busemeyer, J. R., Weg, E., Barkan, R., Li, X., & Ma, Z. (2000). Dynamic and consequential consistency of choices between paths of decision trees. *Journal of Experimental Psychology: General, 129*(4), 530-545.

Carbone, E., & Hey, J. D. (2001). A test of the principle of optimality. *Theory and Decision, 50*, 263-281.

Chersi, F., & Pezzulo, G. (2012). Using hippocampal-striatal loops for spatial navigation and goal-directed decision-making. *Cognitive processing, 13*(1), 125-129.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31*(3), 187-276.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.

Cubitt, R. P., Starmer, C., & Sugden, R. (1998). Dynamic choice and the common ratio effect: An experimental investigation. *The Economic Journal, 108*, 1362-1380.

DeGroot, M. H. (1970). *Optimal statistical decisions*. New York: McGraw-Hill.

Dimperio, E. (2009). *A dynamic model of planning behavior in multi-stage risky decision tasks.* Doctoral dissertation, Retrieved from ProQuest Dissertations and Theses Database. (UMI 3380074)

Dixit, A. K., & Pindyck, R. S. (1994). *Investment under uncertainty*: Princeton University Press.

Dragoi, G., & Tonegawa, S. (2010). Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature, 469*(7330), 397-401.

Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review, 112*(4), 912-931.

Erev, I., Ert, E., & Yechiam, E. (2008). Loss aversion, diminishing sensitivity, and the effect of experience on repeated decisions. *Journal of Behavioral Decision Making, 21*(5), 575-597. doi: 10.1002/bdm.602

Fudenberg, D., & Tirole, J. (1991). *Game theory*: The MIT Press.

Hey, J. D., & Knoll, J. A. (2007). How far ahead do people plan? *Economics Letters, 96*, 8-13.

Hey, J. D., & Knoll, J. A. (2011). Strategies in dynamic decision making - An experimental investigation of the rationality of decision behavior. *Journal of Economic Psychology, 32*, 399-409.

Hirschfeld, L. A., & Gelman, S. A. (1994). *Mapping the mind: Domain specificity in cognition and culture*: Cambridge University Press.

Hotaling, J. M., Busemeyer, J. R., & Li, J. (2010). Theoretical developments in decision field theory: Comment on Tsetsos, Chater, & Usher. *Psychological Review, 117*(4), 1294-1298.

Johnson, A., & Redish, A. D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *The Journal of Neuroscience, 27*(45), 12176-12189.

Johnson, J. G., & Busemeyer, J. R. (2001). Multiple-stage decision-making: The effect of planning horizon length on dynamic consistency. *Theory and Decision, 51*, 217-246.

Katz, L. (1962). Monetary incentive and range of payoffs as determiners of risk taking. *Journal of Experimental Psychology, 64*, 541.

Keeney, R. L., & Raiffa, H. (1976). Decision analysis with multiple conflicting objectives. *Wiley& Sons, New York*.

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. *Decision making in action: Models and methods, 5*(4), 138-147.

Klein, G. A. (1998). *Sources of power: How people make decisions*: The MIT Press.

Ljunggvist, L., & Sargent, T. J. (2000). *Recursive macroeconomic theory*: The MIT Press.

Machina, M. J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature, 27*(4), 1622-1668.

Myers, J. L., & Katz, L. (1962). Range of payoffs and feedback in risk taking. *Psychological Reports, 10*(2), 483-486.

Myers, J. L., & Sadler, E. (1960). Effects of range of payoffs as a variable in risk taking. *Journal of Experimental Psychology, 60*, 306-309.

Myers, J. L., Suydam, M. M., & Gambino, B. (1965). Contingent gains and losses in a risk-taking situation. *Journal of Mathematical Psychology, 2*, 263-370.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal, 7*, 308-313.

Nemeth, C., & Klein, G. (2010). The naturalistic decision making perspective *Wiley Encyclopedia of Operations Research and Management Science*: John Wiley & Sons, Inc.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*: Oxford University Press.

Pezzulo, G., Rigoli, F., & Chersi, F. (2013). The Mixed Instrumental Controller: Using Value of Information to Combine Habitual Choice and Mental Simulation. [Original Research]. *Frontiers in Psychology, 4*. doi: 10.3389/fpsyg.2013.00092

Raiffa, H. (1968). *Decision analysis: introductory lectures on choices under uncertainty*. Reading, MA: Addison-Wesley.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111*(2), 333-367.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review, 108*(2), 370-392.

Sarin, R., & Wakker, P. P. (1998). Dynamic choice and nonexpected utility. *Journal of Risk and Uncertainty, 17*(2), 87-120.

Suydam, M. M., & Myers, J. L. (1962). Some parameters of risk-taking behavior. *Psychological Reports, 10*(2), 559-562.

Suzuki, S., Harasawa, N., Ueno, K., Gardner, J. L., Ichinohe, N., Haruno, M., . . . Nakahara, H. (2012). Learning to simulate others' decisions. *Neuron, 74*(6), 1125-1137.

Von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research* (Vol. 604): Cambridge University Press Cambridge.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology*: Harmondsworth: Penguin.

Wilson, M. A., & McNaughton, B. L. (1993). Dynamics of the hippocampal ensemble code for space. *Science, 261*, 1055-1058.

Appendix

Table A.1. Best fitting parameters of DFT-P for Participant 1 in Experiment 3.

Best Fitting Parameters of DFT-P

Parameter		Description
z	0.000	Response bias in favor of risky alternative
θ_1	2.156	Decision threshold at DN1.
θ_2	20.000	Decision threshold for planned choices at DN2a/b.
θ_3	2.224	Decision threshold for final choices at DN2a/b.
α_1	0.826	Over-simulating bias for extreme outcomes at DN1.
λ_1	0.760	Under-simulating bias for losses at DN1.
α_2	5.000	Over-simulating bias for extreme outcomes at DN2a/b.
λ_2	0.001	Under-simulating bias for losses at DN2a/b.

Table A.2. Best fitting parameters of DFT-P for Participant 2 in Experiment 3.

Best Fitting Parameters of DFT-P

Parameter		Description
z	2.000	Response bias in favor of risky alternative
θ_1	4.104	Decision threshold at DN1.
θ_2	2.630	Decision threshold for planned choices at DN2a/b.
θ_3	12.512	Decision threshold for final choices at DN2a/b.
α_1	0.618	Over-simulating bias for extreme outcomes at DN1.
λ_1	0.494	Under-simulating bias for losses at DN1.
α_2	0.789	Over-simulating bias for extreme outcomes at DN2a/b.
λ_2	1.931	Under-simulating bias for losses at DN2a/b.

Table A.3. Best fitting parameters of DFT-P for Participant 3 in Experiment 3.

Best Fitting Parameters of DFT-P

Parameter		Description
z	1.081	Response bias in favor of risky alternative
θ_1	3.219	Decision threshold at DN1.
θ_2	5.209	Decision threshold for planned choices at DN2a/b.
θ_3	20.000	Decision threshold for final choices at DN2a/b.
α_1	2.716	Over-simulating bias for extreme outcomes at DN1.
λ_1	0.891	Under-simulating bias for losses at DN1.
α_2	2.095	Over-simulating bias for extreme outcomes at DN2a/b.
λ_2	0.233	Under-simulating bias for losses at DN2a/b.