Time Based Reward Maximization

By

Bilgehan Çavdaroğlu

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of Psychology
in Partial Fulfillment of the Requirements for
the Degree of Master of Arts
at the Koç University

JULY 2014

Koc University

Graduate School of Social Sciences and Humanities

This is to certify that I have examined this copy of a master's thesis by

Bilgehan Çavdaroğlu

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

examining committee have been made.

Committee Members:

Asst. Prof. Fuat Balcı

Asst. Prof. İlke Öztekin

Asst. Prof. Gülayşe İnce Dunn

Date:

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 Bilgehan Çavdaroğlu

ABSTRACT

Humans and animals time intervals from seconds-to-minutes with high accuracy but limited precision. Consequently, time-based decisions are inevitably subject to our endogenous timing uncertainty and thus require temporal risk assessment. Study 1 tested the temporal risk assessment ability of humans when participants had to withhold each subsequent response for a minimum duration to earn reward and each response reset the trial time. Premature responses were not penalized in the first group but were penalized (reward/2) in the second group. Participants tried to maximize reward within a fixed session time (over 8 sessions) by pressing a key. No instructions were provided regarding the task rules/parameters. Participants nearly tracked the optimal target inter-response times (IRTs) that changed as a function of the level of timing uncertainty and maximized the reward rate in both experiments. Acquisition of optimal target IRT was rapid and abrupt without any further improvement or worsening. In Study 2 the same behavioral paradigm is used except the penalty was equated to reward, and the minimum withhold duration was halved in the 4^{th} session to measure the adaptation to a new schedule after reaching the steady state. Participants were given trait anxiety scale, Barratt impulsivity scale, behavioral inhibition and behavioral activation scale, and Padua inventory for obsessive compulsive disorder to evaluate the relation between the personality and task performance. Most of the participants were able to adapt to the change in the fourth session. Participants with high behavioral inhibition and high Padua scores are shown to be closer to optimality. These results support the use of non-instructed differential reinforcement of low rates of responding (DRL) as a valuable tool in temporal decision-making literature.

Keywords: decision-making, interval timing, optimality, reward maximization, risk assessment, timing uncertainty

ÖZET

İnsanlar ve hayvanlar saniyelerden dakikalara kadar olan zaman aralıklarında yüksek doğruluğa ancak düşük kesinliğe sahiplerdir. Bundan dolayı zamana bağlı bütün kararlar kaçınılmaz olarak içsel zaman belirsizliğine sahip olmakta ve zamansal risk analizi gerektirmektedir. Birinci çalışmada insanlar ödül alabilmek için bir önceki tepkilerinden sonra bir başka tepki vermeden en az asgari bir sure beklemeleri gereken zamansal risk analizi deneyi ile test edilmişlerdir. İlk grupta asgari süreden önceki tepkiler cezalandırılmamış ancak ikinci grupta bu tepkiler katılımcıların toplam kazançlarının azalmasına neden olmuştur (birim ödül/2). Katılımcılar sekiz seans boyunca sabit oturum süresi içinde ödüllerini azamileştirmeye çalışmışlardır. Deneyin başlangıcında ve devamında katılımcılara yönerge verilmemiş veya deneyin parametreleri hakkında bilgilendirme yapılmamıştır. İki gruptaki katılımcılar da kazanabilecekleri azami ödüle yaklaşmış ve optimal sürelere çok yakın beklemiştirler. Deney kurallarının öğrenimi hızlı ve ani olmuştur, ilk öğrenmeden sonra iyileşme veya kötüleşme olmamıştır. İkinci çalışmada ise aynı deney paradigması kullanılmış ancak erken tepkilere verilen ceza kazanılan ödül miktarına eşitlenmiş ve adaptasyon süreçlerinin incelenebilmesi için dördüncü deneyde asgari bekleme süresi ilk üç oturumun yarısına indirilmiştir. Katılımcılara durumsal kaygı ölçeği, davranışsal inhibisyon ve davranışsal aktivasyon ölçeği, barratt dürtüsellik ölçeği ve Padua obsesif kompülsif bozukluk envanteri verilmiş, skorların deney performansı ile olan ilişkisine bakılmıştır. İkinci deney sonucunda çoğu katılımcının dördüncü oturumdaki azami süre değişikliğine uyum sağlayabildiği gösterilmiştir. Yüksek davranışsal inhibisyon ve yüksek Padua envanteri skorlarına sahip katılımcıların optimale daha yakın olduğu gözlenmiştir. Bu iki çalışmanın sonuçları yönergesiz DRL deneyinin zamansal karar verme literatürü için önemli bir araç olduğunu göstermiştir.

Anahtar Sözcükler: karar verme, aralık zamanlama, optimalite, ödül azamileştirme, riske bağlı karar verme, zamansal belirsizlik

ACKNOWLEDGEMENTS

I am very grateful to my thesis advisor, Assoc. Prof. Fuat Balcı for his guidance, encouragement and support. He has thought me to never limit myself with what is given to me and always look for more. I would not be able to grow into an independent researcher without his challenging supervision, methods, and feedback.

I am also deeply grateful to my other committee members, Assoc. Prof. İlke Öztekin and Assoc. Prof. Gülayşe İnce Dunn for their positive contributions, support and their willingness to teach.

I am greatly indebted to Prof. Zeynep Aycan, for her invaluable advice and guidance during my undergraduate and graduate education. Her encouragement was an important factor for me to choose a research career. I would also like to thank the psychology faculty of Koç University; who has always supported me in my undergraduate and graduate years.

I would like to thank my friends Başak, Filiz and Dilara and other lab members for always being there to genuinely support me during my graduate studies and Rana Kurt, Cana Çağlı and Gözde Ertekin for their help in data collection.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Time is an indispensible component of many daily decisions we face throughout our lives. These common, time-related decisions we make while boiling an egg or catching a bus may seem simple, but the optimal solutions for them require complex statistical analysis of temporal information. Solution to the optimality problem in these seemingly simple tasks requires information about the decision maker's endogenous timing uncertainty and payoff matrix of the task. Both humans and mice were shown to be nearly optimal in temporal decisions in the context of differential reinforcement of low rates of responding (DRL) task (Balcı et al., 2011). But human participants of this experiment were provided with instructions regarding task rules and given the opportunity to experience the minimum wait-time explicitly prior to testing, which limits the generalizability of the task to other animals. The purpose of the first study was to evaluate human participants on the DRL task without providing any instructions or prior experience with the minimum wait-time to constitute a better analogue to animal experiments and introduce penalty parameter to the optimality problem. The purpose of the second study was to investigate the effect of lowering the schedule (that required adaptation of decision strategies) on the optimality of the participants and test the relationship between the task performance and trait anxiety, risk taking, behavioral inhibition and behavioral activation, and obsessive compulsive disorder-related traits.

Humans and other animals are accurate in their timing but they are not precise; their responses are distributed around the timed duration but vary between trials. This variance follows scalar property, which dictates the variability in timed responses (i.e., their standard deviation) should be proportional to the timed duration (Gibbon, 1977). Thus, to optimize performance even in simple daily tasks such as waiting for a bus, humans need to be able to calculate their timing uncertainty. Furthermore they should be able to use it to figure out their expected gain by taking reward contingencies into account based on rather complex mathematical calculations. Decision making tasks that are contingent on timing uncertainty show that both healthy humans and non-human animals are able to optimize decisions according to the variance of their time based responses.(Balcı, Freestone, & Gallistel, 2009; Balcı et al., 2011; Jazayeri & Shadlen, 2010; Kheifets & Gallistel, 2012; Simen, Balcı, Cohen, & Holmes, 2011).

