PROBABILITY LEARNING IN NORMAL AND PARKINSON SUBJECTS: THE EFFECT OF REWARD, CONTEXT, AND UNCERTAINTY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

BURAK ERDENİZ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF COGNITIVE SCIENCE

SEPTEMBER 2007

Approval of the Graduate School of Informatics

Prof.Dr.Nazife BAYKAL
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science/Doctor of Philosophy.

Prof.Dr.Deniz ZEYREK
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.

Assist.Prof.Dr. Bilge SAY	Assist. Prof. Didem GÖKÇAY
Co-Supervisor	Supervisor

Examining Committee Members

Assist. Prof. Didem GÖKÇAY (METU, COGS) ______________

Assist. Prof.Bilge SAY (METU, COGS) _______________

Assist. Prof.Annette HOHENBERGER (METU, COGS) _______________

Assoc. Prof. Hakan GÜRVİT (İU, NEUROLOGY) ______________

Assist. Prof.Mark Ashton SMITH (Bilkent, PSY) _________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this wok.

Name, Last name : Burak Erdeniz

Signature : ____________________

ABSTRACT

PROBABILITY LEARNING IN NORMAL AND PARKINSON SUBJECTS: THE EFFECT OF REWARD, CONTEXT, AND UNCERTAINTY

Erdeniz, Burak
Master, Department of Cognitive Science
Supervisor: Assist. Prof. Dr. Didem Gökçay
Co-Supervisor: Assist. Prof. Dr. Bilge Say

September 2007, 102 pages

In this thesis, the learning of probabilistic relationships between stimulus-action pairs is investigated under the probability learning paradigm. The effect of reward is investigated in the first three experiments. Additionally, the effect of context and uncertainty is investigated in the second and third experiments, respectively. The fourth experiment is the replication of the second experiment with a group of Parkinson patients where the effect of dopamine medication on probability learning is studied. In Experiment 1, we replicate the classical probability learning task by comparing monetary and non-monetary reward feedback. Probability learning behavior is observed in both monetary and non-monetary rewarding feedback conditions. However, no significant difference between the monetary and non-monetary feedback conditions is observed. In Experiment 2, a variation of the probability learning task which includes irrelevant contextual information is applied. Probability learning behavior is observed, and a significant effect is found between monetary and non-monetary feedback conditions. In Experiment 3; a probability learning task similar to that in Experiment 2 is applied, however, in this experiment, stimulus included relevant contextual information. As expected, due to the utilization of the relevant contextual information from the start of the experiment, no significant effect is found for probability learning behavior. The effect of uncertainty observed in this experiment is a replication of the reports in literature. Experiment 4 is identical to Experiment 2; except that the subject population is a group of dopamine medicated Parkinson patients and a group of age matched controls. This experiment is introduced to test the suggestions in the literature regarding the enhancement effect of dopamine medication in probability learning based on positive feedback conditions. In Experiment 4, probability learning behavior is observed in both groups, but the difference in learning performance between Parkinson patients and controls was not significant, probably due to the low number of subject recruited in the experiment. In addition to these investigations, learning mechanisms are also examined in Experiments 1 and 4. Our results indicate that subjects initially search for patterns which lead to probability learning. At the end of Experiments 1 and 4, upon learning the winning frequencies, subjects change their behavior and demonstrate maximization behavior, which makes them prefer continuously one option over the other.

Keywords: Probability Learning, Pattern searching, Parkinson Disease,
Dopamine, Basal Ganglia

ÖZ

ÖDÜLÜN, BAĞLAMIN VE BELİRSİZLİĞİN OLASILIKSAL ÖĞRENMEYE OLAN ETKİSİ

Erdeniz, Burak
Master, Bilişsel Bilimler Bölümü
Tez Yoneticisi: Assist. Prof. Dr. Didem Gökçay
Ortak Tez Yoneticisi: Assist. Prof. Dr. Bilge Say

Eylül 2007, 102 sayfa

Bu tezde uyaran-eylem kalıpları arasındaki olasılıksal ilişki olasılık öğrenme paradigması ile araştırılmıştır. İlk üç deneyde ödülün etkisine bakılmıştır. Ek olarak, sırasıyla bağlam ve belirsizlik ikinci ve üçünçü deneyler aracılığı ile araştırılmıştır. Dördüncü deney ikinci deneyin bir tekrarı olup bir grup dopamin tedavisi gören Parkinson hastasında olasılık öğrenme test edilmiştir. Birinci deneyde klasik olasılık öğrenme paradigması parasal ve parasal olmayan geribildirimler ile test edilmiştir. Olasılık öğrenme davranışı hem ödül olarak para ile geribildirim alan hem de ödül olarak para ile geri bildirim almayan durumlarda gözlemlenmiştir. Buna karşın, bu iki geri bildirim durumu arasında anlamlı bir fark bulunamamıştır. İkinci deneyde, olasılık öğrenme benzeri gereksiz bağlam bilgisi içeren bir görev uygulanmıştır. İkinci deneyde olasılık öğrenme davranışı gözlemlenmiştir ve ödül olarak para ile geribildirimin anlamlı bir etkisi gözlemlenmiştir. Üçüncü deneyde yine ikinci deney gibi olasılık öğrenme benzeri fakat bu sefer gerekli bağlam bilgisi içeren bir görev uygulanmıştır. Beklendiği üzere faydalı bilginin kullanılması neticesinde ne olasılıksal öğrenme ve ne de ödül olarak para geri bildiriminin öğrenme üzerine bir etkisi gözlemlenmemiştir. Bu deneyde bulunan belirsizliğin etkisi literatürdeki raporların bir tekrarıdır. Dördüncü deney ikinci deneyin aynısıdır fakat deneyin uygulandığı örneklemler bir grup dopamin tedavisi gören Parkinson hastası ve aynı yaşlardaki kontrol grubudur. Bu deney literatürdeki dopamin tedavisinin pozitif geri bildirimi ile ilgili önermeleri test etmek için tasarlanmıştır. Dördüncü deneyde, olasılık öğrenme davranışı gözlemlenmiştir fakat Parkinsonlu hastalar ve kontrol grubu arasında olasalık öğrenmede anlamlı bir fark gözlemlenmemiştir bunun nedeni büyük olasılıkla katılımcı sayısının azlığından kaynaklanmaktadır. Bu çalışmalara ek olarak birinci ve dördüncü deneylerde öğrenme mekanizmaları sınanmıştır. Sonuçlarımız deneyin başlangıcında katılımcıların olasılıkları öğrenirken örüntü arama davranışı sergilediğini göstermesidir. Birinci ve dördüncü deneyde frekans öğrenmeye bağlı olarak katılımcılar maksimizasyon davranışı sergilemektedir ve sürekli aynı seçeneği seçmişlerdir.

Anahtar Kelimeler: Olasılık Öğrenme, Örüntü Arama, Parkinson, Dopamin, Basal Ganglia

This work is dedicated to;

MY PARENTS

ACKNOWLEDGMENTS

I want to express my sincere appreciation to my supervisor Assist. Prof. Dr. Didem Gökçay and co-supervisor Assist. Prof. Dr. Bilge Say for their guidance, insight, and financial support throughout my research. Thanks to my close friends for their helpful comments on this thesis and their kind friendship especially Nart Bedin Atalay, Aslı Kılıç, Didem Kadıhasanoğlu, Dicle Dövencioğlu, Eren Dalgıç and Ali Matay. I also want to thank my friends Ayberk Vardar, Dinçer Özoran and Özkan Bayraktar for their useful comments and helping me to finish my thesis. Special thanks to Prof. Dr. Oğuz Tanrıdağ for his financial support on my research by inviting me to a series of Cognitive Neuroscience Congress in Marmaris. I would also express my gratitude to my mentor and friend Assoc. Prof. Hakan Gürvit for his useful comments on my research and letting me to practice my experiments with his Parkinson population in İstanbul University School of Medicine. I also want to thank Bengi Baran and Assist. Prof. Haşmet Hanağası for their guidance and useful comments. Thanks for Assoc. Prof. Fevzi Öztekin for helping me to practice a pilot study with his Parkinson population. Thanks to our departmental secretary Sibel Gülnar, Ali Kantar and Ayşe Ceylan for solving every academic problem in the shortest time. Finally, my endless and profound thanks would go to my father Ali Erdeniz and my mother Zeynep Zafer Erdeniz, I offer sincere thanks for their unconditional love and faith in me.

TABLE OF CONTENTS

CHAPTER

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Decision-making is a high level cognitive function indispensable for humans. According to (some of) the fundamental decision making theories (e.g., Prospect Theory) in economics and psychology, humans use limited cognitive resources such as attention (capacity) or memory while they are making decisions under uncertainty (Simon, 1957; Kahneman, Slovic & Tversky, 1982; Glovic, Griffin & Kahneman, 2002; Gigerenzer & Selten, 2002). Similarly, *bounded rationality*, a general assumption in human economic behavior, indicates that the human decisions are “limited” or “bounded” by the agent’s cognitive capacities (Simon, 1957). In order to understand the cognitive limitation of the human decision making capacity, the empirical research on bounded rationality in psychology has been studied under two paradigms: decision-making and learning.

In the *decision-making paradigm*, the limitations of people’s judgments for single trial choices (Simon, 1957), as well as the systematic biases (confronted) in uncertain decision making situations (Kahneman et al., 1982) are studied.

In the *learning paradigm*, decision making is considered to be a learned behavior (Atkinson, 1964). Therefore, researchers concentrate on the limits of the human learning ability (Rizello, 1999). Psychologists working on this paradigm study the cognitive constraints on learning of probabilities (e.g., Estes, 1961) and they attempt to find a mathematical framework for the ability of learning probabilities (Atkinson, 1964; Bush & Mosteller; 1955; Laming, 1973).

The primary focus of this thesis is on the second paradigm. By investigating the factors involved in probability learning, we specifically aim to understand "how people learn the probabilities to make decisions under uncertainty?". In this paradigm, man is studied cognitively as an "intuitive statistician" (Peterson & Beach, 1967, p.42). To emphasize the importance of probability learning, Estes (1972, p.81) states that "human behavior in many situations involving uncertainty and risk depends on the acquisition of information concerning event probabilities". Regardless of the complexity of people's choices, probability learning is essential. In order to highlight this, Estes (1972, p.81) made a remarkable explanation:

> [A man] … must continually form estimates of probabilities of events as basis for action. He must assess the probabilities of various responses of adversaries to his own actions in economic competition, games, politics or war. He must estimate probabilities of increases and decreases in price as a basis for purchases and investments, of the efficacy of remedies for illness, of changes in weather as a basis of for selecting wardrobe, of the reactions of countless other people to his own behavior during nearly every waking moment.

Therefore, the central issue in probability learning and decision making under uncertainty more specifically focuses on how people make predictions about the probabilities of events considering their short term and long term benefits in the period of learning. For instance, in a probability learning experiment, participants have to decide between two options, but one of them has a higher probability of winning. In the long run, after a trial and error period, people learn to predict on the winning option to be successful.

Although some of the distinct issues like probability learning mechanisms and probability learning strategies will be discussed further in detail in the literature review section, they will be briefly mentioned below. There are at least two learning mechanisms adopted to meliorate our decisions in a probability learning task; *a pattern searching mechanism* and *a frequency learning mechanism* (Figure 1).

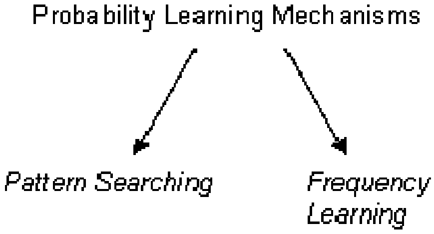

Figure 1 Two learning mechanisms are responsible for probability learning.

The first mechanism that is involved in probability learning experiments is the *pattern searching*. When we consider the amount of information and technology in general that our species was confronted with, it increased exponentially in the last few million years (Kurzweil, 2005; Hawking, 2001). There should be an advantage of a system that is involved in learning of probabilities which *temporally stores* information about the current options. A temporary storage system may help people make their judgments by utilizing their recent memory. The advantage of such a system is that people do not have to search their entire autobiographical memory just to make a simple decision such as deciding A or B. Another advantage of this temporary storage is that it can also cache some sequential information about the outcomes of choices and check these sequences to find whether there is a regular pattern among outcomes or not. This leads to the ability to search for patterns in sequences. It is plausible for a person to use this pattern searching mechanism in a probability learning task, to distinguish the randomness of the sequences of outcomes. According to Clark (2004), we have an innate mechanism to search for patterns, which helps us overcome the uncertainty of a situation.

> "Our survival is likely linked to our ability to effectively identify causal associations, or patterns in our lives, and it is extremely probable for our brain to have developed a specialized function to address such a task." (Clark, 2004, p.17).

In other words, from an evolutionary perspective, this pattern searching mechanism might have helped us as a guide while we are making decisions.

Another mechanism which is involved in probability learning is the *frequency learning* mechanism. In the last decade, in evolutionary psychology, it was postulated that our cognitive capacity is not capable of representing single event probabilities, but we can easily deal with event frequencies which occur by sampling over a period of time. According to Cosmides and Tobby (1996), our hominid ancestors had the ability to share the information of specific events in terms of relative frequencies, which necessitated the existence of a *frequency learning* mechanism. For example, in a hunting situation, hominids might remember which places have more food resources compared to places that have less or no food resources. Hence, Cosmides and Tobby (1996) suggested that our hominid ancestors could not experience single event probabilities, but they must have experienced relative frequencies like which places have more food resources and other places have less. In addition Gigerenzer and Hoffrage (1995, p.26) added that "An evolutionary point of view suggests that the mind is tuned to frequency formats, which is the information format humans encountered long before the advent of probability theory".

Another important issue in this thesis is the distinction between probability learning strategies. For example, in a probability learning experiment, if an option wins 75% of the time and a person choose that option 75% in overall, this is called the “frequency matching strategy”[1]. On the other hand, if a person chooses the winning option more than 75% of the time, this type of strategy is called “overmatching”. A special case of overmatching occurs when the participants choose the winning option 100% of the time. This special case is called “maximizing” (Figure 2).

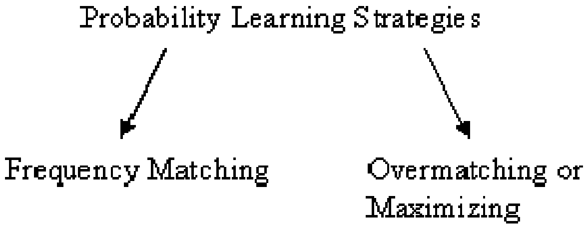

Figure 2 Two learning strategies that are used in probability learning.

Researchers in this field define the strategy in reference to the asymptotic level of the learning curve. In probability learning tasks, one of the options usually wins more than the other. For example, when an option wins 75% of the time and a person follows the frequency matching strategy, this person reaches the asymptotic level of learning curve exactly at the percentage of the winning option. If he uses overmatching strategy, the asymptote of the learning curve will appear above the level of the winning probability (steady state phase occurs above the probability of the winning option, see Figure 3). If the participant uses maximizing strategy, the asymptote is reached well above the level of the winning option.

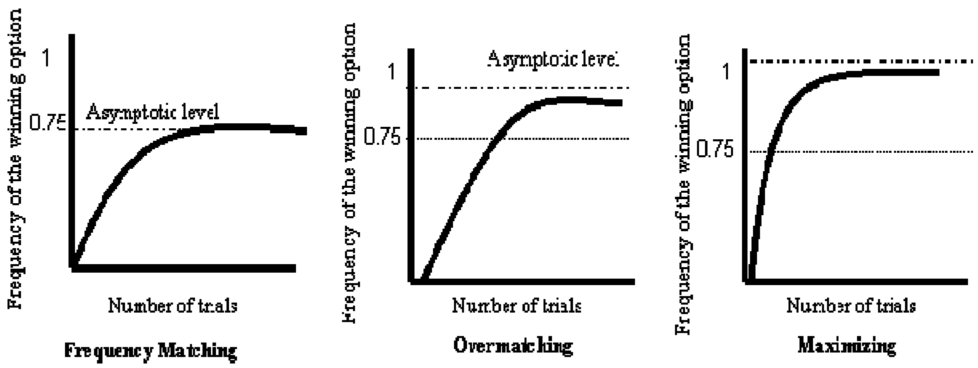

Figure 3 Learning Strategies

---

[1] In the literature probability matching is also used to refer to the frequency matching strategy. We used frequency matching in this thesis to minimize the confusion.

There are several important factors which might directly or indirectly influence frequency matching or maximizing/overmatching strategy in a probability learning task. One of the important factors that influence probability learning strategy is the feedback saliency. There is overwhelming consensus among researchers that rewarding feedbacks such as financial incentives affect the performance positively in a probability learning task (Vulkan, 2000; Hertwig & Ortman, 2001). Using financial incentives in a probability learning environment may have several purposes. First, the general belief among psychologists is that salient payoffs (monetary rewards or punishments) reduce performance variability (Davis & Holt, 1993). Second, it is believed that using monetary feedback provides a framework which is built on the assumption of maximization. According to this framework everybody wants to earn more profit, so the participants may be motivated to use the maximization strategy.

Furthermore, current research in the field of cognitive neuroscience and cognitive psychology has indicated that positive feedback (Elliot et al., 1997; Elliot et al., 2000) and money (Thut et al., 1997) induce increasing blood flow in the brain's reward centers which is facilitated mainly by the dopamine system. Dopamine is a type of neurotransmitter naturally produced by the human body, which is also shown to be active during reward mediated probability learning. On the other hand, among many other diseases such as schizophrenia or attention deficit hyperactivity disorder (ADHD), impairment of the dopamine system affects especially patients with Parkinson Disease (PD), imposing cognitive dysfunction in perceptual category learning (Ashby & Maddox, 2005; Filoteo et al., 2005; Filoteo et al., 2007) as well as probabilistic learning (Knowlton et. al 1996, Frank et. al, 2005).

In this study, motivated by the complexity of the factors influencing probability learning, we investigated the contribution of several factors such as the effect of reward, uncertainty and context. Three experiments are designed to differentiate the effects of reward, uncertainty and context. In the classical probability learning task (CT, Experiment 1), subjects are instructed to predict the winning class out of two options presented as one red and one blue box. In this task, blue wins 80% of the time. In the irrelevant information task (IIT, Experiment 2); subjects see 10 boxes, and in each trial, the proportion of blue and red boxes changes randomly. In this task, again the blue wins 80% of the time regardless of the number of blue/red boxes. In the relevant information task (RIT, Experiment 3); the stimuli are the same as in Experiment 2 (IIT) but in each trial, the probability of winning is proportional to the number of boxes in that trial. If the number of blue boxes is more than the number of red boxes, then the likelihood for blue to win is more. The reason why these three experiments are chosen is as follows: the common point between Experiment 1 and Experiment 2 is the probability of the winning option. Blue wins 80% of the time in these two Experiments. However, the information presented on the screen, context, differs. Thereby, the effects of context in probability learning can be investigated. On the other hand, the common point between Experiment 2 and Experiment 3 is the information presented on the screen (context). However, Experiment 2 and Experiment 3 differ with respect to the uncertainty of winning. In Experiment 2, the uncertainty associated with the winning option is irrelevant with the context, but in Experiment 3 uncertainty can be directly identified and can be derived from the contextual information. This helps us investigate the effects of uncertainty. The other variable, reward, is studied by conducting all experiments twice, with separate subject pools, using either monetary rewarding feedback or non-monetary feedback (win/no-win information as feedback). In addition, in a small set of subjects with PD and an age-matched control group, we investigated dopamine based behavioral differences exhibited during probability learning in Experiment 4, under the same design of the Experiment 2 (IIT).

We hypothesized that when the participants are paid contingently on their correct responses, they will try to avoid making judgment errors. Thereby, monetary rewarding feedback will affect the participant's behavior positively and towards using maximizing strategy during learning. Our second hypothesis is that, when the probability learning experiment is presented within a context, participants will try to use this contextual cue information to guide their choice. Hence, where the context is irrelevant to the task, learning will be harder in comparison to the case when the context is relevant. Third, when the relevant contextual cue information is directly associated with the uncertainty of winning, choice reaction times will reflect this. Finally, when subjects are PD patients who receive dopamine replacement therapy, we hypothesize that this will influence the cognitive functionality and will result in differences in the probability learning performance compared to controls.

The organization of this thesis is as follows: Chapter 2 will give an overview of the probability learning literature. Concepts related to the probability learning framework will be summarized and studies on the probability learning mechanisms will be described. In Chapter 3, the rationale behind studying the effect of reward, context and uncertainty will be explained and behavioral experiments will be outlined. In Chapter 4, the results of the findings will be discussed in the light of previous research and finally, Chapter 5 draws a brief conclusion of the study.

CHAPTER 2

BACKGROUND AND LITERATURE REVIEW ON PROBABILITY LEARNING

Probability learning has been studied extensively to understand learning in humans within a probabilistic context as reviewed in Vulkan (2000). It has been suggested that the probability learning paradigm was first developed for testing the stimulus sampling theory (Bower, 1994; Poon, 1997). According to Estes and Suppes (1974, p.163) stimulus sampling theory stands for the situations

> …that the subject in a learning experiment samples a population of stimuli, or ‘cues’ on each trial, that his probability of making a given response depends on the proportion of sampled stimuli that are ‘conditioned,’ or ’connected,’ to the response, and that the connections between stimuli and response change as a result of reinforcement and non-reinforcement during learning.

Initially, probability-learning tasks were established to understand the principles underlying repetitive decision making situations (Humphreys, 1939). Later on, numerous variations of the probability-learning tasks have been established to test the effect of the payoff value of the feedback (Friedman et. al., 1964) the ratio of the most frequent event (Beach & Shoenberger, 1965), and the effect of the number of trials (Shanks et al., 2002).

In a classic probability-learning task, subjects practice a number of trials that contain a binary choice task (Siegel, 1964). The total number of trials changes from hundreds to thousands. Participants in a classical probability learning task usually see two lights and they are asked to predict which one of these two lights would be illuminated on each of a series of trials (Siegel, 1964). Experimenters usually instruct the participants to try to give as many correct answers as they can. For example, in an experimental design, light 1 has 0.8 probability to illuminate and light 2 has 0.2 probability to illuminate. Therefore, in a task that consists of 200 trials, light 1 illuminates 160 times and light 2 for 40 times on average. The specific task for the participant, then, is to predict which event will occur next. Depending on the subjects' choice, in each trial, the selected item is followed by immediate feedback. There are several different types of feedback in probability-learning tasks such as monetary reward or a written feedback.

In probability learning tasks, types of the stimuli (e.g., color, shape) or positions of the stimuli (e.g., right, left, up or down) may vary. This change occurs probabilistically and is used to facilitate the prediction of the outcome (e.g., correct/false or win/lose). The response of the participant indicates the his/her belief regarding whether a stimulus is winning or losing. In addition, each stimulus comes randomly and in each trial, the correctness of the feedback is independent from the previous feedback. Therefore, in the prediction or decision making stage, subjects have no deterministic information regarding which event will occur in the next trial. The argument is that people equally divide their predictions between options and update their predictions with respect to the outcome of their choices. For this reason, people can trust only the feedback information which comes from the preceding trials (Poon, 1997). According to Estes (1954), and Bower (1994), the feedback information from the previous trials serves as a base rate, which will be used to make a prediction in the forthcoming trials. Base rate refers to the unconditioned probabilities which come from the feedback of the preceding trials.

To sum up, probability-learning tasks examine subjects' predictions throughout the iterating trials in a discrete trial procedure. It has been supposed that the adaptive fashion of response behavior in such iterated tasks with feedback directs the participants to learn the correct probability of each stimulus set (Estes, 1954).

2.1. Underlying Strategies

Probability learning tasks serve as a test base for how the choice behavior changes with increasing experience. Second, the asymptotic level of the learning curve gives information about the *strategy* of the subjects. In the literature of probability learning, there are at least two suggestions for the learning strategy that might be used by decision makers. The first suggestion is the *frequency matching strategy*, which means that the proportion of the subjects' correct choices matches the frequency of the winning option. According to the example in the previous section, light 1 has a probability of 0.8 to be the correct answer and light 2 has a 0.2 probability to be the correct answer. If the participant's ratio of choosing light 1 matches light 1's actual frequency, this means that this participant uses the frequency matching strategy. The second suggestion is the *maximization strategy,* in which subjects do not change their predictions in every trial, but stick to the winning option and constantly predict on the same light.

