DECISION MAKING IN SUBCLINICAL AND CLINICAL POPULATIONS: HOW DRIFT DIFFUSION MODEL VARIABLES RELATE TO OBSESSIVE COMPULSIVE AND RELATED FEATURES

by

CEYLA ERHAN

Submitted to Koç University's Graduate School of Social Sciences and Humanities

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Psychology

Koç University

2017

STATEMENT OF AUTHORSHIP

This thesis contains no material, which has been accepted for any award or any other degree or diploma in any university or other institution. It is affirmed by the candidate that, to the best of her knowledge, the thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

Signature

Ceyla Erhan

ABSTRACT OF THE DISSERTATION

DECISION MAKING IN SUBCLINICAL AND CLINICAL POPULATIONS: HOW DRIFT DIFFUSION MODEL VARIABLES RELATE TO OBSESSIVE COMPULSIVE AND RELATED FEATURES

by CEYLA ERHAN

Dissertation Director:
Fuat Balcı

This thesis investigates how latent decision making processes are affected in subclinical and clinical samples, with the overarching aim of testing and demonstrating the clinical relevance of decision theoretic approaches. Specifically, I investigated how variables that characterize different aspects of decision-making within the framework of Drift Diffusion Model (DDM; Ratcliff, 1978) relate to anxiety, perfectionism, obsessive compulsive (OC) traits and Obsessive Compulsive Disorder (OCD). DDM analyzes response time and accuracy data gathered in two alternative forced choice (2AFC) tasks in a unified manner, leading to explanations for choice behavior based on concepts such as cautiousness, non-decision related slowing, decision biases, and evidence accumulation efficiency.

Chapter I aims to characterize latent decision processes of a non-clinical population who rank on various levels on OC scales. We predicted that those who rank higher on OC traits would be more cautious in their decisions. Indeed, the first study found that higher checking and rumination tendencies as well as high scores in total OC scales predicted more cautious responding (e.g. higher threshold settings).

Chapter II follows up on the first study and investigates latent decision processes in drug-free pediatric OCD patients and healthy controls. Similar to the first

study, we predicted higher caution (e.g. threshold settings) and lower evidence accumulation efficiency (e.g. drift rates) for OCD patients. As hypothesized, OCD patients accumulated evidence less efficiently but exhibited only a tendency to be more cautious than healthy controls. Furthermore, OCD patients became more cautious whereas healthy controls became less cautious after errors compared to after correct responses.

Chapter III gamifies and validates a new 2AFC task to overcome the typical disadvantages of traditional approaches in the study of pediatric and clinical groups and investigates how latent decision processes change for high and low scorers on perfectionism, OC and anxiety traits in this task. We found that the Hierarchical Drift Diffusion Model (Wiecki, Sofer, & Frank, 2013) provided good fits to the data and the latent variables behaved in expected directions as a function of task parameters. Moreover, consistent with the literature, evidence accumulation rates were found to be lower for participants with high anxiety, perfectionism, and obsessive-compulsive trait scores.

Overall, efficiency in accumulating perceptual evidence seems to be a key variable that differentiates both OCD from healthy control populations as well as those who have high, medium and low levels of perfectionism, OC and anxiety traits. Our results point at the advantages of using computational methods in understanding decision making in conditions of clinical relevance.

ÖZET

Bu tez klinik ve subklinik örneklemlerde altta yatan karar verme süreçlerinin nasıl etkilendiğini araştırır, karar teorik yaklaşımlarını klinik alanlarda test etmek ve klinik alanlarla ilintisini göstermeyi amaçlar. Özellikle, hesaplamalı modelleme yöntemleri (sürüklenme-yayılım modelleri; drift diffusion model: DDM, Ratcliff, 1978) dahilindeki karar verme süreçlerinin farklı yönlerini tanımlayan değişkenlerin kaygı, mükemmelliyetçilik, Obsesif Kompulsif (OK) özellikler ve Obsesif Kompulsif Bozukluk (OKB) ile ilişkisi araştırılmıştır. DDM, iki seçenekli zorunlu seçim senaryolarında toplanan tepki süreleri ve kararın doğruluğu verisini bir bütün olarak analiz eder ve tedbirlilik, karar dışı geçen süre, kişinin seçeneklerden birine karşı eğilimi, ve sinyali işlemleme kalitesi gibi karar vermeyle ilgili kuramlar ile ilgili çıkarımlar yapar.

Birinci bölümde klinik teşhisi olmayan ve OK ölçeklerinde farklı seviyelerde yer alan kişilerin altta yatan karar verme mekanizmalarının tanımlanması amaçlandı. Bu çalışmada, OK ölçeklerinde yüksek puan alanların kararlarında daha temkinli olacağını öngördük. Öngördüğümüz gibi kontrol etme ve ruminasyon eğilimi ve toplam OK puanı yüksek olan kişilerin daha temkinli kararlar verdiğini (daha yüksek eşik uzaklıkları olduğunu) bulduk.

İkinci bölüm, birinci çalışmayı takip etti ve ilaç almayan pediyatrik OKB hastaları ve sağlıklı kontrollerin altta yatan karar verme süreçlerini araştırdı. Birinci çalışmaya benzer olarak, OKB hastalarının hem daha temkinli (ör. daha yüksek eşik uzaklıkları) hem de sinyal işlemleme kalitelerinin daha düşük olacağını (ör. sürüklenme hızı) öngördük. Öngördüğümüz gibi, OKB hastaları sağlıklı kontrollere göre sinyali daha az vermimli bir şekilde işlemledi ve daha temkinli olma eğilimi gösterdi. Ayrıca, OKB hastaları doğru yanıt sonrasına oranla, hatalardan sonra daha

da temkinli karar verirken, sağlıklı kontroller hatalardan sonra daha az temkinli kararlar verdi.

Üçüncü bölüm geleneksel iki seçenekli zorunlu seçim görevlerinin klinik ve pediyatrik popülasyonlarda uygulanırken ortaya çıkabilecek dezavantajlarını aşmak için, görevi hem oyunlaştırıp hem yarattığı oyunun geçerliliğini test etti. Bu çalışma aynı zamanda altta yatan karar verme süreçlerinin yüksek ve düşük OK, mükemmelliyetçilik ve endişe seviyesi olan kişilerde nasıl değiştiğini ölçmeyi amaçladı. Hiyerarşik sürüklenme-yayılım modeli (HDDM, Wiecki et al., 2013) tahminleri ampirik veriyle uygunluk gösterdi ve altta yatan karar verme süreçleri, oyunun parametrelerine göre beklenen yönlerde değişti. Ayrıca, literatürle uyumlu olarak, sinyali işlemleme hızı yüksek endişe, mükemmelliyetçilik ve obsesif kompulsif karakteristikleri olan kişiler için daha düşük bulundu.

Özetle, algısal kanıt toplama verimliliği değişkeni hem OKB ve sağlıklı kontrol gruplarını birbirinden ayıran, hem de yüksek, orta ve düşük seviyelerdeki mükemmelliyetçilik, OK ve endişe niteliklerini birbirinden ayıran ana değişken olarak gözükmektedir. Sonuçlarımız klinik yatkınlıklarda hesaplamalı modelleri kullanmanın avantajlarını ortaya koymuştur.

ACKNOWLEDGEMENTS

I want to first and foremost thank my supervisor Fuat Balcı, who pushed me beyond my limits and always believed that I could surpass them. He patiently guided me through all the challenges I faced. His discipline and compassion made this journey possible. For all his help, I am forever indebted.

I also want to thank Sami Gülgöz, who graced me with his wisdom, and allowed me the space and time to realize my own path. His kindness, expertise and sound advice were invaluable.

I also want to thank Banu Çankaya, who soothed me with her calmness and helped me with instrumental feedback throughout my years at Koç University.

I want to thank my committee members, Ali Sarper Taşkıran, Metehan Irak and Aslı Kılıç for their valuable input and warm support during my studies.

I thank The Scientific and Technological Research Council of Turkey (TÜBİTAK-2211) for their financial support throughout my PhD education.

I thank the Science Academy (BAGEP 2013 to Fuat Balcı) and Turkish Academy of Sciences (GEBIP 2015 to Fuat Balcı) for their support in funding the studies.

I thank all my friends and family who literally and metaphorically pulled me through and without whom I would have surely struggled. All my colleagues at Koç and in the lab, I thank for their companionship and support.

I want to further acknowledge some people who made this journey possible. Eda Mızrak provided me the secure space and allowed me to be who I am with no judgment and endless support. Her calmness soothed my worries and her friendship means so much. Esin Türkakın, who is the epitome of logic and whose thirst for science is admirable, has so gently moved me through and still does. Deniz Özkan and

Begüm Dora listened to me patiently anytime I needed a friend; I could not have asked for better companions on this journey. Zeynep Esin reminded me that whatever it is that I felt, it was ok and I was worthy. Ulaş Gülkirpik's attitude in life and instructive math sessions lifted my spirits every time. Gamze Özkan and Dila Betil allowed me to venture out of the process when I had sunk too deep and helped me gain perspective. Meriç, Eren and Eda Yanık, helped me have fun and showed me how precious life actually is.

I also want to thank my family, who so compassionately stood by my side all through my journey. Sinan, Selin, Eliz, Rifat, Arzu, Emre and my lovely grandmother Esin, they are my rock and I love them deeply. I want to especially thank my mom Çulpan Erhan and my dad Faruk Erhan, who allowed me to delve into my passion with no questions asked. Mom washed away my worries when I was so stressed out over the smallest stuff and made me gain a grander perspective. Dad's respect, admiration and love for science instilled in me the same traits, and that passion is what kept me going. I'm thankful beyond words.

And above all, Kaan Karamancı - Any dent at all created with this dissertation is dedicated to him. His love of science was what started me on this journey, and every time I was ready to let go, he stepped in and so gently pushed me along the way. His respect for me and for what I do, and the help he so selflessly and endlessly offered kept me going. I love him eternally.

Author's note: The first and second chapters have been published in The Quarterly Journal of Experimental Psychology and Journal of Affective Disorders respectively. The third chapter has been submitted for publication.

Erhan, C., & Balcı, F. (2016). Obsessive compulsive features predict cautious decision strategies. *The Quarterly Journal of Experimental Psychology*, *0*(ja), 1–29. http://doi.org/10.1080/17470218.2015.1130070

Erhan, C., Bulut, G. Ç., Gökçe, S., Ozbas, D., Turkakin, E., Dursun, O. B., ... & Balcı, F. (2017). Disrupted latent decision processes in medication-free pediatric OCD patients. *Journal of Affective Disorders*, *207*, 32-37. http://dx.doi.org/10.1016/j.jad.2016.09.011

Erhan, C., Karamanci, K., Turkakin, E.& Balcı, F. (submitted) Gamification of Two Alternative Forced Choice: Validation based on Drift Diffusion Model and Individual Differences.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

THESIS INTRODUCTION

This thesis aims to investigate decision processes in relation to various Obsessive Compulsive (OC) trait related conditions. Traditional analyses that separately analyze error rates (ER) and response times (RT) in psychological tasks cannot disentangle concepts such as speed accuracy trade-off, time allocated to non-decision related processes, possible biases to either choice or evidence accumulation efficiency. In other words, one cannot tap into generative processes that lead to decisions by the isolated analysis of accuracy and speed of choice behavior. On the other hand sequential sampling models such as the Drift Diffusion Model (DDM; Ratcliff, 1978) overcome this problem by utilizing a combination of accuracy and RT data to explain such latent decision making processes. The three chapters presented in this thesis, all incorporate the use of two alternative forced choice (2AFC) tasks and a DDM-based approach to the accuracy and response time data with the overarching goal of uncovering differences in the choice behavior of various clinical and subclinical conditions. While the first two chapters use the dot motion discrimination task (a commonly utilized task in decision science), the third chapter incorporates a gamified 2AFC task. Such gamification sought to better engage participants, improve their experience and thereby increase data quality, which are issues that become particularly relevant in various samples (e.g., pediatric, clinical).

In the field of psychiatry, there is the need for a paradigm shift in analytical approaches to data collected from traditional neuropsychological tasks (Abramovitch & Cooperman 2015) to reach at theory-driven and more extensive characterization of clinical conditions. To this end, sequential sampling methods (Wiecki, Poland, & Frank, 2015) and especially DDM (White, Ratcliff, Vasey, & McKoon, 2010a) are argued to be good candidate integral tools for the experimental clinical area, attested by its successful utilization in studies conducted with ADHD (e.g. Metin et al., 2013; Mulder et al., 2010), anxiety (White, Ratcliff, Vasey & McKoon, 2010b), depression (Pe, Vandekerckhove, & Kuppens, 2013), and OCD (Banca et al., 2015).

DDM assumes a noisy evidence accumulation process, starting from a particular point in a decision space (starting point) constrained by two thresholds, the two possible options for the given decision task. The evidence accumulation continues with a steady rate (drift rate) until the amount of evidence reaches one of the two thresholds, at which point the corresponding choice is made. DDM also distinguishes response times that are due to the decision process and processes outside of decision-making (e.g. signal detection, encoding, motor processes; non-decision time) (Ratcliff, 1978; Ratcliff & McKoon, 2008). Each decision variable output in DDM corresponds to a different psychological process. The width of threshold separation refers to how cautious the decision maker is, which in turn determines the speed-accuracy tradeoff adopted by the decision-maker. The drift rate indexes efficiency with which the evidence is accumulated. Starting point bias indexes a prior preference for either of the two choices (i.e., prior belief state) and the non-decision time indexes duration allocated to processes other than decision making.

Incorporating the DDM, the first chapter focuses on latent decision variables mentioned above, for non-clinical participants who differentially rank on the Obsessive Compulsive (OC) scales. The second chapter uses the same task, dot motion discrimination task, and seeks to investigate the latent decision processes in pediatric OCD patients. The third chapter develops and validates a gamified version of the 2AFC task and investigates latent decision making differences between those who rank high, medium and low on scales of perfectionism, anxiety and OC traits. Perfectionism, a transdiagnostic trait, was particularly added given it had not been studied in a DDM paradigm before and that its existence is argued to be a common maintaining factor for both anxiety and OC traits (Egan, Wade & Shafran, 2011).

Altogether this thesis, through the investigation of latent decision variables within the framework of DDM, was able to reveal signatures that set apart both high and low levels of anxiety, OC and perfectionism traits as well as a healthy control and clinical pediatric OCD group. Additionally, it introduced a participant-experience-centered gamified task and validated it as a 2AFC task that can be accounted for by DDM. Studies presented add to an increasing number of explorations using computational decision-theoretic approaches to reveal characterization of behavior in clinical and subclinical populations.

CHAPTER I

Obsessive compulsive features predict cautious decision strategies

Abstract

Introduction: Obsessive Compulsive Disorder (OCD) is occasionally characterized by decision-making deficits. Compared to the isolated analysis of the choice and response times, characterizing decision outputs at the level of latent processes can be a more powerful approach in revealing differences, even in subclinical cases. We hypothesized that participants with higher Obsessive Compulsive (OC) features would set their decision thresholds higher and thus make more cautious decisions.

Method: We used a perceptual two alternative forced choice (2AFC) task (dot motion discrimination) to test this hypothesis in a non-clinical sample (*N*=74). We fit the data with the diffusion model and evaluated the optimality of decision outputs. We also conducted exploratory analyses to reveal which subscales best predicted the differences at the level of latent decision processes.

Results: Higher OC total scores in Maudsley and Padua scales significantly predicted higher threshold settings (cautiousness). The follow-up exploratory analyses with subscale scores showed that checking and rumination tendencies predicted higher threshold settings whereas washing tendency predicted faster non-decision times.

Conclusions: Our primary results showed that participants with higher degrees of OC features exhibited more cautious decision-making. Our exploratory analyses also revealed distinctions based on different types of OC features in both controlled (cautiousness in decision making) and automatic (faster non-decision times) elements of the decision process.

Keywords: Decision Making, Drift Diffusion Model, Obsessive-compulsive disorder, Checking, Rumination

Introduction

Obsessive Compulsive Disorder (OCD) is a debilitating psychiatric condition with the manifestation of obsessions and/or compulsions (American Psychiatric Association, 2013). It is sometimes depicted as a decision making disorder (Sachdev & Malhi, 2005). However, traditional decision-making experiments with OCD patients have yielded inconsistent results. For example, several studies that have used the Iowa Gambling Task (IGT; Bechara, Damasio, Damasio, & Anderson, 1994) (e.g. Nielen, Veltman, De Jong, Mulder, & Den Boer, 2002; Lawrence et al., 2006) found comparable results between healthy controls and OCD patients as a group, in terms of advantageous responding. In contrast, other studies that have also used the IGT reported that OCD patients make more disadvantageous decisions (Cavedini et al., 2002; Rocha, Alvarenga, Malloy-Diniz, & Corrêa, 2011; Starcke, Tuschen-Caffier, Markowitsch, & Brand, 2010).

Similarly, findings regarding executive function and processing speed showed inconsistencies between tasks OCD patients were tested on. In a recent meta-analysis Abramovitch, Abramowitz and Mittelman (2013a) summarized 115 OCD studies and reported a medium mean effect size for both processing speed and executive functions indicating worse performance for OCD patients. However there is much heterogeneity between the findings of individual studies. For instance, in the Go/NoGo task, several studies showed that the number of errors and response times of OCD patients did not differ from those of healthy controls (e.g., Bohne, Savage, Deckersbach, Keuthen, & Wilhelm, 2008). Differently, other studies revealed that OCD patients had slower response times (RTs) and higher number of comission errors than healthy controls (e.g. Abramovitch, Dar, Hermesh, & Schweiger, 2012).

On the other hand, the isolated analysis of RTs and error rates is not always sufficient to reveal possible differences present at the level of latent decision processes. Furthermore, central tendency measures of RTs do not provide information regarding the shape of the distribution resulting in incomplete understanding of the data and the underlying generative processes. Instead, distributional analyses use all the RT information within the experiment providing a much more complete understanding of the participant's performance (Heathcote, Popiel, & Mewhort, 1991). This rich set of RT information combined with the relative densities of correct and error RTs (i.e., accuracy) can be accounted for in a psychomechanistically meaningful fashion by several decision-theoretic approaches such as the Diffusion Model (e.g., Ratcliff, 1978).

The Drift Diffusion Model (DDM) is a random walk model that enables the combined analysis of error rates and RTs and helps explain the processing dynamics that underlie decision-making (Ratcliff, 1978; Ratcliff & Rouder, 1998; Ratcliff & McKoon, 2008). This approach paves the way to psychomechanistic explanations for choice behavior based on notions such as cautiousness, non-decision related slowness, decision biases, and the quality of information processing (e.g., White et al., 2010a).

The DDM assumes that sensory evidence is accumulated in a noisy fashion starting from a particular point in a decision area demarcated by two fixed thresholds (referring to two hypotheses/options in the task). A decision is made when the amount of accumulated evidence reaches one of these two thresholds. The threshold hit determines the choice and the first threshold crossing time is the decision time (Ratcliff, 1978; Ratcliff & McKoon, 2008). The simple form of DDM has four core parameters: boundary separation (a), drift rate (v), starting point (z), and non-decision processing time (T_0).

Each of these parameters refers to different psychologically meaningful elements in the decision process. For instance, the boundary separation indexes how cautious the decision maker is while making a choice (e.g., wider decision boundary suggests a more cautious decision strategy); the drift rate describes the amount of evidence gained per unit time from the stimulus (e.g., higher drift rate indicates more information gained per unit time; higher signal-to-noise ratio); the starting point accounts for response bias (e.g., starting closer to a threshold suggests a pre-existing response bias for the corresponding choice); and the non-decision time indexes delays in signal-detection and/or motor responses not related to the decision process (e.g., higher non-decision time indicates slower signal-detection and/or manifestation of a decision) (White et al., 2010a). The extended form (Ratcliff & Rouder, 1998) of DDM includes trial-to-trial variability parameters in addition to these core parameters, which are variabilities in non-decision time, drift rate, and starting point. With these extra parameters DDM can account for unequal correct and error RTs.

The DDM might prove useful in revealing differences in the choice behavior of different clinical conditions because it enables making inferences regarding latent decision processes. This possibility has recently commanded attention in the field. For instance, the advantages of incorporating diffusion models in the clinical area have been emphasized by White et al. (2010a). In recent years, DDM has been used in the study of decision-making in ADHD (e.g. Metin et al., 2013; Mulder et al., 2010), anxiety (White et al., 2010b), depression (Pe et al., 2013), and recently OCD (Banca et al., 2015).

Most relevant to the current study, Banca et al. (2015) investigated 2AFC decision making in OCD patients, adopting a DDM-based approach (Hierarchical

Drift Diffusion Model - HDDM) to decision analysis. Banca et al. (2015) used the random dot motion discrimination task, in which dots moving in different directions appeared in a circle in the middle of the screen. While some dots moved randomly, introducing uncertainty, a percentage of the dots (referred to as coherence) moved either to the right or left cohesively. Monetary rewards and penalties were used as indicators of correct and incorrect responses. The goal of the participant was to identify in which direction the cohesive dots are moving. HDDM analyses compared performance on three levels of signal to noise ratio (SNR): low (coherences 0.025 and 0.05), medium (coherences 0.15 and 0.25), and high (coherences 0.45 and 0.7). Their findings revealed that OCD patients had higher threshold settings than healthy controls in low and medium SNR levels. In other words, as the signal was more obscure, OCD patients were more cautious and accumulated more evidence than healthy controls before making a choice. In addition, in medium and high SNR levels, OCD patients were less efficient than the healthy controls in gathering perceptual evidence from the visual signal, denoted by lower drift rates. In an alternate condition, where participants were rewarded for faster responses and penalized for slower ones, OCD patients had similar threshold settings and drift rates to the control participants in the low and medium SNR levels, and even lower threshold settings but also drift rates in the high SNR condition. Thus, when induced to respond faster, OCD patients were able to decrease their threshold settings.

In the current study, we aimed to elucidate the differences in the decision making of non-clinical participants with different rankings on the OCD scales at the level of the latent processes within the framework of DDM. More specifically, the current study aimed to characterize the decision-making patterns of those who ranked differentially on the entirety of the OCD scales and explore how different components of obsessions and compulsions (e.g. rumination, checking etc.) contributed to these differences at the level of the latent decision processes. The doubt component is an integral part of indecision attributed to OCD (Sachdev & Malhi, 2005). Repetitive behaviors in the daily lives of OCD patients such as checking the stove or frequent hand washing may be caused by doubting whether or not these tasks were successfully completed previously. The perceptual evidence of having completed a task or feedback regarding the decision state might not be functionally sufficient for OCD patients to make a timely decision to move on. Instead, they repeat the task to be sure (Sachdev & Malhi, 2005).