In the DRL task, only the responses emitted after a fixed amount of time has passed are rewarded. Giving a response terminates the trial, resetting the trial timer. Participants' inter-response times forms an inverse-Gaussian shaped density function. This density function results in non-linear increase in reward probability with increasing time because only the portion of the density that has higher values than the DRL schedule is rewarded (and task can be parameterized such that responses earlier than schedule are penalized) whereas the time cost increases linearly with IRT. Maximization of reward rate in this task ($p(reward)/mean(IRT)$) requires optimal tradeoff between the reward probability and inter-response time (IRT). More importantly, this optimal tradeoff depends on the participants' level of timing uncertainty ($\hat{\omega}$). The tradeoff between the time cost and the reward probability gives DRL task the ability to characterize speed-accuracy tradeoff (SAT) in the temporal decision-making domain. Reward rate in this task can be quantified as shown in the Equation 1 where R is reward magnitude, *T* denotes the DRL schedule, $\hat{t}$ is mean IRT, $\hat{\lambda} \geq 0$ is the Wald shape parameter which is $\hat{\lambda} = \frac{\hat{t}}{\hat{\omega}^2}$.

$$RR(\hat{t}) = t^{-1}\left(R \times \left(1 - waldcdf\left(T, \hat{t}, \hat{\lambda}\right)\right) + waldcdf\left(T, \hat{t}, \hat{\lambda}\right) \times P\right) \quad \text{[Equation 1]}$$

Closed form solution to the optimality problem under timing uncertainty is shown in Equation 2.

$$R - (R + P) \times \left(waldcdf\left(T, \hat{t}, \hat{\lambda}\right) + T \times waldpdf\left(T, \hat{t}, \hat{\lambda}\right)\right) = 0 \quad \text{[Equation 2]}$$

The first major addition of Study 2 was unannounced halving of the minimum withhold duration in the fourth session of the DRL task. This manipulation allowed the measuring of adaptation to the new schedule after the initial steady state DRL performance was reached. This design also made comparisons of steady state performance with different mean wait durations on the same subject possible.

The second major addition of Study 2 was the use of personality questionnaires and Padua Inventory. There has been a surge in the literature that investigates the connections between behavioral task performance, personality traits, cognitive disorders, and genetic factors (Blum, Oscar-Berman, Barh, Giordano, & Gold, 2013; Gizer, Ficks, & Waldman, 2009; Karg, Burmeister, Shedden, & Sen, 2011). While there have been serious endeavors in improving the genetic and clinical sides of these studies, behavioral tasks used are less than perfect in disassociating the theoretical concepts. Non-instructed DRL is a good candidate for classifying the decision-making differences in these disorders and personality groups. Through task performance, it is possible to characterize the acquisition of task representation based on response-outcome contingencies, calculate optimal performance of the participants and measure their sensitivity to the penalty/errors. The DRL performance is also sensitive to biological manipulations and differences. The effect of basal ganglia and limbic system manipulations on DRL performance in animals is well documented (Balcı, Meck, Moore, & Brunner, 2009; Bannerman et al., 1999; Cho & Jeantet, 2010; Costa, Bueno, & Xavier, 2005; MacDonald, Lepage, Eden, & Eichenbaum, 2011; Meck, 1988; Meck, Church, & Olton, 1984; Pellegrino & Clapp, 1971; Yin & Troger, 2011; Young & McNaughton, 2000).

Anxiety is considered as one of the key emotional systems that influences decision-making processes (Bechara & Damasio, 2005; Eysenck, Derakshan, Santos, & Calvo, 2007; Gray & McNaughton, 2003). Individuals with higher anxiety is known to react more to the presence of negative stimuli (Savitsky, Medvec, Charlton, & Gilovich, 1998) and they are more pessimistic in their judgments and predictions of future outcomes (Lerner & Keltner, 2000; Shepperd, Grace, Cole, & Klein, 2005). Dispositional Anxiety is found to be related to risk-aversive decision making (Maner et al., 2007; Maner & Schmidt, 2006) but the exact role of anxiety on behavioral performance is not clear. Studies show that students that have met generalized anxiety disorder (GAD) criteria learned the task contingency in IGT faster compared to control group (Mueller, Nguyen, Ray, & Borkovec, 2010) , and individuals with high trait anxiety perform better (Werner, Duschek, & Schandry, 2009) in some studies while they perform worse in others (Miu, Heilman, & Houser, 2008; Pajkossy, Dezső, & Zoltay Paprika, 2009). These contradictory findings might be caused by the performance measures of IGT, where high score indicates preference for decks that have low immediate gain but high long term gain to the decks that have high immediate gain but low long term gain. Non-instructed DRL task would be a new and possibly a better candidate as a behavioral task to quantify the punishment sensitivity because it is possible for participants to actively avoid penalty by simply waiting longer before responding. Based on the operationalized definition of anxiety in the literature, a positive correlation between the trait anxiety and average IRT is expected. Increased IRT will move the participants away from their optimal target duration, reducing their reward rate.

Behavioral Inhibition System (BIS) and Behavioral Activation System (BAS) scales are created by Carver and colleagues to measure the biopsychological personality dimensions with the same names proposed by Gray (Carver & White, 1994; Gray, 1981, 1982). According to the original theory, BIS is highly related to the punishment sensitivity and dispositional anxiety, measured by Trait Anxiety. On the other hand, the revised version of Gray's Reinforcement Sensitivity Theory (RST) postulates that BIS is related to resolving goal conflict (between approach and avoidance) and anxiety is a byproduct of this process (Corr, 2004; Gray & McNaughton, 2003; McNaughton, 2006). Theoretical assumptions of RST are also supported by behavioral task performance and EEG studies. Behavioral inhibition scores were positively correlated with ERN/Ne amplitudes in flanker task, reflecting enhanced punishment related activity in participants with high BIS scores (Boksem, Tops, Wester, Meijman, & Lorist, 2006). Similar results were found in the Go/No-Go task where participants scoring high on BIS scale showed higher FRN during monetary loss compared to monetary gain (De Pascalis, Varriale, & D'Antuono, 2010). Positive correlations between BIS score and average IRT in steady state is expected based on these findings and similarities between the trait anxiety and BIS definition. On the other hand, it is also possible for participants with high BIS scores to be closer to the optimality if the new conflict resolution conceptualization of BIS is considered as true.

Both Behavioral Activation System (BAS) scale and Barratt Impulsiveness Scale (BIS-11) is widely used to measure impulsivity in clinical and non-clinical groups (Carver & White, 1994; Stanford et al., 2009). BAS scale measures the activity of behavioral approach system proposed by Gray, and it is highly related to impulsivity and reward sensitivity (Carver & White, 1994; Gray, 1981). BIS-11 on the other hand is able to differentiate between three types of impulsivity, namely attentional, motor, and non-planning impulsivity in both clinical and sub-clinical populations (Patton, Stanford, & Barratt, 1995; Stanford et al., 2009). High BAS score is positively associated with increased brain activity in reward regions of ventral-striatum (VS) and orbito-frontal cortex (OFC) for appetizing food (Beaver et al., 2006); nucleus accumbens (NAcc), ventral tegmental area (VTA) and OFC during the processing of reward cues and positive stimuli (Camara, Rodriguez-Fornells, & Münte, 2010;

Carter, MacInnes, Huettel, & Adcock, 2009; Hahn et al., 2009). EEG studies corroborate on these findings, showing higher N1 amplitudes (<150ms) that are related to attention in appetitive stimuli but not in neutral stimuli in participants with high BAS scores (Gable & Harmon-Jones, 2013). This reactivity toward reward might translate into overestimation of the value of positive stimuli in calculations of decision-making circuit. This, based on reward calculations of DRL mentioned above, would shorten the mean IRT. In addition to the increased activity during processing of rewards, higher scores from BAS subscales except fun seeking increase P300 amplitude in oddball task, implying an increased ability to detect irregularity (Nijs, Franken, & Smulders, 2007). This result implies that subjects with higher BAS scores would be better at detecting the schedule change in the fourth session compared to low BAS participants.