The strategies mentioned above are directly measurable from behavioral responses. The tendency to use the frequency matching strategy fascinated investigators because it is a non-optimal strategy which does not allow people to achieve the highest possible number of correct answers (Siegel, 1964). On the other hand, *maximization strategy* which asserts choosing the same option in all trials, allows people to achieve more correct responses than the frequency matching strategy unless p is not equal to 0.5. For example, suppose P1 and P2 be the proportions of predictions of the subject

(P1 for light1 and P2 for light2). Let C1 = .8 and C2 = .2 be the actual proportions of the correct responses for light 1 and light 2, respectively. Then, according to Siegel (1964), the expected proportions of the correct responses (E(x)) would be calculated with the following formula:

$$E(x) = P1\ (C1) + P2\ (C2) \qquad \text{Equation (1)}$$

Because subjects are instructed to do their best and try to predict as much correctly as they can, they try to allocate their predictions in an efficient way. Therefore, over 200 trials the subjects who use the *maximization strategy* should devote their predictions as P1 = 1.0 and P2 = 0 The expected correct response proportion is as follows:

$$E(x) = 1.0\ (.8) + 0\ (.2) = .8 \qquad \text{Equation (2)}$$

However, in the frequency matching strategy, subjects allocate their predictions according to the proportion of the exact frequency of C1 and C2. In this case, the expected correct responses are:

$$E(x) = .8\ (.8) + .2\ (.2) = .68 \qquad \text{Equation (3)}$$

As can be seen from here Equation (3), the maximization strategy is superior to the frequency matching strategy. Interestingly, most people prefer the frequency matching strategy (West & Stanovich, 2003). According to Siegel (1964) one reason for why people use frequency matching strategy might be that, choosing the same option in every trial is boring for participants. But at least one exception is reported in the literature for people’s use of the maximization strategy rather than frequency matching. Shanks, Tunney & McCharty (2002) reported that 6 out of 12 participants used the maximization strategy in the last 150 trials in a total of 1800 trials. Probably, after extensive experience, subjects can derive some statistical information from the feedback information to build up a rule similar to the use of the maximization strategy, which makes them turn to choose the most profitable option in every trial. Animals are also known to exhibit probability learning behavior (Brunswik, 1939; Wilson & Rollin, 1959; Uhl, 1963; Karsh & Suppes, 1964; Kirk & Bitterman, 1965). According to Hinson and Staddon (1983) and Staddon (2003) most species such as rhesus monkeys and rats use maximizing strategy while learning probabilities.

In the context of probability learning, the term “strategy” which is taken from game theory, asserts that people use conscious tactics to win more in a simple game (Siegel, 1964). However, it seems that the amount of reaction time allowed by the experimenter for choosing an option is too short (2 to 5 seconds) for using a conscious strategy (Clark, 2004). Therefore, it has been suggested that probability learning might be a low-level, implicit thought process which embodies a multitude of mechanisms (Reber, 1989; Knowlton et al., 1996; Clark, 2004).

Pattern Searching Mechanism

In the previous section it has been told that there are at least two major learning strategies which are frequency matching and maximizing. These learning strategies are defined by the observed learning behavior. However, it is impossible to see the underlying mechanisms behind these strategies by just looking the learning curves or asymptotic levels as previously showed by the Figure 3.

In the literature it has been suggested that people simply look for patterns in sequences (Jarvik, 1951; Peterson & Ulehla, 1965; Yellott; 1969; Wolford et al., 2004) while learning probabilities. In the context of a binary choice probability learning task, pattern searching refers to “the deduction of future outcomes based on matching patterns recalled from past experiences” (Clark 2004, p.1). For example, in a binary choice probability learning task (Yellott, 1969), several blocks with 50 trials are used such that the actual frequency (i.e. the probability of an option being correct) changed across blocks. In the last block, the task changed such that the sequences of the correct lights are matched to the participants' choice to make the participant predict the outcome correctly in all of the 50 trials of this last block. In this study Yellott (1969) demonstrated that people continue to do frequency matching, even when they predict fully correctly in the last fifty trials. Post hoc analysis of participants' verbal reports showed that they had been looking for a sequence or a pattern and at the final block; they thought that they had found it.

Additional evidence for the ability to search for patterns in a binary choice task comes from the Heuristics and Biases research program of Kahneman et al. (1982). This program revealed that, people make systematic errors while judging probabilities. Kahneman et al. (1972b) demonstrated a common reasoning bias with their representativeness heuristic which was called gambler's fallacy. They showed that participants' predictions were affected by the previous outcomes when people had to guess heads or tails for a sequence of coin tosses. For example, most people regarded the sequence H-T-H-T-T-H to be more likely than the sequence H-H-H-T-T-T. For the participants, the second sequence of head and tails does not appear to be random. Thus, most of the people change their predictions after the same side (e.g., head) appears more than twice. Therefore, the gambler's fallacy bias generally refers to a false belief which asserts that a coin has a memory of previous experiences and if one side of the coin comes too often, then the probability of the other sides coming up will increase. For example, people believe that, after two times heads show up, the probability of tails in the next trial is higher than that of heads. Thus, most people think that fairness of a coin can be represented by the randomness of the sequences (Kahneman et al., 1982). At this point, researchers speculate that the ability to search for patterns causes the gambler's fallacy bias (Wards, 2002; Wolford et al., 2004). In the light of this example and many others, it can be said that people are searching for sequence of short patterns in repeated binary choice tasks.

Furthermore, Hardoon and his colleagues (2001) showed that people are looking for random patterns even when they are filling lottery tickets. They did an experiment with both pathological gamblers (PG) and non-pathological gamblers. The task is ranking 12 lottery tickets according to their possibility of winning, each containing 6 numbers chosen from 49 numbers. The lottery tickets were divided into several categories which were named as pattern (e.g., 5, 10, 15, 20, 25, 30), long sequence (e.g., 1, 2, 3, 4, 5, 6), non-equilibrated (e.g., 3, 5, 9, 12, 15, 17), and random (e.g., 1, 13, 19, 34, 40, 47). Most of the people from both PG and control group rank the random sequence category higher than the pattern, long sequence and non-equilibrated category because most of the participants believe that the degree of winning in the random sequence category was higher than the others. This research clearly showed that people have a false belief -very similar to the gambler's fallacy on the definition of randomness. Hence people are using this false belief in their choices.

Another way to examine the pattern searching mechanism is to measure the effects of previous trials on succeeding trials which is done by calculating the sequential dependencies between the outcomes of trials. Calculating the conditional probabilities between trials indicate some effect of sequential dependencies, thereby indirectly showing some evidence for the pattern searching mechanism. Friedman et al. (1964) provide evidence for the gambler's fallacy bias in a binary choice probability learning experiment. They analyzed the predictions of subjects by comparing their predictions as adjacent couples in a series of trials (first prediction followed by the next prediction such as Light1-Light1, Light1-Light2, Light2-Light1, and Light2-Light2). Their results showed that the conditional probability of choosing different options (Light1-Light2, Light2-Light1) is higher than the conditional probability of choosing the same options (Light2-Light2, Light1-Light1). In agreement with Friedman and colleagues, Lovett (1998) states that the frequency matching strategy is a natural by-product of choice processes which are sensitive to individual past experiences. Similarly, some early studies (Jarvik, 1951; Peterson & Ulehla, 1965) suggested that people look for sequential dependencies or patterns in sequences of events on binary choice probability learning tasks.

According to Wolford et al. (2004) the reason why people have such biases like the gambler's fallacy is not because people lack statistical knowledge or they fail to understand the meaning of independent trials, but it is because most of the time people automatically *search for patterns in random sequences* and at the same time they unconsciously learn the relative frequencies of events. In a binary choice probability learning task, Wolford et. al (2004) showed that, participants' degree of deviation from the frequency matching strategy decreases when a left hemisphere working memory load (such as remembering three digits shown in the previous trials) is introduced. In their experiment, participants in both conditions (working memory load and non-working memory load) learn the frequencies, but participants in the working memory load condition exhibit maximization behavior more than that of the participants in the non-working memory load condition. When subjects are burdened with a working memory task, they abandon frequency matching and switch to the easier maximizing strategy. This experiment showed that working memory is a necessary component of the pattern searching mechanism which directly affects the learning strategy by doing pattern searching.

Briefly, in a probability learning task where the stimuli are randomly ordered, searching for patterns in sequences of events influence people to make mistakes and such mistakes seem to bind the learning curve below the probability of the winning option.

Frequency Learning Mechanism

The frequency learning mechanism was first introduced by Hashner in the framework of memory and attention (for a review see Zacks and Hashner, 2002). The origins of that framework focus on the evidence that people of all ages unintentionally encode information about the relative frequencies of events. In this respect, they learn the relative frequency of words, pictures, events, etc. Thus, people are very accurate at responding to relative frequencies. According to Hasher and Zacks (1984, 2002) learning or encoding the relative frequency of an event is automatic (without intention) and is not affected by age, stress or arousal. Hasher and Zacks (1984, 2002) proposed that frequency learning mechanism is active during the time of decision making. Hence, frequency learning mechanism guides our choices towards the more frequent events. The proposal of Hasher and colleagues on the frequency learning mechanism is also supported by a vast amount of empirical research such as Zajonc, (1968) and Estes (1976a, 1976b). For example, Zajonc (1968) showed that people become sensitive for more frequent items (Chinese symbols) even those symbols have no meaning.

As a consequence of the aforementioned studies, converging evidence accumulates that the internal mechanism to search for patterns result in the frequency matching strategy observed in behavioral responses. On the other hand, when we consider the aforementioned studies, it is reasonable to think that the mechanism for searching for patterns of sequences and the mechanism for learning of relative frequencies might also be active concurrently, because during the search for patterns, frequencies are learned implicitly. Still, the relationship between these two mechanisms is not very clear and needs to be explored further.

2.2. The Effect of Reward in Probability Learning

Rewards have been used to understand probability learning since Thorndike (1911). Stimuli which induce positive conditioning are called rewards. Reward is one of the most powerful motivator that influences animal behavior. For example, if a monkey systematically repeats its behavior after an object is offered, the object is called reward and the behavior is called positive conditioning. Different species can be conditioned to different types of rewards like fruit juice for monkeys and money for humans. According to Thorndike's "law of effect", an animal's response to a stimulus that has a pleasant outcome is more likely to occur again. Conversely, a response to a stimulus with an unpleasant outcome is less likely to occur again when animals are faced with the same situation. In other words, the probability of a correct response with positive feedback tends to increase in time and the probability of a wrong answer with negative feedback tends to decrease in time (Thorndike, 1911). More specifically, there is a systematic relation between an animal's allocation of responses and the ratio of rewards. For example, if an option gives reward twice as more than the other option, then the animal chooses that option twice as more (Herrnstein, 1970). Thereby, rewards make it possible for an animal to learn the frequency of different options.

Reward processing in humans is less well understood because rewards in humans have lots of different forms. For instance, in humans, rewards can take more abstract forms like success, love or money. In general, rewards that are derived by our bodily needs (e.g. water or food) are called primary reinforcer and rewards that take more abstract forms are called secondary reinforcer. There is evidence showing that when feedback given for a choice is a monetary reward rather than a simple verbal feedback, people tend to choose the probable option more. For example, Myers, Fort, Katz, & Suydam (1963) did a binary choice probability learning experiment and compared the learning behavior for three different probability values P= 0.6, 0.7 and 0.8. They demonstrated that when the reward value increased, this made participants choose the highly probable option with higher frequency. This behavior (overmatching) is illustrated in Table 1. In addition, when three different reward conditions where the reward/punishment values changed between ±0 cents, ±1 cent, and ±10 cent were tested, overmatching behavior became more prominent. That is to say, the usage of overmatching strategy increases when the reward value gets increases.

Table 1 The proportions of correct choices in the last 100 trials out of 400 trials. Adapted from Myers et al. 1963.

	Probabilities		
Reward	P = 0.6	P = 0.7	P = 0.8
0 cents	0.624	0.753	0.869
1 cent	0.653	0.871	0.925
10 cents	0.714	0.866	0.951

Interestingly, under some circumstances, increasing the reward value alone may cause overmatching behavior (Lovett, 1998), making it difficult to justify how the two general mechanisms pattern searching and frequency learning, interact with maximization and frequency matching strategy. Some people might not search for patterns, because they might be using only the frequency learning mechanism, i.e., predicting only via a priori choice probabilities (e.g., frequency of options). Rewards may act as a switch between these two alternative paths. However, these comments are highly speculative and need to be tested extensively through future research.

2.3. The Effect of Uncertainty in Probability Learning

In a *binary choice* probability learning study (Friedman et al., 1964), it was demonstrated that people are sensitive to the level of uncertainty of the proportion of options. The experimental condition was to respond over 1,000 choice trials in 3 days. The success probabilities of choosing an option changed as 0.1, 0.2, 0.3, 0.4, 0.6, 0.7, 0.8, and 0.9 across blocks that contain 48 trials. The blocks are randomly ordered. In Figure 4, the average learning curves across different uncertainty levels are shown.

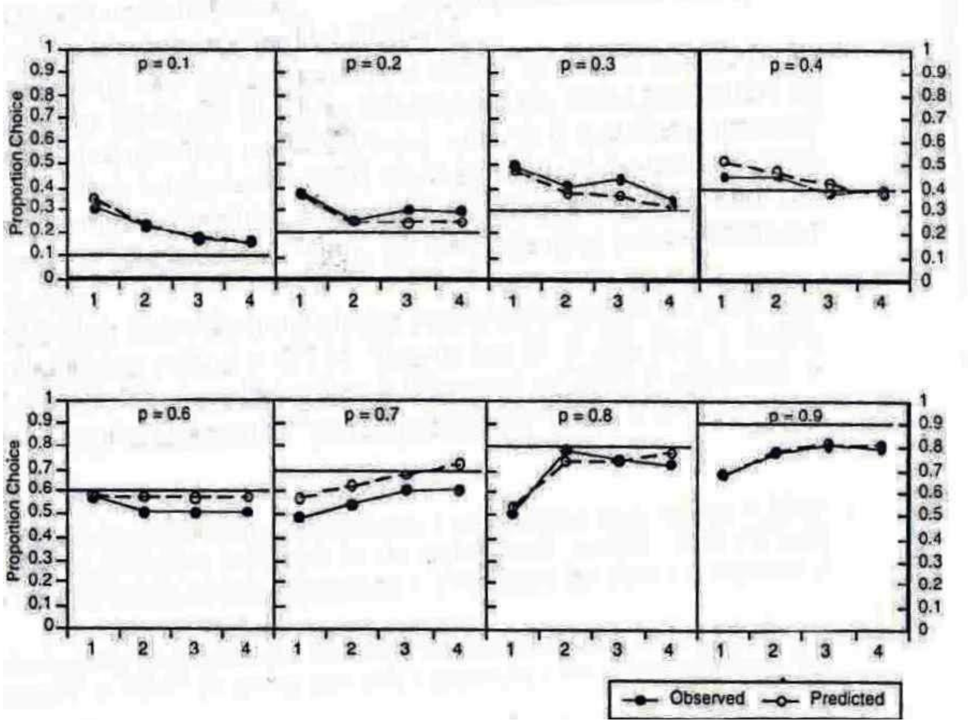

Figure 4 Graphs comparing the results of Friedman et. al (1964) study (observed) with the computational model (ACT-R) of Anderson et. al. (1998). Adapted from Lovett (1998).

In each graph in Figure 4, the horizontal axis shows the block number and the vertical axis shows the participants' proportions of choices for one of the two lights. It indicates that over time, participants' responses tend to match the actual proportion of lights. The degree of learning changes for different probability ratios. The solid lines in each graph represent the asymptotic levels for the learning curves. Thereby, it represents whether the participants learn the underlying frequency or not. In each condition, learning occurs and the proportion of choosing the highly probable option increases systematically. This study showed that people are sensitive to the level of uncertainty of each option if they have been instructed to give as many correct responses as they can.

Some cognitive psychologists apply information theory to the probability learning tasks to investigate the relation between choice reaction times and uncertainty. Information theory in psychology has been used to understand the human information processing capacity in general (Duncan, 2003). Researchers in this field mostly used the analogy of an ideal communication system to human subject (Laming, 1968; 1973). This analogy describes the human sensory system by using the concept of transmitter, channel and receiver. According to Laming (1968; 1973) the speed of the transmission of a message is limited by the capacity of the channel, so in practice, the reaction time is limited to the entropy (uncertainty) of the signal (e.g., message, or stimuli). For Laming (1968; 1973), analyzing the choice reaction time in terms of information theory is only meaningful if the experimenter instructs the subject to be as accurate as possible by not making mistakes and secondly when the subject is instructed to go as fast as possible. Previous studies showed that the probability of winning an option directly affects the choice reaction times (Hick, 1952; Hyman, 1953; Wollen, 1953; Quastler, 1955; Laming; 1973; Norwick, 1993). More specifically, when the probability of winning an option in a binary choice probability learning task is 0.5 the reaction time takes longer. High uncertainty is encoded as high entropy, which is attained when probability of winning option is 0.5. Therefore, it is a useful approach to analyze the choice reaction time data, if there is an uncertainty condition in an experimental situation.

2.4. The Effect of Context in Probability Learning

In the literature of judgment and decision making, “context” has various forms but in general it refers to the circumstances under within an event occurs. In this thesis, “context” refers to a change in stimulus representation. This change occurs only in the information content of a stimulus. For example, stimulus may contain two lights or more than two lights. Previous studies showed that the context of the stimulus directly affects the outcome of the choices (Sloman et al., 2003). In addition, it is believed that the stimulus context defines the premises of the reasoning problem (Held, Knauff, Vosgerau, 2006). For example, the ability to solve some probabilistic reasoning problems using representations in frequency format does not always mean that humans have been adapted just for frequency representations. Sloman et al. (2003) showed that humans perfectly overcome the cognitive bias introduced by a priori choice probabilities when representations of probabilities are presented in the venn diagram format. In agreement with this, according to mental model theory, it is reasonable to think that humans' can draw inferences from different mental representations (Held et al., 2006). According to mental model theory,

> If the organism carries out a "small scale model" of external reality and its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilizes the knowledge of past events in dealing with the present and future, and in every way to react in a much fuller, safer, and more competent manner to the emergencies which face it (Craik, 1943 cited in Held, et al., 2006, p.11).

As a result, it is a challenging question how the probabilistic information is represented in the mind/brain and how the brain evolved by natural selection for such probabilistic representations (Pinker, 1997). Previous studies suggest that probabilistic reasoning is not limited only to frequency representation of probabilities. Indeed, performing better at some probabilistic reasoning problems with the representations in the frequency format do not mean that people cannot make probabilistic reasoning using probability representations in other formats. Sloman et al. (2003) showed that by changing the stimulus context, people can derive probabilistic information and use it while they were reasoning on choices. More specifically, he demonstrated that adding a relevant stimulus dimension increases the accuracy of a decision making situation. Under the light of this evidence we investigate the effect of stimulus context in the second and third experiment. We suggest that in a probability learning task, the irrelevant and relevant stimulus dimension might have such effects on the representations of probabilities. For example, Edgell and Castellan (1996) showed that, using irrelevant cue information in a probabilistic learning task affects the learning performance negatively.

2.5. Neural Correlates of Probability Learning

This section reviews the plausible neural mechanisms underlying the probability learning ability, because by understanding the neural underpinnings of probability learning paradigm, we might develop a better understanding of the cognitive dysfunction of Parkinson patients who have been tested as a part of this thesis in Experiment 4. This section more specifically reviews the basal ganglia dopamine mediated reward system as a relative frequency learning mechanism and the prefrontal cortex working memory system as a pattern searching mechanism. Different methodological findings like lesion studies and imaging studies are discussed in detail.

Underlying strategies

As discussed above, there are at least two mechanisms responsible for probability learning ability (a) the pattern searching mechanism and (b) the frequency learning mechanism. Empirical evidence support the idea that, the pattern searching mechanism has its neural correlates in the prefrontal cortex, specifically in dorsolateral prefrontal cortex (DLPFC) which is associated with the working memory mechanism, and the frequency learning mechanism has its neural correlates in the basal ganglia, specifically in the striatum which is associated with procedural memory.

The idea that there are two separate mechanisms responsible in probability learning ability is inspired from double dissociation studies (Knowlton, Squire & Gluck, 1994). Moreover, Wolford et al. (2000) showed that in a probability learning task, split brain patients who were trained with their right hemisphere showed maximizing behavior. However, when the patients are trained with their left hemisphere, they showed frequency matching behavior. Both patients learn the probabilities, but their strategy changes when they perform the task with different hemispheres. On the other hand, people with dorsalateral prefrontal cortex damage do learn probabilities, but again their strategies differ (most of the time they do frequency matching, Wolford et. al, 2000). It is plausible that basal ganglia is effective on learning of probabilities (frequencies) and it is sensitive to rewarding stimuli but dorsalateral prefrontal cortex (possibly left dorsalateral prefrontal cortex) seems to be only effective in pattern searching. When the neuroanatomical connections between basal ganglia and dorsolateral prefrontal cortex are being considered (Alexander, DeLong & Strick; 1986; Middleton & Strick; 2002), dorsolateral prefrontal cortex has connections with dorsal striatum area of the basal ganglia which consist of caudate nucleus and putamen (Figure 5). The dorsal-striatal pathway is involved generally in cognitive aspects of learning (Seger, 2006). Therefore suggesting that these regions might somehow be involved in probability learning is very reasonable.

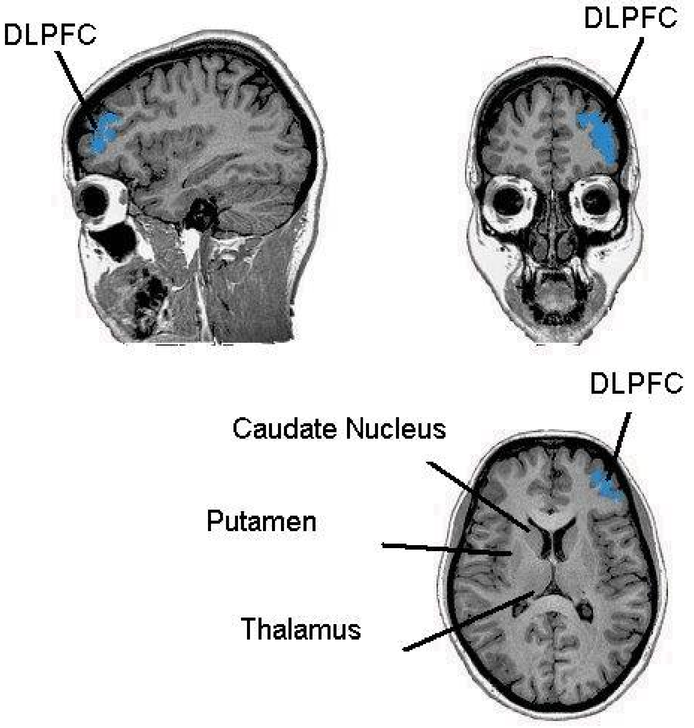

Figure 5 DLPFC and Basal Ganglia Anatomy. Created by Brain Voyager tutor (www.brainvoyager.com)

Neural Correlates of the Pattern Searching Mechanism

Berthoz (2005, p.1) made a useful analogy for the function of prefrontal cortex that is “prefrontal cortex is the center of executive functions like the big companies has their management offices at the top of the skyscrapers”.