We propose that doubt coupled with ideas of rumination, checking, precision seeking and other OC tendencies (captured by the subscales of the OCD scales) push participants to collect more evidence before making decisions. We therefore hypothesized that participants who rank high on OCD scales will have heightened thresholds (owing to requiring more evidence before making a decision) but no difference in the rate of evidence accumulation. We did not have specific predictions regarding the relationship between the OC scores and other DDM parameters (e.g., non-decision time) or the differential predictive value of the OC subscales regarding these dependent variables. To this end, exploratory analyses (e.g., step-wise regression) were conducted to investigate those possible relations.

A wider threshold setting reduces the chances of making a mistake at the expense of longer response times leading to an accuracy bias when evaluated within the framework of optimality (e.g., Balcı et al., 2011; Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Maddox, & Bohil, 1998). Thus, we also evaluated how far participants diverged from optimality (Balcı et al., 2011; Bogacz et al., 2006; Bogacz, Hu, Holmes, & Cohen, 2010) and within the same framework evaluated the level of subjective cost participants attributed to errors in the absence of objective penalties for errors (Balcı et al., 2011; Bogacz et al., 2006; Maddox, & Bohil, 1998).

To this end, the performance of the participants was evaluated with respect to the Optimal Performance Curve (OPC, Equation 1.1) that prescribes the reward rate maximizing relationship between the speed and accuracy of decisions given task parameters do not change within a block (Figure 1.1; dashed curve; Balcı et al., 2011; Bogacz et al., 2006).

$$\frac{DT}{Dtot} = \left(\frac{1}{ER \log \frac{1-ER}{ER}} + \frac{1}{1-2ER} \right)^{-1} \qquad \text{[Eq 1.1]}$$

where *DT* stands for decision time, the time it takes the decision-maker to gather enough information before making a decision and D_{tot} stands for T_0 + *RSI* where T_0 is the time it takes to detect the stimulus and implement the decision (e.g. non-decision time) and *RSI* is the response-to-stimulus interval.

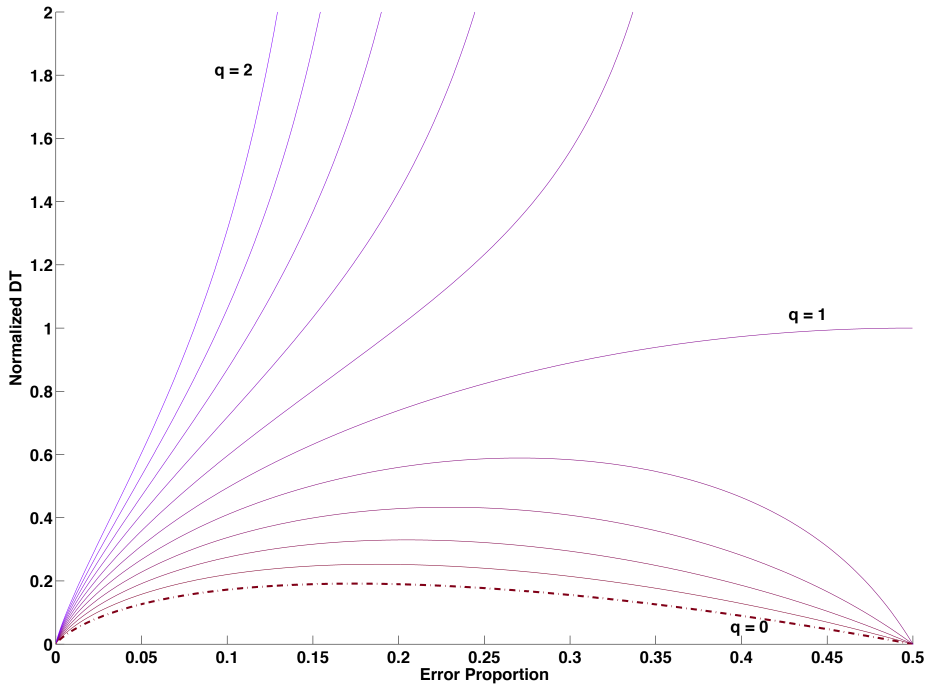

Figure.1.1 Family of Optimal Performance Curves (OPC) for q values ranging from 0 to 2.0 in increments of 0.2. The dashed line is OPC when q=0 (no penalty for errors). This figure was adapted from Balcı et al. (2011).

The OPC [Eq 1.1] yields the reward rate maximizing normalized decision times as a function of the error rates (ER). For example, when the error proportion is 0.5 (when the signal to noise ratio is 0) and when the error proportion is 0 (when the signal to noise ratio is 1), the normalized decision time for maximizing reward rate is virtually 0. For error rates between 0.5 and 0, the reward-rate maximization requires participants to set thresholds to higher values (accumulate more evidence) in order to optimally balance the number of trials (faster response times) and accuracy (Balcı et al., 2011; Bogacz et al., 2006; 2010).

The generalized form of the optimal performance curve that incorporates penalties for errors is defined by Equation 1.2 (Balcı et al., 2011; Bogacz et al, 2006). Equation 1.1 is a special case of Equation 1.2 for when the q (penalty for errors) is 0. The solid curves in Figure 1.1 are the optimal performance curves defined for different levels of objective penalties for errors. When the penalty for an error is above 0 (i.e., a negative gain), the evidence accumulation becomes relatively more important: Participants need to widen their decision boundaries and accumulate more evidence to respond correctly because an incorrect response has a non-zero cost.

$$\frac{DT}{Dtot} = (1+q)\left(\frac{\frac{1}{ER} - \frac{q}{1-ER}}{\log\frac{1-ER}{ER}} + \frac{1-q}{1-2ER}\right)^{-1} \qquad \text{[Eq 1.2]}$$

When there is no objective penalty for errors in the task (as in our task), Equation 1.2 can be used to estimate the relative value assigned to accuracy vs. speed or the subjective penalty of making errors. Penalties require caution because they seem to affect the OCD patients more than they do healthy controls. For example OCD patients compared to healthy controls, made more errors of commission in penalty conditions on a Go/Nogo task (Morein-Zamir et al., 2013). Since we incorporated no objective penalties, the best-fitting q parameter to the empirical data can indirectly show us how much penalty a participant subjectively assigned to an incorrect choice.

We hypothesize that those who rank high on OCD scales will deviate farther from optimality. They will gather more evidence than needed and respond slower, sacrificing from reward rate for accuracy. This hypothesis is in line with our prediction for higher threshold settings for participants with high OCD scale rankings: High threshold settings are indicative of more evidence accumulation and possibly an accuracy bias.

Finally we hypothesize that participants who rank higher on the OCD scales will have more pronounced post error slowing. In accordance with the hypothesis that participants with OC features will have an accuracy bias, we assume that after an error, these participants will be more likely to widen their decision boundaries to minimize further errors. This would be evidenced by a larger difference between pre and post error reaction times (Dutilh et al., 2012a).

Methods

Participants

Seventy four participants from Koç University undergraduates, graduate students, and staff (52 female, 22 male), aged between 18-38 years, were recruited via announcements online at Koç University's website (KUDaily). All participants were naïve to the purpose of the study. The announcement called for participants who "get stuck on details, control things more than needed, and worry often”. The experiment took 1.5 hours (single session). The study was approved by the Institutional Review Board for Human Subjects of Koç University and all participants signed a consent form.

Stimuli and apparatus

White moving dots (3x3 pixels) appeared in the center of a black computer screen within an approximately 3-inch diameter. Participants were asked to remain at a certain distance (~60 cm) from the screen. A portion of dots (12%) moved either to the right or left, randomized for each trial with equal probability while the rest of the dots moved in random directions. MATLAB was used to run the SnowDots framework (developed by Joshua Gold at the University of Pennsylvania) on a 21” LCD screen Mac desktop computer.

Procedure

1) Practice session: Each session started with a 2-min practice session in which 12% of dots moved either to the right or left and participants were asked to press the ‘M’ key if they thought the dots were moving to the right, and the ‘Z’ key if they thought the dots were moving to the left. Participants were asked to respond as quickly and as accurately as possible. Response to stimulus intervals (RSI) with a mean of 2 sec were drawn from a truncated exponential distribution with a lower bound of 1 sec. There were no rewards or penalties associated with choices made in the practice session; the correct decisions were followed by a ‘beep’ sound. No visual or auditory feedback was given after incorrect decisions. To avoid anticipatory responses, a buzzing sound followed all premature responses (RT < 100 ms) including the ones occurring before the stimulus onset. These responses were also followed by a 1-s timeout period.

2) Free response (FR) dot motion discrimination task: Different from the practice block, in the test blocks each correct response was awarded 0.02 TL (equivalent to one US cent). A single coherence of 12% was used throughout the experiment and the task lasted for 8 blocks of 5 minutes each, for a total of 40 minutes.

3) Signal detection (SD) task: Following the FR blocks, 2 blocks of 2 minute-long SD tasks were used to measure the time it takes the participant to detect the dots appearing on the screen and give a contingent response. The stimulus was the same as in FR blocks. However in one block, participants were asked to press 'M' and in the other block 'Z' as soon as they saw the stimulus appear on the screen. Each response was awarded 0.02 TL if it was not premature. Premature responses resulted in a buzzing sound and were penalized with a 1-s timeout period.

4) Scales: Following the completion of the random dot motion discrimination task, participants filled out the Maudsley Obsessive Compulsive Inventory (Hodgson & Rachman, 1977) in Turkish: Maudsley Obsesif Kompulsif Soru Listesi (Erol & Savaşır, 1988) and the 41 item Padua Inventory (Van Oppen, Hoekstra & Emmelkamp, 1995) in Turkish: Padua Envanteri (Beşiroğlu et al., 2005).

The Turkish version of the Padua Inventory revealed a test re-test reliability score of 0.91 for the entirety and 0.81-0.90 for the five subscales of the scale. The internal consistency is good with Cronbach's alpha score of .95 for the entirety and 0.79-0.92 for the five subscales of the scale (Beşiroğlu et al., 2005). The five subscales are: 1) Impulses, 2) Washing, 3) Checking, 4) Rumination, and 5) Precision. The factor analysis for the validity of the Turkish version revealed six factors that explained 62.1% of the total variance. The precision subscale was found to be divided into two factors of three questions in the Turkish form. Beşiroğlu et al. (2005)

concluded that overall, consistent factors with the PI-R are revealed and that reliability and validity of the Turkish version is adequate.

The Turkish version of the Maudsley scale has 37 items formed by adding 7 items to the original form and revealed a test-retest reliability score of .88 for the entirety and .59-.84 for the subscales of the scale (Erol & Savaşır, 1988). The internal consistency (Cronbach's alpha scores) for the entire scale was initially calculated to be .44, after which some items were reworded (for understandability and elimination of mistakes) and internal consistency score was recalculated to be .86. The internal consistency for the subscales revealed Cronbach's alpha scores of .70 for checking, .66 for washing, .31 for slowness and .56 for doubting (Erol & Savaşır, 1988). Rumination subscale has been added on with the additional 7 questions (in addition to the two from the original scale). The scale revealed three factors instead of the four factors in the original form (Erol & Savaşır, 1988). We excluded the slowness subscale from our exploratory analyses given its low internal consistency score ending up with: 1) Checking, 2) Washing, 3) Doubting and 4) Rumination subscales.

Both scales offer a composite score of how strong the OC tendencies are, however they also measure different aspects of OCD with distinct subscales (e.g. checking, precision). Despite the fact that we do not have specific predictions regarding the predictive value of the subscales, given the heterogeneity of OCD, we find it vital to gain information on all possible subscales because they might attest to differences in decision-making.

Data Analysis

One participant was removed from the analyses because he did not complete the scales. Overall there were 3 omitted questions in the Maudsley Scale (out of 2701) and 5 omitted questions in the Padua Scale (out of 2993). The omitted questions were substituted with the statistical mode of the corresponding participant's responses. A repeated measures ANOVA was conducted using test blocks as within-subject variables to characterize when error rates reached steady state. Pairwise comparisons revealed that error rates in the final block were significantly lower than error rates in the first three blocks only. Based on this empirical result, we used data from the last five blocks in the analyses assuming that perceptual learning was over and participants were closer to the steady state of decision-making.

The DDM (e.g. Ratcliff, 1978; Ratcliff & McKoon, 2008; Bogacz et al., 2006) was fit to the choices and response times (RT) (blocks 4-8) using the Diffusion Model Analysis Toolbox (DMAT) (Vandekerckhove & Tuerlinckx, 2008). Two different models with different levels of complexity were fit to the data. In the first model, we fit the data with the reduced form of DDM (pure DDM), allowing no inter-trial-variability in the parameters. The starting point was set to $a/2$, assuming no bias for either of the two choices. In the second model, we allowed inter-trial variability in drift rate, starting point, and non-decision time (extended DDM).

The deviation scores from the OPC per participant were calculated by subtracting the optimal normalized decision time of each participant (given his/her ER - Eq.1) from that participant's empirical normalized decision time. The q parameter (weight that is assigned to accuracy vs. speed) per participant was also estimated by finding the modified OPC (Eq. 2) that fit the empirical normalized DT for the empirical ER.

The RTs on post correct trials, which are at the same time pre error trials were subtracted from RTs on post error trials to quantify post-error slowing (hereon referred to as post error variable; Dutilh et al., 2012a). Post error RTs were compared with the post correct-pre error RTs with a paired samples t-test for those participants who had at least 20 errors in the last five blocks.

The DDM parameters (drift rate, threshold setting, non-decision time), deviation from the OPC, q values, post-error slowing variable, reaction times on the FR blocks, reaction times on the SD blocks, and the accuracy of choices on FR blocks were first regressed on the two total OCD scale scores (Padua and Maudsley) separately. Second, two sets (separate for Maudsley and Padua subscales) of exploratory stepwise multiple regressions were conducted to determine which subscales of Padua and Maudsley best predicted each of these outcome variables. Standardized coefficients were reported. The Pearson correlations among scale and subscale scores were calculated; Benjamini-Hochberg adjusted p values are reported for all correlations (Benjamini & Hochberg, 1995).

Results

The extended DDM outperformed the pure DDM (ΔAIC: 3281.08, ΔBIC: 3255.74) based on fits to the pooled data. Thus, the extended DDM was used in further analyses. The quantile probability plot for the extended DDM fits is presented in Figure 1.2 (note the misfit for the 90th percentile of the response times. A similar misfit was reported in Ratcliff & McKoon, 2008).

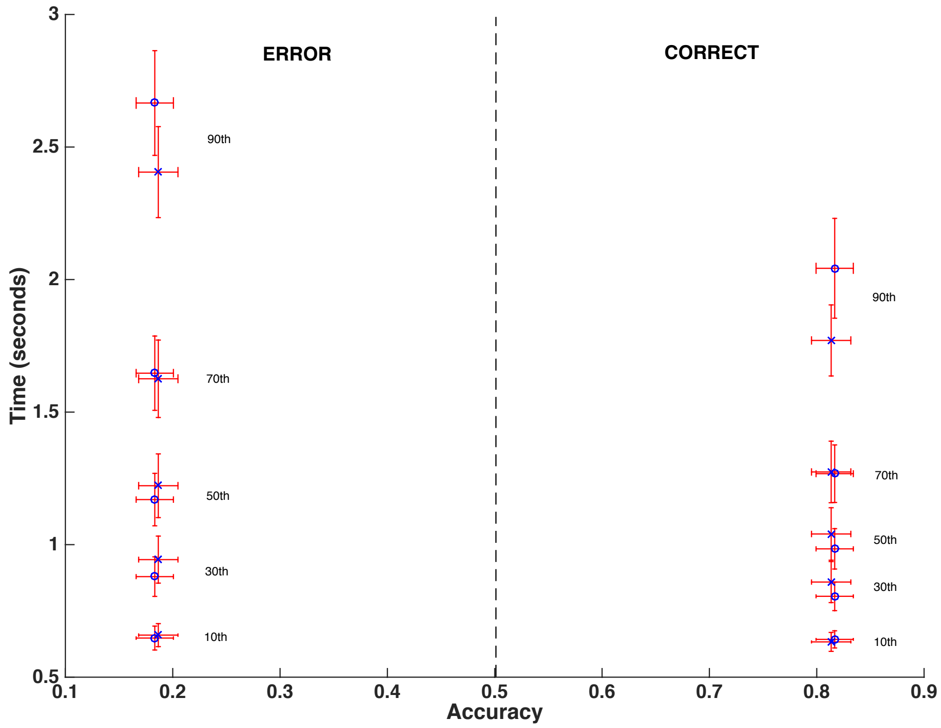

Figure.1.2 Quantile probability plots for the 10th, 30th, 50th, 70th, and 90th percentiles of response times (indicated in the figure). "o" marks the mean observed and "x" marks the mean predicted data. Errors bars indicate S.E.M. Errors are shown on the left and the correct responses are shown on the right. Only those participants who had more than 11 errors in the last four blocks are included in the figure *(N=61)*.

Three subscales were common to both Padua and Maudsley Scales: Washing, Checking and Rumination. The two Checking subscales (r=.63, p<.001), the two Rumination subscales (r=.83, p<.001), the two Washing subscales (r=.64, p<.001) and the total scores of Padua and Maudsley scales (r=.78, p<.001) correlated significantly (Benjamini-Hochberg adjusted). The means and ranges of the scale and subscale scores are presented in Table 1.1 along with the highest and lowest possible scores on these scales. The correlation matrix among all scale scores is presented in Table 1.2.

Table 1.1

Descriptive Statistics of Scale Scores

	N	*Minimum*	*Maximum*	*Mean*	*SE*	*SD*
Padua Total	73	4(0)	116(164)	49.81	2.52	21.54
Padua Washing	73	1(0)	34(40)	12.40	.81	6.90
Padua Rumination	73	0(0)	38(40)	16.07	1.03	8.81
Padua Checking	73	1(0)	23(32)	9.67	.65	5.59
Padua Impulses	73	0(0)	24(28)	7.12	.63	5.40
Padua Precision	73	0(0)	16(24)	4.55	.43	3.70
Maudsley Total	73	2(0)	30(37)	15.11	.70	5.95
Maudsley Washing	73	0(0)	9(11)	3.68	.28	2.36
Maudsley Rumination	73	0(0)	9(9)	4.25	.32	2.73
Maudsley Checking	73	0(0)	9(9)	3.36	.24	2.03
Maudsley Doubting	73	1(0)	7(7)	3.62	.17	1.43

Note. Values in parentheses indicate the minimum and maximum possible values that can be attained in the corresponding scales and sub-scales.

Table 1.2

Pearson Correlation Scores Among the Scale and Sub-Scale Score

	1	2	3	4	5	6	7	8	9	10
1. P. Total										
2. P. Washing	.624**									
3. P. Rumination	.830**	.258*								
4. P. Impulses	.599**	.190	.460**							
5. P. Checking	.753**	.310*	.585**	.243*						
6. P. Precision	.667**	.409**	.416**	.209	.544**					
7. M. Total	.780**	.411**	.732**	.461**	.606**	.442**				
8. M. Checking	.641**	.168	.657**	.349**	.631**	.388**	.817**			
9. M. Washing	.369**	.644**	.088	.118	.190	.280*	.529**	.213		
10. M. Doubting	.517**	.172	.538**	.409**	.335**	.308*	.710**	.502**	.166	
11. M. Rumination	.659**	.158	.833**	.522**	.389**	.205	.772**	.674**	.066	.537**

Note. M stands for Maudsley, P stands for Padua. P values are Benjamini-Hochberg adjusted.
*$p< 0.05$
** $p<0.01$

The regression analyses revealed that OCD scale scores are not significant predictors of response times in FR and SD blocks, or error rates. This indicates that error rates or response times alone did not distinguish between high and low scorers on these scales. However, as expected this was not the case for the analyses conducted on the outputs of the model fits. Both the Maudsley and Padua total scores significantly predicted threshold settings, (β= .30, $t(71) = 2.68$, p=.01) and (β= .24, $t(71) = 2.10$, p=.04), respectively. These findings revealed that an increase in the entirety of OC symptoms, measured by both scales, indicates an increase in cautiousness.

Given that the overall OC scores explained a significant proportion of variance in the threshold settings, we explored the contribution of specific subscales to this parameter. A stepwise multiple regression analysis was conducted to reveal which

Maudsley and Padua subscale scores (conducted separately for each scale) contributed to this relation. Only Maudsley Checking subscale (β= .26, $t(71)$ = 2.29, p=.03) significantly predicted the threshold settings whereas it was only the Padua Rumination subscale that predicted the threshold settings (β= .27, $t(71)$ = 2.40, p=.02). Briefly, increases in checking behavior and ruminative thinking were indicative of increases in cautious decision-making.

Although the simple linear regression analysis with Padua and Maudsley total scores did not reveal a significant relationship with other core DDM parameters (i.e., drift-rate and non-decision time), our exploratory stepwise multiple regression scores revealed that the Padua Washing subscale significantly predicted faster non-decision related delays in responding (β= -.24, $t(71)$ = -2.11, p=.04). This finding indicates that an increase in washing related behaviors signify a fastening in detection of the decision-relevant signal and/or behavioral manifestation of a decision made. Simple linear regressions, conducted for Padua and Maudsley total scores separately (as well as the stepwise multiple regressions conducted for the two sets of subscales separately) did not reveal any significant predictors of the deviations from OPC or q parameters.

Post error slowing was evident in the data calculated for those participants who had at least 20 errors in the last five blocks (N=55). Decisions made after incorrect decisions (M=1.36, SD= .80) took significantly longer than decisions made after correct decisions that were at the same time pre-error trials (M=1.19, SD= .73), $t(54)$=5.96, p<.001, however the scores on scales or subscales did not predict the degree of post error slowing.

Discussion

We investigated the differences in decision-making processes of non-clinical participants with various rankings on the scale and subscales of OCD within the framework of DDM. This constituted the first-time study of an analogue OC sample in a DDM and optimality framework. We expected participants who rank high on OCD scales to have widened decision boundaries (in the absence of differences in drift rates), deviate farther from optimality (calculated by deviations from OPC), and weigh accuracy of decisions over their speed (denoted by higher q values). Our expectations are in line with another subclinical OC study, in which even in a mildly uncertain situation, those who ranked higher on the OCD scales engaged in more checking behavior at the cost of time (Toffolo, van den Hout, Hooge, Engelhard, & Cath, 2013). Our predictions regarding the threshold setting held with the two OCD scales' total scores. Total OC symptom scores were found to predict more cautious decision strategies (widening of the decision boundaries). These findings were also in line with and in support of the recent findings with the OCD patients (Banca et al., 2015).