Obsessive Compulsive Disorder (OCD) is characterized with major symptoms of obsessive thoughts and persistent repetitive behaviors called compulsions. Compulsive behaviors shows similarities to the habitual behaviors and imply a deficiency between the competition of goal-directed and habitual systems (Evans, Lewis, & Iobst, 2004; Gillan et al., 2011). Cognitive deficits are most visible in reversal learning tasks where OCD patients cannot adapt well to the rule change compared to the controls which seems to be caused by the abnormal action in orbito-striatal loop (Gillan et al., 2011; Remijnse et al., 2006). This deficiency in cognitive flexibility might be manifested as slow or incomplete adjustment to the new schedule in the fourth session. In addition to the over reliance to the stimulus-response based habitual responding, OCD patients are more harm avoidant compared to controls (Bejerot, Schlette, Ekselius, Adolfsson, & Knorring, 1998; Kusunoki et al., 2000; Lyoo, Lee, Kim, Kong, & Kwon, 2001; Pfohl, Black, Noyes Jr, Kelley, & Blum, 1990; Richter, Summerfeldt, Joffe, & Swinson, 1996) and more sensitive to punishment (Fullana et al., 2004). They were also found to be better at learning in avoiding negative stimuli than approaching positive stimuli in feedback based learning task (Endrass, Kloft, Kaufmann, & Kathmann, 2011) and showed higher error related negativity during the errors compared to correct responses in EEG (Gehring, Himle, & Nisenson, 2000). This evidence implies that OCD symptoms can be differentiated by punishment sensitivity parameters in our task and participants with higher OCD symptom scores might have impaired acquisition in the fourth session because of their stronger habitual responding tendencies.

The literature that uses dopamine and serotonin-related gene polymorphisms to investigate decision-making and reinforcement learning is rapidly increasing (Frank & Hutchison, 2009; Holmes, Bogdan, & Pizzagalli, 2010). Research with polymorphisms has an advantage of exploring the biological mechanism of human behavior at much lower costs than the alternatives. Depending on the results of these experiments, DRL could be a useful task to differentiate behavioral endotypes for these polymorphisms. This is particularly important because these polymorphisms are also risk markers for many psychological disorders such as depression, addiction and ADHD (Blum et al., 2013; Gizer et al., 2009; Karg et al., 2011). Despite these numerous advantages, relation between dopaminergic/serotonergic gene polymorphisms and temporal decision making in humans is not well-studied. Use of DRL in particular would be useful for classifying persistent patterns of risky genotypes through quantitative methods, and for comparing different alleles in dimensions of optimality and task acquisition.

STUDY 1:

1. Materials and methods

   a. Subjects

Twenty-four adult participants were tested; Group 1 (5 males and 7 females with mean age of 20.50 years, SD= 1.7) and Group 2 (3 males and 9 females with mean age of 19.16 years, SD= 2.1). Participants were recruited through a publically available newsletter published on the Koç University website. The experiment was comprised of eight 50-minute (fixed test duration) daily DRL sessions and one 20-minute working memory task given after DRL testing. Participants received monetary reward based on their performance in each DRL session and fixed monetary compensation for the working memory task. The experiment was approved by the Institutional Review Board at Koç University and all participants provided written consent prior to testing.

   b. Stimuli and apparatus

The visual stimulus consisted of a white square on black background. The square briefly changed its color to red or green to provide feedback after premature responses (errors) and responses emitted after the minimum wait time (correct responses), respectively. The display was generated in MATLAB on a Macintosh computer, using the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997). Responses were collected with a standard computer keyboard.

   c. Procedure

The DRL Task: Each participant was tested with either a 5 or a 10 second DRL schedule over eight sessions. In Group 1, participants earned a reward when they hit the space key after the minimum wait time. There was no penalty for responding earlier than this minimum time. In Group 2, participants earned a reward upon hitting the space key after the DRL schedule; however they were penalized for half of the reward amount if they responded prematurely. The only explicit instruction was to press the designated key for the opportunity to earn a reward and to try to maximize reward earned. In Group 1, participants were told that it was possible for them not to receive a reward upon a key press. In Group 2, participants were told that it was possible to earn reward or suffer penalty upon a key press. Participants were not provided any other instructions regarding the DRL task rules and parameters. On the contrary in the earlier work (Balcı et al., 2011) participants were provided with these critical instructions regarding the DRL task rules and parameters. For instance, participants were told that they would earn reward for their each response emitted following the minimum wait duration since their previous response and that this response would reset the trial clock. They were also told that any response prior to the minimum wait duration would reset the trial clock without the reward delivery. Participants were also provided with prior experience with the critical task parameter prior to DRL testing, namely the minimum withhold duration. Specifically, participants were presented with the time interval that was equal to the DRL schedule. They were allowed to reproduce this interval for 50 times and received parametric feedback regarding the accuracy of each reproduction. This provided prior experience with the DRL schedule in the earlier study, which was absent in the current study.

As in the earlier work (Balcı et al., 2011) participants were asked not to count. A secondary task was used to suppress chronometric counting during the DRL tests. At the beginning of each block, participants were presented with a four-digit number and at the end of that block a single digit number. Participants were asked whether the four digit number presented at the beginning of the block contained the single digit number presented at the end of the block. The total reward earned from the DRL task-related responses was multiplied by the proportion correct from the secondary task. Participants were told that the reward earned from the primary task was going to be multiplied with the proportion correct in the secondary task. Working memory task was the automated version of the operational span task as described in (Unsworth, Heitz, Schrock, & Engle, 2005).

d. Data Analysis:

Cumulative Weibull distribution functions (with an extra scaling parameter) were fit to the inter-response times (IRTs) ordered according to their actual order of occurrence. The onset of steady state responding was defined in terms of the response that corresponded to the mean value plus three times the standard deviation estimated from the best-fit cumulative Weibull distribution. In order to quantify the abruptness of acquisition, the time it took to reach from 25% to 75% of the best-fit scaling parameter was calculated. Acquisition of steady-state performance by one participant in Group 1 (ID:19) and three participants in Group 2 (ID:27, ID:29 and ID:39) exhibited atypical patterns. Specifically, these participants tended to respond after the DRL schedule similar to other participants, but they initially waited much longer than the optimal IRT and then slowly converged on the optimal value gradually by speeding up their responses. Thus, these participants were excluded from the acquisition analysis (for these participants the last three sessions were treated as the steady state data for the other analyses). Note that participants were excluded only from the acquisition analysis and not from the optimality analysis that is described next.

Steady state IRTs were fit with an exponential inverse-Gaussian mixture function that has been previously shown to account for inter-response times in the DRL task (Balcı et al., 2011). There were atypical response patterns of several participants at the session-level during the steady state. Individual sessions with such atypical responses were excluded from the analysis for four participants in Group 1 and two participants in Group 2 (participants were not excluded from the analysis). In order to insure that results gathered were not due to the exclusion of participants from the acquisition analysis and sessions from the optimality analysis we estimated the parameters once more without any exclusions. All of the parameters (Speed and Abruptness of acquisition indices, optimal and empirical IRTs, CVs (coefficient of variance) and maximum possible expected reward rates) were not significantly different from the values obtained without any omissions in either experiments (at alpha level .05, not reported).

Best-fit mean and shape parameter of the inverse-Gaussian portion of the mixture distribution were used to calculate the optimal strategy for the corresponding participant. Mann–Whitney U tests were used for the comparison of the acquisition indices (i.e. rapidness and abruptness), IRTs, and CVs between experiments and schedules within each experiment. Wilcoxon signed ranks test was used for all other analysis. Results based on t-test comparisons revealed the same results (not reported). Alpha level of .05 was used as the significance level in all analyses.