It has also been acknowledged by a multitude of studies that prefrontal cortex is the place where the working memory is (Jonides et al., 1993; Petrides et al., 1993; McCarthy et al., 1994; D'Esposito et al., 1995a; Smith et al., 1995; Cohen et al., 1997; Courtney et al., 1997; Courtney et al., 1998; Courtney, et al., 1998; Miller & Cohen, 2001; Baddeley, 2003; Owen et. al, 2005). According to Baddeley (2003) working memory, which has evolved from the concept of unitary short term memory, is a term that refers to a brain system which provides temporary storage for different types of information such as numbers, words, and spatial places. Another definition supports the view that working memory is also involved in real time (online) monitoring processes for temporal ordering (Shimamura, 1995, Fuster, 1990; Baddeley, 1992; Goldman-Rakic, 1992). In the context of the pattern searching mechanism; working memory operates by correlating subject's judged order of items with the actual presentation order (Milner, 1971, Milner et al., 1991). In addition, findings suggest that patients with prefrontal cortex damage suffer from a lack of learning temporal order (Shimamura, 1995) and dorsalateral prefrontal cortex has been found to be involved in remembering the temporal order or sequence between consecutive trials (Owen et al., 2005). Also, numerous imaging studies reveal that in many of the working memory tasks, the human prefrontal cortex gets activated (for a review see, Miller & Cohen, 2001; Courtney, Petit, Haxby & Ungerleider, 1998).

When we review the literature on animal probability learning, we realize that some low level animals such as rats do not have the ability to search for patterns; however, they show maximizing behavior in a repeated binary choice task (Uhl, 1963). According to Staddon (2003), the reason why lower animals do not search for patterns is because they either have no or limited working memory. It might be that lower animals have no such complex and developed prefrontal cortex as humans (Semendeferi, Lu, Schenker, Damasio, 2002). The human species is the only species that has a highly developed prefrontal cortex which is the major place responsible for working memory.

A novel approach is presented in Huettel et al. (2002) for studying human pattern searching mechanism. The experimental condition that they perform is similar to the "gambler's fallacy", the belief that "chance events form coherent patterns", as described in detail in Ivry & Knight (2002). They instruct the participant to guess between two stimuli, a circle or a square. Experimenters also explicitly instruct the participants that the events are randomly determined. However, in the real task conditions, the series of patterns are not random. Two main patterns occur throughout the experiment: the first pattern is composed of consecutive trials where the two stimuli are altered (circle-square-circle-square-circle) and the second pattern consists of one stimulus followed by a series of the other stimulus (circle-square-square-square). They demonstrate that repetition violations in the outcomes that are similar to the second pattern activated the regions in inferior frontal cortex bilaterally; cingulate cortex, insular cortex and basal ganglia. Huettel and his colleagues link this activity to a pattern searching process of the working memory operations, which is associated with prefrontal cortex (for a detailed discussion see Ivry & Knight, 2002).

Moreover, Miller et al. (2005) also showed activation patterns in the prefrontal cortex including the right dorsalateral, ventrolateral cortex, and the right medial frontal cortex in a classical probability learning task. More specifically, the task was to predict the place of the stimulus by pressing the up or down keys. The stimulus appears on the right side of the screen with a probability of 70% and on the other side with a probability of 30%. They demonstrated that, the activation patterns found in the right dorsolateral, and parietal cortices were significantly altered according to the degree of deviations from the frequency matching strategy. The patterns of activations that Miller et. al (2005) reported were consistent with the previous findings in other types of probability learning tasks (Elliot & Dolan, 1998; Elliot et al., 1999; Paulos et al., 2001; Schubotz & von Cromon, 2002; Volz et al., 2003).

Furthermore, Bogacz et al. (2007) perform a rising optimum task (Egelman et. al 1998; Montague & Berns, 2002) which is a kind of dynamic version of the probability learning task. In this task, participants had to predict between two options where only one of them has a rewarding outcome in each trial. The difference is that, participants see their overall performance and they were instructed to keep their overall performance as high as they can. If they choose the same side too often, they will get more reward in the long period but if they change their prediction frequently, they will get more immediate reward, but less long term reward. The optimal strategy in this task is to predict on the same light in every trial which increases the long-term gain. Bogacz and his colleagues administered this task with three different response time constraints. The result of this experiment shows that when the participants are let to respond slower, they switch between options (frequency matching strategy) more frequently. However, when the constraint to give a response is shorter, they stick to the winning option (maximizing strategy). Bogacz et al. (2007) relate the results of their study with working memory, because increasing the time limit for giving a response, gives opportunity of access to working memory, hence increasing the degree of pattern searching. In addition, they also support the argument that even when the basal ganglia dopamine neurons learn the frequency of each option, a short-term memory component is necessary to compute the "eligibility traces", memory traces which store the temporary information outcomes of a participant's choices.

Frequency Learning Mechanism

Learning of probabilities or event frequencies is an ancient skill (Cosmides et al., 1993), so it is not surprising that it is localized in the phylogenetically older parts of the human brain. The most important brain areas responsible for probability learning are the basal ganglia (Packard and Knowlton, 2002).

It has been suggested that probability learning affects implicit learning of event sequences that are acquired independently of conscious effort, making subjects learn via unconscious intentional strategies (Reber 1989). Evidence shows that the basal ganglia play an important role in most of the implicit and procedural learning tasks (see Packard and Knowlton, 2002; Squire, 2004 for a review). The basal ganglia are a collection of sub cortical structures that have been implicated in skill learning and procedural memory. Procedural memory, which is subsumed under the non-declarative memory branch of multiple memory systems, is also thought to be responsible for habit and skill learning (Squire, 2004). Basal ganglia are connected to several cortical areas by multiple cortico-striatal loops which include divergent projections to prefrontal cortex (Seger, 2006). It is also at the center of the dopamine pathway, producing behavior linked to rewarding stimuli which makes its role critical for learning of probabilities. The most important theory behind the basal ganglia probability learning system is that, the dopamine neurons in the basal ganglia code the prediction error signal[2] (see Schultz, Dayan & Montague, 1997; Schultz, 2006, 2007 for a review) while learning to predict between choices. This is also supported by basal ganglia activity found in imaging studies related to probability learning (see O'Doherty, Dayan, Friston, Critchley, Dolan, 2003; McClure, Berns, Montague, 2003; Seymour, O'Doherty, Dayan, Koltzenburg, Jones, Dolan, Friston, Frackowiak; 2004).

Reward

Neurotransmitters are chemicals which influence and modulate the electrical signals between neurons. Dopamine is the major neurotransmitter in the brain’s reward system (Berridge & Robinson, 1998, 2003; Schultz, 2006). The midbrain dopamine neurons become active both during acquisition of reward related behavior (e.g., action preparation) and at the time of

---

[2] The reward prediction error signal codes the difference between the expected reward and the actual reward, that is this signal works as a teacher which is very similar to the temporal difference reinforcement learning algorithm (Sutton & Barto, 1998).

subjective responses (e.g., hedonic feelings like joy). These two different roles of dopamine in reward conditioning correspond to two major functions: “Wanting”, which relates to action preparation behavior, and “Liking”, which mediates subjective feelings of hedonic pleasure (Berridge & Robinson, 2003). The *liking* function of dopamine neurons contributes to positive emotions. The investigation of liking function is hard to study (Rolls, 2005), therefore, large amounts of literature concentrate on the *wanting* function, which creates the predictive association between a stimulus and a reinforcer. The wanting process can be further separated into sub-processes such as reward prediction: a neural process which occurs before decision making, reward detection: remembering previous rewards, and reward expectation: anticipating the rewards that occurs after decisions. These sub-processes are carried out by several different areas of the brain as can be seen in Figure 6.

According to Berridge & Robinson (1998, 2003), the key role of dopamine neurons is mediating the incentive salience of rewards. As summarized by McClure (2003), “dopamine release assigns incentive or motivational value to objects or behavioral acts”. For example, in mice whose dopamine receptors had been blocked, internal valuation of rewards did not change, but reward seeking actions were impaired (Berridge & Robinson, 1998).

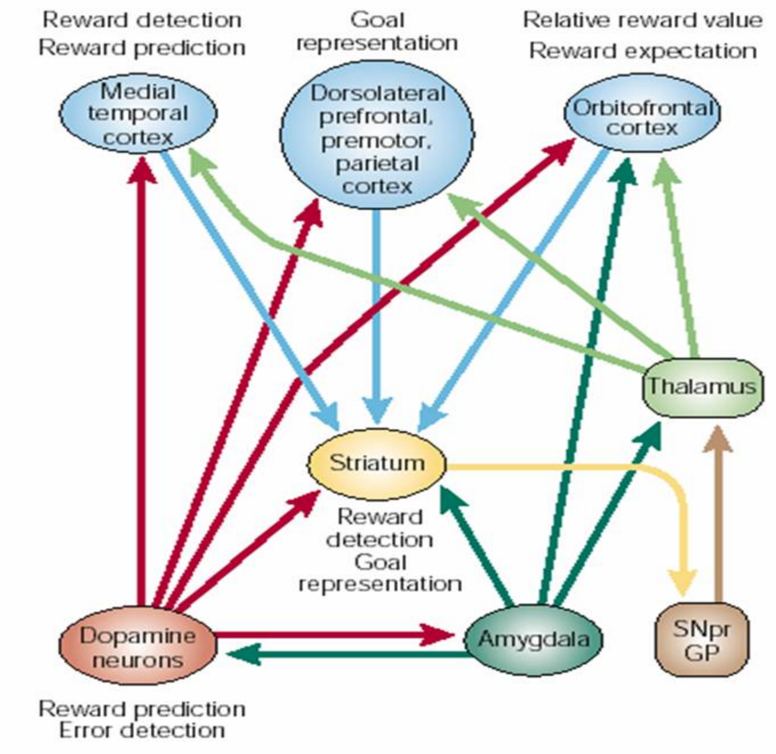

Figure 6 Theoretical network showing multiple brain areas involved in reward related signal (Schultz, 2000 p.3).

Dopamine is produced by the neurons located in the ventral tegmental area and substantia nigra (see Figure 7). The midbrain dopamine neurons, which have cell bodies in the substantia nigra and the ventral tegmental area have divergent projections that have connections to the frontal cortex, dorsal and ventral striatum, and other forebrain regions. Reward processing in the brain is included in a very complex network which is called the “ventral valuation network” (Montague et al., 2006). The ventral valuation network (VVN) refers to a collection of subcortical and cortical brain regions which has connections to nearly every part of the brain.

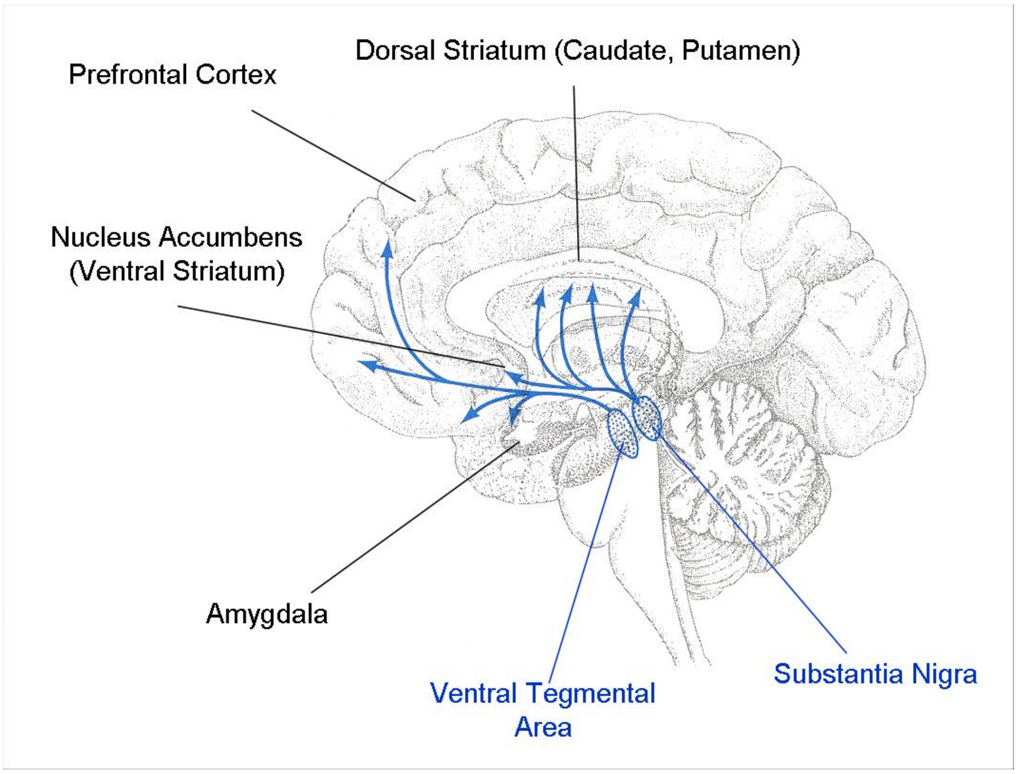

Figure 7 Schematic representation of the human dopamine pathways
Adapted from Yahel (2005).

The VVN is open to different modalities and extracts information from several cortical areas in olfactory, auditory, visual and somatosensory cortex. These connections facilitate multiple forms of rewards such as beautiful scenes, nice smells and even hedonic pleasures such as nicotine[3]. In the midbrain, dopamine neurons show phasic[4] excitatory firing response (activation) to reward predicting stimuli such as food rewards, somotosensory, visual and auditory (Schultz, 2007). These dopamine activations to the reward predicting stimuli occur nearly in 80%-90% of the substantia nigra and in the ventral tegmental area dopamine neurons (Schultz, 2007). Imaging studies of reward processing got started with primary reinforcers (O'Doherty, 2004) and continued with secondary reinforcers. Secondary reinforcers like money (Breiter et al., 2001, Elliot et

---

[3] Dopamine neurons also play an active role in drug addiction (Nestler & Aghajanian, 1997; Chao & Nestler, 2004; Nestler, 2004).

[4] Phasic activity refers to the neural firing which turns to its resting potential when the stimulus still applies.

al, 2003; Knutson et al., 2003), cultural rewards like brand products (Erk et al, 2002; Kawabata & Zeki 2004; McClure et. al 2004; O'Doherty 2006) and even social rewards like trust and love (Bartels & Zeki 2004; King-Casas et. al., 2005; Delgado et al., 2005) were used extensively to understand reward processing in the human brain. Some of the imaging studies showed activation to a variety of rewards in the sub compartments of the striatum which is a part of basal ganglia.

The reward system in humans is also integrated into different functionally localized brain regions including areas for decision making (prefrontal cortex in general), stimulus-response learning (basal ganglia), goal representation (dorsalateral prefrontal cotex, premotor cotex, parietal cortex and striatum), reward detection (medial temporal cortex, striatum), reward prediction (medial temporal cortex, dopamine neurons), reward expectation (orbito frontal cortex) and error detection (dopamine neurons) (Figure 6). Within this framework, in a probability learning task, neural activity is supposed to be in the prefrontal cortex occurs at the prediction stage while activity in the basal ganglia occurs during the period when reward is anticipated.

Dopamine neurons in the midbrain are also sensitive to the value of the expected reward. Tobler et al. (2005) showed that dopamine neurons fire for different combinations of probabilities and magnitudes of rewards depending on the expected reward value. Dopamine activation for an expected reward value has also been reported in magnetic resonance imaging (fMRI) studies (Knutson et al., 2005; Tobler et al., 2007).

An important component of the dopamine mediated reward signal is the prediction error signal. This prediction error activity of dopamine neurons is first reported by Schultz et al. (1997) and repeatedly observed by other researchers (Satoh et al. 2003, Morris et al. 2004, Nakahara et al. 2004, Bayer & Glimcher 2005, Pan et al. 2005). According to the prediction error theory, dopamine neurons do not fire when the reward predicting stimulus fully predicts the reward, but the level of dopamine neuron firing is equal to the difference between the occurred reward and its prediction. For example, after a period of learning, reward omissions produce a depressing neural response which was called negative error prediction. The prediction error signals are described as an indicator of learning. Therefore it can be stated that in the period of learning, the dopamine neurons are programmed to decrease the prediction error signals[5] (Schultz, 1997). Using neuroimaging, the prediction error theory was tested with human subjects, where several studies found activations in the striatum area of the basal ganglia (O'Doherty et al., 2003; Seymour et al., 2004; McClure, Berns, Montague, 2003).

Uncertainty

Striking new evidence on dopamine neurons is that increasing tonic activity[6] is obtained with respect to an increase in stimulus uncertainty (Tobler et al., 2003; Morris et al., 2006). Monkey dopamine neurons are reported to be sensitive for reward predicting stimuli when each stimulus had a different probability of receiving reward. This phenomenon co-exists with the phasic activity discussed before, which specifically predicts the error during the time course of learning of stimulus-action pairs. Therefore, it can be stated that dopamine neurons have different firing patterns for uncertainty and for expected reward. Figure 8 shows the average dopamine neuron firing recorded from two monkeys. Dopamine activity increase when the stimulus uncertainty increases.

---

[5] The equivalent of the prediction error hypothesis in computational reinforcement learning theory is the temporal difference learning algorithm (Sutton and Barto, 1998).

[6] Tonic activity is the slowly increasing ramp when there is a stimulus applied. It is measured by the single neuron electrophysiology studies.

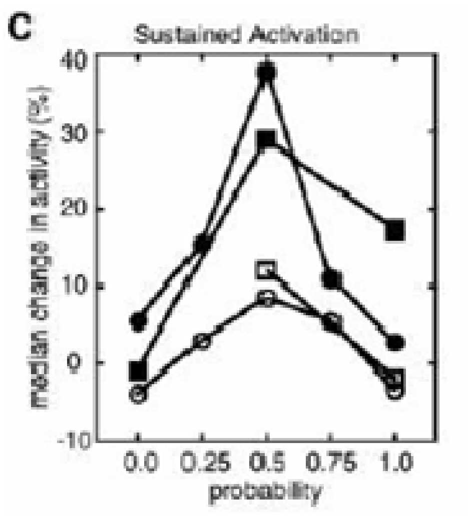

Figure 8 Uncertainty in dopamine neurons. Adapted from Tobler et al. (2003).

Recently, uncertainty in reward processing has also been studied by various researchers using neuroimaging (Volz et al, 2004; Dreher et al., 2005; Preuschoff et al., 2006; Tobler et al., 2007). All of these imaging studies showed activation in the basal ganglia. Interestingly, signals associated with expected reward and uncertainty create different BOLD (blood oxygenated level dependent signal) profiles, indicating that these two processes also differ at the gross level.

2.6. Probability Learning in Parkinson's Disease

The aforementioned studies suggest that the role of dopamine neurons in learning is critical when the context is probabilistic, because the dopamine neurons in the midbrain may code uncertainty and expected reward value at the same time. It is important to keep in mind that these firing patterns of dopamine activity can change due to the requirement of the tasks and the pathological conditions of the participants.

Parkinson disease (PD) is a neurodegenerative disorder which affects the central nervous system. Idiopathic Parkinsonism is generally caused by a reduction in the number of dopamine producing neurons in the substantia nigra and ventral tegmental area. This reduction of dopamine neurons impairs an array of cortico-striatal activity associated with motor as well as cognitive functions. Impairments in the motor functions are tremor (shaking), rigidity (muscle rigidity), akinesia (total loss of movement), and bradykinesia (slowness in physical movement). Cognitive impairments relate to executive functions or memory (both procedural and declarative) as will be explained below.

Cortico-Striatal Cognitive Dysfunctions in Parkinson Disease

Various cognitive functions are linked to synchronized communication of the frontal cortex and the basal ganglia, achieved by the cortico-striatal loops. Cortico-striatal dysfunction may lead to differences in cognition as evidenced by various disorders like the attention deficit hyperactivity syndrome or Parkinson disease (Frank et al., 2005). In PD, a generalized procedural learning deficit (Knowlton, Squire & Mangels, 1994), as well as impairment in ability to learn probabilities (Frank et al., 2006) have been established.

Probabilistic learning deficits have been studied using the weather prediction tasks, originally designed by Knowlton et al. (1994). In such a task, different cues are present on a pair of cards. Participants study the cards, as well as card combinations. In the task, they have to predict whether the cards presented in a given trial are associated with rain or sunshine. Healthy participants learn this probabilistic relationship over multiple trials (Gluck, Shohamy, & Myers, 2002). However, non-medicated Parkinson patients are impaired in the weather prediction task (Knowlton et al. 1994). Researchers think that the basal ganglia dysfunction is responsible from this type of impairment. In addition, according to Maddox (2005), Parkinson patients may be impaired when the learning task is based on trial-by-trial feedback, because the basal ganglia are also implicated in feedback mediated learning. This learning impairment may be due to the deficiency in the dopamine mediated reward signal that drives trial by trial feedback learning (Aron et al., 2004).

An explanation may be that damage to overall dopamine neurons reduces both phasic and tonic activity, thereby, diminishing the effectiveness of the procedural learning system (Frank et. al, 2005). It has been shown that phasic burst are associated with positive error prediction signal whereas phasic dips are associated with negative error prediction signal. These error prediction signals which are critical for learning might not occur in Parkinson patients. On another front, according to Frank et al (2006), phasic burst and dips of dopamine activity during positive and negative feedback may not occur due to modified synaptic plasticity in Parkinson patients. Finally, the impairments may be due to timing problems. Maddox et al. (2003) showed that learning impairments occur if the feedback that creates the reward signal is delayed approximately 2.5 seconds after each response. It is quite possible that dopamine reduction may also affect this circuitry.

The Effect of Parkinson Disease Treatment on Probability Learning

The most common treatment for Parkinson disease is dopamine replacement therapy[7]. In general, motor dysfunction is caused by reduced dopamine activity in the putamen. Although dopamine medication helps increase the dopamine activity to normal levels in putamen, it also increases the dopamine activity in other areas of the brain above the normal level. In the literature, side effects of dopamine medication are not generally reported. Frank et al. (2005) states that "many cognitive studies on Parkinson Disease do not take into account the level of medication administered to patients." Thus, confounding interpretations occur regarding the learning deficits of these patients. Whether the impairments are due to the abnormal dopamine activation caused from the disease or from the treatment is unclear. According to Frank et al. (2005):

---

[7] In general dopamine agonists or dopamine precursors (Levodopa) are used in dopamine replacement therapy.

> “If an effect is found, it is difficult to know if this effect stems from a lack of DA [dopamine] in PD [parkinson disease], or is somehow related to the medication. For instance, medication results in elevated levels of tonic DA in undamaged areas. This may prevent phasic dips from being effective and degrade performance when they are functionally important”.

Frank et al. (2003) demonstrated that Parkinson patients compared to normal controls, Parkinson patients on medication are better in learning relative probabilities with positive feedback. On the contrary, patients off medication are better in learning with negative feedback compared to controls (Frank et. al 2003). Interestingly, Parkinson patients who have dopamine dysregulation syndrome, which is caused by excessive use of dopamine drugs, develop pathological gambling[8] behavior. In addition, it is hypothesized that patients exhibiting pathological gambling behavior do not learn by negative feedback (Ross et al., 2006).

Summarizing this section, there are multiple mechanisms responsible for probability learning and there are several factors effecting these mechanisms, thereby, affecting the probability learning performance (i.e., learning strategy). A graphical summary of this chapter is given in Figure 9.

---

[8] According to DSM-IV (American Psychiatric Association, 1994), pathological gambling is characterized by repetitive and persistent gambling behavior which has negative effects on family, personal, and professional life.

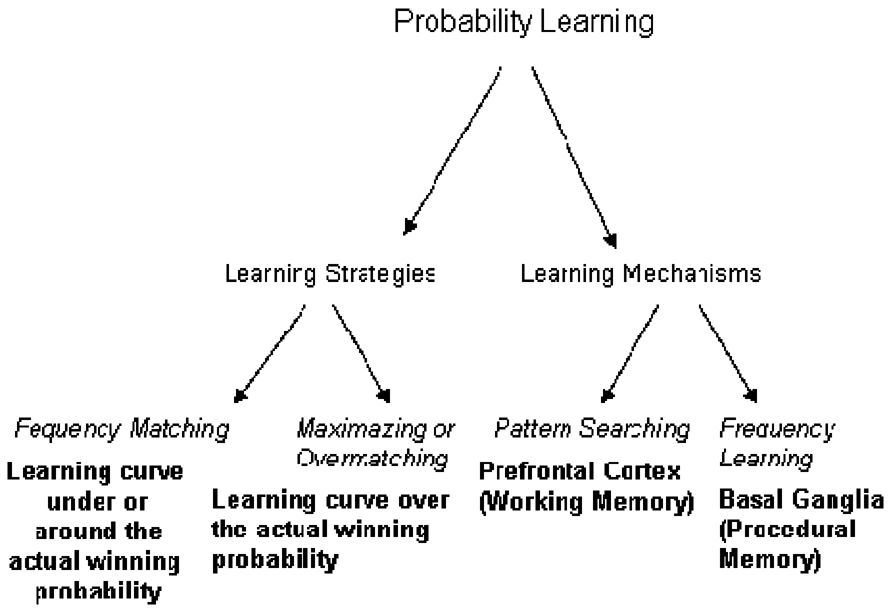

Figure 9 Summaries of the probability learning strategies and mechanisms.