We also conducted exploratory analyses regarding the relationship between subscales of Maudsley and Padua and the DDM parameters. Increases in checking and rumination scores were found to predict more cautious decision strategies. Furthermore, to our surprise increases in washing related behaviors were found to predict faster non-decision related durations (e.g. encoding of stimuli, motor preparation, or both).

Checking and rumination are integral parts of OCD and therefore their relationship with threshold settings is not surprising. The perceptual evidence regarding task completion might not be enough for OCD patients to make/finalize a decision and move on in a timely manner (Sachdev & Malhi, 2005). Consequently, OCD patients might tend to check for more evidence to make sure that their decisions are accurate. Moreover, both checking (Wu & Cortesi, 2009) and rumination (Flett, Madorsky, Hewitt, & Heisel, 2002) are positively correlated with perfectionism. Wanting to reach at perfect outcomes, participants with these features may indeed choose to acquire more evidence before making decisions.

Moreover, the neural correlates of rumination and checking overlap with those of threshold settings. The studies reviewed by Nolen-Hoeksema, Wisco, and Lyubomirsky (2008) point at the ACC as one of the possible neural correlates of rumination. Besides, Mataix-Cols et al. (2004) conclude that when checking symptoms are provoked, OCD patients had higher activations than controls in regions including subthalamic and brainstem nuclei, putamen and globus pallidus, and dorsal cortical areas (including right anterior cingulate). Both the ACC and some of the areas activated in checking are important in the speed accuracy tradeoff, namely in the adjustment of the threshold setting. For instance, a recent review of the neural correlates of perceptual decision-making (Mulder, Van Maanen, & Forstmann, 2014) indicated that threshold level is related to the activations of the pre-SMA, ACC, and the striatal regions. Mulder et al. (2014) discussed that the frontal regions such as the ACC modulate the thresholds through the striatum with the possible contribution of subthalamic nucleus (STN). While a higher activation of striatum is related to speedier responses, higher activation of STN is thought to inhibit fast responding, bringing about caution. Although more research is warranted to understand the interplay between the parts of the basal ganglia and their relationships with the speed accuracy tradeoff (Mulder et al., 2014), given the current findings both checking and rumination seem to share neural correlates with threshold settings.

On the other hand, research into non-decision time is relatively scarce (Mulder et al., 2014). The reason for such scarcity is thought to be the over-general nature of the measure/term; encompassing all process durations within a task that are not stated as decision making times (Mulder et al., 2014). In a clinical study, Karalunas and Huang-Pollock (2013) found that ADHD participants had faster non-decision times than healthy controls. In the current study we also found that washing subscale in the Padua Inventory predicted faster non-decision times. Faster non-decision times can attest to more efficient encoding of stimuli or motor preparation (Karalunas & Huang-Pollock, 2013) and better/faster preparation between a cue and a target (Karayanidis et al., 2010). However it is also possible that the short duration spent in non-decision time can be due to rushing to decision making without sufficient encoding or motor preparation (Karalunas & Huang-Pollock, 2013) possibly ending up hurting the performance.

In our experiment, performances of those who rank higher on the washing subscale were not negatively affected in terms of decision outputs, therefore we can gauge that faster non-decision time here points to more efficient encoding and/or motor function of washers. The reasons for such a relationship can be investigated again by studying the possible neurobiological underpinnings of washing behavior. Mataix-Cols et al. (2004) has shown that the Padua-revised washing subscale scores were positively correlated with activations in several brain areas including the ventrolateral prefrontal cortex (VLPFC), anterior insula and bilateral visual regions. The visual cortex is engaged in early encoding of visual task cues whereas (starting with left intraparietal sulcus) left posterior VLPFC and left anterior VLPFC are engaged in encoding of task rules (Bode & Haynes, 2009). Furthermore, the activation of anterior insula (along with other regions) was at baseline during evidence accumulation, peaking at stimulus recognition (Ploran et al.,2007), which Ploran et al. (2007) concluded is not explicitly related with the decision making process. Recognizing a stimulus and encoding it are parts of non-decision time and they are associated with brain areas that exhibit a correlation with the washing subscale (Mataix-Cols et al., 2004). With these regions activated, there might be more efficient encoding of stimuli and preparation and therefore less time spent on the non-decision parameter.

Banca et al. (2015) conducted a study with OCD patients, which was highly relevant to the current study. Despite our findings regarding thresholds that corroborated their findings, they also found a lower evidence accumulation quality (drift rate) for OCD patients in the medium coherence levels (.15 and .25; similar to our coherence level of .12). We did not find any significant relationship between OC features and the drift rate. It is possible that a lowering in drift rate, an inability to accumulate evidence efficiently from the data, is a distinguishing factor between clinical and subclinical OC individuals. This difference needs to be addressed in future studies.

We followed up on the diffusion model fit-based analysis discussed above with the evaluation of decision-making performance within the framework of optimality. Our analyses did not reveal any significant relationship between deviations from optimality and scores on the OCD scales. Thus, despite OC features predicting higher decision thresholds, these differences did not lead to significant deviations from optimality. Further studies could consider longer training as they might lead to more sensitive outcomes (e.g., Balcı et al., 2011).

Our exploratory analyses revealed that only specific subscales have meaningful associations with DDM constructs. This finding is in line with the heterogeneity of OCD (Mataix-Cols et al., 2004) and with research showing that distruptions of decision making processes in OCD may vary depending on the subtype of the disorder (e.g. Lawrence et al., 2006). Several dimensions of OCD have been identified: a) symmetry/ordering, b) contamination/cleaning, and c) obsessions/checking (Mataix-Cols, do Rosario-Campos, & Leckman, 2005 – in DSM V hoarding is considered as a disorder separate from OCD (American Psychiatric Association, 2013)). Moreover, different dimensions (e.g. checking, washing) are found to be associated with differential activations in fronto-striato-thalamic circuits (Mataix-Cols et al., 2004). The results of the current study relate to the obsessions/checking and contamination/cleaning features of OCD, with no indications for the other dimensions. These differential results are important because a treatment of the sample as a whole (as high or low scorers on the entirety of the scales) would not be able to reach these specific findings. However, it is important to note that we did not have a priori predictions regarding the relations with the subscales and these findings were gathered from our exploratory stepwise regression analyses.

This study, despite its relatively large sample size has a few limitations. Although we have used two psychological scales to make sure we captured most of the variability in the OC trait information we might have still missed some of the variability. Moreover, we studied a non-clinical population, and did not use other clinical scales to screen for psychiatric disorders. Finally, since the study is conducted with a non-clinical sample, we cannot draw inferences for the clinical OCD population.

Although appearing as a limitation, studies conducted with subclinical OC samples, also called ‘analogue’ samples are important resources to better understand the nature of OCD (Abramovitch, Shaham, Levin, Bar-Hen, & Schweiger, 2015a.;

Gibbs, 1996; for a recent review: Abramowitz et al., 2014). OCD is argued to be a dimensional and not a categorical disorder (Abramowitz et al., 2014). The severity of the symptoms ranges from absent to severe on a continuum. Although different in severity, qualitatively the symptoms are mostly comparable in diagnosed OCD patients and non-clinical individuals with OC symptoms (Abramowitz et al., 2014). Used in cognitive studies (e.g. Abramovitch et al., 2015a; Mataix-Cols, 2003), subclinical OC samples have the advantage of capturing the subgroups of the OC symptoms (e.g. checking, precision etc.) (Abramowitz et al., 2014) while for the most part not having medication (Mataix-Cols, 2003) and comorbid conditions as confounds.

Conclusions

Our results overall pointed at differences between people with different degrees of OC features in terms of decision threshold settings. Our exploratory analyses further revealed differences in both controlled (voluntary and requiring attention; threshold settings) and automatic (faster and do not require attention; non-decision time) decision processes (Cohen, Dunbar, & McClelland, 1990). To our knowledge, no studies have been conducted with an analogue OC sample within the decision-theoretic framework of DDM or optimality. Adopting a decision-theoretic approach such as DDM might be beneficial in revealing differences both between clinical disorders as well as between subclinical and clinical populations.

CHAPTER II

Disrupted latent decision processes in medication-free pediatric OCD patients

Abstract

Background**:** Decision-making in Obsessive Compulsive Disorder has typically been investigated in the adult population. Computational approaches have recently started to get integrated into these studies. However, decision-making research in pediatric OCD populations is scarce.

Methods**:** We investigated latent decision processes in 21 medication-free pediatric OCD patients and 23 healthy control participants. We hypothesized that OCD patients would be more cautious and less efficient in evidence accumulation than controls in a two alternative forced choice (2AFC) task.

Results**:** Pediatric OCD patients were less efficient than controls in accumulating perceptual evidence and showed a tendency to be more cautious. In comparison to post-correct decisions, OCD patients increased decision thresholds after erroneous decisions, whereas healthy controls decreased decision thresholds. These changes were coupled with weaker evidence accumulation after errors in both groups.

Limitations**:** The small sample size limited the power of the study.

Conclusions**:** Our results demonstrate poorer decision-making performance in pediatric OCD patients at the level of latent processes, specifically in terms of evidence accumulation.

Keywords**:** decision making; Hierarchical Drift Diffusion Model; Obsessive Compulsive Disorder; children

Introduction

OCD is a debilitating psychiatric condition with the age of onset spanning a range from early childhood to adulthood (Pauls, Abramovitch, Rauch, & Geller, 2014). OCD affects 1-3% of the pediatric/adolescent population (Valleni-Basile et al., 1994; Pauls, Alsobrook, Goodman, Rasmussen, Leckman, 1995; Apter et al., 1996). The condition includes either or both obsessions and compulsions (American Psychiatric Association, 2013), which significantly reduce the quality of the patients' lives.

A number of previous studies have investigated decision-making in adult OCD patients. However there is less research conducted with pediatric OCD populations. The Iowa Gambling Task (IGT; Bechara et al., 1994) has been used in many studies with adult OCD patients. Some studies revealed that OCD patients made more disadvantageous choices (e.g. Cavedini et al., 2002; Rocha et al., 2011; Starcke et al., 2010), however other studies revealed comparable performance to healthy controls (e.g. Nielen, Veltman, De Jong, Mulder, & Den Boer, 2002; Lawrence et al., 2006). In the only IGT study conducted with pediatric OCD patients (n_{OCD}=22; $n_{Control}$=22), Kodaira et al. (2012) found more disadvantageous responding of participants with OCD on the last block of testing and suggested that pediatric OCD patients had impaired decision-making. It is worthy to note that recruiting medication-free OCD patients is challenging and that some patients in the above mentioned studies were on psychiatric medications at the time of the experiment (e.g. Kodaira et al., 2012; Lawrence et al., 2006; Rocha et al., 2011; Starcke et al., 2010).

Much neuropsychological research has been undertaken in adult OCD groups but with highly divergent outcomes (Abramovitch, Abramowitz, & Mittelman, 2013a). A recent meta-analysis with 115 studies revealed an average moderate effect size across domains denoting worse performance for OCD patients, which the authors conclude might not allude to clinical significance (Abramovitch et al., 2013a). There is much less neuropsychological research conducted with pediatric OCD patients. A recent meta-analysis compiling 11 studies investigating executive function, memory, processing speed, visuospatial abilities, and working memory has concluded that there is no evidence for neuropsychological dysfunction in pediatric OCD populations (Abramovitch et al., 2015b). This meta-analysis did find a trend for worse performance in neuropsychological tasks for pediatric OCD patients compared to healthy controls, but the effect sizes were neither statistically nor clinically significant. The authors attributed the lack of statistical significance to the number of available studies and sample sizes across studies, and pointed to a need for further research. Importantly, decision making tasks were not included in the meta-analysis.

In their critical review, Abramovitch and Cooperman (2015) argue that neuropsychological tests, although informative for the psychiatric area, can be improved with some modifications. For instance, because many neuropsychological studies employ commonly used classic experimental procedures and analyses rather than venturing to new methods, the conclusions may become restricted and thereby uninformative. With different approaches in analyses and changes in the established neuropsychological tasks (e.g. adding distractors or manipulating the task load) (Abramovitch & Cooperman, 2015), the area can benefit from more in-depth characterization of behavior.

In the area of computational psychiatry, researchers are also striving to come up with more in depth analyses of psychiatric problems and shift from a symptom-based descriptive understanding of psychiatric disorders to descriptions involving "objective computational multidimensional functional variables" (Wiecki et al., 2015, p. 378). To this end, Wiecki et al. (2015) point to sequential sampling models as important tools for the field of psychiatry. Drift diffusion model (DDM) is a prominent sequential sampling model, which utilizes a combination of accuracy and reaction time data to explain latent decision-making processes (Ratcliff, 1978; Ratcliff & Rouder, 1998; Ratcliff & McKoon, 2008). Through these processes, the model can provide psychological explanations (such as cautious responding, non-decision time, biases in decision-making and evidence accumulation efficiency) to differences in choice behavior (White et al., 2010a).

The DDM assumes that in a decision-making task with two choices, the agent starts at a point (starting point: *z*; initial belief state) between the two alternatives and accumulates evidence from the noisy signal with some rate (drift rate: *v*). As the agent gathers enough evidence to reach one of the thresholds, the corresponding decision is made. The area between the thresholds associated with two alternatives is referred to as the boundary separation (e.g. Ratcliff & McKoon, 2008). The core parameters of DDM are threshold setting (*a*), drift rate (*v*), starting point (*z*), and non-decision time (*Ter*). The more complex version of the model (extended model; Ratcliff & Rouder, 1998) includes variabilities in non-decision time (*St*), drift rate (*eta*), and starting point (*Sz*). Threshold setting indexes speed accuracy tradeoff or the caution with which the decision is made; the higher the threshold setting the more caution the decision maker exercises. Drift rate indexes the rate of evidence accumulation or signal to noise ratio. The starting point indexes the bias towards either of the two choices and the non-decision time indexes the duration of signal detection or motor response (e.g. Ratcliff & McKoon, 2008; White et al., 2010a).

White et al. (2010a) have pointed out the benefits of using DDM in clinical research. In support of this argument, the DDM has indeed been successfully used in studies with populations suffering from ADHD (e.g. Karalunas & Huang-Pollock, 2013; Metin et al., 2013; Mulder et al., 2010), anxiety (White et al., 2010b), depression (Pe et al., 2013), and clinical (Banca et al., 2015) and subclinical (Erhan & Balcı, 2016) OCD.

Banca et al. (2015) used the dot motion discrimination task with three different levels of signal to noise ratios (SNRs) to study decision-making behaviors of mostly medicated adult OCD patients (n_{OCD}=28; $n_{Control}$=35). Monetary rewards and punishments indicated the correct and incorrect responses. Performances on low (coherences .025 and .05), medium (coherences .15 and .25), and high (coherences .45 and .7) levels of SNRs were compared using Hierarchical Drift Diffusion Model (HDDM). The findings revealed higher threshold settings for OCD patients than healthy controls at low and medium SNR levels and lower drift rates than healthy controls in medium and high SNR levels. In other words, OCD patients responded in a more cautious manner and gathered more evidence than controls in lower SNR but accumulated evidence less efficiently in higher SNR scenarios.

Erhan and Balcı (2016) also used the dot motion discrimination task but with a single coherence level (12%) and with healthy adult participants (*N*=74) who rank on various levels on OCD scales. Their findings revealed that increases in rumination and checking tendencies as well as an increase in the entire OC score predicted higher threshold settings. Differing from the clinical OCD study (Banca et al., 2015) subclinical OC traits did not predict drift rates. Authors concluded that a low drift rate could be a signature for clinical OCD populations.

In the current study, we seek to understand the latent decision variables of a medication-free pediatric OCD population (ages 9-16). Even though symptom dimensions of OCD are alike across age groups, pediatric and adult OCD populations seem to have abnormal neural activations in similar brain locations but in reversed directions (Gilbert et al., 2009). Abramovitch, Mittelman, Henin and Geller (2012), in their review on neuroimaging in pediatric OCD, also argued that adult OCD and pediatric OCD can be distinct and that neurodevelopmental factors such as pruning and myelination make it more difficult to pinpoint a common neurobiological basis for OCD across ages.

We sought to fill the empirical gap in the literature regarding the study of decision-making in the pediatric OCD group at the level of latent processes. We hypothesized that pediatric OCD patients, similar to adult OCD patients (Banca et al., 2015) would set higher decision thresholds than control participants due to the checking and doubting nature of the disorder. Based on the argument that sensory-perceptual evidence, which lets most people make rapid decisions, is not sufficient for patients with OCD (Sachdev and Malhi, 2005), we hypothesized that these patients would need higher amounts of evidence before making a decision. We also predicted lower drift rates on the part of OCD patients as an alternative basis for decision-making deficits associated with this group. As an extension of these predictions, we also predicted OCD patients to set higher thresholds after errors compared to after correct responses.

Methods

Participants

Fifty-three participants partook in the study. Participants with OCD diagnosis (n=21) were recruited from the pediatric clinic of a public psychiatric hospital, İstanbul Erenköy Psychiatric Training and Research Hospital. The participants in the control group (n=32) were volunteers from a public school in a nearby neighborhood. The study was approved by the Koc University and Erenköy Psychiatric Training and Research Hospital Ethical Review Boards and related permissions were obtained from the Istanbul Provincial Directorate of National Education Board. All parents signed informed consent forms and all participants gave assent to partake in the study. Subjects were compensated by a fixed amount for participation-related expenses (e.g. travel).

The exclusion criteria for both groups were intellectual disability, major neurological disorders, and use of psychiatric medication within the last 6 months. The inclusion criterion for the clinical participants was an OCD diagnosis by the primary psychiatrist with no comorbid psychiatric disorders. Patients participated in the study immediately after the initial diagnosis without having started medication, thus, their treatment schedule was not delayed.

Inclusion criteria for control participants was the absence of current clinical diagnosis, assessed by The Development and Well-Being Assessment (DAWBA) (Goodman, Ford, Richards, Gatward, & Meltzer, 2000; Dursun et al., 2013) (i.e. less than 3 points in all computer assigned diagnostic criteria). For stringency, presence of any psychiatric symptom was a reason for exclusion from the control group, even if the symptoms did not amount to a full clinical diagnosis. The symptoms for exclusion were ascertained by the psychiatrists in our group, who personally assessed the DAWBA results in addition to the computerized system. Seven participants were removed for having psychiatric symptoms (note that the inclusion of these participants in the control group, overall, led to similar findings). An additional two participants were removed because their parents did not complete the DAWBA. The final control group was composed of 23 participants.

Procedure

All participants completed the Dot Motion Discrimination Task and the block design and vocabulary subsections of Wechsler Intelligence Scale for Children Revised, in Turkish (WISC-R; Wechsler, 1974; Savasir & Sahin, 1995). The children in the clinical group were also administered the CY-BOCS (Scahill et al., 1997; Yucelen, Rodopman-Arman, Topcuoglu, Yazgan, & Fisek, 2006). All parents in both patient and control groups filled the socio economic status forms and completed the DAWBA Interview for Parents with the researcher's instructions. Participants were also provided passwords for DAWBA and were asked to complete it in their own time.

DAWBA is a valid computerized diagnostic package formed of various scales and open-ended questions (Goodman et al., 2000). The Turkish translation and validation of this tool has been undertaken by Dursun et al. (2013). The interrater reliability score in Dursun et al. (2013) was shown to be excellent and the validity score to be good to excellent. CY-BOCS is a valid and reliable 10-item semi-structured clinical interview to measure the severity of OCD symptoms in children and adolescents (Scahill et al., 1997). The Turkish version demonstrates good interrater reliability scores and researchers report that the translated CY-BOCS can be used in clinical research settings (Yucelen et al., 2006).

Dot Motion Discrimination Task

Dot motion discrimination task (DMDT) is a commonly used visual perceptual decision-making task with several adjustable parameters. In the current study, white dots appeared in a circular space with a diameter of 3 inches in the middle of a black screen. While some percentage of the dots (12%) moved cohesively to either right or left, the rest of the dots were displaced randomly within the circular space. Response to stimulus intervals (*M*=2 sec) were sampled from a truncated exponential distribution with a 1 sec lower bound. The task was to identify to which direction the dots were moving and press the corresponding keys. The correct responses were followed by a beep sound and were worth one point each. The incorrect responses neither had feedback nor penalty. On every 10th response total points earned appeared on the screen. The experiment was run on MATLAB, using the SnowDots (2012) developed at the University of Pennsylvania by Joshua Gold.

Each session was formed of a practice session of 2 minutes, followed by a free response session of 24 minutes (8 blocks of 3 minutes) and a signal detection session of 3 minutes. Different than the first two sessions, in the signal detection session, participants were asked to press the corresponding keys immediately as they saw the dots emerge on the screen, with no consideration for direction of movement. In all sessions a buzzing sound followed all anticipatory responses (RT < 100 ms), which were penalized with 1 sec timeout. Participants were told to accumulate as many points as possible.

Data Analysis

We used both Bayesian and frequentist independent samples t-tests to check if the clinical and control group were comparable with regards to age and IQ levels. The difference between accuracy, reaction times and signal detection times between the clinical and control group were also assessed by independent samples t-tests.

HDDM (Wiecki et al., 2013) was fit to the response times (RT) and choices (correct-incorrect) using the software developed by Wiecki et al. (2013) (http://ski.clps.brown.edu/hddm_docs/). Our model allowed for changes in both drift rate and threshold setting. On all models we drew 5000 posterior samples using the Markov Chain Monte Carlo (MCMC) algorithm discarding first 20 as burn-in. Wiecki et al. (2013) suggest mixture models, where a certain percentage of trials are outliers that come from a uniform distribution and are generated by processes other than the DDM. We set the outlier ratio to be 1% for the between-subject models and 5% for the within-subject models. To assess model convergence, we visually observed the trace, the autocorrelation, and the marginal posterior. To assess individual fit qualities we visually inspected the posterior predictions. To test our hypotheses for the models, we compared the posterior probabilities of parameters, which leads to a Bayesian probability measure (*P*). The P value is the probability that one variable's estimate is larger than that of the other, based on their posterior probability distributions.

We used paired samples t-tests to compare post-error and post-correct trials in terms of reaction times (PES; post error slowing) as well as accuracy. A within-subjects HDDM was fit to post error and post correct responses for each group (e.g. Dutilh et al., 2012b), allowing both drift rate and threshold setting to vary.