2. Results

e. Acquisition

Acquisition of the DRL responding was characterized first, which was particularly relevant given the lack of instructions and multiple session testing in the experiment. To that end, two different measures of acquisition; rapidness and abruptness of the acquisition of optimal/steady state IRTs was calculated.

(i) Rapidness (speed) of acquisition:

Average onset of steady state occurred around the 19th (SEM=5.87, median 18.90, IQR=33.00) minute of Group 1 (with no penalty for errors) and 6th (SEM=2.86, median 0.60, IQR=14.30) minute of Group 2 (with penalty for errors) and this difference was statistically significant, (Z=2.17, p<.05). Rapidness of acquisition did not differ between the DRL schedules (5 and 10 seconds) in either Group 1(Z=1.28, p=.25) or Group 2 (Z=.98, p=.41). There was a significant negative correlation between working memory span and onset of steady state in Group 1, $r(9)$=-.59, p<.05. Although, in the same direction this relation was not significant in Group 2, $r(7)$=-.11, p=.38.

(ii) Abruptness of acquisition

The mean abruptness index normalized by the schedule was 14.28 (SEM=6.39, median 4.54, IQR=13.70) in Group 1 while it was 8.32 (SEM=3.99, median=2.21, IQR=15.11) in Group 2. This difference between two groups was not significant (Z=.65, p=.52). There was no significant correlation between working memory span and abruptness in either Group 1, $r(11)$=-.33, p<.16 or Group 2, $r(9)$=.05, p=.45.

f. Steady State Responding:

IRTs after acquisition (Mean+3×SD) were evaluated for possible trend towards slowing or speeding. In both Group 1 and Group 2, the group average slope of the linear regression fits to post-acquisition data points were 0.00 (SEM=0.00, median 0.00, IQR=0.00) suggesting that IRTs remained very steady after the acquisition took place. An exponential-inverse Gaussian mixture distribution function was fit to steady state IRTs in Group 1 (mean omega squared =.91, SEM=.03, median=.95, IQR=.09) and in Group 2 (mean omega squared =.91, SEM=.02, median=.94, IQR=.08). When an exponential-Gaussian mixture distribution function was fit to the same dataset, the omega squared values decreased to (mean=.88, SEM=.03, median=.93, IQR=.11) and (mean=.85, SEM=.03, median=.88, IQR=.09) for Group 1 and Group 2, respectively. Wilcoxon signed ranks test showed that this difference was significant for Group 1 Z= 2.90, p<.01 and Group 2 Z=3.06, p<.01.

g. Optimality Analysis:

The optimality analysis of steady state responding in both experiments showed that participants aimed for the optimal IRT that was parameterized by the payoff structure and participants' timing uncertainty level. Figure 2A and 2B depict the performance of each participant tested with 5 or 10 second schedules in Group 1 and Group 2, respectively. These figures show the heat map of the expected reward rates (for normalized DRL schedules) expressed over a parameter space composed of target IRT and the level of timing uncertainty (CV). Ridges of these two "surfaces" are indicated by the black curves, namely the optimal performance curves for the DRL task with two different payoff structures. Optimal performance curves indicate how long participants should aim to wait for (normalized by DRL schedule) before responding again given their level of timing uncertainty and payoff structure.

Participants earnings compared to how much they could maximally earn given these endogenous and exogenous parameters was calculated. In Group 1, participants achieved 99.1% (SEM=.41%, median=99.8%, IQR=1.26%) of the maximum possible expected reward rate for their level of timing uncertainty. In Group 2, this value was 98.6% (SEM=.62%, median=99.7%, IQR=1.95%). An analysis adopting a more conservative approach was also conducted. This analysis consisted of dividing the difference between the empirical expected reward rate and the expected reward rate if targeting the schedule (ER($\hat{t}$)-ER(T)) by the difference between the maximum possible reward rate and empirical expected reward rate if targeting the schedule (ER($\hat{t}_o$)-ER(T), where $\hat{t}_o$ is the optimal IRT). The average of this conservative estimate of proportion of the maximum possible reward rate was 97.9% (SEM=.95%, median= 99.5%, IQR=3.62%) in Group 1 and 98.0% (SEM=.86%, median=99.6%, IQR=2.74%) in Group 2. The proportions of maximum expected reward rates gathered without excluding the sessions with atypical response patterns were nearly identical to the values reported above (means ranging between 97.8%-99.1% and medians ranging between 99.2%-99.7%). Briefly, participants nearly maximized their rewards in both experiments. Participant's reward rates were significantly higher than the reward rates they would attain if their mean IRT was equal to the DRL schedule (if they were targeting the DRL schedule) in Group 1 (Z =3.06, p<0.01) and in Group 2 (Z=3.06, p<.01). There was no significant difference between the two schedules in terms of the percentage of maximum expected reward rate attained in either Group 1 or 2; (Z=.80, p=.48) and (Z=1.44, p=.18), respectively. Note that this comparison was conducted after normalization by the corresponding DRL schedule.

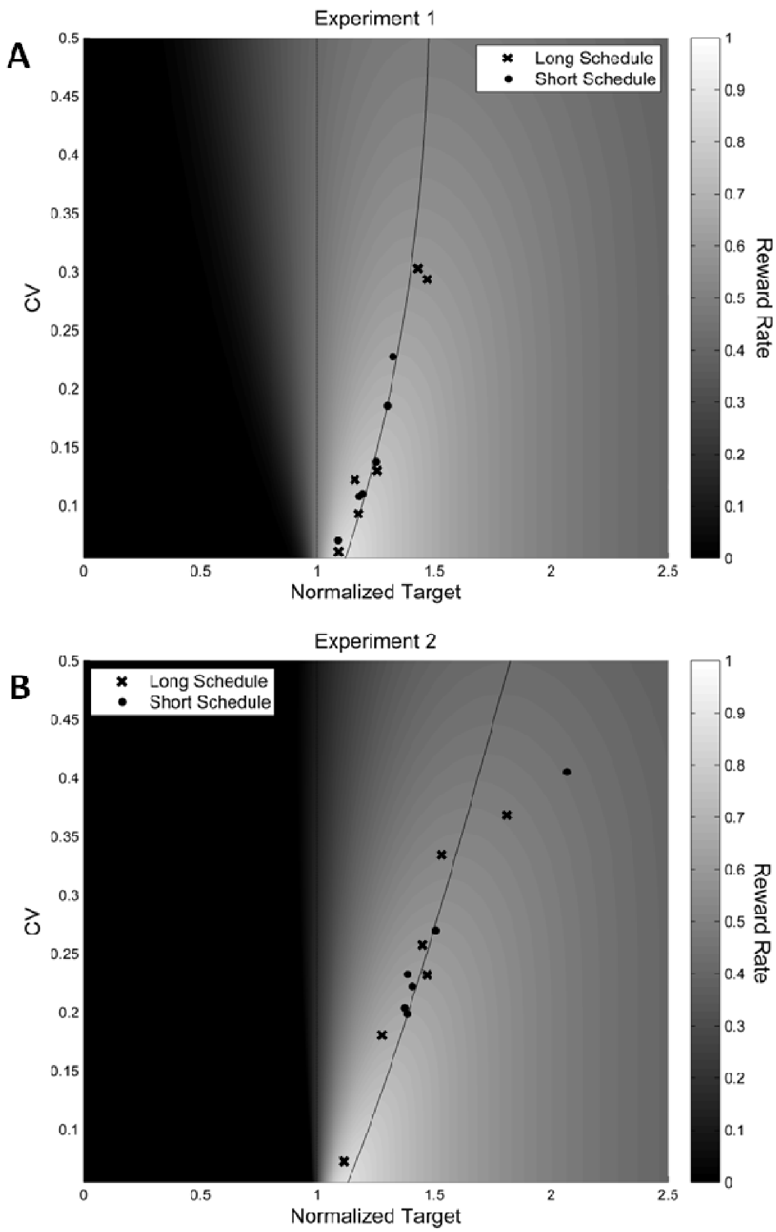

Figure 1. Heat map of expected reward rates for Group 1 (A) and Group 2 (B) for normalized DRL schedule. Curves denote the ridge of these surfaces indicating the optimal normalized target IRTs for different levels of timing uncertainty for no penalty (A) and penalty (B) conditions separately. Each data point corresponds to a participant and each symbol corresponds to a different schedule. Note that empirical IRTs were normalized by the corresponding DRL schedule. Short Schedule: 5s; Long Schedule: 10s.