CHAPTER 3

BEHAVIORAL EXPERIMENTS

Performance in a probability learning task could be influenced by a variety of factors (reward value, context & uncertainty). In this thesis, three behavioral experiments are conducted to investigate the effects of reward, relevance and irrelevance of context, and uncertainty on learning of probabilities. Furthermore, a small group of Parkinson patients are examined to investigate behavioral differences caused by medication.

3.1. Overview of the Experiments

We recorded data from four experiments. In the first classic probability learning experiment (CT), subjects have to choose between two colored boxes (red and blue). We expected that participants will learn to predict to choose the more rewarding color. The rewarding outcome between these two colors is distributed probabilistically and only one color has a rewarding outcome in a trial. Overall, in each session, blue option wins 80% of the time and red option wins 20% of the time. There are two conditions. The first condition gives monetary feedback after each correct trial and the second condition gives only positive non-monetary feedback after each correct trial. We compare learning behavior between these two conditions.

For the non-monetary feedback condition, participants receive “You won” and “You did not win” phrases for the positive and negative feedbacks respectively. For the monetary payoff condition we use “You won 1 YTL” for the positive feedback condition and “You did not win” for the negative feedback condition. It is important to note that, in all feedback conditions there is no punishment like losing money or decreasing in overall performance. We did not add a punishment condition because it has been claimed that experiments and theories underlying learning with punishment are different from theories learning with rewards (Seymour, Singer & Dolan, 2007).

In the second experiment, there are ten colored boxes shown on the computer screen. Each box is either red or blue; however the number of boxes shown is irrelevant for winning in that trial. Similar to the Experiment 1, blue choice wins 80% and red choice wins 20% of the time. This is the reason why we named the Experiment 2 as irrelevant information task (IIT). In the period of learning, subjects have to learn to choose the blue response and suppress the irrelevant information. This approach will lead us to understand how people learn the probabilistic information by suppressing the irrelevant information. Again in the second experiment, we have two reward conditions for correct feedbacks (monetary feedback & non-monetary feedback).

In the third experiment, the cue information which is represented by the number boxes is relevant. In this experiment there are ten colored boxes on the screen and the probabilities of the reward outcomes are represented explicitly according to the number of boxes. For example, if there are nine blue boxes and one red box, blue has a rewarding outcome with a probability of 0.9 and red has a rewarding outcome with a probability of 0.1. Subjects have to learn to weigh their subjective probability of winning in a trial according to the number of blue and red boxes. This is the reason why we name this task, relevant information task (RIT). In the third experiment, the proportion of blue and red winning is balanced in the total of 200 trials such that blue choice wins 50% and red wins %50 overall. Similar to

Experiment 1 and 2, there are two reward conditions: monetary and non-monetary.

The common points and differences across these three experiments are summarized as follows:

- In Experiment 1 (CT) and Experiment 2 (IIT), overall probability of winning for blue is the same which is 80%. However the context which is represented by the stimulus on screen is different.
- In Experiment 2 (IIT) and Experiment 3 (RIT) the context is the same but overall probability of winning for blue is 80 % and % 50 percent respectively. Furthermore, the probability of winning is unrelated to the context in Experiment 2, but directly related to the context in Experiment 3

Using data from these series of experiments the effects of reward, uncertainty and context are investigated as follows.

- In each Experiment 1, 2 and 3, the effect of reward is studied by comparing two subject pools, one receiving monetary feedback and the other receiving non-monetary feedback.
- The effect of context is studied by comparing behavioral data from Experiment 1 and 2, for which winning probabilities for both choices are exactly the same, but only the stimulus presented on the screen differs.
- The effect of uncertainty is studied by using data from Experiment 3, for which the uncertainty of choosing the winning option is directly related to the proportion of the number of colored boxes on the screen.

The effect of dopamine replacement therapy in probability learning is studied by analyzing data from Experiment 4 which consists of the same test in Experiment 2 (IIT). In Experiment 4 the only difference from Experiment 2 is the subject population which is a group of medicated Parkinson patients and age matched controls. Finally, in an explorative study we investigated the role of probability learning strategies using data from Experiment 1 (CT) and Experiment 4 (IIT). The experiments and analyses are summarized in the table below.

Table 2 Experiments and investigated factors in probability learning.

Experiment Number	**Reward**	**Uncertainty**	**Context**	**Learning Strategy**
1	✓		✓	✓
2	✓		✓	
3	✓	✓		
4				✓

Our hypotheses are as follows:

1. Monetary rewarding feedback will affect the participant's behavior positively and towards using maximizing strategy during learning.
2. When the probability learning experiment is presented within a context, participants will try to use this contextual cue information to guide their choice. Therefore, in IIT, learning will be harder in comparison to the RIT.
3. In RIT, choice reaction times will reflect the uncertainty associated with winning.
4. In PD patients who receive dopamine replacement therapy, probability learning performance will differ in comparison to controls who have less dopamine in their system.

3.2 EXPERIMENT 1: Classic Probability Learning Task (CT)

3.2.1 Method

Participants

Forty-four participants were all Middle East Technical University students or employees. They had been recruited by assigning posters to a variety of department bulletin boards. All participants were randomly allocated to one of the two experimental conditions: the first condition monetary feedback (n=22, 11 males, 11 females, 22.3 mean age, SD 2.8) or the second condition non-monetary feedback (n=22, 10 males, 12 females, 21.4 mean age, SD 2.1). Participants in the monetary feedback condition received financial payoff according to their performance in the experiment, and the participants in the non-monetary feedback condition received nothing at the end of the experiment.

Stimuli

The stimuli in Experiment 1 were two differently colored boxes. The spatial locations of the colored boxes are randomly distributed and counter balanced in the total of 200 trials. In each trial participants saw these colored boxes as shown in Figure 10.

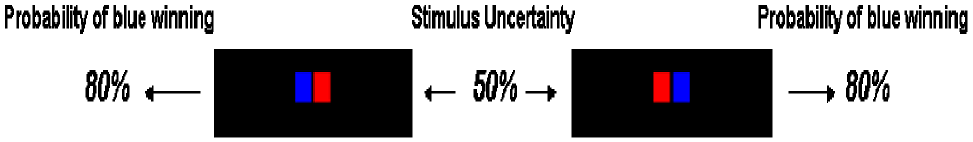

Figure 10 Stimulus uncertainties[9] are 0.5 but the probability of blue winning is 0.8.

---

[9] Stimulus uncertainty refers to the proportion of blue and red boxes.

Procedure

The data is collected by using a desktop computer. Each participant sat facing a computer screen. Participants gave their responses by pressing one the two designated response keys “M” for “blue” and “K” for “red” on the keyboard. Participants took the experiment in the Experimental Psychology Laboratory of the Department of Cognitive Science, METU. SuperlabPro[10] experiment design and management software was used for stimulus presentation. Before the experimental session started, participants signed an informed consent form (Appendix-A), then they were briefed (Appendix-B) and after finishing the experimental session they were debriefed (Appendix-C).

The experiment began with a test session in which participants practiced 5 trials. In the practice session, participants learned to press the keys. In all 5 practice trials the feedback is correct (positive). They had been told that if they did not respond to a stimulus in 5 seconds after it was presented, they would get a negative feedback. They were instructed to give as many correct answers as they could.

The practice session was followed by the test session. The test session was the same as the practice session, except that responses given in the experimental session were included in the analysis. Responses and reaction times were recorded in all 200 trials. Once the participant has pressed on a key, the computer proceeded to select a winning color independent of the participant’s choice. Then the participants were provided with feedback on the actual outcome regarding their choice and the reward amount. The feedback was presented for 3 seconds and after the feedback is presented, the next trial started in 2 seconds (Figure 11).

---

[10] More information can be obtained from http://www.**superlab**.com/

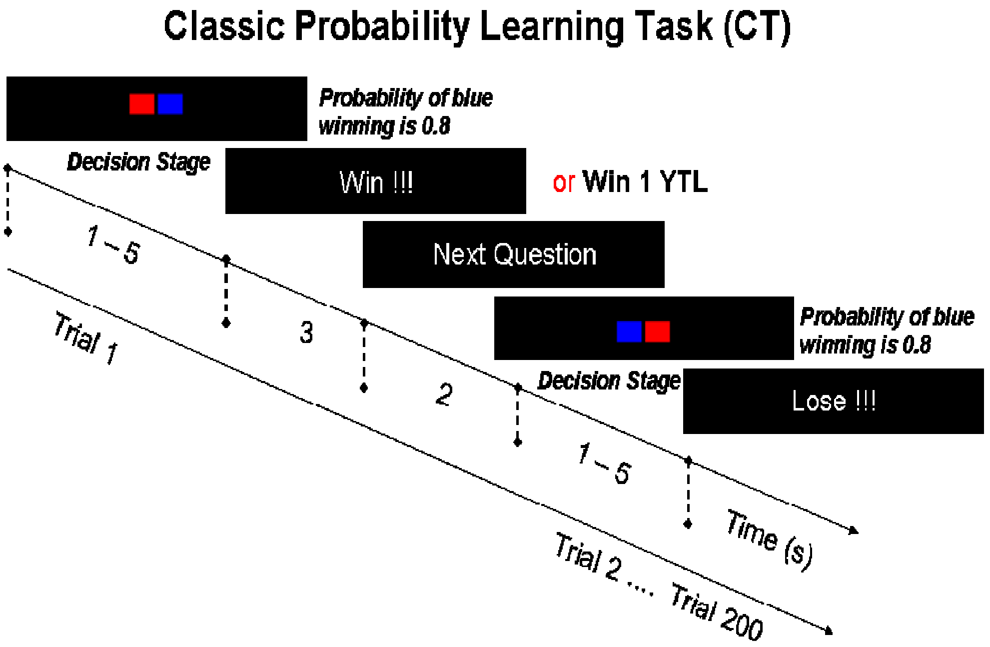

Figure 11 The experimental procedure for the first experiment.

3.2.2 Results and Discussion

We calculated the percentage of "blue" responses over 5 blocks which contained 40 trials each. Blocks were averaged for all forty trials and for all subjects. A two-way mixed ANOVA with the factors of group (between participants, 2 levels) and block (within participants, 5 levels) revealed an insignificant main effect of group, $F(1, 42) = 0{,}779$ $Ms_e = 251.023$ $p > .01$ $\eta^2 = 0.018$, a significant main effect of block, $F(4,168) = 36.987$, $Ms_e = 3775.8$ $p < .01$, $\eta^2 = 0.468$ and an insignificant group/block interaction, $F(4, 168) = 0{,}460$, $Ms_e = 46.974$ $p > .01$. $\eta^2 = 0.011$.

There is a significant increase in the level of blue responses across blocks in the classic probability learning task (CT) which shows that learning occurs gradually by increasing experience. However, we found no significant difference between the monetary and non-monetary feedback conditions (see Figure 12). It is important to note that as seen in Figure 12 participant's exhibit overmatching behavior starting from block 3.

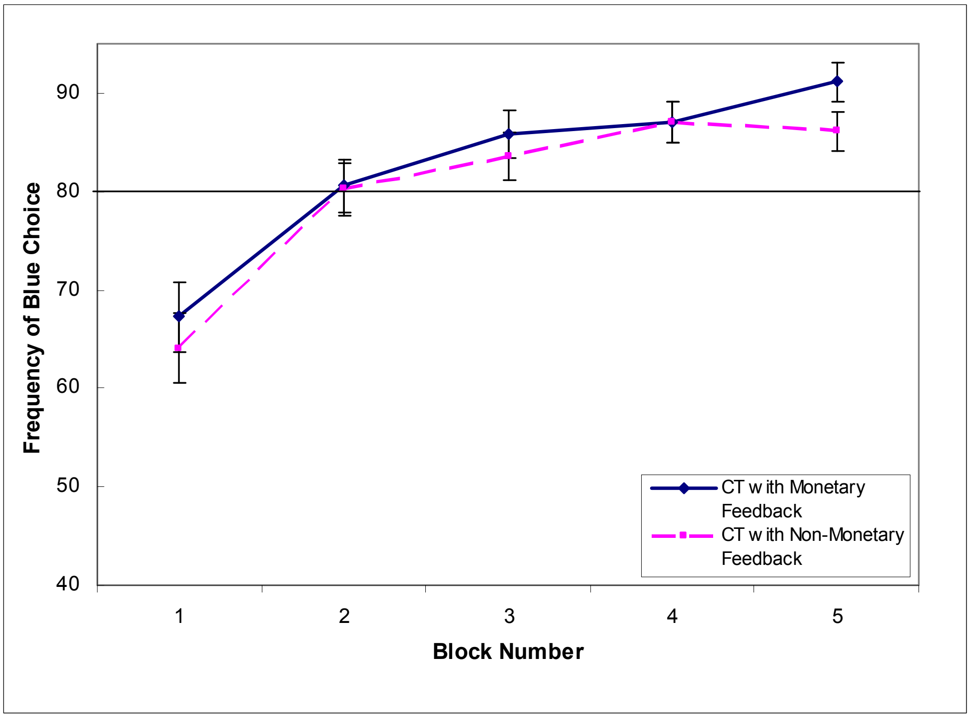

Figure 12 Frequency of blue response graphs for the two experimental conditions in the classic probability learning task (CT).

Figure 12 demonstrates that the average choice of blue box in the first block for both conditions start over 0.6, which is not surprising since subjects start learning even in the first 40 trials. In the first 40 trials, they learn to choose the blue option more. Figure 12 also demonstrates that subjects performance in the first 40 trials is more than chance (probability= 0.5).

In order to examine the average number of correct responses between these two conditions, block scores were calculated due to the value of percent correct responses (Figure 13). A two-way mixed ANOVA with factors of group (between participants, 2 levels) and block (within participants, 5 levels) revealed a insignificant main effect of group, $F(1,42) = 0.079$, $Ms_e =$ $p > .01$ $\eta^2 = 0.018$, a significant main effect of block, $F(4,168) = 13.815$, $Ms_e = 1287.03$ $p < .01$ $\eta^2 = 0.248$, an insignificant group/block interaction, $F(4, 168) = 0.444$, $Ms_e = 37.598$ $p > .01$. $\eta^2 = 0.014$.

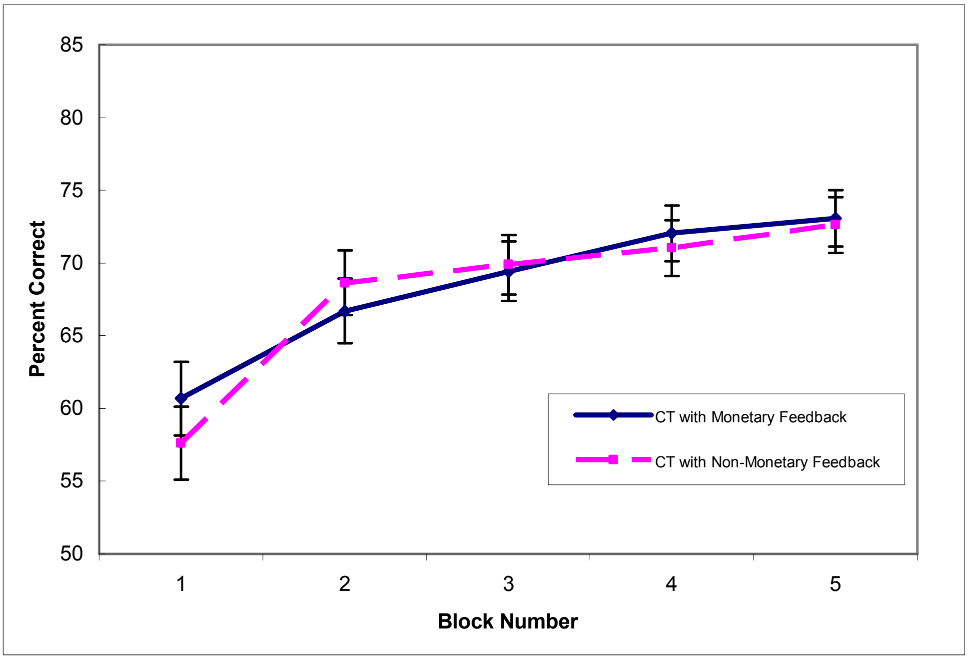

Figure 13 Average correct response graph for the two feedback conditions.

The average reaction time (RT) for all responses was calculated for both feedback conditions (monetary and non-monetary feedback) for each block (Figure 14). A two-way mixed ANOVA with factors of group (between participants, 2 levels) and block (within participants, 5 levels) revealed an insignificant main effect of group, $F(1,42) = 0.277$, $Ms_e = 118161.2$ $p > .01$, $\eta^2 = 0.007$ a significant main effect of block, $F(4,168) = 23.130$, $Ms_e = 1022970.1$ $p < .01$, $\eta^2 = 0.355$ and insignificant group/block interaction, $F(4, 164) = 0.833$, $Ms_e = 36843.6$ $p > .01$ $\eta^2 = 0.019$. As seen from Figure 14 reaction time shows a consistent decrease over blocks, which is an indication of learning in both monetary and non-monetary feedback conditions.

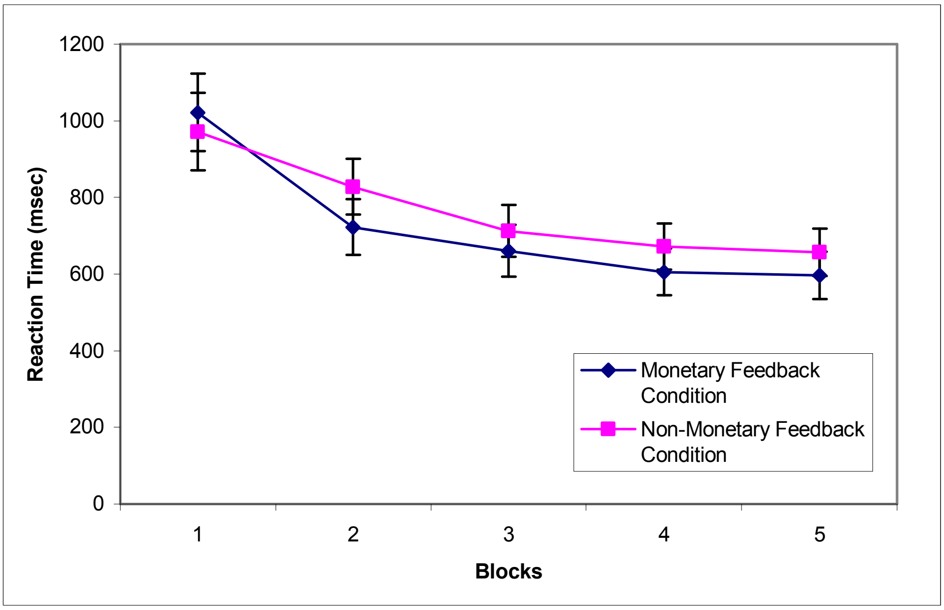

Figure 14 The average reaction times for the classic probability learning task.

The overall results revealed that people in both monetary and non-monetary feedback group started to use overmatching strategy after the second block. Statistical analyses showed that there is no significant difference found for the learning performance between these two groups (both group showed overmatching behavior). This insignificant main effect of group is also same for the average correct responses and average reaction times. This unexpected result may be due to the feedback condition which includes non-losing negative feedback.

3.3 EXPERIMENT 2: Probability Learning with Irrelevant Information Task (IIT)

3.3.1. Method

Participants

Forty-four participants were Middle East Technical University students or employees. They have been collected by assigning posters to a variety of department bulletin boards. All participants were randomly allocated to one of the two experimental conditions: the first, monetary feedback condition (n=22, 12 males, 10 females, 22.6 mean age, SD 1.7) or the second, non-monetary feedback condition (n=22, 10 males, 12 females, 22.8 mean age, SD 2.2). The participants in the monetary feedback condition received money according to their performance in the experiment, and the participants in the non-monetary feedback condition received nothing at the end of the experiment.

Stimuli

There were 10 colored boxes. In each trial, the proportion of these colors changed but in the total 200 trials, their number was balanced. The proportions of colors have no association with the prediction of correct choice (see Figure 15).

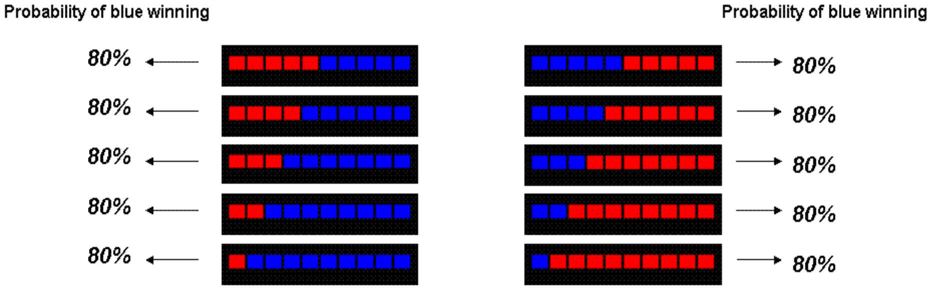

Figure 15 The stimulus uncertainty change but probability of blue winning is fixed.

Procedure

The apparatus and practice session was same as in the previous experiment. In the experimental session of the irrelevant information task (IIT), rewards at the outcome phase of the experiment were not contingent on the proportion of blue and red boxes in a trial and the probability of blue winning was not effected by stimulus uncertainty (see Figure 16). The probability of being correct in each trial was independent from the number of blue and red boxes seen on screen. Therefore, we can say that the information gained from the proportion of blue and red boxes (context) of that trial is irrelevant. To be successful, participants had to focus on the type of color and suppress the irrelevant information. In every trial, there is a 0.8 probability that blue boxes have a winning outcome and 0.2 probability of a red boxes have a winning outcome (see Figure 16). Participants have to learn to choose the blue boxes over time. Each participant is tested with a total of 200 trials. The demand for the second task is then to suppress the irrelevant stimulus dimension (proportion of red and blue boxes) and generate a rule for the color dimension.

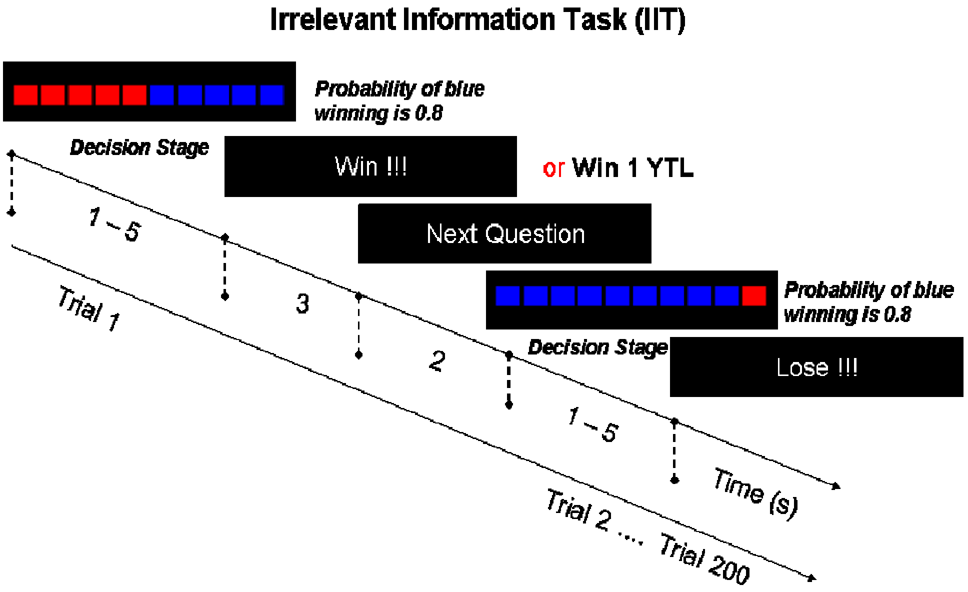

Figure 16 Experimental procedure for the 2nd experiment.