Results

The ages were matched for the clinical (M =12.00, SD =1.90) and control (M=12.46, SD=1.14) groups (t(32.15)=0.98, p=.34, BF_{01}=2.25). The vocabulary (v) and block design (b) subtests of the WISC-R were also matched for the clinical (M_v=11.55, SD_v=2.09; M_b=12.00, SD_b=3.42) and control (M_v=12.00, SD_v= 1.41; M_b=13.61, SD_b=3.43) groups (t(32.68)=0.81, p_v=.42, BF_{01}=2.51; t(42)=1.56, p_b=.13, BF_{01}=1.28). In the clinical group one participant could not complete the vocabulary subtest due to time constraints. Of the included participants, all parents completed DAWBAs. In the clinical group out of 21 participants, 12 completed and 4 partially completed the DAWBAs. In the control group out of 30 participants, 16 completed and 1 partially completed the DAWBAs.

Paired samples t-test scores revealed that both the control and clinical group had post-error slowing. In the control group, participants responded faster after correct (M=1.44 sec, SD= 0.49 sec) than after incorrect trials (M=1.65 sec, SD=0.77 sec), $t(22)=2.52$, $p=.02$. Same was true for the clinical group such that participants responded faster after correct trials (M=1.78 sec, SD=0.64 sec) than they did after incorrect trials (M=2.24 sec, SD=1.08 sec), $t(20)=2.32$, $p=.03$.

Paired samples t-tests also revealed that accuracy after correct responses was higher than accuracy after error responses for both groups. In the clinical group post correct accuracy (M=.74, SD=.10) was higher than post error accuracy (M=.62, SD=.08), $t(20)=7.78$, $p<.001$. In the control group as well, post correct accuracy (M=.79, SD=.13) was higher than post error accuracy (M=.71, SD=.12), $t(22)=6.25$, $p<.001$.

Overall the control group (M=1.48, SD=0.55) had faster reaction times than the clinical (M=1.91, SD=0.72) group, $t(42)= 2.26$, $p=0.03$, $BF_{01}=0.46$. The clinical (M=.71, SD=.10) and control (M=.78, SD=.14) groups did not differ in their accuracy ($t(42)=1.82$, $p=0.08$, $BF_{01}=0.91$). Signal detection times for the control (M=0.40 sec, SD=0.06 sec) and clinical (M=0.41 sec, SD=0.10 sec) groups did not differ ($t(42)=0.45$, $p=.65$, $BF_{01}=3.09$).

HDDM analyses comparing the two groups revealed that OCD patients had lower drift rates than controls (P=.99). Moreover, a tendency for higher threshold settings was observed in the clinical group (P=.93). The within subject HDDM analyses revealed that the post error responses had lower drift rates than the post correct responses for both the clinical (P=1.0) and control group (P=1.0). However a difference between the groups emerged in terms of post-error threshold setting. The post error responses had higher threshold settings than the post correct responses for the clinical group (*P*=1.0) whereas the post error responses had lower threshold settings than the post correct responses for the control group, (*P*=1.0).

Discussion

We investigated the latent decision-making processes of a pediatric OCD sample in comparison to a healthy control sample in a two alternative forced choice task using HDDM. We hypothesized that OCD patients would have higher threshold settings and lower drift rates than control participants. We also predicted that OCD patients would set higher thresholds after errors compared to after correct responses. Our findings provided support for our predictions: OCD patients displayed lower drift rates and showed a strong tendency for higher threshold settings than healthy controls. The patients may have compensated their lower evidence accumulation efficiency (drift rate) with increased caution, possibly explaining their slower RTs but comparable error rates with the healthy control group.

These results are in support of the findings of the recent clinical OCD study (Banca et al., 2015) and in partial support of the findings of the subclinical OC study (Erhan & Balcı, 2016). The subclinical OC study found that higher total OC scores and higher rumination and checking scores predicted higher threshold settings, with no relationship between OC scores and drift rates. The clinical study (Banca et al., 2015) on the other hand found that patients with OCD had both higher threshold settings and lower drift rates than the healthy controls in medium SNR conditions (which best represents the SNR used here). In the current study, we found lower drift rates for OCD patients with a tendency for higher threshold settings. Taken together these findings indicate that lower drift rates might be a signature of clinical OCD as was originally suggested by Erhan and Balcı (2016).

The threshold setting differences in the pediatric OCD group vs. the healthy controls are not as distinct as those reported by Banca et al. (2015) for adults. According to recent approaches that rely on signal detection theory, it is reasonable to be more cautious when signal detection capacity is low (akin to walking slower and more cautiously (e.g. higher threshold setting) in a dimly lit room (e.g. low drift rate)) (Lynn & Barrett, 2014). One reason why the pediatric population does not show as increased a caution despite low drift rates, as adults do, could be the differences in used coping responses possibly due to age (Aldwin, 1994). Overall, people with high OC traits have low distress tolerance (e.g. Blakey, Jacoby, Reuman, & Abramowitz, 2015); however adults might have compensated with better emotion regulation and coping strategies. We did not originally incorporate emotion regulation or distress tolerance into our interpretation of the decision-making behavior. However, emotions have been argued to effect decisions strongly (Lerner, Li, Valdesolo, & Kassam, 2015), and changing emotional reactions through emotion regulation strategies also change decisions (Phelps, Lempert, & Sokol-Hessner, 2014). Future studies can focus on the relations of emotion regulation with latent decision variables.

Our findings regarding the post correct and post error behaviors indicate that both groups have lower drift rates while responding after errors than when they do after correct trials. Moreover, in line with our hypothesis, OCD patients have higher threshold settings responding after errors than after correct trials. Interestingly, this pattern is reversed in the healthy control participants for whom threshold settings are lower in the post error trials than post correct trials. Our findings are very similar to those of White et al. (2010a)'s study, in which participants with high and low anxiety scores performed a recognition memory task. Akin to our study, highly anxious participants had significantly increased threshold settings after errors whereas participants with lower anxiety scores had reduced (although non-significant) post error thresholds. White et al. (2010a) also found that after errors, discriminability (difference in drift rate of familiar and novel stimuli) decreased for both groups. Although they worked with students with high and low anxiety scores, our findings with a clinical pediatric population are highly comparable.

Results in our study are however different than Dutilh et al. (2012b)'s findings, which suggest that post error slowing almost solely results from an increase in threshold settings after errors. What we find across groups, a decrease in drift rates, point to a dampened evidence accumulation efficiency, possibly due to distraction (Dutilh et al, 2012b). This difference in the explanation of post error slowing at the level of latent processes could be due to age, and can benefit from more research.

The symptomatology of OCD as put forth in DSM-V (American Psychiatric Association, 2013) matches with the observed latent variables. Time-consuming obsessions, compulsions, and efforts to suppress them behaviorally manifest as inattention and distraction in individuals with OCD (Abramovitch, Dar, Mittelman, & Schweiger, 2013b). This is also what our outcome in terms of latent variables show. The pediatric OCD group was found to be less efficient in accumulating evidence compared to healthy controls, possibly indicating that they were distracted and less attentive to the stimuli. Although OCD patients would also be expected to require more evidence before making a decision (higher threshold) based on DSM-V's characterization of this disorder (e.g. checking), albeit strong, this was only a trend in the data.

Moreover, the proposed neural correlates of drift rate and threshold settings match with the proposed neural correlates of OCD. The frontoparietal (e.g. de Vries et al., 2014; Melloni et al., 2012) and corticostriatal pathways (e.g. Burguiere, Monteiro, Mallet, Feng, & Graybiel, 2015) are implicated in OCD. Similar neurobiological mechanisms have been implicated for drift rate and decision threshold. For instance, in a recent review Mulder, Van Maanen, and Forstmann (2014) concluded that threshold settings are associated with activations in the frontobasal ganglia network and changes in drift rates are associated with activations in frontoparietal network.

Limitations

The relatively low number of participants is a limitation of this study, preventing investigation of within-group effects due to factors such as age. In order to assure a healthy and homogenous control group devoid of any psychiatric symptoms, we further reduced the sample size of our control group from 30 to 23. The sample size of our OCD group was limited as we excluded patients with medication use and comorbidities. However, the smaller sample sizes were compensated by the homogeneity of the groups, which decreases the possibility of confounds and within-group variability, and bolsters the validity of our inferences regarding the corresponding populations. Another limitation is the lack of a structured diagnostic assessment tool for the OCD group. However, the lack of such assessment tool was compensated with the diagnostic clinical experience of the assessing psychiatrist.

Conclusion

Our study revealed that pediatric OCD patients have lowered evidence accumulation efficiency and a trend in increased caution, in comparison to healthy controls. Moreover, while errors cause distraction in the subsequent trial for both groups, the two groups react differently in terms of caution; OCD patients become more cautious after they make an error whereas healthy controls become less cautious. Our findings add to a series of studies that emphasize the importance of computational decision-theoretic approaches for characterizing latent processes in the study of clinical populations. By investigating the latent processes through HDDM, we were able to identify differences that match the symptomatology and neurobiology of OCD and were not evident in the isolated analysis of the reaction time and accuracy data. We also observed indications that comparing latent processes may reveal a signature that sets clinical OCD apart from high OC traits alone. The minor differences between our findings with a pediatric population and those of studies with adult populations point to a need for further research that incorporates development and emotion regulation into studies of latent processes in decision making.

CHAPTER III

Gamification of Two Alternative Forced Choice: Validation based on Drift Diffusion Model and Individual Differences

Abstract

Two alternative forced choice (2AFC) paradigms, coupled with the unified analysis of accuracy and response times within specific decision theoretic frameworks, have provided a wealth of information regarding decision-making processes. One problem of associated experimental tasks is that they are typically not engaging and do not contain stimuli or task representations familiar to participants, resulting in contaminants in the data. Furthermore, when investigating decision strategies, use of noisy stimulus attributes result in undesired variance in perceptual process complicating the analysis and interpretation of results. To address these fundamental issues, we developed a 2AFC soccer game in which participants' task is to score goals by making leftward or rightward shots after observing the trajectory of the goalkeeper within a trial. The goalkeeper's location is repeatedly sampled from a normal distribution with a constant variance with a mean either to the left or right of the midpoint. We tested participants on three difficulty levels parameterized by the distance between the two means. We also biased the ball placement to test its effect on bias in decisions. We expected rate of evidence integration to decrease with increasing difficulty and participants to be biased in accord with ball placement. Drift-diffusion model provided good fits to data and their outputs confirmed our primary predictions outlined above. Furthermore, consistent with earlier findings, evidence integration rates were lower after errors and for those who scored higher on anxiety, perfectionism, and obsessive compulsive trait scales.

Keywords: gamification, HDDM, decision making, perfectionism, anxiety

Introduction

Two Alternative Forced Choice (2AFC) tasks have been vastly used in the behavioral and neurophysiological study of decision making (e.g., Balcı et al., 2011; Gold & Shadlen, 2007). In these tasks, participants simply choose one option or the other based on the information available to them in a given trial. These studies have greatly benefited from the application of computational models that allow the estimation of decision process parameters (Forstmann, Ratcliff, & Wagenmakers, 2016). One of these signature models is the Drift Diffusion Model (DDM) (Ratcliff, 1978; Ratcliff & Rouder, 1998; Ratcliff & McKoon, 2008). In DDM, a decision is conceptualized as evidence accumulated in a noisy fashion over time (drift), until the accumulated evidence passes one of the two thresholds that represent one of the options (decision boundary) (Ratcliff & Rouder, 1998). DDM allows a deeper understanding of the decision process, by utilizing the basic elements of the performance (e.g. response time (RT) and responses) in a unified fashion (in contrast to their isolated analyses) to characterize the decision performance in terms of the latent variables of the presumed generative process: cautiousness (i.e threshold setting), evidence gathering efficiency/difficulty of task (i.e. drift rate), prior tendency towards one of the responses (i.e. decision bias), and delay unrelated to the decision process (i.e. non-decision time) (Forstmann et al., 2016).

Many of the experiments in experimental psychology require participants to complete long, repetitive, and relatively dull tasks with the purpose to collect enough data for reliable estimates and model fits described above. On the other hand, monotonous tasks usually lead to disengagement and boredom, and consequent poor and variable task performance (Hawkins, Rae, Nesbitt, & Brown, 2013). Disengagement and boredom related performance issues might be more prevalent among pediatric groups and certain clinical disorders (i.e., Attention Deficit and Hyperactivity Disorder - ADHD) for instance due to associated lapses in attention. These task-engagement related issues might in turn confound the analyses that aim to characterize putative differences in the decision process itself. A potential and contemporary way of overcoming these problems is embedding the task rules into games or adding game-like elements to the task.

Gamification is defined as "the use of game design elements in non-game contexts" (Deterding, Dixon, Khaled & Nacke, 2011, p.11), and such practice is increasingly applied in various domains (Ninaus, et al., 2015) including experimental psychology (Hawkins et al., 2013). For example, in their working memory training task, Ninaus et al. (2015) included game elements such as a progress bar, level marker, and a theme. The group who completed the gamified version performed better than the group who participated in the original (non-gamified) version of the task. Moreover both Shaw, Grayson, & Lewis (2005) and Prins, Dovis, Ponsioen, Ten Brink, and Van der Oord (2011) tested participants diagnosed with ADHD in gamified tasks. Prins et al. (2011) trained one group of ADHD patients for working memory in the game format and another group in a non-game format for three weeks. The game format group had both higher motivation and performed better than the control group. Similarly, Shaw et al. (2005) showed that ADHD patients performed better on a gamified continuous performance task (CPT II) than the original version.

The commonly used 2AFC tasks (e.g., dot motion discrimination task) are also monotonous and it is difficult to sustain the participants' task focus. In the same vein with prior efforts in the domain of working memory and attention, we developed a gamified 2AFC task that involved shots made in the context of soccer game. The advantages of our gamified task over already existing 2AFC tasks are that 1) the task, its aim, and reward structure are familiar to the participants, 2) it is relatively engaging and fun, 3) it is possible to induce bias based on stimulus properties (independent of prior history of stimulus frequencies), and 4) the momentary evidence (which varies from one frame to the next due to noise) for two alternatives is not only observable to the participants but also to the experimenters during data analysis. Due to this last advantage, it is possible to reconstruct the evidence accumulation in each trial for an ideal observer (e.g., see Figure 3.3).

Aims:

There were three primary aims of this work. The first aim was to test the basic form of the game with three difficulty levels to investigate their effect on decision outputs (i.e., accuracy and RT) as well as latent decision parameters. This test is essential for the validation of the task as a 2AFC paradigm within the framework of DDM.

The second aim was to observe differences in the task performance of the participants who ranked differently on various scales that measure: Obsessive compulsive (OC) traits (41 item Padua Inventory; Van Oppen et al., 1995; validated in Turkish by Beşiroğlu et al., 2005), trait anxiety (STAI; State-Trait Anxiety Inventory trait form; Spielberger, Gorsusch & Lushene, 1970; validated in Turkish by Öner & Le Compte, 1985), and perfectionism (Frost Multidimensional Perfectionism Scale - FMPS; Frost, Marten, Lahart, & Rosenblate, 1990; validated in Turkish by Kağan, 2011). Pertaining to non-clinical populations, decision-making in anxiety (White et al., 2010b) and subclinical OCD (Erhan & Balcı, 2016) have been previously studied within the framework of DDM, however perfectionism has not been studied in this context before. Perfectionism is a transdiagnostic trait that is argued to be a maintaining factor for OCD as well as a risk and maintaining factor for anxiety disorders. Perfectionism thwarts the treatment process, however when it is treated symptoms of clinical disorders lessen (Egan, Wade & Shafran, 2011). The close relationship of perfectionism with both anxiety and OCD (e.g. Frost & DiBartolo, 2002) warrants its investigation within the same decision-theoretic framework.

The final aim of the current work is to provide an adaptable method (i.e., gamification of paradigms) to be used for a wide range of research areas in psychology in general and in specific to provide the developed game to other researchers, with easy documentation to facilitate the integration of a gamified task into their ongoing research. We believe that this will improve both the participants' experience during experimentation (an endpoint typically overlooked in experimental psychology) and the quality of data collected in the experiments.

Hypotheses:

We expected hierarchical drift diffusion model (HDDM; Wiecki et al., 2013) to provide acceptable fit quality for the data gathered from the game. Primarily, we predicted error rates and RTs to increase and the drift rate to decrease with increasing task difficulty and starting point to be biased in accordance with the non-centralized placement of the ball on the screen. We also predicted that there will be slowing and increased error rate accompanied by lower drift rates after errors compared to post-correct responses.

As for our second order predictions, we hypothesized people who ranked high on OC, anxiety, and perfectionism scales to have higher threshold settings than those who ranked lower on these scales (for these relations in subclinical OC traits, see Erhan & Balcı, 2016). Since Banca et al. (2015) working with adult, and Erhan et al. (2017) working with pediatric OCD patients have shown that those diagnosed with OCD generally have lower drift rates than controls, we also predicted a residual decrease in drift rates from low to medium to high rankers on OC traits, perfectionism, and anxiety scales.

Methods

Shoot, 2AFC Game:

The primary task in “Shoot” is taking a kick toward the left or right of the goalkeeper in the goal area to score as many goals as possible during the experiment. The task presents the visual images of a goal, goalkeeper, and ball on a soccer field. The goalkeeper continuously moves laterally in the goal area until the participant takes the shot. The decision is to which side of the keeper (left or right) to shoot the ball at. Immediate feedback is provided after each kick.

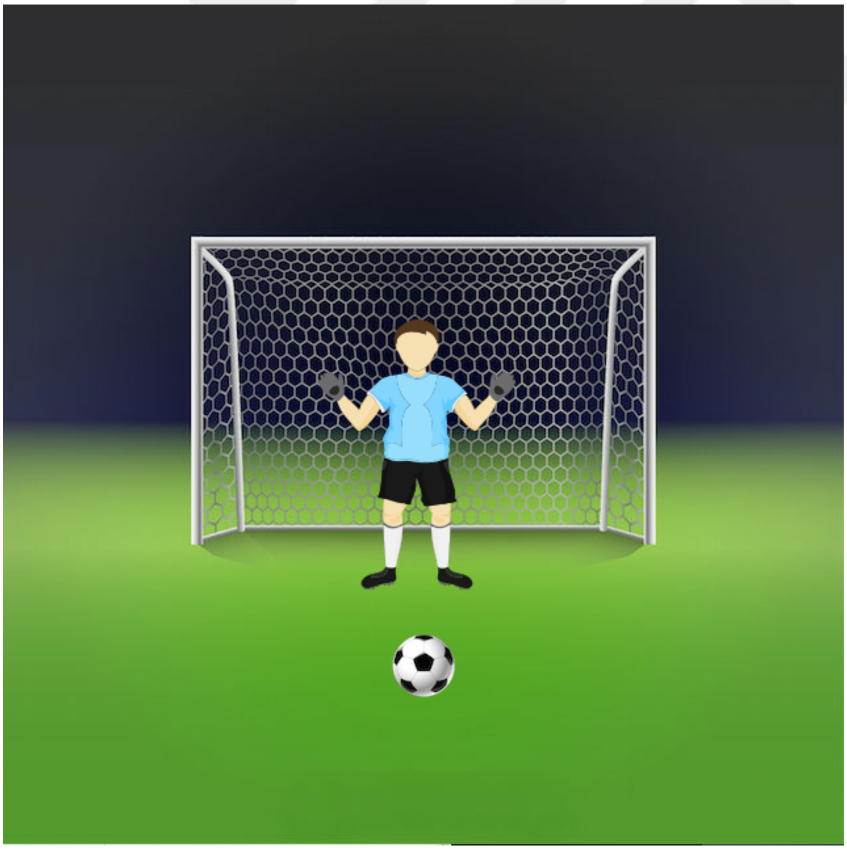

Figure 3.1. A screenshot from the task: The goalkeeper moves continuously toward left and right. The player’s task is to decide shooting the ball to the left or right of the goalkeeper.

Goalkeeper Motion:

The task in "Shoot" is to observe the perceptually continuous motion of the goalkeeper and to decide which side of the goal area it *favors* (spends more time on) so that shot is taken in the opposite direction. The goalkeeper moves in front of the goal area in a lateral fashion (Figure 3.1) and in each trial, the goalkeeper spends more time either on the left or right of the goal area. This constitutes the evidence that the observer accumulates from one frame to the next.

The location of the goalkeeper on the horizontal axis is a random variable sampled from a Gaussian distribution with mean slightly to the left or right of the center (Figure 3.2). This slight deviation of the mean from the center results in the on average biased behavior of the goalkeeper. Thus, there are two hypotheses for the observer to test given the behavior of the goalkeeper: the goalkeeper's behavior comes from a normal distribution with a mean either to the left (x_{center}-Δx) or to the right of the center (x_{center}+Δx), with the task becoming more difficult as the distance between the two distributions decreases. An ideal observer would compute the likelihood of a given location under these two hypotheses (two distributions with different means and constant variance) and estimate the cumulative sum of these log likelihoods with each change in the location of the goalkeeper. The refresh rate of the location varies greatly from sample to sample, but averages to about 1 per 100ms.

As the goalkeeper moves from one location to another, the observer builds more evidence regarding which side of the goal the goalkeeper favors. The variability in the sampling process of goalkeeper positions leads to a random walk with a bias toward the correct threshold when the relative likelihoods of the two partially overlapping distributions are summed over time. Note that this makes the evidence accumulation process in each trial observable to the experimenter.