Empirical normalized IRTs of the participants were compared with the corresponding optimal normalized IRTs using Wilcoxon signed rank test; there were no significant differences between empirical and optimal values in either Group 1 (Z=.86, p=.39) or Group

2 (*Z*=.71, *p*=.48). Normalized empirical IRTs in Group 2 were significantly longer than the normalized IRTs in Group 1 (*Z*=2.80, *p*<.005). Coefficient of variations obtained in Group 2 were significantly higher than the CVs obtained in Group 1 (*Z*=2.28, *p*<.05).

Deming regression fits revealed a significant relation between the optimal and empirical IRTs for both Group 1 (Figure 3A) and Group 2 (Figure 3B): $F(1,10)= 188.4$, $p<.001$, *Slope* =1.34 in Group 1 and $F(1,10)=61.46$, $p<.001$, *Slope*=1.77 in Group 2. When we excluded the participant that had empirical IRT longer than two standard deviations from the mean in Group 2, the slope decreased to 1.38, $F(9)=75.74$, $p<.001$. Note that best-fit regression lines crossed over the identity line in both Groups. This observation suggests that participants had a tendency to respond earlier than the optimal when the optimal IRT was closer to the DRL schedule and later than the optimal when the optimal IRT was farther from the DRL schedule. This pattern suggests an over-adjustment of IRTs in relation to the level of timing uncertainty. These small biases however resulted in only negligible costs in terms of the reward rate attained.

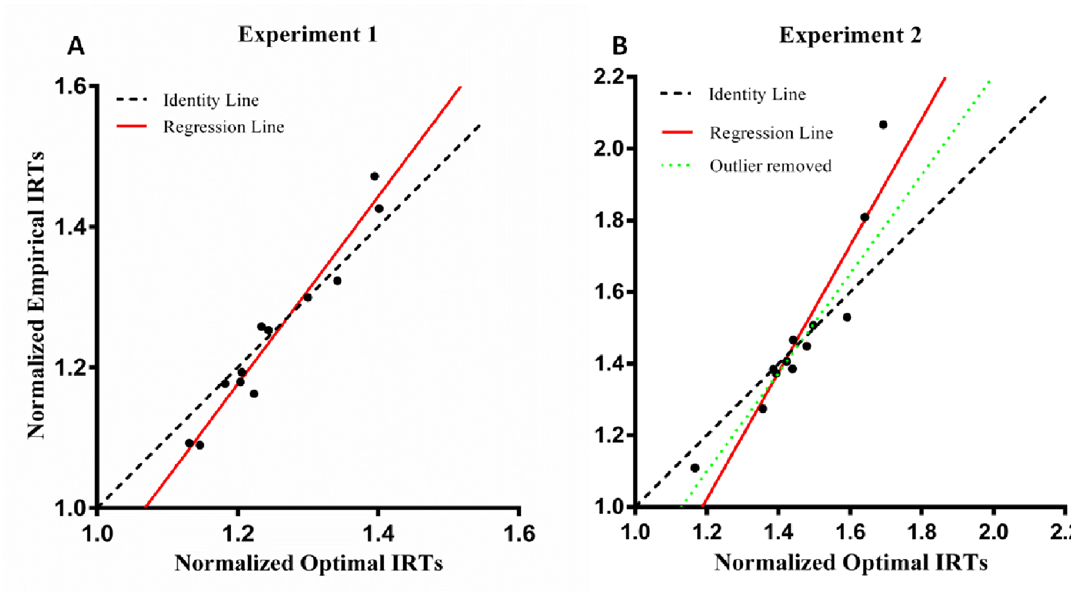

Figure 2. Deming regression fits to empirical and optimal IRTs in Group 1 (A) and Group 2 (B). Data from Group 2 were fit twice: 1- full data set and 2- after removing one participant based on mean -/+ two standard deviation exclusion rule.

3. Discussion

This study aimed to expand the scope of temporal decision-making research by addressing novel questions to bridge the gap between interval timing and decision-making fields. To that end, temporal risk assessment performance of human participants in the DRL task with two different schedules (5 and 10 seconds) and payoff structures (i.e., penalty vs. no penalty for premature responses) was investigated and evaluated it within the framework of optimality based on the statistical decision theory.

The results indicated that humans can maximize reward rate by taking normative account of their endogenous timing uncertainty even when instructions regarding the task rules were absent, when the exact minimum wait time itself was never experienced, and when premature responses were explicitly penalized. Observed performance was comparable, if not closer to optimality when compared with earlier single-session DRL experiments with instructions, prior experience of the minimum wait-time, and no penalty for errors (Balcı et al., 2011). Overall, the findings corroborated optimal (reward maximizing) performance of human and/or non-human animals in other timing tasks where time was an explicit component of the decisions; switch task (Balcı, Freestone, et al., 2009; Kheifets & Gallistel, 2012), temporal bisection task (Çoşkun, Sayali Ungerer, Emine, & Balci, In press 2014), beat-the-clock task (Simen et al., 2011), temporal reproduction (Jazayeri & Shadlen, 2010), and motor timing (Landy, Trommershäuser, & Daw, 2012). Different from these discrete-trial tasks however, the experiment demonstrated the optimality of temporal decisions when the task rules imposed a trade-off between the "speed" (i.e., response time) and accuracy (i.e., probability of reward) of temporal decisions, a relation that has been shown to adaptively guide decisions in other free-response non-temporal tasks (e.g., (Balcı et al., 2011)). Thus, findings also expanded the scope of optimal speed-accuracy tradeoffs to the domain of interval timing and temporal decision-making.

Acquisition of the optimal DRL responding in the absence of instructions occurred early during training (mostly within the first session) in an abrupt fashion. This constituted one of the fundamental differences from the DRL acquisition pattern of rats, which is typically gradual. After the acquisition took place (as captured by cumulative Weibull fits), the performance was very steady and nearly optimal given the DRL schedule, payoffs, and level of participants' timing uncertainty. These findings support a model-based guidance of human timed behavior in the DRL task and constitute a challenge for gradual reinforcement learning-based accounts of performance (see also (Simen et al., 2011)).

Earlier acquisition of timed responding when participants suffered monetary penalty (penalty>0) for premature responding was expected as this penalty would locally/transiently or globally motivate longer wait-times, which would presumably facilitate the acquisition of the DRL task. This is exactly what was observed; participants learned the optimal wait-time earlier in the second experiment (penalty>0) compared to the first experiment (penalty=0). It is of interest to examine if similar manipulations (e.g., in the form of time-out period for premature responses) would lead to same findings in non-human animals. There were no differences between schedules in terms of the rapidness of the acquisition of DRL responding, which does not corroborate the view that acquisition scales with the degree of temporal uncertainty (for review see (Gallistel & Gibbon, 2000)). This discrepancy can be attributed to the peculiar features of the DRL task such as the acquisition of response inhibition and the absence of discrete timing cues.