3.3.2. Results and Discussion

In this section, we calculated the percentage of "blue" responses over 5 blocks which contained 40 trials each. Blocks are averaged for all forty trials and for all subjects. A two-way mixed ANOVA with the factors of group (between participants, 2 levels) and block (within participants, 5 levels) revealed a significant main effect of group, $F(1,42) = 3.961$ $Ms_e = 1350.11$ $p < .05$, one-tailed, $\eta^2 = 0.086$, a significant main effect of block, $F(4,168) = 31.912$, $Ms_e = 3164.247$ $p < .05$, $\eta^2 = 0.\ 432$, an insignificant group/block interaction, $F(4,\ 168) = 1.942$, $Ms_e = 192.514$ $p > .01$. $\eta^2 =$

0.044. The results reveal that there is a significant increase in the level of blue responses across block in the IIT which shows that learning occurs (see Figure 17). In addition the effect of monetary feedback is observed in this experiment because the subjects who receive monetary feedback demonstrated significantly different overmatching behavior with in the last two blocks as observed from Figure 17.

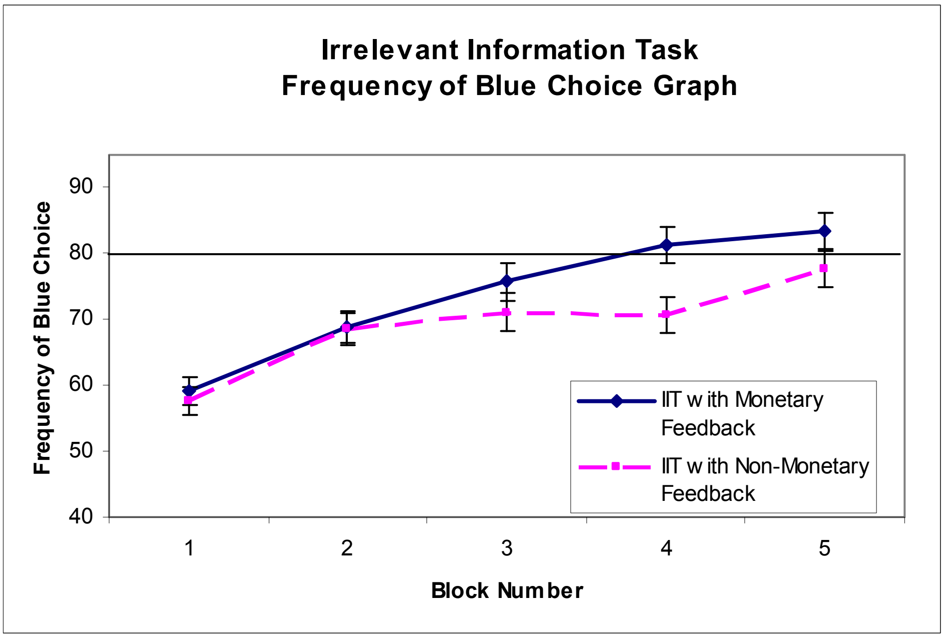

Figure 17 Frequency of blue response graph for the irrelevant information task (IIT).

Next, block scores were calculated due to the value of percent correct responses (Figure 18). A two-way mixed ANOVA with factors of group (between participants, 2 levels) and block (within participants, 5 levels) revealed a significant main effect of group, $F(1,42) = 2.169$, $Ms_e = 248.091$ $p < .05$, one-tailed, $\eta^2 = 0.049$, a significant main effect of block, $F(4,168) = 17.192$, $Ms_e = 1448.6$ $p < .01$, $\eta^2 = 0.290$, an insignificant group/block interaction, $F(4, 168) = 1.375$, $Ms_e = 119.034$ $p > .05$ $\eta^2 = 0.32$. The prediction proportions start at over 0.5 for all two treatments, which is not surprising since subjects were not given the probabilities of the two events;

learning occurs in the first 40 trials. There is a significant increase in the level of correct responses across blocks in IIT due to learning.

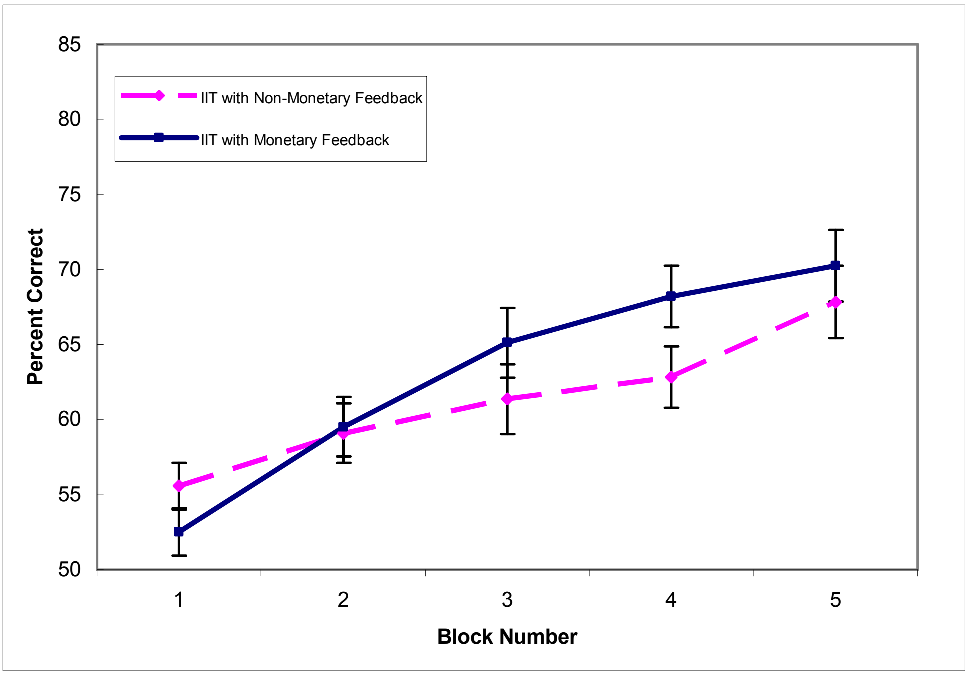

Figure 18 Average correct response graph for the irrelevant information task.

The average reaction time (RT) for all responses was calculated for the two feedback conditions (Figure 19). The average RTs for each block and for each subject were calculated. A two-way mixed ANOVA with factors of group (between participants, 2 levels) and block (within participants, 5 levels) revealed an insignificant main effect of group, $F(1,42) = 0.02$ $Ms_e = 1638.039$ $p > .05$, $\eta^2 = 0$, a significant main effect of block, $F(4,168) = 13.036$, $Ms_e = 1172816.9$ $p < .01$, $\eta^2 = 0.237$, and insignificant group/block interaction, $F(4, 168) = 1.012$, $Ms_e = 91.9013$ $p > .05$. $\eta^2 = 0.24$.

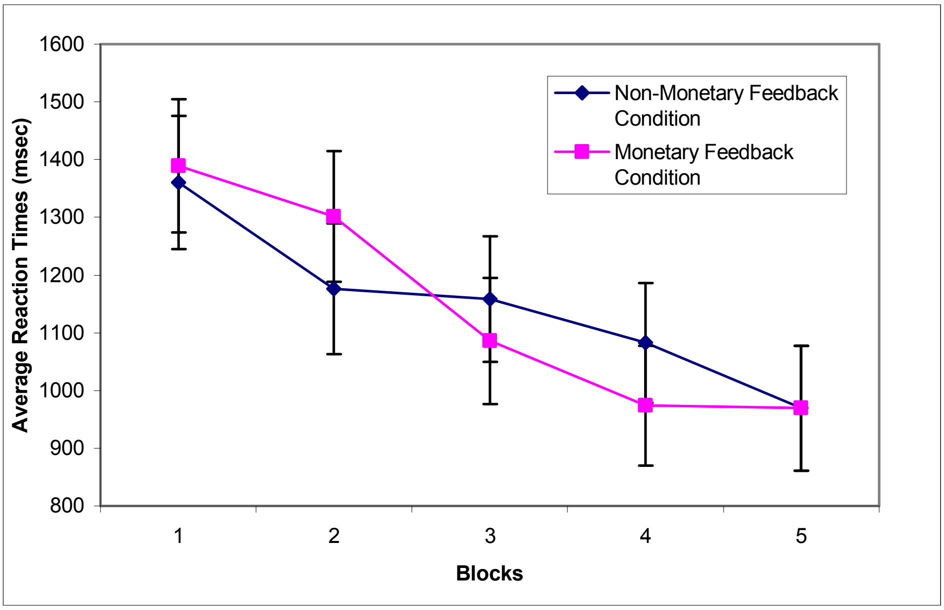

Figure 19 Average reaction time graph for the irrelevant information task (IIT).

The analyses revealed that the monetary feedback group significantly performs better than non-monetary group in learning to choose the blue option. A part of this effect could be explained by the difference in the last three blocks. On the other hand, when the Experiment 2 is compared with Experiment 1, it should be said that monetary feedback is more effective when the task includes irrelevant information which may indicate that task complexity is important.

3.4 EXPERIMENT 3: Probability Learning with Relevant Information Task (RIT)

3.4.1. Method

Participants

Forty-four participants were Middle East Technical University students or employees. They have been collected by assigning posters to a variety of department bulletin boards. All participants were randomly allocated to one of the two experimental conditions: the first condition monetary feedback (n=22, 11 males, 11 females, 22.3 mean age, SD 2.8) or the second condition non-monetary feedback (n=22, 10 males, 12 females, 21.4 mean age, SD 2.1). The participants in the monetary reward condition received money according to their performance in the experiment, and the participants in the positive feedback condition received nothing at the end of the experiment.

Stimuli

In Experiment 3 the stimuli was same with the Experiment 2, however their probability of being correct is different (Figure 20).

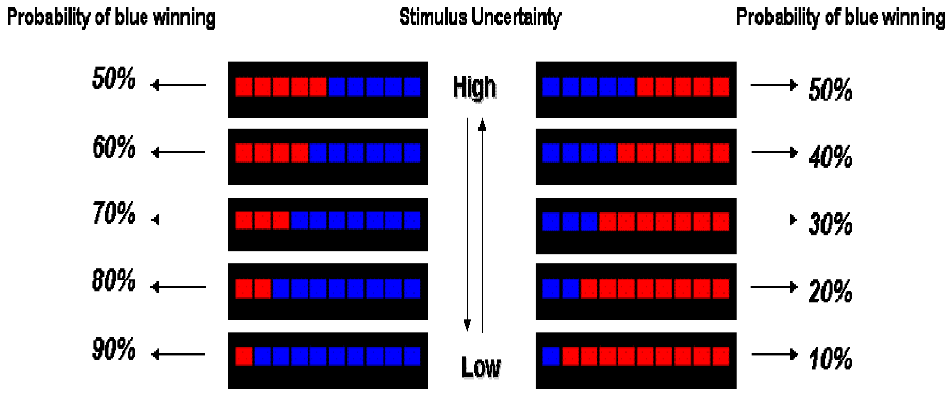

Figure 20 The stimulus uncertainty and probability of winning change according to the distribution of the number of colored boxes.

Procedure

In the RIT, we use the same stimuli as in the IIT, but the proportions of red and blue boxes change in each trial. The proportion of blue/red boxes is directly related to the probability of winning a positive feedback. Therefore, the variability of uncertainty of each choice changes according to the proportion of blue and red boxes (Figure 21). For example, one red box and nine blue boxes in a total of ten boxes means that there is one over ten chances that the red box has a rewarding outcome and nine over ten chances for the blue box to have a rewarding outcome. The level of uncertainty is maximized when there are equal numbers of red and blue boxes corresponding to 0.5 probability.

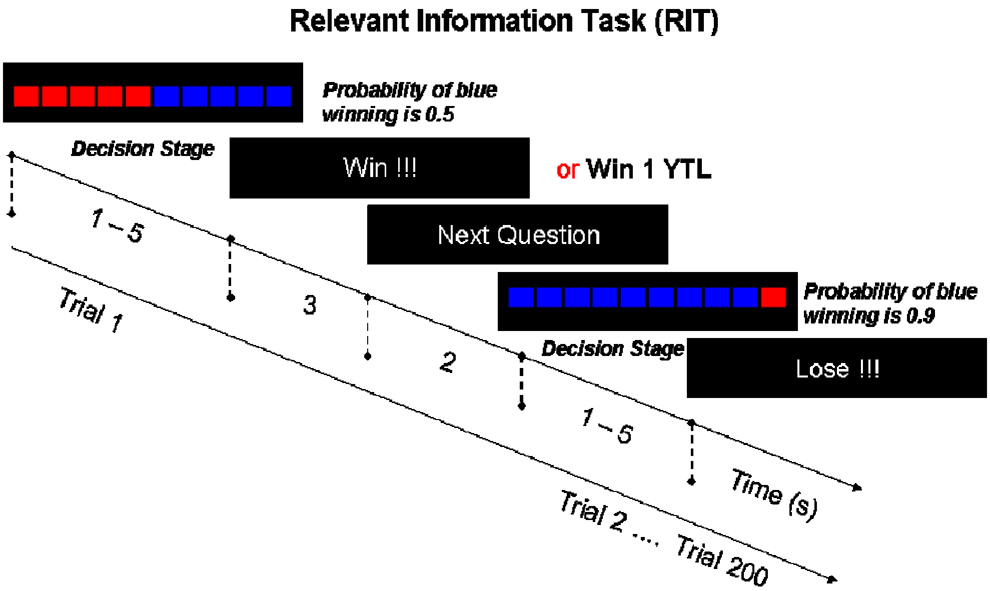

Figure 21 Proportions of colors have direct effect on winning in that trial.

3.4.2 Results and Discussion

Percentages of "blue" responses were calculated over 5 blocks which contained 40 trials each. Blocks are averaged for all forty trials and for all subjects. A two-way mixed ANOVA with factors of group (between participants, 2 levels) and block (within participants, 5 levels) revealed an insignificant main effect of group, $F(1,42) = 0.462$ $Ms_e = 47.756$ $p > .05$, $\eta^2 = 0.11$, an insignificant main effect of block, $F(4,168) = 1.016$, $Ms_e = 58.532$ $p > .05$, $\eta^2 = 0.024$, an insignificant group/block interaction, $F(4, 168) = 0.124$, $Ms_e = 7.131$ $p > .05$. $\eta^2 = 0.003$ In the RIT, the frequency of winning of blues versus reds across the experiment is the same, so the subject presses blue half of the time (see, Figure 22). It is worth noting the contrast between Figure 22 and Figure 17. The effect of learning in Figure 17 has disappeared in Figure 22. This is probably because in RIT, the initial assumption of the participants regarding the frequencies involved in each trial is valid trough out the experiment.

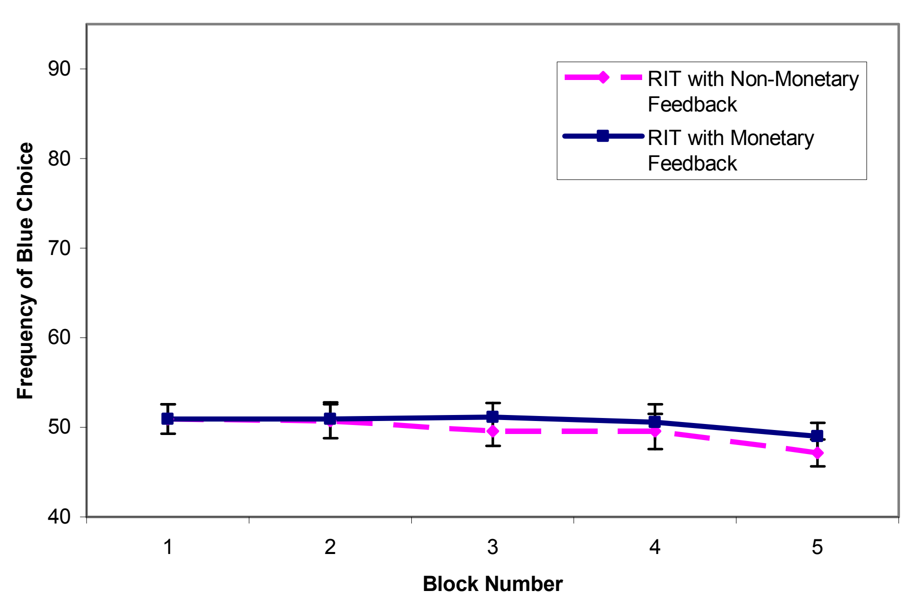

Figure 22 Frequency of blue choice in the relevant information task (RIT).

Next, block scores were calculated due to the value of percent correct responses (Figure 23). A two-way mixed ANOVA with factors of group (between participants, 2 levels) and block (within participants, 5 levels) revealed a insignificant main effect of group, *F*(1,42) = 0.558, $Ms_e = 89.091$ $p > .05$, $\eta^2 = 0.013$, a significant main effect of block, *F*(4,164) = 3.778, $Ms_e = 239.389$ $p < .01$, $\eta^2 = 0.83$ and insignificant group/block interaction, *F*(4, 164) = 0.488, $Ms_e = 30.923$ $p > .05$. $\eta^2 = 0.11$.

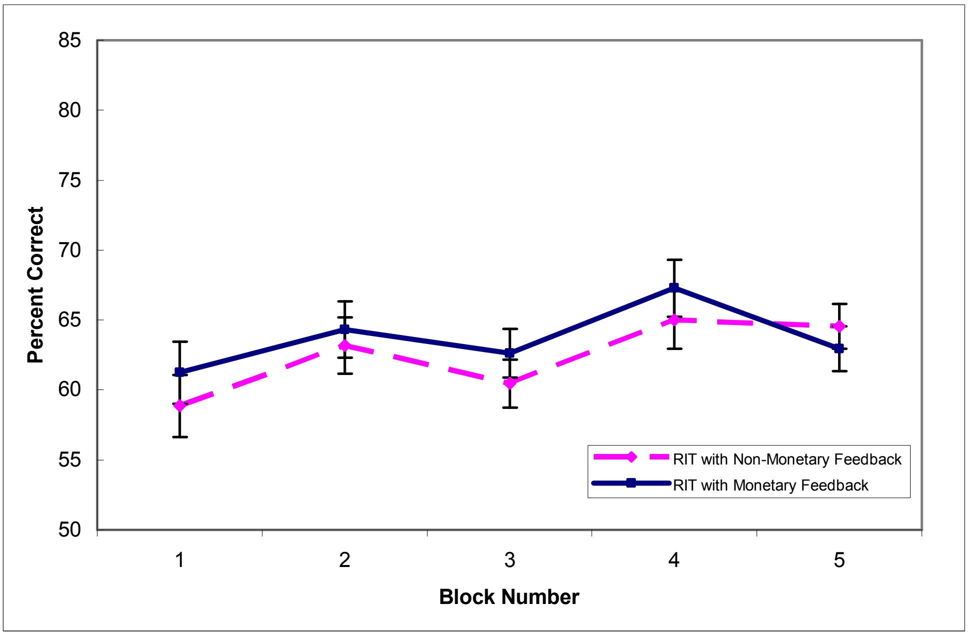

Figure 23 Average correct responses for the relevant information task.

Jittering occurs in the RIT which means no learning occurs (see, Figure 23). This is expected because even at the beginning the subjects have an intuition to choose the color associated with the larger number of boxes. The probability of winning is 0.9, 0.8, 0.7, 0.6, 0.5, which is evenly distributed across both colors and blocks. The mean of these probabilities is 0.7; therefore the subjects' performance of about 70 % correct in the RIT is expected.

The average RTs for each block and for each subject were calculated (Figure 24). A two-way mixed ANOVA with factors of group (between participants, 2 levels) and block (within participants, 5 levels) revealed an insignificant main effect of group, $F(1,42) = 0.02$ $Ms_e = 73.545$ $p > .05$, $\eta^2 = 0.013$ a significant main effect of block, $F(4,168) = 13.036$, $Ms_e = 144621.3$ $p < .01$, $\eta^2 = 0.079$ and an insignificant group/block interaction, $F(4, 168) =$ 1,012, $Ms_e = 8818.5$ $p > .05$. $\eta^2 = 0.005$

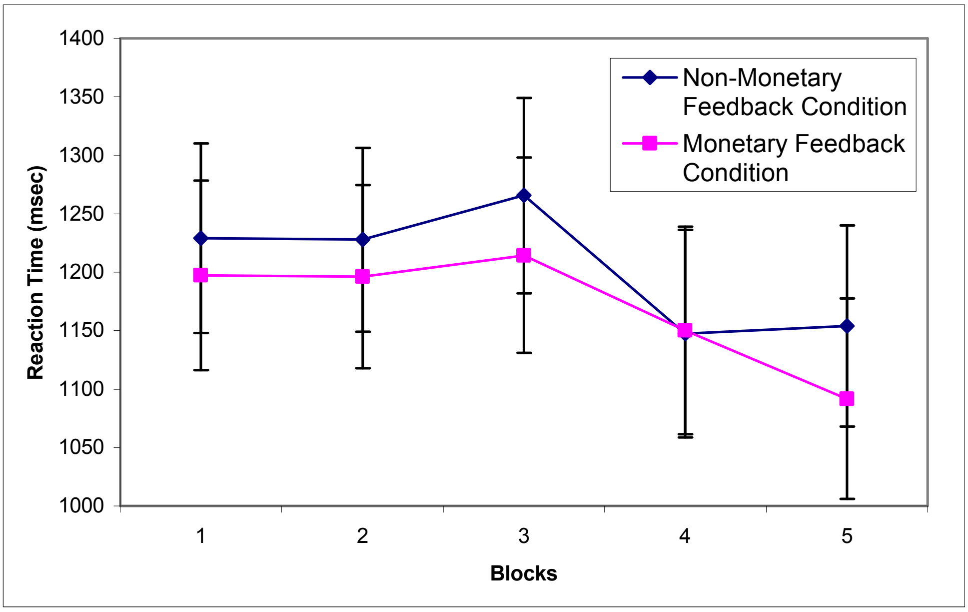

Figure 24 Average reaction time graph for the relevant information task.

Experiment 3 showed that when the stimulus includes relevant information for making a choice people automatically start to use this information. Due to effective usage of this relevant information, no learning effect was seen from the Figure 22 as expected by the hypothesis. In addition the reduction in reaction times towards the end of the experiment might be interpreted as a side effect of practice instead of learning.

3.5 EXPERIMENT 4: Probability Learning with Irrelevant Information Task in a Group of Parkinson Patients

Experiment 4 is a replication of Experiment 2 except we did not examine the monetary feedback condition because of time constraints on the present study. In the Experiment 4, probability learning behavior of Parkinson patients (PD) and normal controls was compared with the Irrelevant Information Task.

3.5.1. Method

Participants

Twenty seven participants were recruited in this experiment. Thirteen of them were idiopathic Parkinson patients whose mean age was 57 (range, 45-65 years). One PD patient was excluded due to his low score on a mini mental status examination (Güngen, et al., 2002), and another PD patient was excluded due to his treatment condition (deep brain stimulation). Participants were recruited in this experiment in the İstanbul Faculty of Medicine, Neurology department under professional supervision[11]. Patients were non-demented (average Mini Mental Status Examination-MMSE under 26), and non-depressed (Geriatric Depression Scale, under 13, Ertan et al., 2005). The mean motor stage (Hoehn and Yahr, 1965) was 1.5. All patients were L-dopa and DA agonist treated. The average symptomatic disease duration of PD was 4 years (range, 1-7 years). Fourteen healthy controls were age matched between 45 and 65 years and recruited to the same experiment.