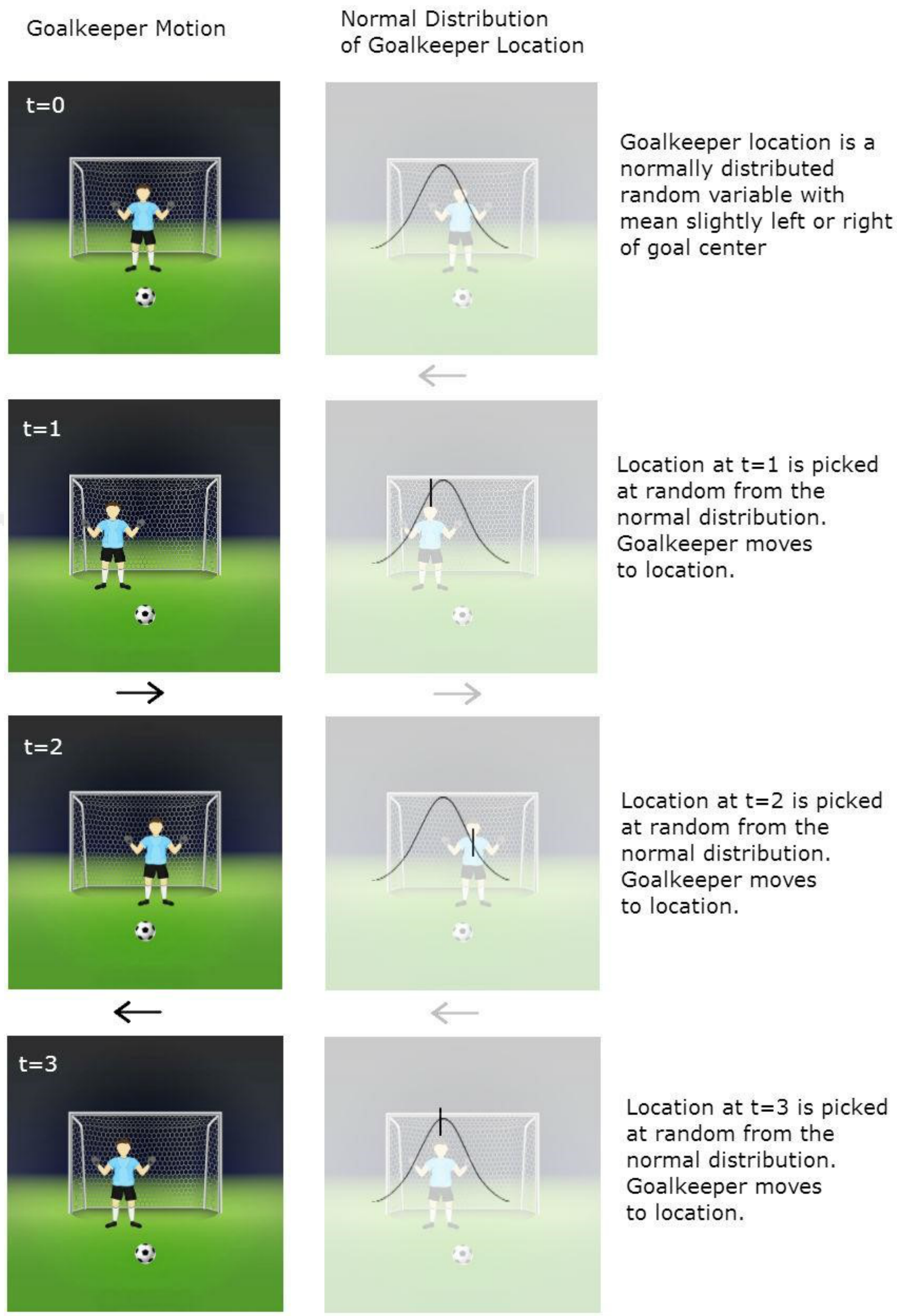

Figure 3.2. The goalkeeper motion is a series of locations sampled from one of the two distributions in a given trial.

Goalkeeper Motion Simulation:

In order to validate the separation of sample location likelihoods, between the left-deviated normal distribution and the right-deviated normal distribution, we ran a simulation, which was set up to sample from distributions that generated the goalkeeper motion in the task. Each of the easy, medium, and hard task conditions were simulated for both 50 sample and 1000 sample runs. The 50 sample run demonstrates the divergence up to 5 seconds of decision time in experiment (as sampling rate of task is approximately 1 per 100ms), while 1000 samples (very long evidence accumulation process) demonstrates eventual divergence in all conditions. The code for the simulation is provided in Supplement II.

The simulation computes and aggregates the likelihood ratios for the two options (i.e., likelihood ratio test) given the movement of the goalkeeper. This computation results in a random walk that resembles the sequential probability ratio test (Wald, 1947) and DDM evidence accumulation process when the sampling becomes continuous. Figure 3.3 illustrates the results of these simulations; the observed trajectory of the decision variable shows clear divergence as the number of samples increase. The rate of divergence also closely matches the decline in accuracy from easy to hard tasks.

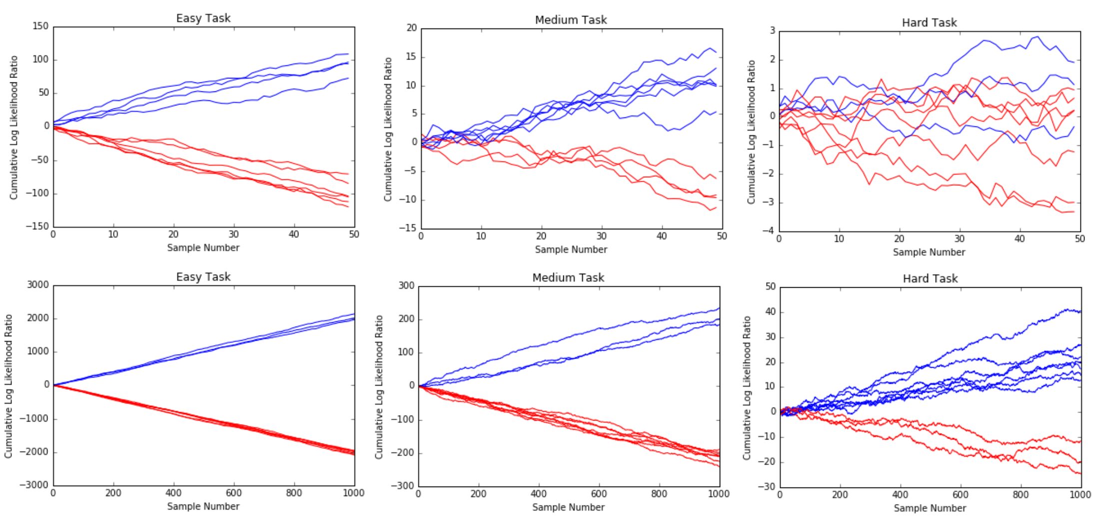

Figure 3.3. Simulation results illustrate cumulative likelihood ratio separation of the left-deviated and right-deviated normal distributions. Note that the number of trials for right-correct and left-correct trials differed due to randomness with the asymptotic probability of .5 as the number of trials becomes infinite. Note that since different y-scales are used for easy, medium, and hard task, the rate of evidence accumulation appears similar.

Player Shot Choice and Feedback:

The player must choose to shoot either to the left or right by pressing the corresponding key on the keyboard. This choice is followed by a shot animation in the chosen direction with either a save by the goalkeeper or a score depending on the underlying distribution. After the shot is resolved, feedback is provided via text above the screen as illustrated in Figure 3.4. Further rewards (e.g. points, monetary reward, etc.) may be provided depending on the area that Shoot is being used for. Scoreboard visuals may also be added to the game screen.

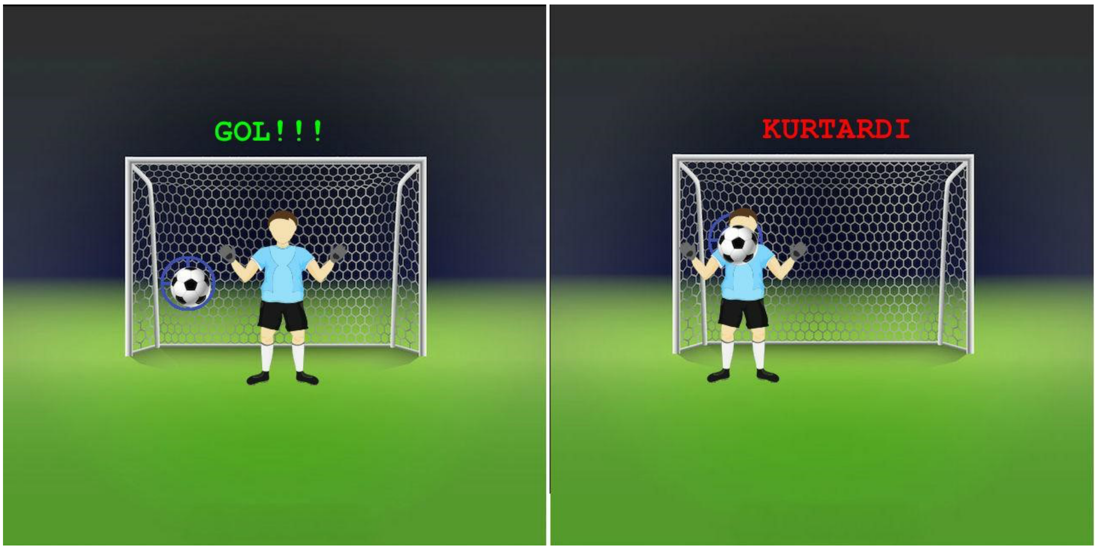

Figure 3.4. In two separate trials, the player shoots to the left. In the first trial, this was the correct choice and a goal is scored with appropriate feedback (*Gol* means *Goal* in Turkish). In the second, it was the incorrect choice, and the goalkeeper saved the shot with the appropriate feedback (*Kurtardı* means *Saved* in Turkish).

We also aimed to induce bias in the choice of right vs. left sides by biasing the location of the ball with respect to the center of the goal in order to test if this resulted in a bias parameter in DDM. Note that graphs provided in this manuscript provide a non-biased location for the ball, which is ideal for experiments in which decision bias is not a critical factor.

Implementation:

The game was built on the OpenSesame toolkit (Mathôt, Schreij, & Theeuwes, 2012), PyGame engine and Python code. OpenSesame is a well-documented toolkit with tutorials on functionality readily available on their website (http://osdoc.cogsci.nl/) and original paper (Mathôt et al., 2012). This infrastructure has several layers of complexity. For experimenters with no coding background, there is the User Layer. The OpenSesame toolkit provides simple graphical customization of the core experiment. For experimenters with coding experience, the Code Layer exposes Python code underlying the game for full customization. Along with the technical details outlined in this paper, the fully commented ready to use code is provided to allow for a quick and clear understanding of the game rules and architecture. The source code and installation instructions for this implementation, along with a few other experimental setups, may be found in the BitBucket repository at *https://bitbucket.org/kkaramanci/shoot_exp*.

User Layer:

The experimental setup consists of 3 sections: Introduction, Warmup, and Experiment (Figure 3.5).

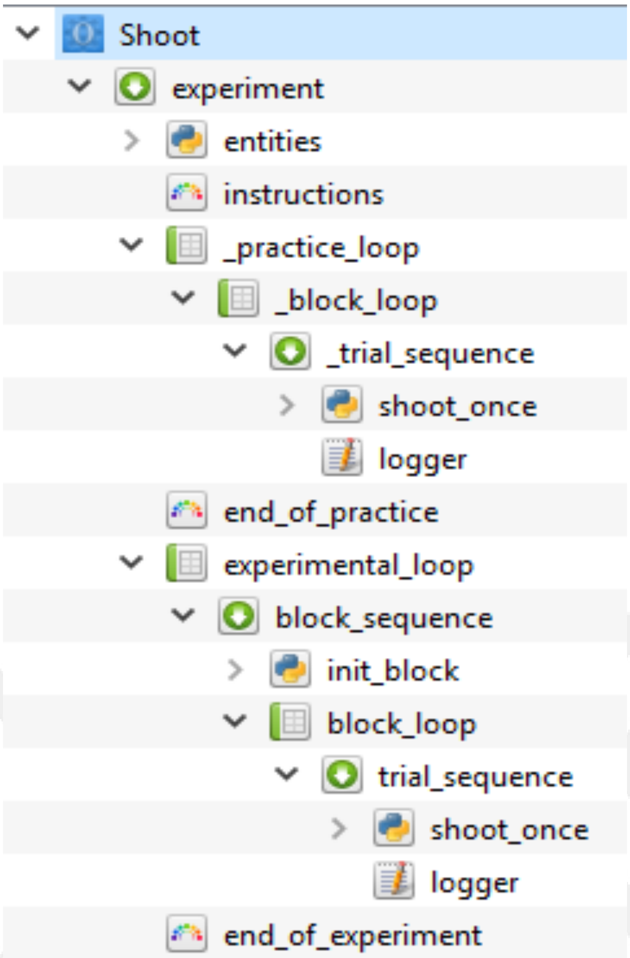

Figure 3.5. Experimental components setup on OpenSesame.

The introduction section consists of a Python block, which sets up all the in-game objects and configurations, and the instructions block, which provides the gameplay instructions to participants. The Warmup section is a section with a small number of trials, aimed at providing the participant with initial practice trials, before the actual experiment begins. Except for the number of trials, and a "trial" flag in data recording, it is identical to the experiment section as outlined below.

The Experiment section contains the bulk of the functionality and parameters. It consists of several components. The first component is the experimental_loop. While this is a loop component, which could run several times, it is only run once and is used mainly to initialize the first set of parameters – speed (goalkeeper speed), difficulty, and shot speed. These parameters are set in English here, (easy, hard etc.) while the translation to actual numeric values is done in code in the entities code block. The experimental_loop can be altered to run several iterations with various combinations of these settings if desired.

The block_sequence is a container for components to run in sequence. Conditionals may be set for components contained in the sequence. Init_block is a python code block, which initializes any variable used in this iteration of the experiment. In this case, it is used to initialize the start time for the experiment, so that experiment time may be tracked by the program and end the experiment when the time is up.

The block loop is the actual loop of trials run in the experiment. Here we define the different variations of experimental parameters to be run in this block. Specifically, we define absolute values for the deviation of the mean of the normal distribution (for goalkeeper location) from the center of the goal area. The mean of the distribution deviates from the center of the goal area by the function goal_area_width / deviation. The larger the number, the smaller the deviation and harder to detect the side favored by the goalkeeper.

In the block_loop, we define 3 separate values for the deviation, thereby forming three difficulty conditions. In every trial, one of these values is used at random for a total of 500 times each. While 1500 total trials may seem like a large number, we also define a break condition for this loop, which is set to break when the experiment has been running for 20 minutes.

The trial_sequence is the final container component that contains the python code, which executes the goalkeeper motion and shot trial, and the logger item, which logs each of the experimental variables.

Code Layer:

The code layer contains the Python code that sets up the graphical objects and defines their "behavior" in the experiment. While this is done through distribution of 3 separate block components in the OpenSesame setup, it is important to note that they are executed continuously in one run of the experiment, and are not isolated from each other.

Each Python block contains 2 pages of code, Prepare and Run. All of the Prepare code is executed first, and before the presentation of any stimuli to players (Mathôt et al.,2012). In this way, all the heavy lifting is done before the timed portion of the experiment, ensuring that only a small amount of code necessary to run the experiment is executed during time logging. This is necessary to minimize computational delays during experimentation.

The entities block sets up all the objects used in the experiment, as well as configurations for parameters, and graphical asset locations. Update functions for objects, which are run once per frame and responsible for the motion of graphics, are defined here. The nested object location coordinates within the objects they are nested in are also set here.

The shoot_once block is the code that initializes the primary Game object, and runs through one iteration of the game. In the Prepare code, all the objects defined in the entities block are initialized to their initial settings and coordinates on the screen. The screen is defined as a 640x640 square, which ensures that almost all screens commercially available today should be able to run it in the center of the screen, with no resolution differences or incompatibilities. Alternatively, if a different resolution is desired, the resolution setting may be changed. However, all the initialized positions for in-game objects must also be redefined to fit the new screen resolution.

The shoot_once block also defines the main function routine. In this routine, the framerate is checked to ensure 60 fps (consistent timing), each of the update functions for graphical assets are called, and the keyboard is checked for input by the user and –if provided- the input is evaluated. The update routine runs once per frame and the game routine runs at 60 fps. Further validation of timing accuracy may be found in the original OpenSesame paper (Mathôt et al., 2012). The Run code of the shoot_once block runs the main function routine, which starts the presentation of the game. Fully commented code for each of the blocks can be found in the Supplement.

Experimental Validation:

We tested 50 undergraduates from Koc University. Participants received course credit for their participation in the experiment. The experiment was approved by the Koc University Ethical Review Board and all participants signed a consent form. Participants played Shoot for 20 minutes in which they were instructed that being both accurate and fast was important. Afterwards they filled out Padua, Trait anxiety form of STAI and FMPS scale to measure OC, anxiety, and perfectionism traits, respectively. FMPS scores were calculated omitting the Organization subscale as suggested by Frost et al. (1990) and Kağan (2011). We excluded 3 participants from our analyses; one due to premature termination of the task (13 minutes instead of 20), and two due to random responding, spotted through RT data.

Data analysis:

We used one-way repeated measures ANOVA to assess the effect of difficulty on behavioral outcomes of the task such as RTs and accuracy. We used paired samples t-tests to compare RT and accuracy after correct and error responses. Participants were assigned to one of the three rank groups (i.e., “high”, “medium” and “low”) for each scale with a tertiary split (e.g. White et al., 2010b). In the cases where participants with the same scores fell in two different groups, the participants in the lower groups were automatically re-assigned to the upper group. We used one-way ANOVA to assess differences between RTs and accuracy of high, medium and low scale score groups. In order to test for bias in keys pressed, we compared the percentage of left key presses at each difficulty condition to chance level (0.5) using one-sample t-tests. We also used Pearson correlations to assess the linear relationships between scale scores. All tests were conducted using both frequentist and Bayesian methods.

HDDM (Wiecki et al., 2013) was fit to RT and accuracy data to estimate and compare latent decision variables such as drift rate, threshold, and starting point for different difficulty levels and in post-error compared to post-correct trials. We simulated data using the model outcomes to test whether comparable parameters could be recovered.

We used a within-subjects HDDM design to investigate the main effects of task difficulty and main effects of errors in previous trials (i.e. post-error effects) on drift rate, threshold, and starting point. In addition, we investigated how scores in OC, trait anxiety, and perfectionism scales were related to latent variables at different difficulty levels by adding the scale groups as between subjects factors.

Trials with RTs faster than or equal to 100 ms were discarded (amounting to less than 1% of all trials) as they were too fast to have come from a decision process (premature responding), and thus not relevant for the HDDM analyses. All models had 5000 samples with the first 20 draws discarded as burn-in. The models assumed that 5% of the trials were outliers generated by processes not related to DDM and came from a uniform distribution. We assessed model convergence using the trace, autocorrelation, and marginal posteriors. In order to ensure convergence, the models involving difficulty levels had the drift rate and starting point as group-only variables, meaning these variables were estimated only on the group level and not on the subject level. The post-error models had only the starting point as a group-only variable, as drift rate convergence was not a problem for these models. We relied on the separation of posterior probabilities *(P)* of parameters for hypothesis testing.

Results

Participants completed 503 trials (*SD*=92) on average in 20 minutes, with a minimum of 229 trials and a maximum of 668 trials. As expected, RTs got slower as task difficulty increased from easy (*M*=843ms, *SD*=38ms), to moderate (*M*=1155ms, *SD*=686ms), to hard (*M*=1277ms, *SD*=885ms) conditions, $F(1.098, 50.50)=26.95$, $p<0.001$, $BF_{10}>100$. Again as expected, accuracy levels dropped as task difficulty increased from easy (*M*=90.7%, *SD*=7.75%) to moderate (*M*= 72.6%, *SD*=7.97%) to hard (*M*=58.2%, *SD*=4.64%) levels, $F(2, 92)=697.51$, $p<0.001$, $BF_{10}>100$. We observed no difference between RTs or accuracies of different ranked scale groups for any scale ($p>0.05$ for all comparisons, $BF_{01}>1$ for accuracy differences among OC trait groups, $BF_{01}>3$ for all other comparisons).

Overall, participants responded more slowly after errors (*M*=1178ms, *SD*=821ms) than after correct (*M*=1070ms, *SD*= 618ms) responses, $t(46)=2.45$, $p=0.02$, $BF_{10}=2.29$. Moreover, accuracy was higher after correct responses (*M*=74.7%, *SD*=5.52%) than after errors (*M*=71.9%, *SD*=8.06%), $t(46)=3.95$, $p<0.001$, $BF_{10}>30$.

Averaging over difficulty levels, we observed a bias towards responding “left” (*M*=55.3%, *SD*=5.2%, $t(46)=6.99$, $p<0.001$; $BF_{10}>100$). While left-responding was not above chance level at the easy level (*M*=51.3%, *SD*=5.3%, $t(46)$ =1.72, p=0.09; BF_{01}=1.62), participants increasingly pressed left from moderate level (*M*=56.4%, *SD*=6.5%, $t(46)$=6.77, p<0.001; BF_{10}>100), to hard level (*M*=58.3%, *SD*=7.3%), $t(46)$=7.88, p<0.001; BF_{10}>100. There was stronger bias in error trials (*M*=60.3%, *SD*=9%) than correct trials (*M*=53.6%, *SD*=4.7%), $t(46)$=6.48, p<0.001, BF_{10}>100.

Perfectionism, OC and trait anxiety scores were all positively correlated with each other. Perfectionism had linear relationships with both OC traits (r=0.43, p<0.01, BF_{10}>10), and trait anxiety (r=0.48, p=0.001, BF_{10}>30). There was also a linear relationship between OC traits and trait anxiety, r=0.54, p<0.001, BF_{10}>100.

The within-subject HDDM with drift rate, threshold, and starting point values that vary according to difficulty levels revealed that as task difficulty increased, drift rates decreased (M_{v_easy}=1.24, $M_{v\_moderate}$ = 0.49, M_{v_hard} = 0.16; P = 1.0 for all comparisons) but threshold levels (M_{a_easy}= 2.12, $M_{a\_moderate}$ = 2.07, M_{a_hard} = 2.06; P < 0.75 for all comparisons) and the starting points (M_{z_easy}=0.531, $M_{z\_moderate}$=0.534, M_{z_hard}=0.536; P < 0.90 for all comparisons, $P_{z>0.5}$=1.0 for all difficulty levels) remained stable across difficulty levels (Figure 3.6). All parameters were successfully recovered from data simulated with the outcome of the model. The within-subject model in which drift rate, threshold and starting point were allowed to vary only according to post-error and post-correct status indicated that post-error trials had lower drift rates than post-correct trials (M_{v_pe} = 0.47, M_{v_pc} = 0.59; $P_{v\_pc>v\_pe}$ = 1.0), but the threshold parameter (M_{a_pe} = 2.05, M_{a_pc} = 2.01; $P_{a\_pc>a\_pe}$ = 0.39) and the starting point parameter (M_{z_pe}=0.54, M_{z_pc}=0.53; $P_{z\_pe>z\_pc}$=0.94, $P_{z>0.5}$= 1.0 for both) remained the same (Figure 3.7). Note that starting point values in the plots (Figures 3.6 and 3.7) are the output of the HDDM Regression package, where z is estimated in the range of plus and minus infinity and the non-biased midpoint is 0. The starting point means reported in the text are converted from these values (with the inverse logit function) to the conventional scale where z is bound between 0 and 1 and the midpoint is 0.5. All variables in both models had satisfactory convergence.