Since the acquisition of task representation depended on experienced response-outcome contingencies, higher working memory capacity was expected to facilitate the acquisition of optimal/steady state responding. This expectation was confirmed by Group 1; participants with higher working memory span acquired time response inhibition earlier. However, this relation did not hold for Group 2. It is possible that differences in the effect of penalty on acquisition (e.g., due to differential penalty processing) masked the effect of working memory span on the same measure in Group 2. It is also possible that the sample size was adequate for the timing study but not for the working memory task. Future studies can use larger sample sizes to investigate this relation.

Different from steady state performance in non-human animals, human participants exhibited lower proportion of untimed responses despite the purely experiential nature of the current experiments. These responses are typically emitted very shortly after the previous response. Balcı et al (Balcı et al., 2011) argued that non-timed responses might help agents detect beneficial alterations in environmental statistics (e.g., shift to a shorter DRL schedule). This would constitute a long-term adaptive strategy in unstable environments particularly given the minimal time cost exerted by these responses (Wearden & Culpin, 1998). The difference between rats and humans in terms of the frequency of non-timed responses might be related to their different expectations regarding the stability of the environmental conditions. Humans might be less willing to explore possible changes in the schedule based on their prior belief that task parameter values remain stable throughout the experiment. On the other hand, environmental statistics are less stable in nature, and for non-human animals and smaller organisms. Alternatively, these results can be explained by the superior inhibitory control of humans compared to the rats. These issues can be addressed empirically with experiments in which the DRL schedule (minimum wait-time for reward) unpredictably shortens and lengthens without signaling.

The current study has a number of methodological advantages over the previous studies with instructions and prior experience of the DRL schedule. Acquiring the task relevant parameters purely based on experienced response-outcome contingencies with no instructions or experience of task parameters minimizes the likelihood of adopting task-related top-down auxiliary processes (at least early in training). Investigations of optimal temporal decision-making often consider steady state performance. Our methodological approach on the other hand also allowed the characterization of acquisition of timed response inhibition in relation to optimality. Finally, lack of instructions minimized the gap between the human and non-human animal versions of the task, increasing the interspecies generalizability of the conclusions and emphasizing the translational nature of temporal risk-assessment.

Optimal performance under uncertainty appears to be a common feature of human and animal time-based decision-making (Balcı et al., 2011). This is not surprising as time is a determinant of the amount of reward earned in many biologically critical situations humans and non-human animals have faced in their evolutionary history. Given that interval timing is a primitive and fundamental function observed in many different species with similar psychophysical properties, its noise characteristics might have indeed been well-integrated into decision-making mechanisms over the course of evolution. In other words, the nervous system of many vertebrates might be pre-wired to parametrically convert the endogenous timing noise into an adaptive bias signal during decision-making when reward maximization requires it.

DRL task's usefulness in psychopharmacology and its ability to characterize impulsive behavior have attracted the attention of behavioral neuroscientists. The most prominent neuroanatomical target of these studies had been the limbic system, which has been also shown to be important region for timing (Balcı, Meck, et al., 2009; Bannerman et al., 1999; Cho & Jeantet, 2010; Costa et al., 2005; MacDonald et al., 2011; Meck, 1988; Meck et al., 1984; Pellegrino & Clapp, 1971; Yin & Troger, 2011; Young & McNaughton, 2000). For instance, hippocampus (Bannerman et al., 1999; Cho & Jeantet, 2010; Costa et al., 2005; Pellegrino & Clapp, 1971; Young & McNaughton, 2000) and amygdala (Pellegrino & Clapp, 1971) lesions have been shown to result in clear impairments of DRL performance in the form of increased responses per reinforcer. Unfortunately, many of these reports did not present the complete response curves, which does not enable the characterization of observed deficiencies differentially as being due to an increase in frequency of non-timed responses (Exponential portion of our mixture model) or due to a leftward shift of the timed response curve (Inverse-Gaussian portion of our mixture model) or both.

Few studies that presented the response curves revealed that lesions of dentrate gyrus, a subregion of hippocampal formation (Costa et al., 2005) resulted in a leftward shift in the timed portion of the response curve. Bilateral hippocampus lesions showed similar effects but they also increased the frequency non-timed responding (see Figure 3 in (Cho & Jeantet, 2010)). The leftward shift in IRT curves is consistent with the assumed role of hippocampus in timing accuracy (Balcı, Meck, et al., 2009; Meck et al., 1984; Yin & Troger, 2011). For instance, hippocampal lesions have been observed to result in leftward shifts in bisection curves (Meck et al., 1984) and peak response curves (Balcı, Meck, et al., 2009; Meck, 1988; Meck et al., 1984). Single cell recordings in DRL (Young & McNaughton, 2000) and other tasks with time gaps (MacDonald et al., 2011) also revealed time-dependent activity of hippocampal cells. For instance, a subgroup of hippocampal cells gradually decreased their firing rate over the course of the trial until the emission of the response and went back to basal levels after responding in the DRL task (Young & McNaughton, 2000).

On the other hand, the effect of hippocampal lesions on DRL performance might not be due to its effects on timing processes itself but primarily due to its disruption of inhibitory control over timed anticipatory responding. To that end hippocampus might affect DRL responding by modulating nucleus accumbens (NAc) activity via gating its cortical inputs (O'Donell, Greene, Pabello, Lewis, & Grace, 1999). Based on previous studies, NAc core and its target structures indeed appear as possibly crucial structures in timed response inhibition. For instance, lesions of NAc core but not NAc shell cause a leftward shift in DRL response curve, which become more apparent with longer DRL schedules (Pothuizen, Jongen□Rêlo, Feldon, & Yee, 2005). Biological and pharmacological manipulations of downstream regions of NAc core, such as blocking the NAc core-ventral pallidum (VP) GABAergic pathway (Wheeler, 2009) and lesions of subthalamic nucleus (STN) (Uslaner & Robinson, 2006) also impair DRL performance. The effect of other downstream regions on DRL performance is not well investigated but entopeduncular nucleus' (or GPi in humans) inhibitory role on action (and increase in premature responses with its inactivation) has been shown using lesions and inactivation methods (Baunez & Gubellini, 2010).

Excitatory and inhibitory connections between these basal ganglia structures and their effect on thalamus activity suggest that NAc core might modulate the output/manifestation of dorsal striatal temporal processing at multiple levels (Buhusi & Meck, 2005; Meck, 2005) (GPi/SNr and STN). Prefrontal inputs (e.g., anterior cingulate cortex, dorsal agranular insular areas, prelimbic cortex) to Nac core (Dalley, Cardinal, & Robbins, 2004; Pennartz, Ito, Verschure, Battaglia, & Robbins, 2011) are on the other hand potential candidates as the source of adaptive bias signal and its parameterization by timing uncertainty. Future neuroimaging studies can provide clues regarding the precise role of this network in adaptive timed response inhibition.

In summary, the results provided strong evidence for optimal temporal risk assessment performance of humans in a task that imposed a trade-off between the “speed” and “accuracy” of timed responses. This work further extended the scope of optimal temporal risk assessment performance to those conditions in which errors (i.e., premature responses) were penalized. These findings overall pointed at the robustness of reward maximization in the context of temporal decision-making. Importantly, optimal performance was observed in the absence of instructions and pre-training with task parameters, and purely based on experienced response-outcome contingencies constituting a better analogue of animal studies. Acquisition of optimal performance was rapid and abrupt; speed of acquisition was further facilitated by penalizing errors and higher working memory span. Future studies can investigate the neural correlates of timed response inhibition and its interaction with timing uncertainty focusing on the afferent and efferent projections of ventral striatum.