Stimuli

Stimuli were the same as in Experiment 2.

Procedure

The irrelevant information task (see Experiment 2 for details) with non–monetary feedback condition was administered to both groups. All experimental procedures were the same as in Experiment 2, except we additionally collected verbal reports of each subject at the end of each session. The analyses of verbal reports were planned as part of a future research project on the medicated PD patients.

---

[11] Participant had been screened for their regular visiting by Neurologist Assoc. Prof. Dr. Hakan Gürvit, Assist. Prof.Dr.Haşmet Hanağası before the experimental sessions started.

3.5.2. Results and Discussion

Percentages of "blue" responses were calculated over 5 blocks which contain 40 trials each (Figure 25). Blocks were averaged for all forty trials and for all subjects. A two-way mixed ANOVA with factors of group (between participants, 2 levels) and block (within participants, 5 levels) revealed an insignificant main effect of group, *F*(1,23) = 1.728 Ms_e =696.826 $p > .05$, $\eta^2 = 0.070$, an significant main effect of block, *F*(4,92) = 19.022, Ms_e = 1914.42 $p < .01$, $\eta^2 = 0.453$, and an insignificant group/block interaction, *F*(4, 92) = 0.111, Ms_e = 11.210 $p > .05$. $\eta^2 = 0.005$.

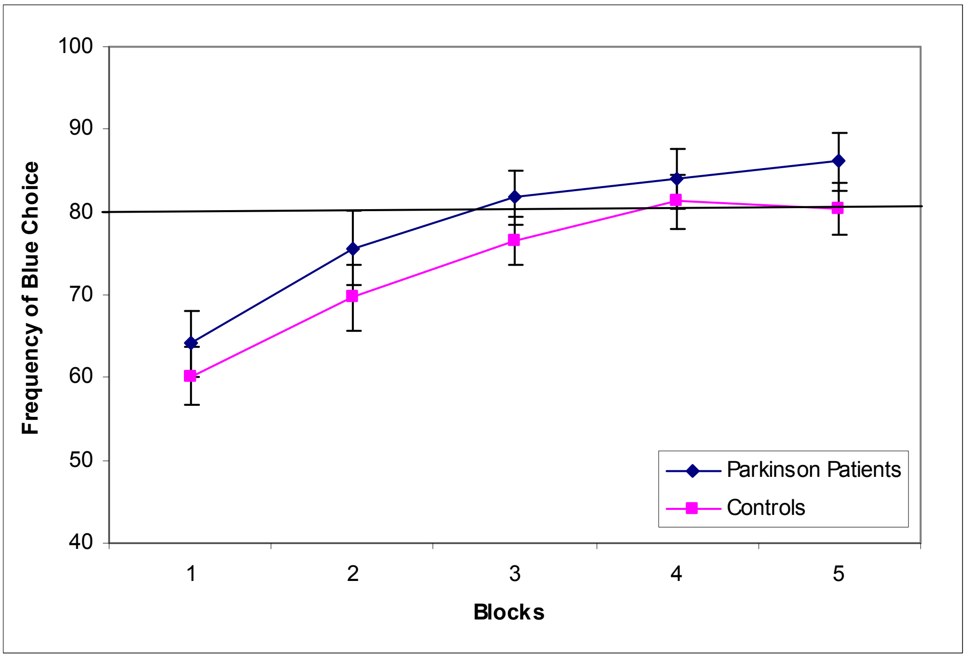

Figure 25 Frequency of blue choice graph between Parkinson patients and healthy controls.

Percentages of correct responses were calculated over 5 blocks. Blocks were averaged for all forty trials and for all subjects (Figure 26). A two-way mixed ANOVA with factors of group (between participants, 2 levels) and block (within participants, 5 levels) revealed an insignificant main effect of group, *F*(1,23) = 0,417 Ms_e =10.755 $p > .05$, $\eta^2 = 0.018$, a significant main effect of block, *F*(4,92) = 7.813, Ms_e = 127.782 $p < .01$, $\eta^2 = 0.254$, an insignificant group/block interaction, *F*(4, 92) = 0,319, $Ms_e = 5.222$ $p > .05$. $\eta^2 = 0.014$.

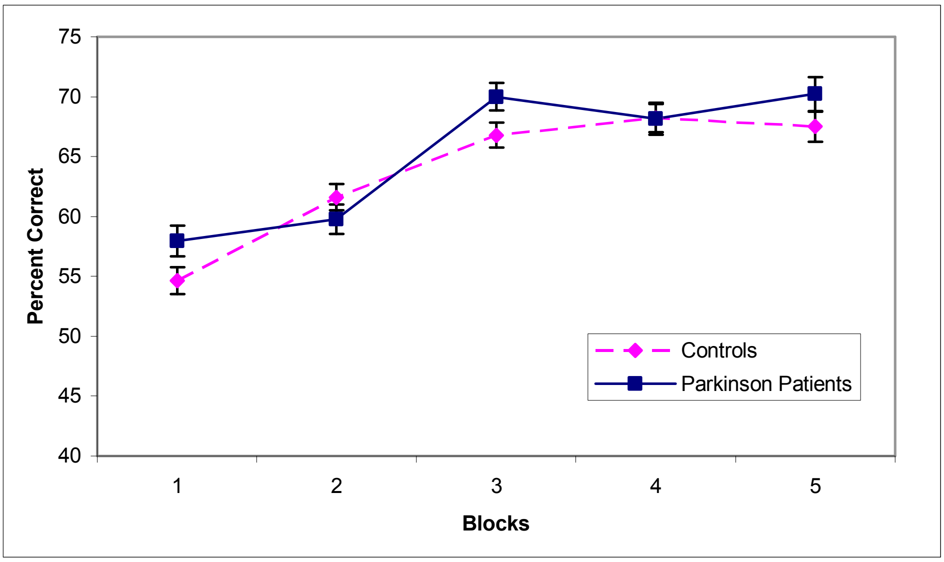

Figure 26 Average percent correct responses for Parkinson patients and healthy controls.

The average reaction time (RT) for overall responses was calculated for the two groups. The average RTs for each block and for each subject were calculated. A two-way mixed ANOVA with factors of group (between participants, 2 levels) and block (within participants, 5 levels) revealed an insignificant main effect of group, *F*(1,23) = 1,621 $Ms_e = 887149.3$ $p > .05$, $\eta^2 = 0.066$ a significant main effect of block, *F*(4,92) =6.934 , $Ms_e = 372285.2$ $p < .01$, $\eta^2 = 0.232$ and a significant group/block interaction, *F*(4, 92) = 2.908, $Ms_e = 156127.5$ $p > .05$, $\eta^2 = 0.112$.

As seen from the percent correct graph in Figure 26, both Parkinson patients and normal controls exhibit learning behavior. It is worth noting that Parkinson patients showed overmatching behavior earlier than control group (Figure 25).

Although RT did not differ significantly between PD and Controls, the reaction time in the PD group considerably less during the beginning and middle phases of the task (Figure 27). This may be due to the impulsiveness of the patient population. Similar impulsive behavior is also observed in pathological gambling who are known to have dopamine dysfunction.

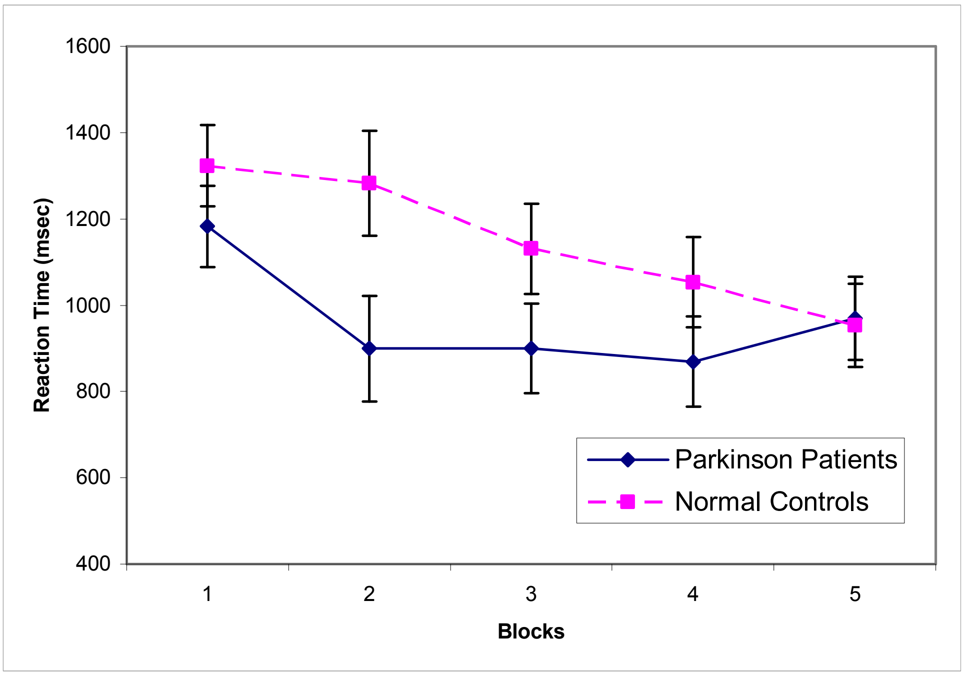

Figure 27 Average reaction time graph between Parkinson patients and healthy controls.

3.6 Further Analysis

In the standard analyses performed above, effects of monetary feedback for the classic probability learning task, irrelevant information task, and relevant information task is investigated. In addition, further analyses were carried in order to show the effect of uncertainty and context in detail as described below.

3.6.1. Effect of Uncertainty

Experiment 3

In the literature review section it has been told that cognitive psychologists such as Laming (1968) apply information theoretic entropy to choice reaction times to find a correlation between reaction times and uncertainty. The correlation between the information theoretic entropy[12] (Shannon, 1948) and the average choice reaction times were calculated for the number of blue boxes for each stimulus in Experiment 3 (RIT). The statistical analysis is carried out by using the average reaction time and information entropy. The results showed that the average reaction time and information theoretic uncertainty are positively correlated $P < .05$, $r = 0.745$ (see Figure 29)

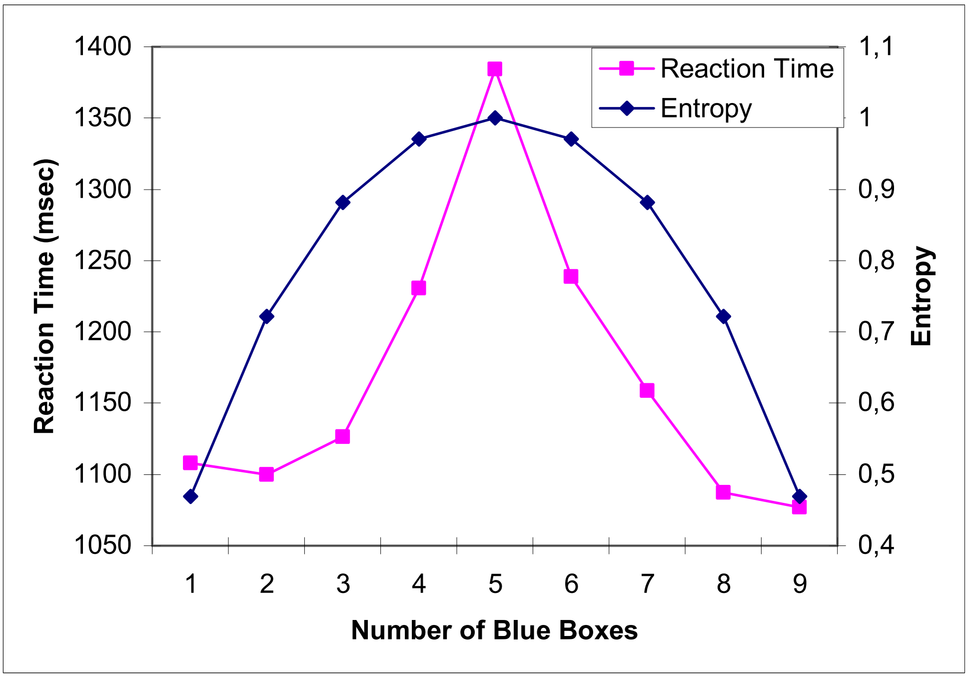

Figure 28 Relation between information entropy and choice reaction times.

Figure 28 demonstrates that the average reaction times take longer when the contextual information represents 0.5 probabilities (where the stimulus represents 5 blue, 5 red boxes). Although entropy and reaction time were correlated the shapes of these graphs were different. Our attempt was just to establish a relationship between entropy and reaction times but not develop a model.

---

[12] Information theoretic entropy = -plog(p)-(1-p)log(1-p). P stands for probability of winning.

3.6.2. Effect of Context

Experiment 2

In order to show that there was an interaction between the proportions of blue responses and the learning experience, we took the first (first 50 trials) and last block (last 50 trials) of Experiment 2 (IIT) to calculate the average blue responses with respect to proportions of blue boxes (see Figure 29). The first 50 and last 50 trials were used because we believe that after a period of learning, people learn to suppress the belief about the proportion of the boxes.

For the non-monetary feedback condition, a two-way repeated ANOVA with factors of block (within participants, 2 levels) and proportions of blue boxes (within participants, 9 levels) revealed a significant main effect of block, $F(1,21) = 37.693$, $Ms_e = 3.785$ $p > .05$, $\eta^2 = 0.642$, a significant main effect of number of blue boxes, $F(8,168) = 8.081$, $Ms_e = 0.610$ $p < .01$, $\eta^2 = 0.278$ and a significant effect of proportion of blue boxes/block interaction, $F(8,168) = 4.096$, $Ms_e = 0.193$ $p > .05$ $\eta^2 = 0.163$.

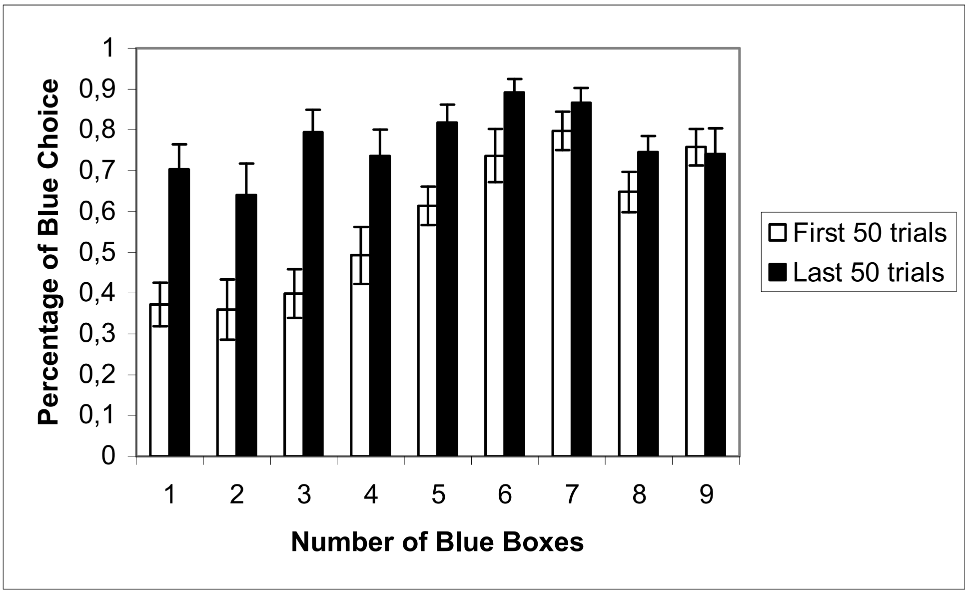

Figure 29 Changes in blue choice with in the first and last fifty trials for the non-monetary feedback group.

Figure 29 demonstrates that in the initial stages (first 50 trials) of the experiment, participants use the contextual cue information but after learning occurs (last 50 trials) they suppress this information and learn by feedback.

Comparison of Reaction Times between Experiment 1 and Experiment 2

It was previously mentioned that irrelevant information affects probability learning negatively, which is observed through an increase in reaction times or decrease in performance. In order to understand the effect of irrelevant information effect on the choice reaction times (RT), we compared the reaction times between Experiment 1 (CT) and Experiment 2 (IIT) by merging overall responses for both monetary and non-monetary groups. For each block and for each subject average reaction times were included in data analysis.

A two-way mixed ANOVA with factors of group (between participants, 2 levels) and block (within participants, 5 levels) revealed a significant main effect of group, $F(1,86) = 25.988$ $Ms_e = 393554792{,}1$ $p < .01$, $\eta^2 = 0.870$ a significant main effect of block, $F(4, 344) =32.010$, $Ms_e = 2145762.6$ $p < .01$, $\eta^2 = 0.271$ and a insignificant group/block interaction, $F(4, 344) = 0.746$, $Ms_e =50024.24$ $p > .05$ $\eta^2 = 0.009$.

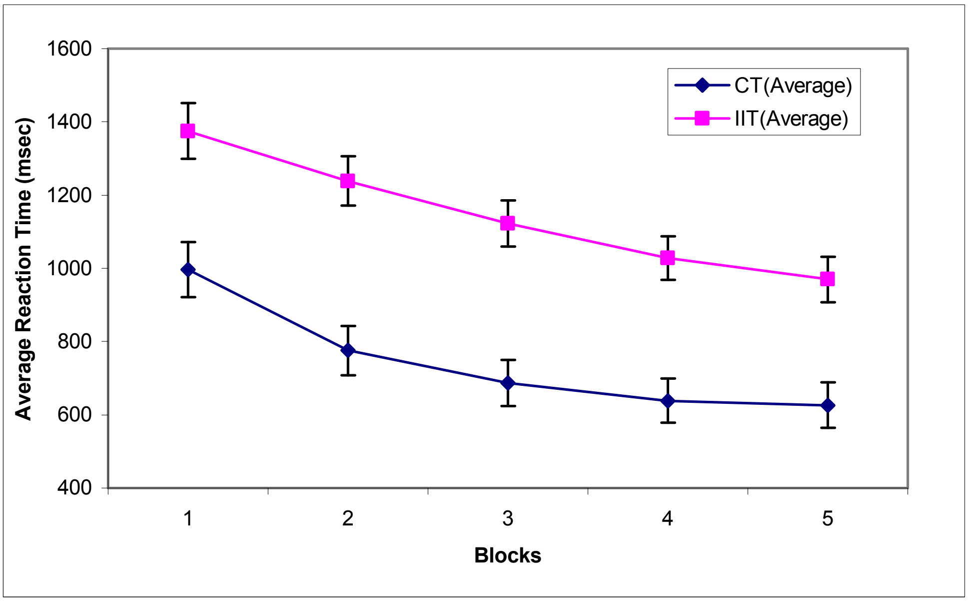

Figure 30 Average choice reaction times between Experiment 1 and Experiment 2.

As seen from Figure 30, there is a prominent effect of the context in RT. Although the probability of blue and red choices are the same between CT and IIT tasks, the information presented on screen is more complex and irrelevant in IIT. Hence, this complexity is directly reflected as significantly higher reaction times.

3.6.3. Probability Learning Strategy

Frequency Matching Strategy in CT

As mentioned in chapter 2, one possible way for the subjects to learn frequencies is by a pattern search. In order to evaluate frequency matching versus maximization behavior, the response sequences of subjects have to be analyzed individually. In this exploratory analyses, we used data from the first (CT) and fourth experiment (PD-IIT). An ad-hoc method is devised to investigate how participants estimate the current winning option based on the previous options.

Two types of patterns are identified: alternating patterns and fixed patterns. Alternating patterns correspond to frequency matching behavior whereas fixed patterns correspond to maximization behavior. Alternating patterns are defined as sequences of blue (B) or red (R) choices such that B and R are distributed evenly in the sequence. The reason why choices are evenly distributed in alternating patterns is because according to gambler's fallacy, alternating patterns seem more random, hence preferable than fixed sequences. The minimum length sequence which can hold a meaningful alternating pattern should be 4. A shorter sequence with 0.5 probability for 2 options would have been of length 2, but then, within such a short sequence, only the random patterns BR, RB, BB, RR could have been produced, which we found to be inadequate to represent a consistent pattern search. With 4 digits, patterns with 0.5 probability of winning for blue and red could have been represented by: BBRR, RRBB, BRBR, RBRB, BRRB, and RBBR. On the other hand, fixed patterns of length 4, BBBB, and RRRR would represent a consistent behavior of maximization. Within the 200 trials, the occurrence of these patterns is reported[13] separately for the first and last blocks (40 trials). While searching for these two groups of patterns, each pattern is reported by the index of the trial which starts with the pattern, and these indices are bundled into block one if the value is between 1-40, and block five, if the value is between 160-200. While reporting the indices, a moving window is used to search for pattern occurrences. Once a pattern is found, the window is advanced to the next non-overlapping index for further search.

For the Experiment 1, a two-way mixed ANOVA with factors of pattern (within participants, 2 levels alternating-fixed) and block (within participants, 2 levels first 50- last 50 trials) revealed a significant main effect of block, $F(1,43) = 50,031$, $Ms_e = 162,27$ $p < .01$, $\eta^2 = 0,537$ and a

---

[13] The author would like to express his gratitude to Özkan Bayraktar for helping to write a program for the current analyses.

significant pattern/block interaction, $F(1, 43) = 42{,}049$, $Ms_e = 684{,}142$ $p < .01$. $\eta^2 = 0{,}494$.

As seen in Figure 31, in the first fifty trials participants choose alternating patterns more than fixed patterns. This indicates the usage of frequency matching strategy in the beginning of the experiment. On the contrary, in the last fifty trials higher usage of fixed patterns by most of the participants leads to maximization strategy.

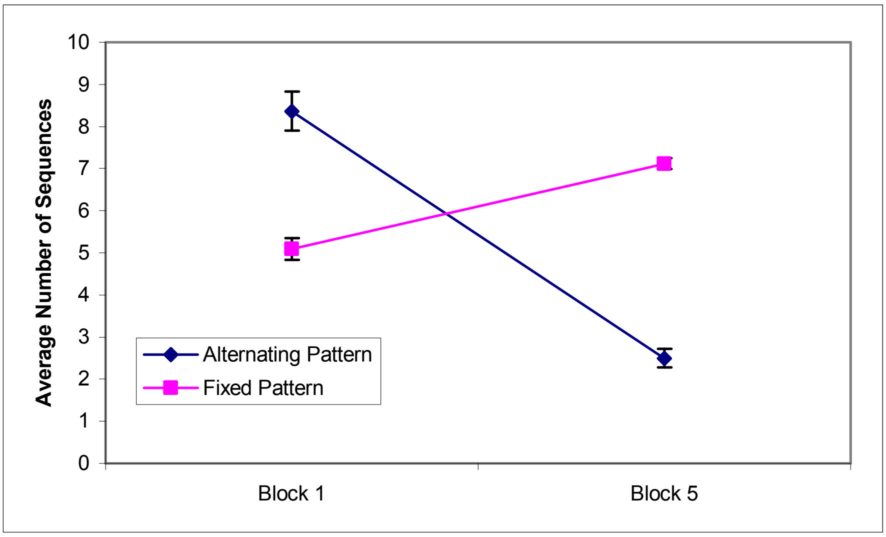

Figure 31 Average numbers of patterns between blocks in the Experiment1.

Parkinson Patients and Maximization Strategy

For the Experiment 4, a two-way mixed ANOVA with factors of group (between participants, 2 levels PD vs Control) x block (within participants, 2 levels first 50- last 50 trials) x pattern (within participants, 2 random-nonrandom) levels revealed a significant main effect of block, $F(1,23) = 20.621$, $Ms_e = 91.646$ $p < .01$, $\eta^2 = 0.473$, a significant pattern /block interaction, $F(1,23) = 40{,}10$ $Ms_e = 488.93$ $p < .01$. $\eta^2 = 0{,}636$ and an insignificant effect of group $F(1,23) = 0.755$ $Ms_e = 3.617$ $p > .01$. $\eta^2 = 0{,}032$.

As seen from Figure 32 in general PD patients show less frequency matching but more maximization behavior in comparison to controls.

However, this difference was not significant; possibly due to the number of participants.