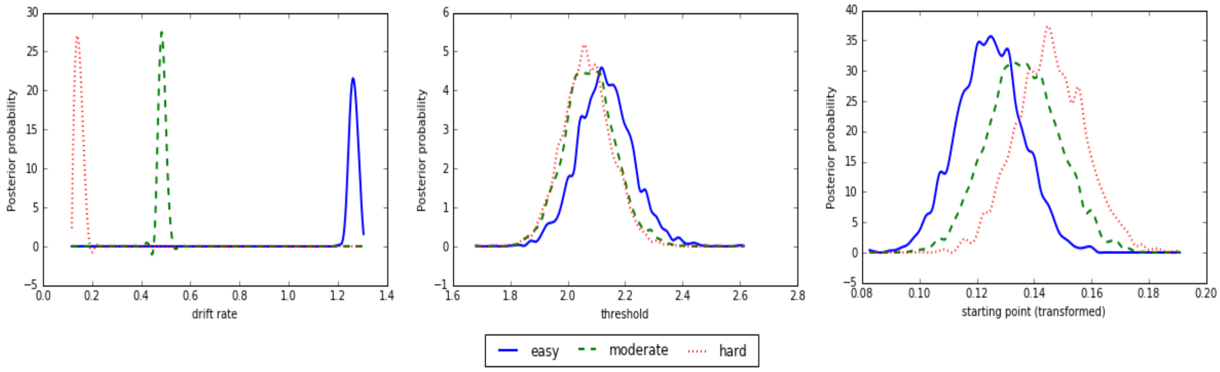

Figure 3.6. Group mean posterior distributions of drift rates, thresholds, and starting points at three difficulty levels.

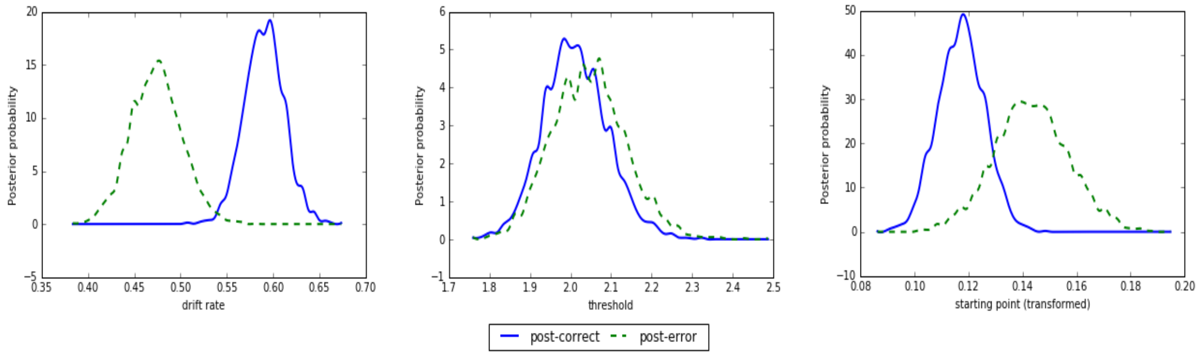

Figure 3.7. Group mean posterior distributions of drift rates, thresholds and starting points in trials after correct and error responses.

We investigated drift rates for participants grouped in low, medium and high perfectionism, OC, and trait anxiety scale scores separately for difficulty levels (Figure 3.8). In the easiest level, as expected, low scale scorers had the highest drift rates whereas high scale scorers had the lowest drift rates. All drift rate differences between scale groups at the easy level were significant (P>0.95). At the moderate difficulty level, low scorers on the OC scale had higher drift rates than high scorers

($P_{low>high}$>0.99), with a trend for a difference between the medium level with either extreme ($P_{low>medium}$=0.92; $P_{medium>high}$=0.94). Also at the moderate difficulty level, low scorers on perfectionism had higher drift rates than both the medium and high scorers (P>0.99 for both), but there was no difference between the moderate and high score groups ($P_{medium>high}$=0.34). There was no difference between the groups in the trait anxiety scale at the moderate difficulty level. We observed no difference between high, medium and low scorers for any scale at the hardest level of difficulty, but there was a trend for high scorers in the OC scale to have lower drift rates than medium and low scorers ($P_{low>high}$=0.93; $P_{medium>high}$=0.92). Thresholds did not change at any difficulty level for any scale group. Starting points varied non-systematically in a narrow range slightly above the midpoint (0.527 - 0.546, $P_{z>0.5}$>0.95 for all). Although a few of the differences were statistically significant, they did not point to a systematic relationship between starting points, difficulty levels, and scale scores.

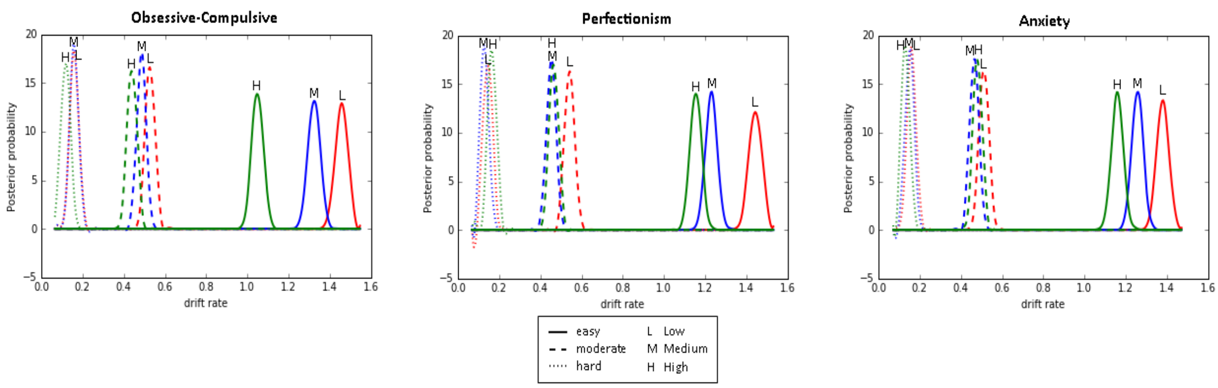

Figure 3.8. Group mean posteriors of drift rate as modulated by difficulty level and scale score groups. H=High, M=Moderate, L=Low scale score groups.

Discussion

We developed a gamified 2AFC task in which participants made repeated decisions based on observed evidence (i.e., movements of the goalkeeper) about which side of the goal to shoot at for a score. We manipulated the task difficulty by parameterizing the distribution that guided the movements of goalkeeper from one frame to the next, and introduced a perceptual biasing signal by placing the ball closer to one end of the goal. The aim of this study was threefold: 1) to observe whether behavioral outputs (e.g. RT and accuracy) as well as latent decision variables estimated from model (HDDM) fits behaved in predicted directions and thereby validated the “Shoot” as a 2AFC task; 2) to investigate how anxiety, perfectionism and OC traits are related to DDM variables; 3) to provide an easy to use gamified task for researchers to integrate into decision-making studies.

Regarding the first facet of the study, we expected response times to increase and accuracy and drift rates to decrease with increasing task difficulty (e.g., see Balcı et al., 2011) and HDDM to provide good fits to the data. Regarding the second facet of the study and based on our previous results reported in Erhan and Balcı (2016), we expected those who ranked higher on OC, anxiety, and perfectionism scales to have higher threshold settings than those who ranked lower on these scales. Since Banca et al. (2015) and Erhan et al. (2017) generally found a decrement in drift rates of OCD patients we also expected those who scored high on these scales to have lower drift rates. Our prediction/intention regarding the third facet of the study can only be tested over the next years.

Our predictions regarding the basic behavioral and model outcomes were confirmed. Specifically, response times and error rates increased whereas drift rates decreased with increasing task difficulty as well as after errors compared to after correct responses. We did not predict different threshold settings for different task difficulties since difficulty levels were randomly assigned for each trial instead of being presented in separate blocks (Bogacz et al., 2006). In line with this theory-driven rationale, thresholds remained stable across different difficulties. We also predicted participants to exhibit bias in the starting point in line with the biased placement of the ball during testing. In line with this prediction we observed biased starting points in the direction of the ball placement. However, note that we did not have a control condition for this test and thus results regarding starting point should be interpreted with caution. Even though behaviorally participants increasingly pressed left as the task got harder and that starting point bias was observed on all difficulty levels, no systematic difference of starting points was observed between difficulty levels. This finding suggests that participants tried to process task relevant information irrespective of its signal to noise ratio.

Our predictions regarding the relationship between threshold settings and scores on the scales were not confirmed however our predictions regarding the relationship between scale scores and drift rates were confirmed. The relationship between drift rates and scale scores was more apparent in the easy level, possibly because the signal to noise ratio on moderate and hard levels were too low, resulting in lower between-subject variance due to floor effect. Specifically, we observed a negative relationship between drift rates and OC traits; as OC trait scores increased drift rates decreased. This finding is in line with the results of Banca et al. (2015; see also Erhan et al., 2017) and Metin et al. (2013) as drift rates were observed to be lower in both OCD and ADHD groups compared to healthy controls. On the other hand, it is in contrast with the results of Erhan and Balcı (2016), who found that threshold settings were increased in participants with higher OC features without any apparent differences in drift rates.

Based on these differential results gathered with subclinical OC and clinical

OCD, Erhan and Balcı (2016) argued that changes in drift rates might be a diagnostic signature of clinical OCD. Consequently, it is possible that the analytical approach adopted in the current study (dividing scores into three groups) increased the sensitivity to differences in drift rates even in a subclinical sample. Alternatively, the idiosyncratic features of our task (e.g., higher familiarity, single data point per frame) might have bolstered the possible differences in the ability to integrate evidence. Furthermore, frequent change in difficulty levels from one trial to the next might have thwarted differences in threshold settings between participants of various ranks on the scales. Further studies are needed to elucidate the factors that might have led to differential results between different tasks.

Importantly, trait anxiety as well as perfectionism, argued to be associated with both OC and anxiety traits (Frost & DiBartolo, 2002) also followed the pattern of the OC traits; as perfectionism and trait anxiety scores increased drift rates decreased. Interestingly, scores on these scales were related to drift rates in a task that presumably minimized the between-subject variability in drift rates by minimizing differences in perceptual processing of sensory stimulus (due to clear indication of the momentary evidence at each frame). Thus, the observed between-group differences are more likely due to differences in evidence integration ability itself. For instance, each piece of information might be weighted less in participants who score high on OC, perfectionism, and anxiety. An alternative explanation relates to the width of the window over which information is integrated by participants or the decrement in the contribution of previous bits of evidence to the current estimates of the state of the world; this window might be shorter or the contribution of the previous evidence might be lower for high scorers due to factors such as working memory deficits (e.g. de Vries et al., 2014; Eysenck, Derakshan, Santos & Calvo, 2007).

An important advantage of the presented game is the ability to record the momentary state of evidence as observed by the participants, which allows reconstructing the exact trajectory of the belief state of an ideal observer. This perfectly tractable information within each trial might prove useful in reducing the number of parameters in the models developed/fit to explain the generative processes that underlie two alternative forced choice behavior.

We believe that this work provides a good exemplar to other researchers for participant-experience-centered experiment design to improve task engagement, experience of the participants, and the quality of data. This becomes particularly beneficial in studies with pediatric and clinical samples. We find that conventional tasks with novel stimulus (such as random dot motion discrimination) increase the risks of not-understanding task rules, require longer practice to familiarize participants with stimulus and task rules, and lead to boredom during testing. One of these factors or their combination poses difficulties for gathering high quality behavioral data and leads to the contamination of task performance due to factors not directly relevant for research questions. Secondly, we believe that improving the quality of participant experience is a factor that should be observed during the development of experimental tasks for ethical reasons and to minimize attrition in studies that require repeated testing. Finally, gamification of conventional tasks paves the way to large-scale and high-throughput testing outside the lab setting for instance through mobile devices.

THESIS DISCUSSION

The chapters included in this thesis focused on the characterization of latent decision making processes in subclinical populations and clinical Obsessive Compulsive Disorder (OCD) by utilizing the Drift Diffusion Model (DDM). Obsessive Compulsive (OC) traits were the common focus of all three chapters. The first chapter focused on a non-clinical population with various levels of OC traits, the second chapter focused on a pediatric OCD sample in comparison to healthy controls, and the third chapter focused on differences between high, medium and low levels of trait anxiety, perfectionism, and OC traits in a non-clinical population using a novel gamified 2AFC task. Thus, the third chapter offers a novel methodological approach to the study of choice behavior by using the advantages of gaming.

The first chapter aimed to understand not only how the entirety of OC scale scores predicted latent decision processes but also how various OC components (e.g. checking, rumination) contributed to these decision-making variables. Moreover, divergence from optimality and subjective cost assigned to errors were also studied with respect to OC scale and subscale scores. While OC scores were not predictors of traditional behavioral measures (RTs, error rates, post-error slowing), increases in entirety of OC scores, rumination, and checking, predicted higher threshold settings (e.g. more cautious responding). This chapter was the first to incorporate DDM to understand latent variables of decision making in an analogue OC sample. Analogue OC samples are important given the argument that OCD is not a categorical but a dimensional disorder, and that symptom severity ranges from none to very high on a continuum (Abramowitz et al., 2014). Such range deems studying the entire spectrum important for both understanding and using the information gained with a non-clinical sample to prevent against the disorder. This chapter unlike the study of Banca et al.

(2015) did not find a relationship between evidence accumulation efficiency and OC traits. Therefore it concludes with the question of whether low drift rates were a signature of clinical OCD.

The second chapter followed up on the first chapter picking up from this question. It studied latent decision making in a pediatric OCD population using dot motion discrimination task and Hierarchical Drift Diffusion Model (Wiecki et al., 2013). It aimed to understand how latent decision making variables differed between the clinical and the healthy control groups, trying to answer the question of if it is drift rate that sets clinical and healthy groups apart and if low drift rates are indeed a signature of clinical OCD as compared to healthy controls. Moreover, it also aimed to further investigate the topic of post-error behaviors in OCD. The first chapter found that post error slowing was not predicted by OC traits. In this chapter we modeled post-decision choices and RTs to investigate if post-error and post-correct decisions differ in terms of the related latent decision processes and whether these processes differed between clinical and healthy groups. Overall, the pediatric OCD group had significantly lower drift rates and slightly higher threshold settings than the healthy control group. Such finding is similar to that of adult OCD patients (Banca et al., 2015), however the pediatric group seems to not have as high increases in threshold settings as their adult counterparts. In addition, post-error latent processes differed for the clinical and healthy groups. While errors pave way to distraction in the following trial for both groups, errors in comparison to correct decisions lead to more caution in the OCD group and less caution in the healthy control group.

The second chapter closes by raising two important questions. The first one is why the pediatric OCD population does not show as increased a caution as adults in the Banca et al. (2015) study do, despite both groups having lower drift rates than the corresponding healthy controls. The chapter points to possible differences in emotion regulation due to age. Such differences in emotion regulation could be due to the number of years the participants lived with the disorder and developed coping strategies, as well to their medication history. Given decisions can change by way of emotion regulation techniques (Phelps et al., 2014), effects of emotion regulation on latent decision processes should be a topic of further investigation. The second question is whether lower evidence accumulation efficiency is indeed a signature for clinical OCD as was suggested in the first chapter. This chapter found that indeed HDDM can differentiate the healthy control and clinical OCD groups based on drift rates, which, given OCD is argued to be a dimensional disorder, paves way to the investigation of just how high and low OC symptoms can be differentiated based on drift rates. A reliable differentiation of high and low risk groups in non-clinical samples can potentially be used in preventative treatments. Given neurobiological bases of latent decision processes are already being studied, these brain regions could be targeted in potential research and treatments. Moreover, psychological treatments to improve evidence accumulation efficiency as a way to lower decision deficits associated with OC symptoms and thereby prevent against a clinical diagnosis can be investigated.

Two necessities emerged from the second chapter. First was the need for more engaging tasks. Such need becomes more urgent when running studies with pediatric and clinical populations. Second necessity was the need to study anxiety and perfectionism alongside OC traits, given the three traits' close relationship and the transdiagnostic property of perfectionism (Egan et al., 2011). The third chapter covered both necessities and further sought an answer for whether high and low levels of OC, perfectionism and anxiety traits in non-clinical populations can also be differentiated by drift rates.

In the third chapter we produced a new gamified 2AFC task and investigated whether drift rates and/or threshold settings could differentiate participants who rank on high, medium and low levels of OC, trait anxiety and perfectionism scales. For the task, we developed a 2AFC soccer game with three levels of task difficulty. As expected, RTs got slower and accuracy and drift rates decreased with increasing task difficulty. Moreover HDDM provided good fits to the data. As expected drift rates differentiated participants who rank on high, medium and low levels of OC, perfectionism and anxiety scales: drift rates significantly decreased from low to medium to high rankers on these scales. It is possible that the gamified task with its familiar aim (e.g. scoring a goal) and single data point per frame evidence, while decreasing the variance in perceptual processing between groups made the variance in evidence integration abilities between low and high rankers on the scales more visible. The evidence accumulation efficiency difference between the groups might indeed be due to the evidence integration ability, a part of drift rate, which was not possible to differentiate in the first chapter.

To conclude, the first two chapters had raised the question of whether low drift rate is a signature of clinical OCD that sets it apart from healthy populations. The third chapter however found that drift rate differences are also apparent between high, medium and low rankers on OC trait scale in a non-clinical sample in a more familiar task representation. Overall, the chapters point to evidence accumulation efficiency as being the most reliable latent decision making variable that sets both OCD apart from healthy controls, as well as those with high, medium and low levels of OC, perfectionism and anxiety traits from each other. Threshold setting differences that were predicted by OC symptoms and that were slightly increased in pediatric OCD patients in comparison to healthy controls, although seem to be related to OC traits and OCD, need further investigation. This thesis also produced a participant-experience-centered 2AFC soccer game and validated it by way of DDM. Overall, the three chapters add to a rapidly growing literature on mathematical models used to better characterize decision making in clinical disorders, especially OCD.

References

Abramovitch, A., Abramowitz, J. S., & Mittelman, A. (2013a). The neuropsychology of adult obsessive–compulsive disorder: A meta-analysis. *Clinical Psychology Review*, *33*(8), 1163–1171. https://doi.org/10.1016/j.cpr.2013.09.004

Abramovitch, A., Abramowitz, J. S., Mittelman, A., Stark, A., Ramsey, K., & Geller, D. A. (2015b). Research Review: Neuropsychological test performance in pediatric obsessive–compulsive disorder – a meta-analysis. *Journal of Child Psychology and Psychiatry*, *56*(8), 837–847. https://doi.org/10.1111/jcpp.12414

Abramovitch, A., & Cooperman, A. (2015). The cognitive neuropsychology of obsessive-compulsive disorder: A critical review. *Journal of Obsessive-Compulsive and Related Disorders*, *5*, 24–36. https://doi.org/10.1016/j.jocrd.2015.01.002

Abramovitch, A., Dar, R., Hermesh, H., & Schweiger, A. (2012). Comparative neuropsychology of adult obsessive-compulsive disorder and attention deficit/hyperactivity disorder: Implications for a novel executive overload model of OCD. *Journal of Neuropsychology*, *6*(2), 161–191. https://doi.org/10.1111/j.1748-6653.2011.02021.x

Abramovitch, A., Dar, R., Mittelman, A., & Schweiger, A. (2013b). Don't judge a book by its cover: ADHD-like symptoms in obsessive compulsive disorder. *Journal of Obsessive-Compulsive and Related Disorders*, *2*(1), 53–61. https://doi.org/10.1016/j.jocrd.2012.09.001

Abramovitch, A., Mittelman, A., Henin, A., & Geller, D. (2012). Neuroimaging and neuropsychological findings in pediatric obsessive–compulsive disorder: a review and developmental considerations. *Neuropsychiatry*, *2*(4), 313–329. https://doi.org/10.2217/npy.12.40

Abramovitch, A., Shaham, N., Levin, L., Bar-Hen, M., & Schweiger, A. (2015a). Response inhibition in a subclinical obsessive-compulsive sample. *Journal of Behavior Therapy and Experimental Psychiatry*, *46*, 66–71. https://doi.org/10.1016/j.jbtep.2014.09.001

Abramowitz, J. S., Fabricant, L. E., Taylor, S., Deacon, B. J., McKay, D., & Storch, E. A. (2014). The relevance of analogue studies for understanding obsessions and compulsions. *Clinical Psychology Review*, *34*(3), 206–217. https://doi.org/10.1016/j.cpr.2014.01.004

Aldwin, C. M. (1994). Developmental Studies of Coping In *Stress, coping and development: An integrative perspective* (pp. 217-239). New York, NY: Guilford Press.

American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Disorders, 5th ed American Psychiatric Publishing, Arlington, VA.

Apter, A., Fallon, T.J., King, R.A., Ratzoni, G., Zohar, A.H., Binder, M., Cohen, D.J. (1996). Obsessive-compulsive characteristics: from symptoms to syndrome. *Journal of American Academy of Child and Adolescent Psychiatry, 35(7),* 907–912.

Balcı, F., Simen, P., Niyogi, R., Saxe, A., Hughes, J. A., Holmes, P., & Cohen, J. D. (2011). Acquisition of decision making criteria: reward rate ultimately beats accuracy. *Attention, Perception, & Psychophysics*, *73*(2), 640–657. https://doi.org/10.3758/s13414-010-0049-7

Banca, P., Vestergaard, M. D., Rankov, V., Baek, K., Mitchell, S., Lapa, T., … Voon, V. (2015). Evidence Accumulation in Obsessive-Compulsive Disorder: the Role of Uncertainty and Monetary Reward on Perceptual Decision-Making

Thresholds. *Neuropsychopharmacology*, *40*(5), 1192–1202. https://doi.org/10.1038/npp.2014.303

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1–3), 7–15. http://doi.org/10.1016/0010-0277(94)90018-3

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Beşiroğlu, L., Ağargün, M. Y., Boysan, M., Eryonucu, B., Güleç, M. & Selvi, Y. (2005). Obsesif-Kompulsif belirtilerin değerlendirilmesi: Padua Envanteri'nin Türk toplumunda geçerlilik ve güvenilirliği. *Türk Psikiyatri Dergisi, 16*, 179-189.