STUDY 2:

1. Materials and methods

   a. Subjects

Eighty-seven adult participants were tested and 75 of these participants had lowered schedule on their fourth session. Participants were recruited through a publically available newsletter published on the Koç University website. The experiment was comprised of four 50-minute (fixed test duration) daily DRL sessions. Participants received monetary reward based on their performance in each DRL. The experiment was approved by the Institutional Review Board at Koç University and all participants provided written consent prior to testing.

   b. Measures

The following questionnaires were completed by participants.

i. State-Trait Anxiety Inventory-Trait Form (STAI-T), an anxiety questionnaire that measures the dispositional anxiety independent from the transient daily conditions (Spielberger, Gorsuch, & Lushene, 1970). Item structure of the questionnaire was maintained in the Turkish version (Öner & Le Compte, 1985).
ii. BIS/BAS Scale, a questionnaire that measures the behavioral inhibition system, and behavioral approach system that underlies avoidance and approach behaviors (Carver & White, 1994; Gray, 1981, 1982). The factor structure of the scale and the drive, reward and fun seeking subscales of BAS was preserved in the Turkish version (Şişman, 2012).
iii. Barratt Impulsiveness Scale-11 (BIS-11), a widely used impulsiveness scale that is validated in many different clinical populations (Patton et al., 1995; Stanford et al., 2009). The factor structure of attentional impulsivity, motor impulsivity, and

non-planning impulsivity was preserved in the Turkish version (Güleç et al., 2008).

iv. Padua-R Inventory (PI-R), an OCD symptom questionnaire that was initially created with 60 items (Sanavio, 1988) but later revised to a shorter, 41-item, form (Oppen, Heokstra, & Emmelkamp, 1995). Turkish version uses the shorter form with impulses, washing, checking, rumination, and precision subscales, albeit with few item differences (Beşiroğlu et al., 2005).

c. Stimuli and apparatus

The visual stimulus consisted of a white square on black background. The square briefly changed its color to red or green to provide feedback after premature responses (errors) and responses emitted after the minimum wait time (correct responses), respectively. The display was generated in MATLAB on a Macintosh computer, using the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997). Responses were collected with a standard computer keyboard.

d. Procedure

The DRL Task: Participants were tested with 11.2 second DRL schedule over three sessions and 5.6 second schedule for the final fourth session. Participants earned a reward upon hitting the space key after the DRL schedule; however they were penalized for the reward amount if they responded prematurely. There were no further changes from the procedure of the first study.

e. Data Analysis:

%3.5 of the data from the personality questionnaires was missing. These missing values were replaced by the values obtained by using regularized expected maximization imputation method (Schneider, 2001).

Cumulative Weibull distribution functions (with an extra scaling parameter) were fit to the inter-response times (IRTs) ordered according to their actual order of occurrence similarly to the Study 1. But unlike the Study 1; steady state responding was defined as the start of the second session to minimize the effect of the atypical patterns in the first session. Nevertheless, the acquisition parameters were calculated using Cumulative Weibull fits. Dynamic range is calculated as the points between 10% and 90% of the best-fit scaling parameter. Abruptness is quantified as the dynamic range (Gallistel, Fairhurst, & Balsam, 2004). This measure of abruptness was normalized with the schedule (11.20) to be able to compare it with the abruptness in the acquisition of the fourth session. Onset of acquisition is calculated as the total duration between the beginning of the session and the start of dynamic range, while rapidness of acquisition is calculated as the total duration passed to reach end of the dynamic range. Figure 3A depicts an exemplar acquisition pattern, with significant variables shown. Acquisition of steady-state performance by one participant in Group 1 (ID:118) exhibited atypical patterns, this participant was removed from the acquisition analysis.

Change point algorithm was used to characterize acquisition of the task contingency in the fourth session (Gallistel et al., 2004). The data from the first three sessions and the data from the fourth session was pooled and fed into the change point algorithm. First change point in the fourth session with the rate smaller than the mean rate of the steady state was set as the onset of learning. The average rate following the onset of learning was considered as the asymptote. For the calculations of abruptness, asymptote was first subtracted from the mean of steady state in the first step (referred to as the Difference). In the second step, the

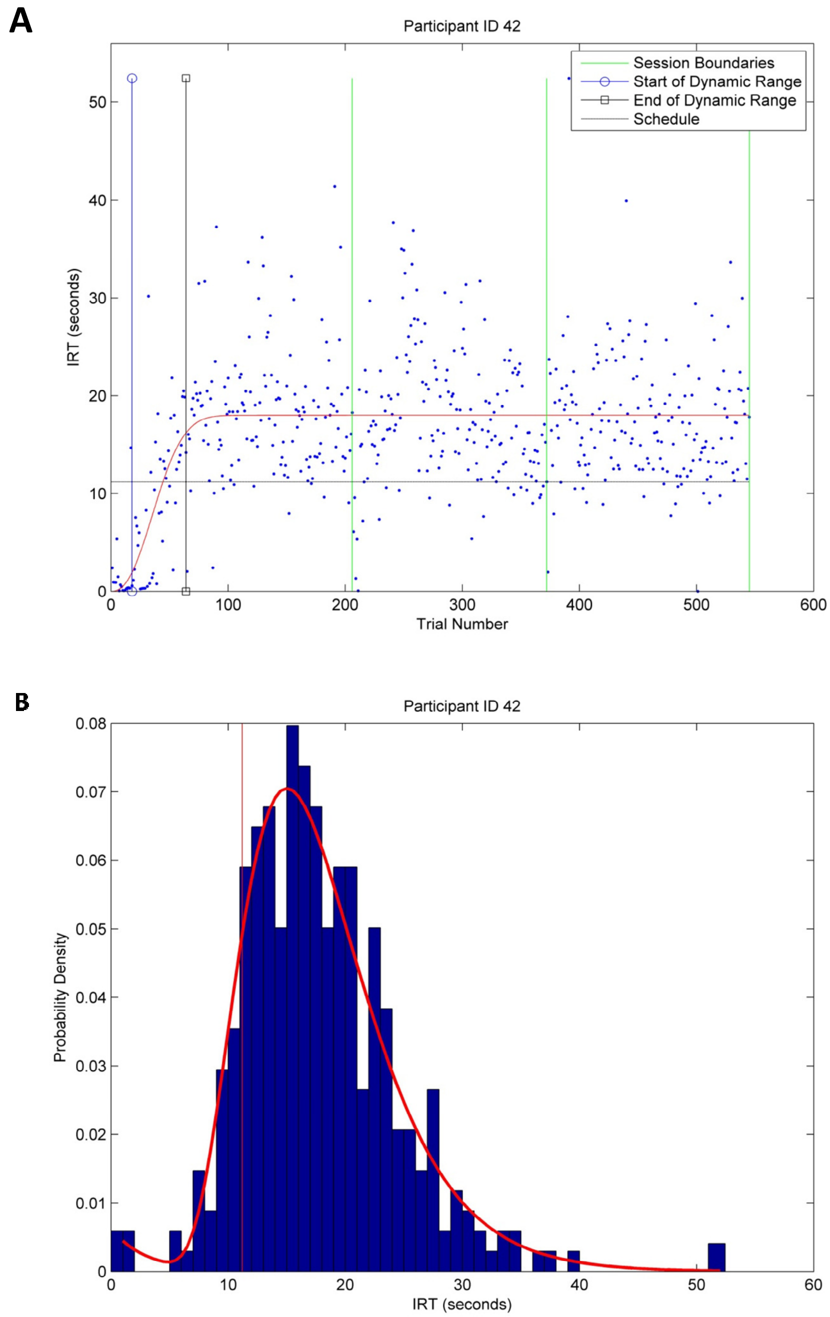

Figure 3: Acquisition (A) and Steady State (B) of an exemplar participant in the first three sessions obtained by fitting cumulative Weibull function. Red curves denote the best fitting cumulative Weibull function (A) and best fitting exponential inverse-Gaussian mixture distribution (B).

first change point smaller than steady state mean- %10*Difference was set as the start of the Dynamic range and the first change point smaller than the steady state mean- %90*Difference was set as the end of the dynamic range. Abruptness was calculated as the duration of the dynamic range and it was normalized by dividing it to the schedule of the fourth session (5.6). Trials following the end of the dynamic range were set as the fourth session steady state performance. Total duration from the start of the fourth session to the start of the dynamic range is set as the onset of acquisition and the duration to reach end of dynamic range is set as the rapidness of acquisition. Figure 4A depicts an exemplar acquisition pattern with the parameters mentioned above. Six participants were removed from the acquisition and steady state analysis of fourth session, because they were not able to learn task contingency (ID: 48, 59, 89, 95, 100, 121). One participant (ID:101) was removed from all analysis based on his/her highly atypical response patterns that can be only explained by counting.