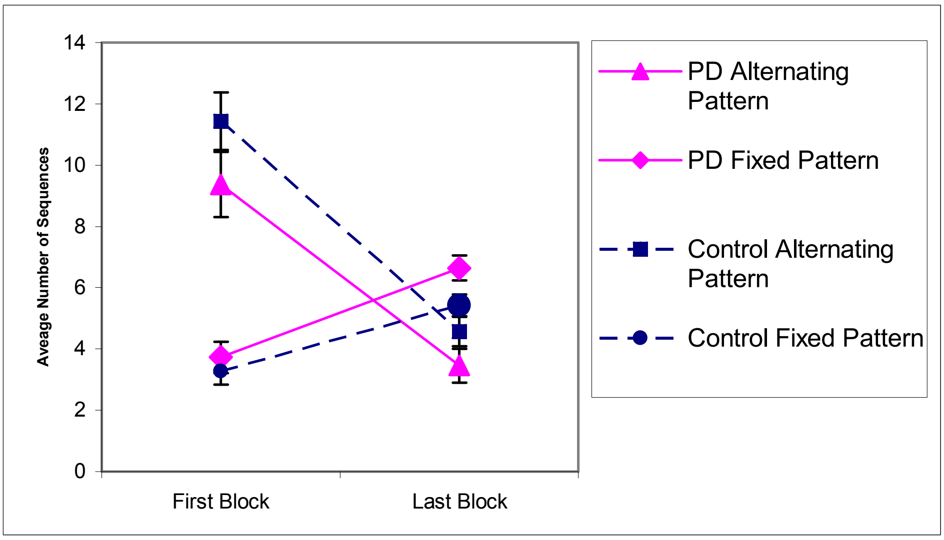

Figure 32 Average numbers of patterns between blocks in the Experiment4.

In Figure 33, the behavioral response of a PD patient is shown over 200 trials. Until now, the behavior of the participants has been illustrated through graphs that average data from all participants. As seen clearly from Figure 33, there is the high contrast between frequency matching behavior at the beginning of the experiment and maximizing behavior at the end.

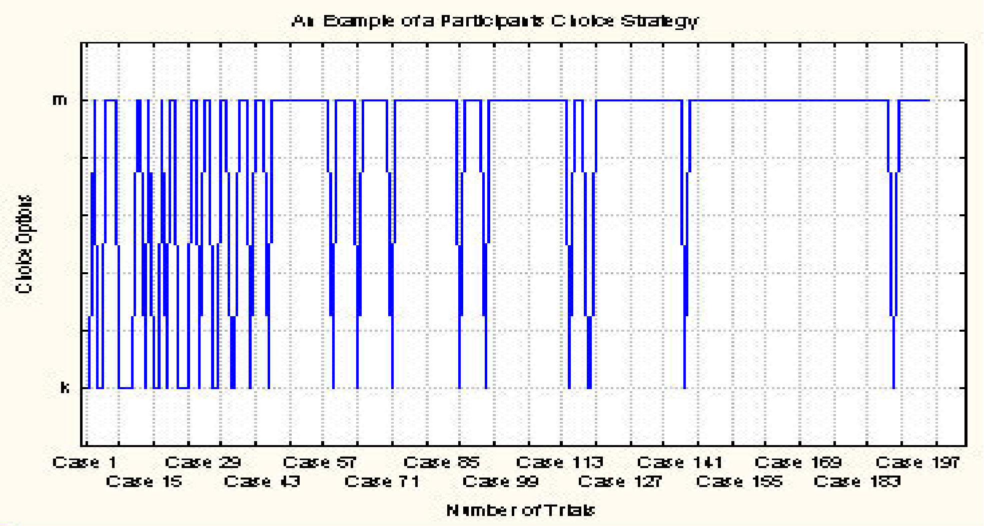

Figure 33 One participant’s blue choice graph.

CHAPTER 4

GENERAL DISCUSSION

The difference between monetary feedback and non-monetary positive feedback was examined in the first three experiments. A barely significant effect of monetary feedback was found in the second experiment (IIT). Additionally, in the second and third experiment the effect of context (irrelevant & relevant) was tested to find how people utilize the cue information. It was found that initially people automatically use the cue information to guide their choices, but after a period of trial and error, people start using the feedback information. In the fourth experiment, medicated Parkinson patients performed the Irrelevant Information Task. Their performance was compared with normal controls. No significant difference was found between the control group and Parkinson patients. On the other hand, both in Experiment 2 and Experiment 4, it was found that in the initial stages of the experimental session participants search for patterns of random sequences but later on, they give up searching and exhibit maximization behavior.

4.1. Effect of Monetary Feedback on Probability Learning

In Experiment 2 (IIT), a significant effect of monetary reward was found. However, in Experiment 1 (CT) and Experiment 3 (RIT) no significant effect of monetary reward was found. We had hypothesized that monetary reward would have a significant effect, because in previous studies it was found that monetary feedback increases subjects performance positively (Edward, 1956; Myers et al., 1963, Shanks et al, 2002)

There are several possible explanations for these controversial results. One explanation is that the monetary feedback that we used was not enough for the participants to be effective. Participants can only eat one lunch in the University campus with that money if they had showed average performance. However, studies (Edward, 1956; Myers et al., 1963, Shanks et al, 2002) showed a clear difference between monetary feedback and non-monetary feedback with similar or lower amount of money than we used. Therefore, the conclusion that there is no effect of monetary feedback due to the inadequacy of reward value is not convincing.

An alternative explanation may be that monetary feedback is more effective if there was a loss condition in the experimental procedure. It is possible that the effect of monetary feedback is more prominent in the presence of a negative feedback such as loss (Siegel & Goldstein, 1959). In the previous studies, both reward and punishment were used in the same experimental procedure where the positive feedback condition was winning and the negative feedback condition was losing (Edward, 1956; Myers et al., 1963; Shanks et al., 2002).

On the other hand, we can probably say that the monetary feedback became effective when the task complexity increased. Experiment 2 (IIT) demonstrates that subjects in the monetary feedback condition showed overmatching behavior but non-monetary feedback group showed under matching behavior in the last 80 trials. Interestingly, learning performance was not affected by the monetary payoff in Experiment 1 (CT), because both groups do show overmatching behavior. We can suggest that when the task complexity is high (such as in Experiment 2 (IIT)), subjects need monetary feedback to be motivated and perform better. When the task complexity is low (such as in Experiment 1 (CT)), subjects perform well regardless of monetary feedback.

4.2. Effect of Uncertainty in Probability Learning

We studied the contribution of uncertainty in the third experiment (RIT). Previous studies on the psychology of information processing showed that when the number of possible outcomes are equally probable (P=0.5, maximum entropy) choice reaction time is at its maximum level (Wollen, 1963). When we analyzed the choice reaction time data of Experiment 3 (RIT) it has been clearly seen that, the longest reaction time is for the 0.5 probability condition. The reaction times for other probability values decrease due to a decrease in probabilistic uncertainty (Wollen, 1963). We investigated the correlation between the information theoretic entropy and choice reaction times. Thereby, we replicated the earlier findings regarding the effect of uncertainty in probabilistic learning.

4.3. Effect of Context in Probability Learning

By comparing the reaction time between Experiment 2 (IIT) and Experiment 1 (CT), we investigated the effect of irrelevant probabilistic context on learning performance in a probability learning task. We hypothesized that subjects in the second experiment (IIT) will respond slower than the subjects in the first experiment (CT) due to the complexity introduced by irrelevant context. Experiments 1 and 2 are identical, except that in Experiment 1 (CT) subjects are presented with simpler stimulus screen consisting of one blue and one red box, whereas in Experiment 2 (IIT), the subjects are presented with a set of red and blue boxes. The multiplicity of the boxes in Experiment 2 has no relationship with the winning probabilities, and in both experiments, the winning probability of colors is the same (blue winning 80 percent of the time). We hypothesized that the subjects need to resolve the effects of the irrelevant information in Experiment 2, so this would increase the time for the response to occur. Our findings on reaction time replicated the previous findings that irrelevant information increases response latencies (Edgell & Castellan, 1996).

In Experiment 3 (RIT), subjects start the experiment using the probabilistic cue information and continue to use this information till the very end. Because of this, learning effects are not seen in all 5 blocks in the frequency of blue choices and percent correct answers. It seems that people automatically make probabilistic predictions from the cue information (proportions of blue boxes). This raises the question "Do people have probabilistic templates in their minds for some perceptual cues?".

The presence of a plausible mental model which draws probabilistic inference from the ratio of blue and red boxes presented in the stimulus is a reasonable opinion. For example, our findings indicate that participants think a priori that the proportion of blue boxes to red boxes represents the probability of choosing a specific color. This argument is reasonable, only if participants have a mental representation for every box (blue/red) in which the probability scale is divided evenly between the boxes. For example, because there are 10 boxes presented in Experiment 3 (RIT), subjects seem to automatically believe that each box represents a probability value of 0.1. Therefore when they see 9 blue boxes and 1 red box, they infer that blue is the rewarding category, and interestingly, this inference is present even at the beginning of the experiment. Such mental representations are common in Gestalt psychology (Kahneman, 2002). For example, Gestalt similarity effect will lead people to automatically group physically similar items. Kahneman (2002) also suggested that some percepts are more easily accessible that others and people might use these easily accessible information to guide their choices. Combining our results with these ideas, we can state that the proportion of blue and red boxes (cue information) in the second and third experiment may be perceived as an easily accessible information for the probability areas of the brain. Thereby, in experiment 2, it takes time to suppress this automatic inference to the mental representation and learn via the feedback signal.

4.4. Mechanisms of Probability Learning

We observed learning effects in Experiment 1, 2, and 4, as revealed by the significant difference across blocks for both percent of correct and frequency of blue choice values. The reaction time data also indicates significant differences across blocks. On the other hand, in Experiment 3, only the reaction time graph revealed a significant effect of block in choice reaction times. On average, in Experiment 3, subjects start with 1220 msec of reaction time and decrease to 1100 msec at the end of the experiment. This effect may be due to the practice (Speelman and Maybery, 1998), a point which needs further investigation. The absence of learning effect in Experiment 3 (RIT) is not surprising. In this experiment subjects used the relevant cue information from the very beginning because the probability of winning is directly related to the context, there is no further information to learn.

4.4.1. Implicit (Procedural) versus Explicit Learning

Implicit learning has been categorized as an unconscious learning process. On the contrary, in explicit learning, there is conscious monitoring of learning strategies (e.g., finding a rule). It has been shown that implicit learning is unlimited and stores all contingencies between stimulus variables (O'Brien-Malone and Maybery, 1998). By contrast, explicit learning has been showed to have limited capacity which is maintained by working memory (O'Brien-Malone and Maybery, 1998). According to Reber (1992), implicit learning is more robust and show less population variance compared to explicit learning.

In the debriefing period at the end of each experimental session, we let participants ask questions. Many of the participants in Experiments 1, 2, 4 said that they had searched for a pattern, but they had failed to find it.

Afterwards, they realized that blue wins more than the red color and they decided to choose blue more frequently. Additionally, they reported that they did not have a specific rule-based strategy. In Experiment 4, we also collected verbal reports at the end of the experiment, similar to those of Maia and McClelland (2004). Our brief judgments of these reports indicate that at the end of the experiments, most subjects guessed the proportion of blue winning correctly. On average, the participants in both PD group and control group reported a winning percent close to 0.8. At this point, because we do not have any statistics, we can only speculate that probability learning is an implicit thought process. At the end of the experiments, subjects mostly guess the true proportion of blue winning, but do not have a specific rule-based strategy. In order to be able to state this objectively, we need to analyze the verbal reports as a part of our future studies.

4.4.2. Frequency Matching versus Maximization

In the literature chapter, we discussed that there are at least two mechanisms responsible for probability learning: a pattern searching mechanism and a frequency learning mechanism. We were able to show that in the initial stages of Experiments 2 and 4, people searched for alternating patterns across trials. However, in the later stages of the experiments, subjects gave up searching for patterns and at the end of the experiments, subjects preferred to choose the winning option (blue) more frequently. Therefore, most of the subjects choose maximization behavior towards the end of the experiments. However, this is not directly observable from the graphs presented in the previous chapter because these graphs show average response of all participants. Interestingly, in Experiment 4, the PD patients showed less pattern searching behavior and more maximizing behavior than controls but it needs more patients to confirm this trend. Whether these behavioral differences were influenced by the dopamine replacement therapy is an intriguing question.

In the context of the pattern searching mechanism in probability learning, a working memory system seems to be necessary to integrate previous information to predict future actions. Working memory might be acting like a buffer which stores previous patterns of feedback information and compares them with the newly learned ones in order to execute new action strategies. According to this hypothesis, if working memory could not find any pattern information from the preceding trials, procedural memory is still there to learn the relative frequency of outcomes. Interestingly, our findings indicate that pattern search occurs at the beginning which leads to probabilistic learning of the frequencies of choices. Once the frequencies are learned, pattern search behavior is replaced with maximization behavior to achieve better winning performance. In the light of these findings, we project that at the beginning of the experiments, activity occurs at DLPFC and striatum due to pattern searching and frequency learning. Towards the end of the experiments, however, learning would have been completed, so reduced activity only due to error prediction will be observed at striatum exclusively. Nevertheless, this projection must be verified by neuroimaging experiments in future studies.

4.5. Performance of the Medicated Parkinson Disease Patients

Experiment 4 is conducted to investigate the probability learning performance of dopamine medicated Parkinson patients. Frank and his colleagues (2004) reported that patients who are on medication learn better with positive feedback than un-medicated patients. In Experiment 4, we examined the PD patients on the most complex probability learning task, IIT. When the percent blue choice graphs are considered, medicated Parkinson patients showed better performance than healthy controls in the irrelevant information task (IIT). Hence our findings are similar to Frank’s (Frank et al., 2004), but due to the low number of PD patients, statistical significance could not be obtained in Experiment 4. We anticipate that increasing the number of Parkinson patients in this study might increase the statistical significance, leading to the replication of Frank’s study.

Unfortunately, we were unable to study the behavior of unmediated patients because of the difficulty of obtaining this type of subject pool. In the future it would be interesting to complement our research with un-medicated patients as well.

4.6. Limitations of the Study

One limitation of our experiments is that in all experiments the stimuli and the feedbacks were presented randomly which makes it harder to analyze the response data in terms of sequences of patterns. If the sequences of the feedbacks have been controlled we could have differentiated the random responses and non-random responses more accurately.

Another limitation is the number of trials in our experiments. In the literature, probabilistic learning experiments have been conducted with 100 - 3000 trials extending from 30 min to a few days. Due to technical and timing limitations, we administered one 30 min session consisting of 200 trials to each participant. We speculate that if we had a chance to study more trials, the statistical significance of monetary versus non-monetary reward comparisons might have changed.

Finally, the number of Parkinson patients recruited for Experiment 4 is low because we applied a high standard to control for dopamine medication, level of disease, on/off condition for accepting a patient into our experiment. Increasing the number of patients might change the significance of the results.

CHAPTER 5

CONCLUSION

The present study investigated the effects of reward, context, and uncertainty in probability learning behavior in a series of four experiments. Reward was investigated by changing the feedback saliency, to compare the effect of monetary and non-monetary feedback. In Experiment 1 and Experiment 3 we found no significant effect of monetary reward on probability learning, but in Experiment 2 the performance of the subjects who are in the monetary feedback condition was significantly better than the non-monetary feedback group. It seems that monetary feedback is more effective when there is irrelevant information, or when task complexity is higher. Therefore, we can say that when the variability in the data is increased money can explain part of this variability.

On the other hand, when we compared the overall reaction time data of the subjects in both groups of Experiment 1 and Experiment 2, overall reaction time in the Experiment 2 is significantly higher than the overall reaction time in Experiment 1. Therefore, we can say that the irrelevant information increases the complexity of the context to be processed, hence increasing the response latency.

The effect of uncertainty was measured by the frequency of blue winning with respect to stimulus type in Experiment 3. Because blue wins half of the time, no learning effect was expected in this experiment. The overall reaction time data for both the monetary and non-monetary group shows that when there is high uncertainty, the reaction time takes longer. This finding replicates the literature which indicates that uncertainty embodied in the task is similar to the entropy calculated from the probabilities, and reaction time correlates with the entropy but not perfectly.

Frequency matching behavior is studied for Experiment 1 and Experiment 4. It seems that in the initial trials of the probability learning experiments, subjects look for sequential patterns, but after a period of learning, they show maximizing behavior preferring to choose the more frequent option, which leads to optimal performance. We suggest that this finding supports the claims in the literature which indicate pattern search is a behavioral choice for learning the frequency of outcomes implicitly.

Experiment 4 was designed to test the hypothesis on the effects of dopamine medication. Whether dopamine replacement therapy improves or impairs probability learning ability, it was not explicitly known in the literature. We found no significant behavioral difference between normal controls and medicated Parkinson patients, although the learning curves indicated more overmatching behavior for Parkinson patients. The observed overmatching result actually goes along with our hypothesis that positive feedback causes dopamine medicated Parkinson patients to learn better than controls.

Overall, this thesis helped us make the following observations regarding probability learning: 1. the positive effect of reward is not as pronounced as reported in the literature; 2. The uncertainty embodied in the task has a direct effect on reaction times; 3. The effect of the complexity of context in the stimuli is reflected in reaction times, but does not affect performance reported as percentage of correct responses; 4. Dopamine medication may change the behavior of subjects during and at completion of probability learning, but to conclude whether this effect is significant, a larger subject pool is needed.

As part of our future studies, we would like to develop computational models of the probability learning behavior, and investigate neural correlates of this behavior using neuroimaging techniques. We also would like to extend our experiments of the dopamine medicated Parkinson population by increasing the number of participants and further investigating the effects of treatment.

REFERENCES

Aharon, I., Etcoff, N., Ariely, D., Chabris, C.F., O'Connor, E., Breiter, H.C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron 32:* 537-551.

Alexander, G.E., DeLong, M.R., Strick, P.L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience 9*: 357–381.

Angus W. MacDonald, Jonathan D. Cohen, V. Andrew Stenger, and Cameron S. Carter (2000). Dissociating the Role of the Dorsolateral Prefrontal and Anterior Cingulate Cortex. *Cognitive Control Science, 288,* 1835 - 1838.

Anderson, J. R. & Lebiere, C. (1998). *The atomic components of thought.* Mahwah, NJ: Erlbaum.

Anderson, J.R (2005) *Cognitive Psychology (2005 Ed.).* W H Freeman & Co.

Aron, A., Fisher, H., Mashek, D.J., Strong, G., Li, H., Brown, L.L. (2005). Reward, motivation, and emotion systems associated with early-stage intense romantic love. *Journal of Neurophysiology, 94,* 327–337.

Ashby, F.G., Maddox, W.T. (2005) Human Category Learning. *Annual Review of Psychology*, *56*, 149-78.

Ashby, F.G. & O'Brien, J.B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Science, 2*, 83-89.

Atkinson, R. C. (1964). *Studies In Mathematical Psychology*. Stanford University Press.

Baddeley, A. (1992) Working memory**.** *Science, 255,* 556-559.

Baddeley, A.D. (2003) Working Memory: Looking back and looking forward. Nature Reviews Neuroscience, 4 (10): 829-839.

Beach L.R. & Shoenberger R.W. (1965) Event Salience and Response Frequency in a Ten-Alternative Probability Learning Situation. *Journal of Experimental Psychology, 69,* 312-6.

Bechara, A., Tranel, D., Damasio, H., Damasio, A.R. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex, 6,* 215-225

Berns, G.S., McClure, S.M., Pagnoni, G., Montague, P.R. (2001). Predictability modulates human brain response to reward. *Journal of Neuroscience, 21,* 2793-2798.

Berridge K.C., Robinson T.E. (2003). Parsing reward. *Trends in Neuroscience, 26*(9)*,* 507–13.

Berridge, K.C. & Robinson, T.E. (1998). What is the role of dopamine in reward: Hedonics, learning, or incentive salience? *Brain Research Reviews, 28*(3), 308-367.

Berthoz, A. (2006). *Emotion and Reason. The Cognitive Neuroscience of Decision Making*. Oxford University Press.

Birnbaum, M. H., & Wakcher, S. V. (2002). Web-based experiments controlled by JavaScript: An example from probability learning. *Behavior Research Methods, Instruments, & Computers, 34,* 189-199.

Blood A. J., Zatorre R. J. (2001) Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of National Academy Sciences (USA), 98,* 11818-11823.

Bogacz, R., McClure, S.M., Li, J., Cohen, J.D., Montague, P.R. (2007). Short-term memory traces for action bias in human reinforcement learning. *Brain Research, 1153,* 111-121.

Botvinick M. M., Cohen J.D., & Carter C.S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences, 12,* 539-546.

Bower, G. H. (1994) Turning Point in Mathematical Learning Theory. *Psychological Review, 101,* 290-300.

Brunswik (1939). Probability as a determiner of rat behavior. *Journal of Experimental Psychology, 25.*

Bush, R.R, Mosteller F.R. (1955). *Stochastic Models for Learning.* John Wiley & Sons

Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Cohen, J.D. (1998). Anterior Cingulate Cortex, Error Detection, and the Online Monitoring of Performance. *Science, 280,* 747 - 749.

Chao, J. and Nestler, E.J. (2004). Molecular neurobiology of drug addiction. *Annual Review of Medicine, 55*, 113-132.

Clark, Rose Kimberly (2004) *Time to Decide: Probability Guessing Dynamics Under Varying Time Constraints*. PHD thesis.

Cohen, J.D., Peristein, W.M., Braver, T.S., Nystrom, L.E., Noll, D.C., Jonides, J., Smith, E.E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature, 386,* 604-608.

Cosmides, L. & Tooby, J. (1996). Are humans good intuitive statisticians after all?: Rethinking some conclusions of the literature on judgment under uncertainty. *Cognition, 58*, 1-73.

Courtney, S.M., Ungerleider, L.G., Keil, K., Haxby, J.V. (1996). Object and spatial visual working memory activate separate neural systems in human cortex. *Cerebral Cortex, 6,* 39-49.

Courtney, S. M. Petit, L. Haxby, J.V. Ungerleider, L. G (1998). The Role of Prefrontal Cortex in Working Memory: Examining the Contents of Consciousness. *Philosophical Transactions of the Royal Society, 353,* 1819-1828.

Craik, Kenneth (1943). *The Nature of Exploration.* Cambridge, England, Cambridge University Press.

Davis, D. D. & Holt, C. A. (1993). *Experimental economics.* Princeton University Press.

Delgado, M.R., Miller, M.M, Inati, S., Phelps, E.A. (2005). An fMRI study of reward-related probability learning. *NeuroImage*, *24*(3), 862-73.

D'Esposito, M., Detre, J.A., Alsop, D.C., Shin, R.K., Atlas, S., Grossman, M. (1995). The neural basis of central execution systems of working memory. *Nature, 378*, 279-281.

De Quervain, D.J., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., Fehr, E. (2004). The neural basis of altruistic punishment. *Science, 305*, 1254-1258.

Dreher, J-C. Kohn, P. Berman, K.F. (2006). Neural coding of distinct statistical properties of reward information in humans. *Cerebral Cortex, 16,* 561-73.

Edwards, W. (1956). Reward probability, amount, and information as determiners of sequential two-alternative decision. *Journal of Experimental Psychology, 52,* 177-188.

Edwards, W. (1961). Probability Learning in 1000 trials, *Journal of Experimental Psychology, 4,* 385-394.

Elliott, R., Frith, C.D., Dolan, R.J. (1997). Differential neural response to positive and negative feedback in planning and guessing tasks. *Neuropsychologia, 35*, 1395-1404.

Elliott, R., Friston, K.J., Dolan, R.J. (2000). Dissociable neural responses in human reward systems. *Journal of Neuroscience, 20*, 6159–6165.

Erk, S., Spitzer, M., Wunderlich, A.P., Galley, L., Walter, H. (2002). Cultural objects modulate reward circuitry. *Neuroreport,* 13, 2499-2503.

Estes, W. K. (1961). A descriptive approach to the dynamics of choice behavior. *Behavioral Science*, *6*, 177-184.