Blakey, S. M., Jacoby, R. J., Reuman, L., & Abramowitz, J. S. (2015). The Relative Contributions of Experiential Avoidance and Distress Tolerance to OC Symptoms. *Behavioural and Cognitive Psychotherapy*, *FirstView*, 1–12. http://doi.org/10.1017/S1352465815000703

Bode, S., & Haynes, J. D. (2009). Decoding sequential stages of task preparation in the human brain. *NeuroImage*, *45*(2), 606–613. https://doi.org/10.1016/j.neuroimage.2008.11.031

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765. https://doi.org/10.1037/0033-295X.113.4.700

Bogacz, R., Hu, P. T., Holmes, P. J., & Cohen, J. D. (2010). Do humans produce the speed–accuracy trade-off that maximizes reward rate? *The Quarterly Journal of Experimental Psychology*, *63*(5), 863–891. https://doi.org/10.1080/17470210903091643

Bohne, A., Savage, C. R., Deckersbach, T., Keuthen, N. J., & Wilhelm, S. (2008). Motor inhibition in trichotillomania and obsessive–compulsive disorder. *Journal of Psychiatric Research*, *42*(2), 141–150. https://doi.org/10.1016/j.jpsychires.2006.11.008

Burguiere, E., Monteiro, P., Mallet, L., Feng, G., & Graybiel, A. M. (2015). Striatal circuits, habits, and implications for obsessive–compulsive disorder. *Current opinion in neurobiology*, *30*, 59-65. http://dx.doi.org/10.1016/j.conb.2014.08.008

Cavedini, P., Riboldi, G., D'Annucci, A., Belotti, P., Cisima, M., & Bellodi, L. (2002). Decision-making heterogeneity in obsessive-compulsive disorder: ventromedial prefrontal cortex function predicts different treatment outcomes. *Neuropsychologia*, *40*(2), 205–211. http://doi.org/10.1016/S0028-3932(01)00077-X

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*(3), 332–361. https://doi.org/10.1037/0033-295X.97.3.332

de Vries, F. E., de Wit, S. J., Cath, D. C., van der Werf, Y. D., van der Borden, V., van Rossum, T. B., … van den Heuvel, O. A. (2014). Compensatory Frontoparietal Activity During Working Memory: An Endophenotype of Obsessive-Compulsive Disorder. *Biological Psychiatry*, *76*(11), 878–887. https://doi.org/10.1016/j.biopsych.2013.11.021

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining “Gamification.” In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9–15). New York, NY, USA: ACM.

Dursun, O. B., Guvenir, T., Aras, S., Ergin, C., Mutlu, C., Baydur, H., … Goodman, R. (2013). A new diagnostic approach for Turkish speaking populations DAWBA Turkish Version. *Epidemiology and Psychiatric Sciences*, *22*(03), 275–282. http://doi.org/10.1017/S2045796012000479

Dutilh, G., van Ravenzwaaij, D., Nieuwenhuis, S., van der Maas, H. L. J., Forstmann, B. U., & Wagenmakers, E.-J. (2012a). How to measure post-error slowing: A confound and a simple solution. *Journal of Mathematical Psychology*, *56*(3), 208–216. https://doi.org/10.1016/j.jmp.2012.04.001

Dutilh, G., Vandekerckhove, J., Forstmann, B. U., Keuleers, E., Brysbaert, M., & Wagenmakers, E.-J. (2012b). Testing theories of post-error slowing. *Attention, Perception, & Psychophysics*, *74*(2), 454–465. https://doi.org/10.3758/s13414-011-0243-2

Egan, S. J., Wade, T. D., & Shafran, R. (2011). Perfectionism as a transdiagnostic process: A clinical review. *Clinical psychology review*, *31*(2), 203-212. http://dx.doi.org/10.1016/j.cpr.2010.04.009

Erhan, C., & Balcı, F. (2016). Obsessive compulsive features predict cautious decision strategies. *The Quarterly Journal of Experimental Psychology*, *0*(ja), 1–29. http://doi.org/10.1080/17470218.2015.1130070

Erhan, C., Bulut, G. Ç., Gökçe, S., Ozbas, D., Turkakin, E., Dursun, O. B., ... & Balcı, F. (2017). Disrupted latent decision processes in medication-free pediatric OCD patients. *Journal of Affective Disorders*, *207*, 32-37. http://dx.doi.org/10.1016/j.jad.2016.09.011

Erol, N., & Savaşır, I. (1988). Maudsley Obsesif-Kompulsif Soru Listesi. *24. Ulusal Psikiyatri ve Nörolojik Bilimler Kongresi Bilimsel Çalışma Kitabı*, 104-114.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion*, *7*(2), 336-353. http://dx.doi.org/10.1037/1528-3542.7.2.336

Flett, G. L., Madorsky, D., Hewitt, P. L., & Heisel, M. J. (2002). Perfectionism Cognitions, Rumination, and Psychological Distress. *Journal of Rational-Emotive and Cognitive-Behavior Therapy*, *20*(1), 33–47. https://doi.org/10.1023/A:1015128904007

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E. J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual review of psychology*, *67*, 641-666. doi: 10.1146/annurev-psych-122414-033645

Frost, R. O., & DiBartolo, P. M. (2002). Perfectionism, anxiety, and obsessive-compulsive disorder. In Flett, Gordon L. (Ed); Hewitt, Paul L. (Ed). (2002). Perfectionism: Theory, research, and treatment , (pp. 341-371). Washington, DC, US: American Psychological Association, xiv, 435 pp.

Frost, R. O., Marten, P., Lahart, C., & Rosenblate, R. (1990). The dimensions of perfectionism. *Cognitive Therapy and Research*, *14*(5), 449–468. https://doi.org/10.1007/BF01172967

Gibbs, N. A. (1996). Nonclinical populations in research on obsessive-compulsive disorder: A critical review. *Clinical Psychology Review*, *16*(8), 729–773. https://doi.org/10.1016/S0272-7358(96)00043-8

Gilbert, A. R., Akkal, D., Almeida, J. R. C., Mataix-Cols, D., Kalas, C., Devlin, B., Birmaher, B., & Phillips, M. L. (2009). Neural Correlates of Symptom Dimensions in Pediatric Obsessive-Compulsive Disorder: A Functional Magnetic Resonance Imaging Study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(9), 936–944. http://doi.org/10.1097/CHI.0b013e3181b2163c

Gold, J. I., & Shadlen, M. N. (2007). The Neural Basis of Decision Making. *Annual Review of Neuroscience*, *30*(1), 535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038

Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The Development and Well-Being Assessment: Description and Initial Validation of an Integrated Assessment of Child and Adolescent Psychopathology. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *41*(5), 645–655.

Hawkins, G. E., Rae, B., Nesbitt, K. V., & Brown, S. D. (2013). Gamelike features might not improve data. *Behavior Research Methods*, *45*(2), 301–318. https://doi.org/10.3758/s13428-012-0264-3

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin*, *109*(2), 340–347.

Hodgson, R. J., & Rachman, S. (1977). Obsessional-compulsive complaints. *Behaviour Research and Therapy*, *15*(5), 389–395. https://doi.org/10.1016/0005-7967(77)90042-0

Kağan, M. (2011). Frost Çok Boyutlu Mükemmelliyetçilik Ölçeğinin Türkçe formunun psikometrik özellikleri. *Anatolian Journal of Psychiatry/Anadolu Psikiyatri Dergisi*, *12*(3).

Karalunas, S. L., & Huang-Pollock, C. L. (2013). Integrating Impairments in Reaction Time and Executive Function Using a Diffusion Model Framework. *Journal of Abnormal Child Psychology*, *41*(5), 837–850. http://doi.org/10.1007/s10802-013-9715-2

Karayanidis, F., Jamadar, S., Ruge, H., Phillips, N., Heathcote, A., & Forstmann, B. U. (2010). Advance preparation in task-switching: converging evidence from behavioral, brain activation, and model-based approaches. *Frontiers in Psychology*, *1*, 25. https://doi.org/10.3389/fpsyg.2010.00025

Kodaira, M., Iwadare, Y., Ushijima, H., Oiji, A., Kato, M., Sugiyama, N., Sasayama, D., Usami, M., Watanabe, K., & Saito, K. (2012). Poor performance on the Iowa gambling task in children with obsessive-compulsive disorder. *Annals of General Psychiatry*, *11*(1), 1–6. http://doi.org/10.1186/1744-859X-11-25

Lawrence, N. S., Wooderson, S., Mataix-Cols, D., David, R., Speckens, A., & Phillips, M. L. (2006). Decision making and set shifting impairments are associated with distinct symptom dimensions in obsessive-compulsive disorder. *Neuropsychology*, *20*(4), 409–419. http://doi.org/10.1037/0894-4105.20.4.409

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and Decision Making. *Annual Review of Psychology*, *66*(1), 799–823. http://doi.org/10.1146/annurev-psych-010213-115043

Lynn, S. K., & Barrett, L. F. (2014). “Utilizing” Signal Detection Theory. *Psychological Science*, *25*(9), 1663–1673. http://doi.org/10.1177/0956797614541991

Maddox, W. T., & Bohil, C. J. (1998). Base-rate and payoff effects in multidimensional perceptual categorization. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, *24*(6), 1459–1482. https://doi.org/10.1037/0278-7393.24.6.1459

Mataix-Cols, D. (2003). Declarative and procedural learning in individuals with subclinical Obsessive-Compulsive symptoms. *Journal of Clinical and Experimental Neuropsychology*, *25*(6), 830–841. https://doi.org/10.1076/jcen.25.6.830.16477

Mataix-Cols, D., do Rosario-Campos, M. C., & Leckman, J. F. (2005). A multidimensional model of Obsessive-Compulsive disorder. *American Journal of Psychiatry*, *162*(2), 228–238. https://doi.org/10.1176/appi.ajp.162.2.228

Mataix-Cols, D., Wooderson, S., Lawrence, N., Brammer, M. J., Speckens, A., & Phillips, M. L. (2004). Distinct neural correlates of washing, checking, and hoarding symptom dimensions in Obsessive-compulsive disorder. *Archives of General Psychiatry*, *61*(6), 564–576. https://doi.org/10.1001/archpsyc.61.6.564

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

Melloni, M., Urbistondo, C., Sedeño, L., Gelormini, C., Kichic, R., & Ibanez, A. (2012). The extended fronto-striatal model of obsessive compulsive disorder: convergence from event-related potentials, neuropsychology and neuroimaging. *Frontiers in Human Neuroscience*, *6*, 259. https://doi.org/10.3389/fnhum.2012.00259

Metin, B., Roeyers, H., Wiersema, J. R., van der Meere, J. J., Thompson, M., & Sonuga-Barke, E. (2013). ADHD performance reflects inefficient but not impulsive information processing: A diffusion model analysis. *Neuropsychology*, *27*(2), 193–200. https://doi.org/10.1037/a0031533

Morein-Zamir, S., Papmeyer, M., Gillan, C. M., Crockett, M. J., Fineberg, N. A., Sahakian, B. J., & Robbins, T. W. (2013). Punishment promotes response control deficits in obsessive-compulsive disorder: evidence from a motivational go/no-go task. *Psychological Medicine*, *43*(2), 391–400. https://doi.org/10.1017/S0033291712001018

Mulder, M. J., Bos, D., Weusten, J. M. H., van Belle, J., van Dijk, S. C., Simen, P., … Durston, S. (2010). Basic Impairments in Regulating the Speed-Accuracy Tradeoff Predict Symptoms of Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry*, *68*(12), 1114–1119. https://doi.org/10.1016/j.biopsych.2010.07.031

Mulder, M. J., Van Maanen, L., & Forstmann, B. U. (2014). Perceptual decision neurosciences–a model-based review. *Neuroscience*, *277*, 872-884.

Nielen, M. M. A., Veltman, D. J., de Jong, R., Mulder, G., & den Boer, J. A. (2002). Decision making performance in obsessive compulsive disorder. *Journal of Affective Disorders*, *69*(1–3), 257–260. http://doi.org/10.1016/S0165-0327(00)00381-5

Ninaus, M., Pereira, G., Stefitz, R., Prada, R., Paiva, A., Neuper, C., & Wood, G. (2015). Game elements improve performance in a working memory training task. *International Journal of Serious Games, 2*, 3-16.

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science*, *3*(5), 400–424. https://doi.org/10.1111/j.1745-6924.2008.00088.x

Öner, N., & Le Compte, A. (1985). *Durumluk – Sürekli Kaygı Envanteri El Kitabı*: İstanbul Bogaziçi Üniversitesi Yayınları.

Pauls, D. L., Abramovitch, A., Rauch, S. L., & Geller, D. A. (2014). Obsessive–

compulsive disorder: an integrative genetic and neurobiological perspective. *Nature Reviews Neuroscience*, *15*(6), 410–424. http://doi.org/10.1038/nrn3746

Pauls, D. L., Alsobrook, J. P. II., Goodman, W., Rasmussen, S., Leckman, J. F., (1995). A family study of obsessive–compulsive disorder. *The American Journal of Psychiatry,* 152*(1), 76–84.*

Pe, M. L., Vandekerckhove, J., & Kuppens, P. (2013). A diffusion model account of the relationship between the emotional flanker task and rumination and depression. *Emotion*, *13*(4), 739–747. https://doi.org/10.1037/a0031628

Phelps, E. A., Lempert, K. M., & Sokol-Hessner, P. (2014). Emotion and Decision Making: Multiple Modulatory Neural Circuits. *Annual Review of Neuroscience*, *37*(1), 263–287. http://doi.org/10.1146/annurev-neuro-071013-014119

Ploran, E. J., Nelson, S. M., Velanova, K., Donaldson, D. I., Petersen, S. E., & Wheeler, M. E. (2007). Evidence Accumulation and the Moment of Recognition: Dissociating Perceptual Recognition Processes Using fMRI. *Journal of Neuroscience*, *27*(44), 11912–11924. https://doi.org/10.1523/JNEUROSCI.3522-07.2007

Prins, P. J. M., Dovis, S., Ponsioen, A., ten Brink, E., & van der Oord, S. (2010). Does Computerized Working Memory Training with Game Elements Enhance Motivation and Training Efficacy in Children with ADHD? *Cyberpsychology, Behavior, and Social Networking*, *14*(3), 115–122. https://doi.org/10.1089/cyber.2009.0206

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., & Rouder, J. N. (1998). Modeling Response Times for Two-Choice Decisions. *Psychological Science*, *9*(5), 347–356. https://doi.org/10.1111/1467-9280.00067

Rocha, F. F. da, Alvarenga, N. B., Malloy-Diniz, L., & Corrêa, H. (2011). Decision-making impairment in obsessive-compulsive disorder as measured by the Iowa Gambling Task. *Arquivos de Neuro-Psiquiatria*, *69*(4), 642–647. https://doi.org/10.1590/S0004-282X2011000500013

Sachdev, P. S., & Malhi, G. S. (2005). Obsessive–compulsive behaviour: a disorder of decision-making. *Australian and New Zealand Journal of Psychiatry*, *39*(9), 757–763. https://doi.org/10.1080/j.1440-1614.2005.01680.x

Savasir, I., & Sahin, N. (1995). Wechsler cocuklar icin zeka olcegi (WISC-R). *Turkish Psychological Association, Ankara*

Scahill, L., Riddle, M. A., Mcswiggin-Hardin, M., Ort, S. I., King, R. A., Goodman, W. K., Cicchetti, D., & Leckman, J. F. (1997). Children's Yale-Brown Obsessive Compulsive Scale: Reliability and Validity. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(6), 844–852. http://doi.org/10.1097/00004583-199706000-00023

Shaw, R., Grayson, A., & Lewis, V. (2005). Inhibition, ADHD, and Computer Games: The Inhibitory Performance of Children with ADHD on Computerized Tasks and Games. *Journal of Attention Disorders*, *8*(4), 160–168. https://doi.org/10.1177/1087054705278771

SnowDots [MATLAB code]. (2012). Retrieved from http://www.med.upenn.edu/goldlab/Goldlabcode.shtml

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the state-trait anxiety inventory*. Palo Alto, CA: Consulting Psychologists Press.

Starcke, K., Tuschen-Caffier, B., Markowitsch, H. J., & Brand, M. (2010). Dissociation of decisions in ambiguous and risky situations in obsessive–compulsive disorder. *Psychiatry Research*, *175*(1–2), 114–120. https://doi.org/10.1016/j.psychres.2008.10.022

Toffolo, M. B. J., Hout, M. A. van den, Hooge, I. T. C., Engelhard, I. M., & Cath, D. C. (2013). Mild Uncertainty Promotes Checking Behavior in Subclinical Obsessive-Compulsive Disorder. *Clinical Psychological Science*, 2167702612472487. https://doi.org/10.1177/2167702612472487

Valleni-Basile, L. A., Garrison, C. Z., Jackson, K. L., Waller, J. L., McKeown, R. E., Addy, C. L., & Cuffe, S. P. (1994). Frequency of Obsessive-Compulsive Disorder in a Community Sample of Young Adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, *33*(6), 782–791. https://doi.org/10.1097/00004583-199407000-00002

Van Oppen, P., Hoekstra, R. J., & Emmelkamp, P. M. G. (1995). The structure of obsessive-compulsive symptoms. *Behaviour Research and Therapy*, *33*(1), 15–23. https://doi.org/10.1016/0005-7967(94)E0010-G

Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods*, *40*(1), 61–72. https://doi.org/10.3758/BRM.40.1.61

Wald, A. (1947) Sequential analysis. Wiley, NewYork

Wechsler, D. (1974). Manual for the Wechsler Intelligence for Children-Revised. *New York: Psychological Corporation*.

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010a). Anxiety enhances threat processing without competition among multiple inputs: A diffusion model analysis. *Emotion*, *10*(5), 662–677. https://doi.org/10.1037/a0019474

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010b). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology*, *54*(1), 39–52. https://doi.org/10.1016/j.jmp.2010.01.004

Wiecki, T. V., Poland, J., & Frank, M. J. (2015). Model-Based Cognitive Neuroscience Approaches to Computational Psychiatry Clustering and Classification. *Clinical Psychological Science*, *3*(3), 378–399. https://doi.org/10.1177/2167702614565359

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: hierarchical bayesian estimation of the drift-diffusion model in python. *Frontiers in Neuroinformatics*, *7*(August), 1–10. https://doi.org/10.3389/fninf.2013.00014

Wu, K. D., & Cortesi, G. T. (2009). Relations between perfectionism and obsessive–compulsive symptoms: Examination of specificity among the dimensions. *Journal of Anxiety Disorders*, *23*(3), 393–400. https://doi.org/10.1016/j.janxdis.2008.11.006

Yucelen, A. G., Rodopman-Arman, A., Topcuoglu, V., Yazgan, M. Y., & Fisek, G. (2006). Interrater reliability and clinical efficacy of Children's Yale-Brown Obsessive-Compulsive Scale in an outpatient setting. *Comprehensive Psychiatry*, *47*(1), 48–53. http://doi.org/10.1016/j.comppsych.2005.04.005

Supplement I – “Shoot” game code

For the open source software please visit http://www.scipy.org/ (Jones, Oliphant & Peterson, 2001)