Estes, W.K. (1954). Individual Behavior in Uncertain Situations: An Interpretation in Terms of Statistical Association Theory. In R.M. Thrall, C.H. Coombs, and R.L. Davis (Eds.), *Decision Processes* (pp. 127-137). Wiley, New York.

Estes,W.K.(1972) Research and Theory on the Learning of Probabilities *Journal of the American Statistical Association*, Vol. 67, No. 337 (pp. 81-102

Estes, W. K., Suppes, P. (1974). Foundations of stimulus sampling theory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary Developments in Mathematical Psychology, Vol. 1: Learning, memory, and thinking*. San Francisco: Freeman, 163-183.

Estes, W. K. (1976). The cognitive side of probability learning. *Psychological Review*, *83*, 37-64.

Estes, W. K. (1976). Some functions of memory in probability learning and choice behavior. In Bower, G.H. (ed.), *The Psychology of Learning and Motivation*, vol. 10, Academic Press, New York.

Ertan, F.S., Ertan, T., Kiziltan, G., Uyguçgil, H. (2005) Reliability and validity of the Geriatric Depression Scale in depression in Parkinson's disease. *J Neurol Neurosurg Psychiatry*. Oct;76(10):1445-7.

Filoteo, J.V., Maddox, W.T., Ing, A.D., Song, D.D. (2007). Characterizing Rule-Based Category Learning Deficits in Patients with Parkinson's Disease. *Neuropsychologia, 45*, 305-320.

Filoteo, J.V., Maddox, W.T., Ing, A.D., Zizak, V., & Song, D.D. (2005). The Impact of Irrelevant Dimensional Variation on Rule-Based Category Learning in Patients with Parkinson's Disease. *Journal of the International Neuropsychological Society*, *11*, 503-513.

Fiorillo, C.D., Tobler, P.N., Schultz, W. (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science, 299,* 1898-1902.

Frank, M.J., Seeberger, L. & O'Reilly, R.C. (2004). By carrot or by stick: Cognitive reinforcement learning in Parkinsonism. *Science, 306,* 1940-1943.

Frank, M.J. (2005). Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism. *Journal of Cognitive Neuroscience, 17,* 51-72.

Frank, M.J. & Claus, E.D. (2006). Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making and reversal. *Psychological Review, 113,* 300-326.

Friedman, M. P., Burke, C. J., Cole, M., Keller, L.,Millward, R. B., & Estes, W. K. (1964). Two-choice behavior under extended training with shifting probabilities of reinforcement. In Atkinson, R.C. (Ed.), *Studies in mathematical psychology* (pp. 250-316). Stanford, CA: Stanford University Press. MA: MIT Press.

Fuster, J.M. (1990). Prefrontal cortex and the bridging of temporal gaps in the perception-action cycle. In (Diamond A., Ed), *The development and neural bases of higher cognitive functions*, (pp 318-336). New York: New York Academy of Sciences.

Gallagher, M., McMahan, R.W. & Schoenbaum, G. (1999). Orbitofrontal cortex and representation of incentive value in associative learning. *Journal of Neuroscience, 19*, 6610−6614.

Gardner, R. (1957). Probability learning with two and three options,. *American Journal of Psychology, 70,* 174-185.

Gibbard, P. & van Kolfschoten, T. (2004). The Pleistocene and Holocene Epochs. In Gradstein, F. M., Ogg, James G., and Smith, A. Gilbert (Eds.), *A Geologic Time Scal.* Cambridge University Press, Cambridge.

Gigerenzer, G. & Selten, R. (Eds). (2001). *Bounded Rationality: The Adaptive Toolbox.* The MIT Press.

Gigerenzer, G., Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102(4)*, 684-704.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review, 103*, 592-596.

Glimcher, P (2004) *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics*. The MIT Press.

Glovich, T., Grifin, D., Kahneman, D. (2002). *Heuristics and Biases: The Psychology of Intuitive Judgment.* Cambridge University Press

Gluck, M. A., Shohamy, D., & Myers, C. E. (2002). How do people solve the “weather prediction” task?: Individual variability in strategies for probabilistic category learning. *Learning and Memory*, *9*, 408-418.

Goldman-Rakic, P.S. (1992). Working memory and the mind. *Scientific American, 267,* 111-117.

Güngen, C., Ertan, T., Eker., E. (2002) Standardize Mini Mental Test'in Türk toplumunda hafif demans tanısında geçerlilik ve güvenilirligi. *Türk Psikiyatri Dergisi* 2002; 13: 273-281.

Hardoon, K. K., Baboushkin, H. R., Derevensky, J. L., & Gupta, R. (2001). Underlying cognitions in the selection of lottery tickets. *Journal of Clinical Psychology*, 57(6), 749-763.

Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist, 39,* 1372-1388.

Hawking, S.W. (2001). *The Universe in a Nutshell (1st Ed.).* Bantam.

Held, C., Knauff, M. & Vosgerau, G. (eds.): *Mental Models and the Mind: Current Developments in Cognitive Psychology, Neuroscience, and Philosophy of Mind.* Elsevier, Oxford, pp. 255-275.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13*, 243-266.

Hertwig, R. & Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists?. *Behavioral and Brain Sciences, 24*, 383–403.

Hinson, J.M., Staddon, J.E.R. (1983). Matching, maximizing and hill-climbing. *Journal of Experimental Analysis of Behavior, 40,* 321-331.

Humphreys, L.G. (1939). Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *Journal of Experimental Psychology, 25,* 294-301.

Hoehn, M.M., Yahr, M.D. (1967) Parkinsonism: Onset, progression, and mortality. Neurology: 17: 47-51.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology, 45,* 188-196.

Ivry, R., Knight, R.T. (2002) Making order from chaos: The misguided frontal lobe. *Natural Neuroscience, 5*, 394–396.

Jarvik M.E. (1951). Probability learning and a negative recency effect in a serial anticipation of alternative symbols. *Journal of Experimental Psychology*, *41*, 291-297.

Jonides, J., Smith, E.E., Koeppe, R.A., Awh, E., Minoshima, S., Mintun, M.A. (1993). Spatial working memory in humans as revealed by PET. *Nature, 363,* 623-625

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Kahneman, D. & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*, 430–454.

Kahneman, D. & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103*, 582-591.

Kahneman, D. (2002). *Maps of Bounded Rationality A Perspective on Intuitive judgments and Choice. Nobel Prize Lecture*. Retrieved August 27, 2007, from http://nobelprize.org/nobel_prizes/economics/laureates/2002/kahneman-lecture.html

Karsh, B.E. & Suppes, P. (1964). Probability learning of rats in continuous-time experiments. *Psychonomic Science*, *1*, 361-362.

Kirk, K.L. & Bitterman, M.E. (1965). Probability learning by the turtle. *Science, 148*(3676), 1484-1485.

Knowlton, B.J., Squire, L.R. and Gluck, M. (1994). Probabilistic classification learning in amnesia. *Learning and Memory, 1*, 106-120.

Knowlton, B.J., Mangels, J.A., and Squire, L.R. (1996). A neostriatal habit learning system in humans. *Science, 273*, 1399-1402.

Knutson B, Fong GW, Bennett SM, Adams CM, Hommer D. (2003) A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid eventrelated fMRI. *Neuroimage* 18:263–72

Knutson, B., Wimmer, G. E. (2007). Reward: Neural circuitry for social valuation. In Harmon-Jones, E., Winkielman, P. (Eds.) Social Neuroscience, New York: Guilford Press, pp. 157-175.

Kurzweil, Ray. (2005*). The Singularity Is Near: When Humans Transcend Biology* (Hardcover)**.** Viking Adult

Lamberts, K. Goldstone, R. (2005) *Handbook of Cognition*. Sage Publications

Laming, D (1973) *Mathematical Psychology*. Academic Press.

Lovett, M. C. (1998). Choice. In J. R. Anderson, & C. Lebiere (Eds.). *The atomic components of thought,* 255-296. Mahwah, NJ: Erlbaum.

Maia, T. V. & McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: what participants really know in the Iowa gambling task. *Proceedings of National Academy of Science USA, 101***,** 16075–16080.

McCarthy, G., Blamire, A.M., Puce, A., Nobre, A.C., Boch, G., Hyder, F., Goldman-Rakic, P., Shulman, R.G. (1994). Functional magnetic resonance imaging of human prefrontal cortex activation during a spatial working memory task. *Proceedings of Nationall Academy of Sciences USA, 91,* 8690-8694.

McClure, S.M., Daw, N.D., Montague, P.R. (2003). A computational substrate for incentive salience. *Trends in Neurosciences*, *26*, 423-428.

McClure, S.M., Berns, G.S., Montague, P.R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, *38*, 339-346.

McClure, S.M. (2003). *Reward prediction errors in human brain*. Ph.D. *Dissertation*.

Middleton, F.A. & Strick, P.L. (2002). Basal-ganglia 'Projections' to the Prefrontal Cortex of the Primate. *Cerebral Cortex, 12*, 926-935.

Miller, E.K., Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24,* 167-202.

Miller, M.B., Valsangkar, M., Newman, S., Dumont, H., Wolford, G. (2005). Brain activations associated with probability matching. *Neuropsychologia, 43*(11)*,* 1598-608.

Mobbs, D., Greicius, M.D., Abdel-Azim, E., Menon, V., Reiss, A.L. (2003). Humor modulates the mesolimbic reward centers. *Neuron, 40,* 1041–1048.

Montague, P.R., King-Casas, B., Cohen, J.D. (2006). Imaging valuation models in human choice. *Annual Review of Neuroscience, 29*, 417-448.

Myers, J. L., Fort, J. G., Katz, L., & Suydam, M. M. (1963). Differential monetary gains and losses and event probability in a two-choice situation. *Journal of Experimental Psychology, 66*, 521-522.

Neimark, E. & Shufod, E. (1959). Comparison of predictions and estimates in a probability learning situation. *Journal of Experimental Psychology, 57*, 294-298.

Nestler, E.J. and Aghajanian, G.K. (1997). Molecular and cellular basis of addiction. *Science, 278*, 58-63.

Nicks D. C. (1959). Prediction of sequential two-choice decisions from event runs. *Journal of Experimental Psychology*, *57*, 105-114.

Niv, Y., Duff, M.O. & Dayan, P. (2004). Dopamine, Uncertainty and TD Learning. In: *CNS2004 - The Thirteenth Annual Computational Neuroscience Meeting,* Presentation notes, Baltimore, Maryland.

O'Brien-Malone, A. & Maybery, M**.** (1998). *Implicit Learning: Implicit and Explicit Mental Processes*. K. Kirsner, C. Speelman, M. Maybery, A. O'Brien-Malone, M. Anderson, C. MacLeod (Eds.). Mahwah, New Jersey, Lawrence Erlbaum Associates, pp 37-56

O'Doherty, J.P., Dayan, P., Friston, K., Critchley, H., Dolan, R.J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron, 28,* 329-337.

O'Doherty J.P., Buchanan TW, Seymour B, Dolan RJ. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron, 49*(1), 157–66.

O'Doherty J.P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current Opinion in Neurobiology*, *14*(6), 769–76

O'Doherty, J.P., Critchley, H., Deichmann, R. & Dolan, R.J. (2003). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *Journal of Neuroscience, 23*, 7931−7939.

Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Bulletin, 68*, 29-46.

Peterson, C.R., Ulehla, Z.J. (1965). Sequential patterns and maximizing. *Journal of Experimental Psychology, 69*, 1-4.

Petrides, M., Alivisatos, B., Evans, A.C., Meyer, E. (1993). Dissociation of human mid-dorsolateral from posterior dorsolateral frontal cortex in memory processing. *Proceedings of National Academy Science USA, 90*, 873-877.

Petrovic, P., Dietrich, T., Fransson, P., Andersson, J., Carlsson, K., Ingvar, M. (2005). Placebo in emotional processing-induced expectations of anxiety relief activate a generalized modulatory network. *Neuron, 46*, 957–969.

Pinker, S. (1999). *How the Mind Works(New Ed. Edition).* Penguin Press Science.

Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Creso-Moyano, J., Myers, C. E. & Gluck, M. A. (2001). Interactive memory systems in the brain. *Nature, 414,* 546-550.

Poon, S. K. (1997). *Seeing Leads to Knowing: Mentalistic bias in choice and probability judgment under uncertainty.* Phd Thesis.

Reber, A.S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*, 219-235.

Reber, A. S. (1992). An evolutionary context for the cognitive unconscious. *Philosophical Psychology, 5*, 33-51.

Redgrave P.(2007) Basal Ganglia. Scholarpedia, p.13246

Rescorla R. A., Wagner A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In Black, A.H. & Prokasy, W.F. (Eds.), *Classical Conditioning II: Current Research and Theory* (pp. 64-99). New York: Appleton Century Crofts.

Rolls, E. (2005). *Emotion Explained*. Oxford University Press.

Ross, D., Spurrett, D., Vuchinich, R. (2006) The Behavioral Economics and Neuroeconomics of Disordered Gambling. http://www.uab.edu/Philosophy /ross.html Last Checked [9.22.2007]

Rubinstein, I. (1959). Some factors in probability matching. *Journal of Experimental Psychology*, *57*, 413-416.

Salvatore, R. (1999). *The Economics of the Mind*. Edward Elgar.

Schultz , W. (2007) Reward. Scholarpedia, p.10861

Schultz, W. (2007). Multiple dopamine functions at different time courses. *Annual of Review of Neuroscience, 30,* 259-288.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology, 57*, 87-115.

Schultz, W., Dayan, P. & Montague, P. (1997). A Neural Substrate of Prediction and Reward. *Science, 275,* 1593-1599.

Semendeferi, K., Lu, A., Schenker, N., and H. Damasio. 2002. Humans and great apes share a large frontal cortex. *Nature Neuroscience*. 5:272-276.

Seymour B, O'Doherty J, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS. (2004). Temporal difference models describe higher-order learning in humans. *Nature*, *429*, 664-7.

Shanks, D. R., Tunney, R. J., & McCarthy, J. (2002). A Re-examination of Probability Matching and Rational Choice. *Journal of Behavioral Decision Making, 15*, 233-250.

Shohamy, D., Myers, C. E., Onlaor, S., & Gluck, M. A. (2004). The role of the basal ganglia in category learning: How do patients with Parkinson's disease learn? *Behavioral Neuroscience, 118 (4),* 676-686.

Siegel, Sidney (1964). *Choice, Strategy, and Utility.* McGraw-Hill

Siegel S. & Goldstein, D. A. (1959). Decision making behaviour in a two-choice uncertain outcome situation. *Journal of Experimental Psychology*, *57*, 37-42.

Simon, H. (1957). *Models of man: Social and Rational.* John Wiley & Sons

Sloman, S.A., D. Over, L. Slovak, and J.M. Stibel. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes*, *91,* 296–309.

Smith, E.E., Jonides, J., Koeppe, R.A., Awh, E., Schumacher, E.H., Minoshima, S. (1995). Spatial versus object working memory: PET investigations. *Journal of Cognitive Neuroscience, 7*, 337-356.

Soliveri, P., Brown, R., Jahanshahi, M., Caraceni, T., & Marsden, C. (1997). Learning manual pursuit tracking skills in patients with Parkinson's disease. *Brain, 120,* 1325–1337.

Speelman, C. & Maybery, M. (1998). *Automaticity and Skill Acquisition, Implicit and Explicit Mental Processes*. K. Kirsner, C. Speelman, M. Maybery, A. O'Brien-Malone, M. Anderson, C. MacLeod (Eds.) pp. 79-98. Mahwah, New Jersey, Lawrence Erlbaum Associates.

Squire, L.R. (2004). Memory Systems of the Brain: A brief history and current perspective. *Neurobiology of Learning & Memory, 82*, 171-177.

Staddon, J. E. R. (2003) *Adaptive behavior and learning.* New York: Cambridge University Press. Second (internet) edition.

Thut, G., Schultz, W., Roelcke, U., Nienhusmeier, M., Maguire, R.P., Leenders, K.L. (1997). Activation of the human brain by monetary reward. *NeuroReport*, *8*, 1225-1228

Tulving, E. (1998): Study of memory: processes and systems. In J.K. Foster & M. Jelicic (Eds.), Memory: Systems, Process, or Function? (pp. 11-30). Oxford: Oxford University Press.

Tversky, A. & Edwards, W. (1965). Information versus Reward in Binary Choices. *Journal of Experimental Psychology*, 71.

Uhl, C. N. (1963) Learning of interval concepts: 1. Effects of differences in stimulus weights. *Journal of Experimental Psychology* 66:264-273. 641.

Volz, K. G., Schubotz, R. I., von Cramon, D. Y. (2004) Predicting events of varying probability: Uncertainty investigated by fMRI. *NeuroImage*, 19 (2): 271-280

Vulkan, N. (2000). An Economist's Perspective on Probability Matching. *Journal of Economic Surveys 14 (1), 101–118.*

Wallis J.D. (2007). Orbitofrontal Cortex and Its Contribution to Decision-Making. *Annual Review of Neuroscience, 30,* 31-56.

West, R. F., & Stanovich, K. E. (2003). Is probability matching smart? Associations between probabilistic choices and cognitive ability. *Memory & Cognition*, *31*, 243-251.

Wilson, W. A. Jr., & Rollin, R. A. (1959). Two choice behavior of rhesus monkeys in a noncontingent situation. *Journal of Experimental Psychology, 58, 2*, 174-180.

Wolford, G.L., Miller, M. B. & Gazzaniga, M. (2000). The left hemisphere's role in hypothesis formation. *Journal of Neuroscience*, *20(RC64),* 1-4.

Wolford, G., Newman, S., Miller, M., Wig, G. (2004). Searching for patterns in Random sequences. *Canadian Journal of Experimental Psychology*, 58(4):221-8.

Wollen, K.A. (1963). Relationships between choice time and frequency during discrimination training and generalization tests. *Journal of Experimental Psychology, 66,* 474-84.

Yellott, J.I. (1969). Probability learning with noncontingent success. *Journal of Mathematical Psychology, 6*, 541-575

Zacks, R. T., & Hasher, L. (2002). Frequency processing: A twenty-five year perspective. In P. Sedlmeier & T. Betsch (Eds.), *Frequency processing and cognition* (pp. 21-36). New York: Oxford University Press.

Zajonc, R. B. (1968) Attitudinal Effects of Mere Exposure. *Journal of Personality and Social Psychology, 9*(2)*,* 1-27.

APPENDICES

APPENDIX A - A Sample Inform Consent Form

Teyit

Verilerinizin ve kişisel bilgilerinizin saklı tutulacağını ve korunacağını teyit ediyorum. Bu çalışmanın sonuçları isminizi veya katılımcıları belirleyecek herhangi bir bilgi içermeyecektir. Bu veriler üçüncü kişilere verilmeyecek ve sadece araştırma amaçlı kullanılacaktır. Bu deneyde öğrenmenin ve karar vermenin doğası araştırılmaktadır. Katılımcılara bir grup farklı renge sahip şekil gösterilecektir. Katılımcılardan ödülün hangi rengin altında olduğunu bulmaları istenmektedir. Her doğru tahmin sonunda da bir miktar para ödülü kazanılacaktır.

Burak Erdeniz

Rıza Onay Formu

Ben,____________________________, Orta Doğu Teknik Universitesi Enformatik Enstitüsüne bağlı olan Bilişsel Bilimler bölümü Master öğrencisi Burak Erdeniz denetimindeki öğrenme ve karar verme ile ilgili bu deneye gönüllü olarak katılmayı kabul ediyorum.

Bu çalışmada ekranda bazı figürler gösterileceğinin farkındayım.Bunun devamında bazı sorulara cevap vermem gerektiğini biliyorum.

Çalışma sonuçlarında ismimin kullanılmayacağını ve ismimin benden toplanan verilerle ilişkilendirilmeyeceğini biliyorum.

Kimliğimi belirleyecek bir bilgi olmaksızın benden toplanan verilerin araştırma sonuçlarını paylaşma amaçlı bir dergide yayınlanmasına veya konferansta sunulmasına izin veriyorum.

Çalışmanın ortalama 30 dakika süreceğinin ve istedeğim zaman bırakabileceğimin farkındayım. Çalışmaya gönüllü olarak ve hiç bir etki altında kalmadan katılıyorum ve çalışmayı tamamladığım takdirde deneydeki performansıma göre belli bir miktar para almaya hak kazanabileceğimin farkındayım.

İmza_________________________ Tarih_______________

APPENDIX B - A Sample Instructions Form

Yönerge – PRFL01

- Öncelikle deneye katıldığınız için teşekkürler. Bu deneyde sizden bir takım seçimler yapmanız beklenmektedir. Yapacağınız bu seçimler sonucunda para kazanacak veya hiç birşey kazanamayacaksınız. Size 10 adet kare şekil göstereceğiz. Bunlarda bazıları **mavi** bazıları ise **kırmızı** olacak. **Kırmızıların** veya **mavilerin** altında bir **sarı** kare şekil saklı olacak ve sizden deney esnasında **sarı** karenin yerinin mavilerin mi yoksa **kırmızıların** mı altında saklı olduğunu bulmanız beklenecek. Seçimlerinizi klavye kullanarak yapacaksınız. Eğer **sarı** karenin mavilerin altında saklı olduğunu düşünüyorsanız klavyeden "**M**" tuşuna, kırmızıların altında olduğunu düşünüyorsanız "**K**" tuşuna basınız. **Sarı** karenin yerini her doğru bildiğinizde 50 kuruş kazandığınızı gösteren bir ekran göreceksiniz. Yanlış bildiğinizde ise kazanamadınız yazısı çıkacak.
- Öncelikle 5 denemeden oluşan bir pratik yapıcaksınız ondan sonra gerçek deneye geçilecek. Bu deneme esnasında seçiminiz ne olursa olsun kazancaksınız. Yalnız bilmeniz gereken birşey var eğer 5 saniye içinde bir seçim yapmazsanız bilgisayar bunu yanlış tahmin

olarak kabul edecek. Neticede sizden bütün denemelerde bir seçim yapmanız beklenmektedir.

- Birazdan 180 denemeden oluşan deneye başlayacağız. Lütfen bütün sorulara 5 saniye içinde cevap vermeye çalışın. Tahminimizce deney en fazla 15 dakika sürecektir. Deneyin sonunda ne kadar para kazandığınızın hesaplandığını söyleyen bir ekran karşınıza çıkacaktır. Ancak kazancınız bir katsayı ile çarpılarak size ödenecektir. Bu ekranı gördügünüzde lütfen deneyi yapan kişiye deneyin bittiğini söyleyin.
- Son olarak bu deneyde sizden beklenen mümkün olduğu kadar doğru tahmin yapmaya çalışmanızdır. Teşekkürler.

APPENDIX C - A Sample Debriefing Form

AÇIKLAMA

Lütfen arkadaşlarınıza bugün katıldığınız deney hakkında hiç bir şey söylemeyin. Deney hakkında vereceğiniz herhangi bir bilgi, diğer deneklerin karar verme sürecini etkileyeceğinden yapılan bütün çalışmanın boşa gitmesine sebep olabilir.

Bu deneyde insanlara renklerle kodlanmış bir bilgi sunulduğunda karar vermenin nasıl gerçekleştiğini araştırmaktayız. Farklı öğrenme yöntemlerini test etmek için deneyde pekiştirmeli öğrenme benzeri deneme ve yanılma içeren ödül tabanlı bir test uygulanmıştır. Bu deney sonucunda elde edilecek bilgiler ışığın daha sonra benzeri çalışmalar hasta populasyonlarında da denenebilecektir.

Katılımınız için çok teşekkür ederiz. Eğer daha sonra, yapılan bu çalışma hakkında başka sorunuz olursa veya genel bir sorunuz olursa lütfen Bilişsel Bilimler yüksek lisans öğrencisi Burak Erdeniz ile bağlantıya geçiniz (MM410). Tekrardan hatırlatalım, lütfen arkadaşlarınıza bugün katıldığınız deney hakkında hiç bir şey söylemeyin. Deney hakkında vereceğiniz her hangi bir bilgi diğer deneklerden toplanan verileri bozucu bir etkiye neden olacaktır.