entities_run.py

```
import pygame, math, random, scipy
from scipy.stats import norm as normaldist
from pygame.sprite import Sprite
from openexp.keyboard import keyboard

####### Entities_Run ##############################################################
# Declerations of experiment object classes
# and settings/configurations.
# This code is run in the "run" loop of the experiment.
# It occurs before experiment presentation.
# Therefore, it does not hinder experiment runtime performance.
#
# Object Classes declared:
# Target - Shows the location to where the shot is headed
#
# Goal - The goal area and graphics
#
# Goalkeeper - The goalkeeper object, its motion properties and checking
# if goals are scored. Nested under the goal object
#
# Feedback - Presentation of textual feedback on the top of the screen
#
# Ball - Ball object which travels towards target location at set velocity
#
# Vector - Generic object for vector calculations. Useful in motion.
#
###### CONFIGURATION SETTINGS ###################################


# Path of the experiment graphics folder. Must include the final '/' for the adress.
# Linux or Windows filepaths supported. kkaramanci will be replaced by the
experimenter's username
path = "/Users/kkaramanci/Documents/Shoot/assets/"

# RGB color codes for colors used in experiment for easy setup and reference
red = (255,0,0)
green = (0,255,0)
blue = (0,0,255)
darkBlue = (0,0,128)
white = (255,255,255)
black = (0,0,0)


###### EXPERIMENT PARAMETER SETTINGS ##############################################
# Settings dictionaries that translate worded settings to parameter values in code.
# Parameters may be altered to fit experiment or new ones may be defined
# But dictionaries should be used in setting up experiment sequences & trials
# rather than writing directly in code to ensure proper referenced use throughout
# experiment and standardization over users of the experiment.
##################################################################################
DIFFICULTY = {
    'easy' : 1,
    'medium' : 2,
    'hard' : 3
    }
SPEED = {
    'slow' : 1,
```

```
    'medium' : 3,
    'fast' : 5
    }

SHOOT_SPEED = {
    'slow' : 5,
    'medium' : 7,
    'fast' : 10
    }

##### Initialize global experiment variables###############
# Initialize the previous agency variable for the goalkeeper.
# This defines the level of control the subject has on scoring
# a goal in the previous trial
exp.set('prev_agency', -1)

#The time limit after which the experimental session ends
exp.set('time_limit', 1200000) #20 mins


class Target(Sprite):
"""
A target object that is displayed to indicate where a shot will land.
"""
    def __init__(self, position, *groups):
    """
    Object is initialized at the 0,0 location and is hidden.
    """
        Sprite.__init__(self, *groups)
        self.image = pygame.Surface([0, 0])
        self.image.set_colorkey(black)
        self.rect = self.image.get_rect()
        self.rect.x = position[0]
        self.rect.y = position[1]

    def hide(self):
        self.image = pygame.Surface([0, 0])

    def show(self):
    """
    Reveals the object at the current location as set in the object Sprite
coordinates.
    """
        self.image = pygame.image.load(path + "target.png").convert()
        self.image.set_colorkey(black)
        self.rect = self.image.get_rect()


    def update(self):
        pass

class Goal():
"""
This is the goal object which displays the goal graphics and serves as an anchor
container for other objects.
Contains:
- A Target object: Target operates within the bounds of the goal rectangle
- A Goalkeeper object: Goalkeeper operates within the bounds of the goal rectangle

Arguments:
- position: x,y coordinates of the goals location on screen
- agency: Defines the agency of the subject. Does the subject have control over the
shot outcome? (1-agency) is the probability of a random result (goal, nogoal)
occurring, instead of the subject’s choice.
- deviation: Sets the deviation from the center of goal, of the contained goalkeeper
object
```

```
Other Notes:
- rect: This is the Sprite argument that defines the rectangular bounds of the goal
area


"""
    def __init__(self, position, agency, deviation, *groups):
    """
    Initializes the Target and Goalkeeper objects at the relevant location and
settings.
    Target is initialized hidden in the goal rectangle
    Goalkeeper is initialized with a set agency (whether subject has control on
outcome), at the central x,y
    coordinates of the goal area with set deviation from the center of the goal
    """
        self.rect = pygame.Rect(0,0,320,200)

        #self.image = pygame.Surface([0, 0])
        #self.image.set_colorkey(black)
        #self.image.fill(white)
        #self.rect = self.image.get_rect()

        self.rect.x = position[0]
        self.rect.y = position[1]

        self.agency = agency
        self.deviation = deviation

        self.target = Target(position, *groups)
        self.goalkeeper = Goalkeeper([self.rect.centerx,self.rect.centery], self.rect,
self.agency, self.deviation, *groups)

    def paint_target(self, new_pos = ()):
    """
    Reveals the target at new_pos or at a random location within the goal area if
new_pos not supplied

    Arguments:
    - new_pos: x,y coordinates of the location to show the Target
    """
        self.target.show()
        #If new_pos exists, set x coordinate to x coordinate of new_pos. y coordinate
is random
        if new_pos:
            new_x = new_pos[0]
            new_y = self.rect.top + random.randrange(self.rect.height -
self.target.rect.height)
        #If new_pos does not exist, set x,y to random positions in goal rect area
        else:
            new_x = self.rect.left + random.randrange(self.rect.width -
self.target.rect.width)
            new_y = self.rect.top + random.randrange(self.rect.height -
self.target.rect.height)
        self.target.rect.x = new_x
        self.target.rect.y = new_y


    def update(self):
        pass


class Feedback(Sprite):
    """
    Feedback object displays text messages on the screen.
    """
```

```
    def __init__(self, position, *groups):
    """
    Initializes the text position and color

    Arguments:
    - position: (x,y) coordinates of position on screen
    """
        Sprite.__init__(self, *groups)
        self.position = position
        self.font = pygame.font.SysFont("monospace", 36, bold=True)
        self.image = pygame.Surface([0, 0])

        # initial color
        self.image.fill(black)
        self.rect = self.image.get_rect()
        # position and size of initial feedback
        self.rect.x = self.position[0]
        self.rect.y = self.position[1]
        self.rect.width = self.image.get_width()
        self.rect.height = self.image.get_height()

    def showText(self, text, color):
    """
    Shows feedback text
    Arguments:
    - text: the text message to display
    - color: text color to display
    """
        self.image = self.font.render(text, 1, color)
        self.rect.width = self.image.get_width()
        self.rect.height = self.image.get_height()
        pass


    def hideText(self):
    """
    Hides the current feedback message
    """
        self.image = pygame.Surface([0, 0])
        pass

    def update(self):
        pass

class Ball(Sprite):
    """
    Ball object displays a ball on screen and handles the smooth motion of the ball
    """
    def __init__(self, position, *groups):
    """
    Initializes the ball position and initial velocity & destination/target to 0.
    Arguments:
    - position: initial position of the ball
    """
        Sprite.__init__(self, *groups)
        self.velocity = 0
        # load the image of the ball
        self.image = pygame.image.load(path + "sball.png")
        self.rect = self.image.get_rect()
        # set the initial position of the ball
        self.rect.x = position[0]
        self.rect.y = position[1]
        # initial velocity set to 0
        self.velocity_vec = [0, 0]
        self.position_vec = [0, 0]
        self.position_init = [0, 0]
```

```
        self.position_vec[0] = position[0]
        self.position_vec[1] = position[1]
        self.position_init[0] = position[0]
        self.position_init[1] = position[1]
        # initial target destination for ball is empty
        self.target = None

    def update(self):
    """
    Update function runs once per frame. At each frame a new position for the ball is
calculated to simulate ball movement
    """
        # Determine the direction from current position to the ball target
        self.dir = self.get_direction(self.target)
        # before updating position check to see the ball is not at the target and has
a velocity and direction
        if self.dir and self.velocity and not self.in_target():
            self.position_vec[0] += (self.dir[0] * self.velocity) # calculate velocity
vector x component from direction and base velocity
            self.position_vec[1] += (self.dir[1] * self.velocity) # calculate velocity
vector y component from direction and base velocity
            self.rect.center =
(round(self.position_vec[0]),round(self.position_vec[1])) # move the center of the
object for 1 time unit by the x,y vectors
        else:
            # if ball is not moving or at target/destination, reset velocity to 0 and
target to None
            self.velocity = 0
            self.target = None


    def shoot(self, shoot_to = None):
    """
    Ball is assigned a new target/destination and velocity
    """
        if shoot_to:
            self.target = shoot_to
            self.velocity = configuration['shoot_speed']

    def in_target(self):
    """
    Determines if the ball is at the target location. Returns True/False
    """
        return self.target.contains(self.rect)



    def get_direction(self, target):
    """
    Get the vector direction from object to target
    """
        if self.target:
            # create a vector from center x,y value for object
            position = Vector(self.rect.centerx, self.rect.centery)
            # create a vector from center x,y value for target
            targetv = Vector(target.centerx, target.centery)

            # get total distance between target and position
            self.dist = targetv - position
            # normalize
            direction = self.dist.normalize()
            return direction

    def reset(self):
    """
    Reset the ball position, vector and target to init values
```

```
    """
        self.velocity = 0
        self.velocity_vec = [0, 0]
        self.rect.x = self.position_init[0]
        self.rect.y = self.position_init[1]
        self.position_vec[0] = self.position_init[0]
        self.position_vec[1] = self.position_init[1]

        self.target = None


class Goalkeeper(Sprite):
    """
    Goalkeeper object displays and moves the goalkeeper on screen.
    """
    def __init__(self, position, goal_rect, agency, deviation, *groups):
    """
    Initialize goalkeeper parameters
    Arguments:
    - Position: x,y coordinates of goalkeeper
    - goal_rect: the goal rectangular area of tended goal
    - agency: Control subject has over shot outcome (the goalkeepers ability to save)
    - deviation: deviation of position from center of goal
    """
        Sprite.__init__(self, *groups)
        # load goalkeeper image
        self.image = pygame.image.load(path + "gk2.png")
        self.rect = self.image.get_rect()

        #position goalkeeper in goal area
        self.rect.centerx = position[0]
        self.rect.centery = position[1] + 30
        self.goal_rect = goal_rect
        self.velocity_vec = [0, 0]
        self.position_vec = [0, 0]
        self.position_init = [0, 0]
        self.position_vec[0] = position[0]
        self.position_vec[1] = position[1]
        self.position_init[0] = position[0]
        self.position_init[1] = position[1]
        self.target = None
        self.saving = False

        # randomly decide left or right deviation and alter center position by that
amount
        self.center = random.choice([self.goal_rect.centerx - (self.goal_rect.width /
deviation), self.goal_rect.centerx + (self.goal_rect.width / deviation)])
        #generate a normally distributed random variable with mean at deviated center
and std at 1/4 the width of goal
        self.norm = normaldist(self.center, self.goal_rect.width/4)
        # pick first target location from the normally distributed random variable
self.target = pygame.Rect(self.norm.rvs(1)[0], self.rect.centery,5,5)
        # set initial velocity for goalkeeper
        self.velocity = 5


    def update(self):
    """
    At every frame, update the goalkeepers position depending on target location and
velocity
    """
        # before a shot is taken, this section enables the goalkeeper to wander
        # according to a normal distribution
        if not self.saving:
            # get direction from current position to target location
            self.dir = self.get_direction(self.target)
```

```
            # check if velocity exists and not arrived at target
            if self.dir and self.velocity and not self.in_target():
                #set direction of horizontal velocity
                self.velocity = math.copysign(self.velocity, self.dir[0])
                # update new horizontal position by moving 1 time unit in direction at
velocity
                self.position_vec[0] += (self.velocity)
                self.rect.centerx = round(self.position_vec[0])

                # if goalkeeper arrives at or passes by target position, pick new
location
                # from normally distributed random variable
                if self.velocity < 0 and self.target.left > self.rect.left:
                    self.target.x = self.norm.rvs(1)[0]
                if self.velocity > 0 and self.target.right < self.rect.right:
                    self.target.x = self.norm.rvs(1)[0]

            else:
                self.target.x = self.norm.rvs(1)[0]

        # after the shot, if the goalkeeper will save the shot, he moves towards the
target location of the ball for the save
        # if he doesn't save, he stands still
        else:
            self.dir = self.get_direction(self.target)
            if self.dir and self.velocity and not self.in_target():
            # calculate speed from direction to move and speed constant
                self.position_vec[0] += (self.dir[0] * self.velocity)
                self.rect.centerx = round(self.position_vec[0])
            else:
                self.velocity = 0
                self.target = None

    def move(self, move_to = None):
    """
    Assigns a target location for the goalkeeper to move to and sets motion velocity
    """
        if move_to:
            self.save = True
            self.target = move_to
            self.velocity = 10


    def get_direction(self, target):
    """
    Get the vector direction from object to target
    """
        if self.target:
            # create a vector from center x,y value for object
            position = Vector(self.rect.centerx, self.rect.centery)
            # create a vector from center x,y value for target
            targetv = Vector(target.centerx, target.centery)

            # get total distance between target and position
            self.dist = targetv - position
            # normalize
            direction = self.dist.normalize()
            return direction


    def in_target(self):
    """
    Determines if the goalkeeper is at the target location. Returns True/False
    """
        return math.fabs(self.target.centerx - self.rect.centerx) <=
math.fabs(self.velocity)
```

```
class Vector():
    '''
    Generic vector object to handle direction, position, and speed
    '''
    def __init__(self, x, y):
        self.x = x
        self.y = y

    def __str__(self):
    # printing vectors in console
        return "(%s, %s)"%(self.x, self.y)

    def __getitem__(self, key):
        if key == 0:
            return self.x
        elif key == 1:
            return self.y
        else:
            raise IndexError(str(key)+" is not a vector key")

    def __sub__(self, o):
    # vector subtraction
        return Vector(self.x - o.x, self.y - o.y)

    def length(self):
    # get vector length
        return math.sqrt((self.x**2 + self.y**2))

    def normalize(self):
    # normalize vector by length
        l = self.length()
        if l != 0:
            return (self.x / l, self.y / l)
        return None
```

init_block_run.py

```
# experimental block start time is recorded for timed experiments. Quit signal is
reset.
exp.set('block_starttime', self.time())
exp.set('exp_exit_signal', 0)
```

shoot_once_prepare.py

```
####### shoot_once_prepare
###############################################################
# This section initializes objects, configurations and graphics during the preperation
of the experiment.
# Occurs before running a trial so that processing times do not affect experimental
timing
#
# Initializes the main Game singleton object, as well as the individual graphic assets
# Assigns initial positions to all assets on screen, as a function relative to screen
size
# The main run routine is also declared, which updates the asset positions, once per
frame
# The update() functions of each object is called once per frame
# Subject actions are also defined in the main routine
```

```
#


###### EXPERIMENT PARAMETERS & CONFIGURATION ###################################

# Load experiment parameters
# These are read in from trial block loop spreadsheets in the OpenSesame experiments
# may vary per trial or be fixed throughout experiment as setup in the experiment
i_speed = self.get('i_speed')
i_difficulty = self.get('i_difficulty')
i_agency = 1
i_shoot_speed = self.get('i_shoot_speed')
i_deviation = self.get("i_deviation")

# configurations to define default screen size
#(should be small enough to ensure compatibility across screens)
# configurations to locate settings dictionaries as declared in the entities_run ile
configuration = {
        'screen_size': (640,640),
        'speed': SPEED[i_speed],
        'difficulty': DIFFICULTY[i_difficulty],
        'shoot_speed': SHOOT_SPEED[i_shoot_speed],
        'agency': 1,
        'deviation': i_deviation,
        'bar_size':200,
        'sound_nogoal': path + 'nogoal.wav',
        'sound_goal': path + 'goal.wav',
        'sound': True
    }

class Game(object):
"""
Main singleton Game class which contains the entire experiment setting and graphics
"""
    def __init__(self, configuration, input_state):
    """
    Initialize all settings and the initial graphics environment. Indicate the
locations of
    each of the graphics assets
    """
        # initialize game clock
        self.game_clock = pygame.time.Clock()


        # initialize game states
        self.shooting = False
        self.scored = False
        self.not_scored = False
        self.cleanup = False
        self.correct = None

        self.configuration = configuration
        self.input_state = input_state
        # load background image
        self.background = pygame.image.load(path + "bg.jpg").convert()
        # initialize all sprites (game graphics assets which update once per frame)
        self.sprites = pygame.sprite.OrderedUpdates()
        # Position the goal object and add it to the Game sprites list
        self.goal = Goal((configuration['screen_size'][0]/4,
                         configuration['screen_size'][1]/3),
                         configuration['agency'],
                         configuration['deviation'],
                         self.sprites)
        # Position the feedback object and add it to the Game sprites list
        self.feedback = Feedback((configuration['screen_size'][0]*7/18,
```

```
                                configuration['screen_size'][1]*2/10), self.sprites)
        # Position the ball object and add it to the Game sprites list
        self.ball = Ball((self.goal.rect.centerx,
                                configuration['screen_size'][1]*3/4), self.sprites)
        self.ball.rect.centerx = self.goal.rect.centerx

        self.agency = configuration['agency']
        self.sound_nogoal = pygame.mixer.Sound(configuration['sound_nogoal'])
        self.sound_goal = pygame.mixer.Sound(configuration['sound_goal'])
        self.reset_game(random.random()<0.5)
        self.running = True


    def play_sound(self, sound):
    """
    Function that plays a sound
    """
        if self.configuration['sound']:
            sound.play()

    def reset_game(self, serveLeft=True):
        pass

    def update(self):
    """
    Main update function for the game. Runs once per frame.
    Triggers all other object update functions.
    """
        # If game is in cleanup state
        if self.cleanup:
            self.running = False
            self.shooting = False
            self.scored = False
            self.not_scored = False
            self.cleanup = False
            self.feedback.hideText()
            #self.slider.reset()
            #self.ball.reset()
            #self.goal.paint_target(self.slider.get_pos())

        # If the game is in a shooting state, we will check for when it arrives at
target.
        # Game will be reset 1s after arriving at target at the conclusion of the
trial.
        elif self.shooting:
            if (self.ball.in_target()):
                self.shooting = False
                now = pygame.time.get_ticks()
                self.release_time = now + 1000

        # If the game is in a scored state, the appropriate feedback is shown and
correct states are set
        # Once the 1s timer expires (which was set at the end of shooting state), the
trial is reset
        elif self.scored:
            self.feedback.showText("GOL!!!", green)
            #self.sound_goal.play()
            if(pygame.time.get_ticks() > self.release_time):
                self.scored = False
                self.running = False
                self.cleanup = True

        # If the game is in a not scored state, the appropriate feedback is shown and
correct states are set
        # Once the 1s timer expires (which was set at the end of shooting state), the
trial is reset
```

```
        elif self.not_scored:
            self.feedback.showText("KURTARDI", red)
            #self.sound_nogoal.play()
            if(pygame.time.get_ticks() > self.release_time):
                self.not_scored = False
                self.cleanup = True

        # If a keyboard input is entered and is one of the correct inputs, game enters
"shooting" state
        elif self.input_state['key'] == pygame.K_a or self.input_state['key'] ==
pygame.K_d:
            #now = pygame.time.get_ticks()
            #self.release_time = now + 1500

            self.shooting = True

            #check to see if goalkeeper position allows for a save, given the keyboard
input
            saved = self.saved()

            # If not saved, set the appropriate states
            if not saved:
                self.goal.goalkeeper.saving = True
                self.ball.shoot(self.goal.target.rect)
                self.scored = True
                self.correct = True
                self.goal.goalkeeper.move(pygame.Rect(self.goal.rect.centerx,
self.goal.rect.centery,1,1))

            # If saved, set the appropriate states and set the target of the
goalkeeper to match ball trajectory
            else:
                self.goal.goalkeeper.saving = True
                self.ball.shoot(self.goal.target.rect)
                self.goal.goalkeeper.move(self.goal.target.rect)
                self.not_scored = True
                self.correct = False


        # Trigger update functions of each object contained in the Game object. Occurs
once per frame
        self.sprites.update()

    def draw(self, display_surface):
        """
        This function draws all graphics onto the game screen in given positions and
settings
        """
        self.sprites.clear(display_surface, self.background)
        return self.sprites.draw(display_surface)

    def saved(self):
        """
        Checks if the input keyboard action is the right one, given the deviation of the
goalkeeper from the center
        Returns True/False
        """
        if self.input_state['key'] == pygame.K_a:
            self.goal.paint_target([self.goal.rect.x+25, 0])
            return self.goal.goalkeeper.center < self.goal.rect.centerx
        else:
            self.goal.paint_target([self.goal.rect.x+self.goal.rect.width - 50, 0])
            return self.goal.goalkeeper.center > self.goal.rect.centerx

    def is_correct(self):
        """
```

```
    Returns true if the keyboard action was the correct one for the trial
    """
        return self.correct


# initialize pygame module
pygame.init()
# initialize global clock
clock = pygame.time.Clock()

# initialize color RGB values
red = (255,0,0)
green = (0,255,0)
blue = (0,0,255)
darkBlue = (0,0,128)
white = (255,255,255)
black = (0,0,0)

# initialize display surface global variables and load background image
display_surface = win
output_surface = display_surface.copy()
output_surface = pygame.image.load(path + "bg.jpg").convert()

# initialize keyboard input state
input_state = {'key': None}
# create game singleton object
game = Game(configuration, input_state)

# Main routine that runs the experiment. Is called once per trial
def run():
    # initialize response time and t1 to help calculate response time
    rt = 0
    t1 = 0
    timestamp = 1

    #runs forever while game has not exited or trial not concluded
    while game.running:
        # Ensures that game runs at 60 frames per second (fps). Delays if faster.
        clock.tick_busy_loop(60)

        #Calculate and display fps every 2 seconds
        now = pygame.time.get_ticks()
        if timestamp > 0 and timestamp < now:
            timestamp = now + 2000
            print clock.get_fps()

        # Handler for game quit and keyboard press events
        for event in pygame.event.get():
            # quit game if quit event triggered
            if event.type == pygame.QUIT:
                game.running = False
            # trigger quit event if escape key is pressed
            elif event.type == pygame.KEYDOWN and event.key == pygame.K_ESCAPE:
                game.running = False
                exp.set('exp_exit_signal', 1)
            # ignore key press event if pressed keys are not relevant to experiment
            elif event.type == pygame.KEYDOWN and not(event.key == pygame.K_a or
event.key == pygame.K_d):
                pass
            # record the key that was pressed if relevant to the experiment. Calculate
and record response time
            elif event.type == pygame.KEYDOWN and not(input_state.get('key')):
                rt = self.time() - t1
                input_state['key'] = event.key
```

```
        # call update functions (occurs once per frame)
        game.update()
        display_surface.fill(black)

        # draw graphics (or redraw altered graphics) in the second (hidden) screen
        game.draw(output_surface)
        display_surface.blit(output_surface, (0,0))
        # present the hidden screen
        pygame.display.flip()
        # set the start time of the experiment, if not set. Used to calculate rt
        if (t1 == 0):
            t1 = self.time()

    # at the conclusion of the experiment, return the response time and result
    return rt, game.is_correct()
```

shoot_once_run.py

```
####### shoot_once_prepare
##############################################################
# This section runs the trial. Assets were previously initialized in preparation.
# Has minimal code to ensure timing accuracy.
#
#

# Set previous trial agency. Used in agency experiments.
exp.set('prev_agency', i_agency)

# Run the main trial function. Response time and accuracy are returned.
rt, cor = run()

# Response keypress, response time and accuracy are recorded
exp.set_response(response=input_state['key'], response_time=rt,correct=cor),
```

Reference:

Jones, E., Oliphant, E., Peterson, P. (2001) SciPy: Open Source Scientific Tools for Python, http://www.scipy.org/.

Supplement II - Simulation code

shootsim.py

```
"""
Likelihood Divergence Simulation of Samples for Shoot Task

This simulation samples from a normal distribution that matches the experiment
task distribution, deviated either to the right or left of center.
It then plots the sum of log likelihood ratios for the pdf of left
and right deviated normal distributions for each sample obtained.
"""

import numpy as np
import scipy.stats as st
import matplotlib.pyplot as plt



# mean and std for normal distributions match the experimental distributions
# for each of the easy, medium and hard settings of the experiment
meanL = [240,294,312]
stdL = [80, 80, 80]
meanR = [400,346,328]
stdR = [80, 80, 80]
taskLabel = ['Easy','Medium','Hard']
taskOrder = [0, 1, 2]

# number of samples and experimental runs to simulate
sampleCount = 100
roundCount = 10



# for each of easy, medium and hard tasks
for task in taskOrder:

# setup the normal distribution variables
    normL = st.norm(meanL[task], stdL[task])
    normR = st.norm(meanR[task], stdR[task])

# Generate 10 rounds at random to left or right of goal. 0 for L, 1 for R

    LorR = np.random.randint(2, size=roundCount)

# For each shooting round (with goalie to the L or R)
    for pick in LorR:

# Generate values for normal distribution to the left and right of goal
        simL = normL.rvs(sampleCount)
        simR = normR.rvs(sampleCount)

        if pick:      #1 is R, 0 is L

#plot the cumulative sum of log likelihood ratios for L and R normal  distributions
#likelihood accumulation resembles evidence accumulation model
               likeDistL = normL.logpdf(simR)
               likeDistR = normR.logpdf(simR)
               plt.plot(np.cumsum(likeDistR - likeDistL), color='blue')
        else:
               likeDistL = normL.logpdf(simL)
               likeDistR = normR.logpdf(simL)
               plt.plot(np.cumsum(likeDistR - likeDistL), color='red')

#Plot for each of easy, medium and hard tasks
```

```
plt.xlabel('Sample Number')
plt.ylabel('Cumulative Log Likelihood Ratio')
plt.title(taskLabel[task] + ' Task')
plt.savefig(taskLabel[task] + '.png')
plt.show()
```