Steady state IRTs were fit with an exponential inverse-Gaussian mixture function that has been previously shown to account for inter-response times in the DRL task (Balcı et al., 2011). Best-fit mean and shape parameter of the inverse-Gaussian portion of the mixture distribution were used to calculate the optimal strategy for the corresponding participant. Figures 3B and 4B depict the representative steady state histograms with the best fitting exponential inverse-Gaussian mixture function. Difference from optimal IRT measure was created by subtracting participants average IRTs in the steady state from their optimal IRTs. Values farther away from zero signified that participants were moving away from optimality; negative values pointing speed bias while positive values pointing an accuracy bias. Finally perceived penalty measure was created by finding the penalty/reward ratio that would make the participants’ steady state IRTs equal to the optimal IRT with their current timing uncertainty.

Robust regression with iteratively reweighted least squares with a bisquare weighting function was used for the analysis of the relationship between personality traits and task

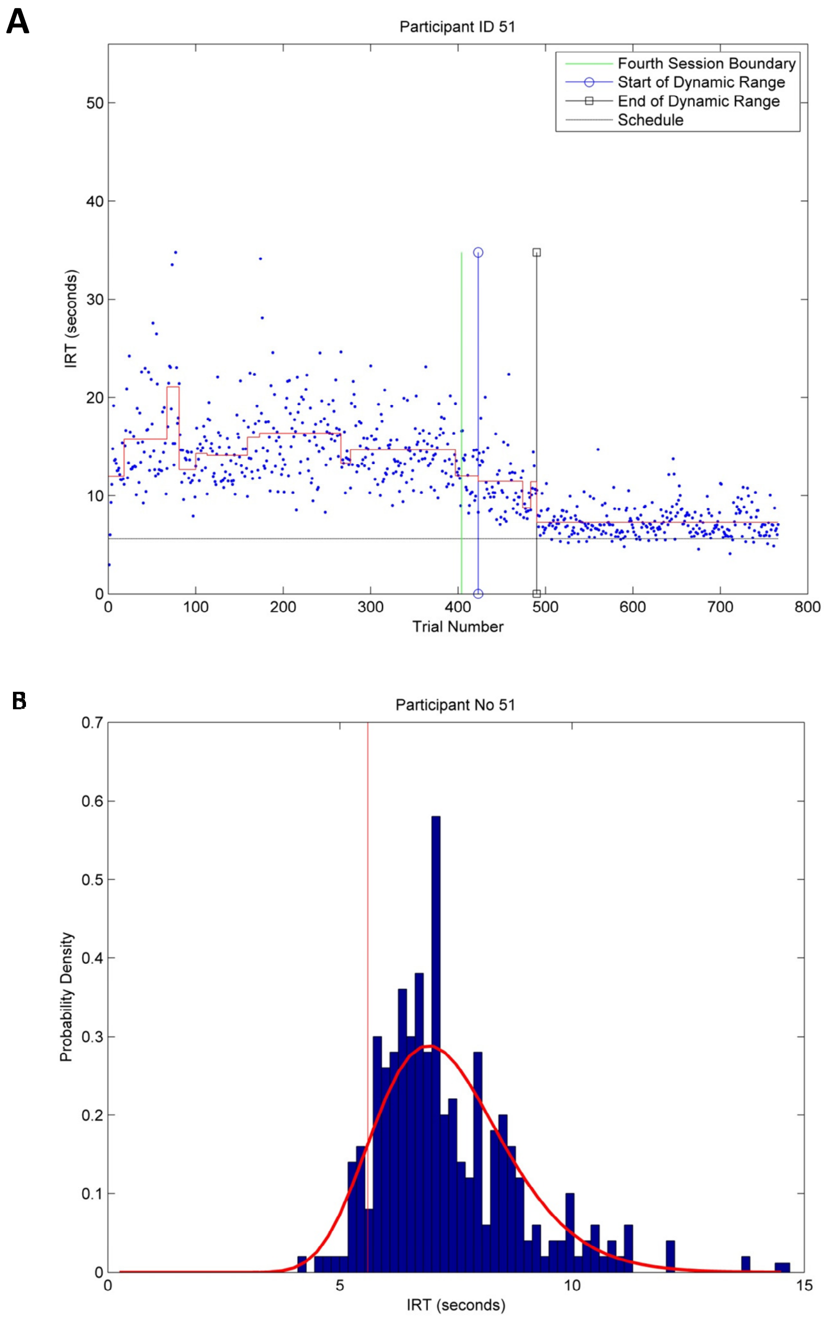

Figure 4: Acquisition (A) and Steady State (B) of an exemplar participant in the first three sessions obtained with change point algorithm. Red lines denote the average rates calculated between the change points in Figure 4A. Red curve in figure 4B denotes the best fitting exponential inverse-Gaussian mixture distribution.

parameters. The main motivation behind using robust regression was sensitivity of least square derivation methods to outliers and heteroscedasticity. Robust regression analysis for the percentage of maximum expected reward rate were also done by grouping participants according to bias type (speed or accuracy) to reveal any hidden associations.

2. Results

a. Questionnaires

Personality questionnaires were found to be highly correlated with each other. Correlations between the questionnaires and their subscales can be found in Table 1.

b. Acquisition

Acquisition of the DRL responding was characterized for both first three sessions and fourth session. Rapidness and abruptness of the acquisition of optimal/steady state IRTs were calculated.

(i) Rapidness (speed) of acquisition:

Average onset of steady state occurred around the 8^{th} (SEM=1.18, median 3.66, IQR=8.15) minute in the first three sessions and 8^{th} minute (SEM=1.19, median 4.28, IQR=12.30) in the fourth session. This difference was not statistically significant; $t(67)=-.91$, $p=.36$. There was no significant association between the personality trait scores and rapidness of acquisition in the first three sessions. In the fourth session, participants with higher scores in reward subscale of BAS had earlier onset of acquisition while participants with higher precision (counting) scores had later onset. Rapidness of acquisition on the fourth session was

Table 1. Correlation matrix for personality variables

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	Trait Anxiety																
2	Behavioral Inhibition	.610**															
3	Drive (BAS)	.054	.132														
4	Fun Seeking (BAS)	.238*	.278**	.452**													
5	Reward (BAS)	.246*	.451**	.497**	.621**												
6	BAS Total	.217*	.347**	.792**	.831**	.855**											
7	Barratt	.570**	.466**	.126	.490**	.346**	.387**										
8	Attentional (Barratt)	.474**	.345**	.166	.570**	.314**	.421**	.832**									
9	Motor (Barratt)	.332**	.197	.306**	.537**	.428**	.513**	.728**	.639**								
10	Non-planning (Barratt)	.569**	.494**	-.034	.338**	.224*	.211	.858**	.632**	.461**							
11	Padua	.478**	0.20246	-.079	0.10911	-.029	.000	0.17673	0.18182	.041	.074						
12	Rumination (PI)	.667**	.335**	-.258*	0.15749	-.055	-.078	.378**	.334**	.107	.363**	.814**					
13	Washing (PI)	0.20945	.072	.007	.092	.071	0.07365	-.027	-.022	-.037	-.142	.765**	.425**				
14	Checking (PI)	.381**	.134	-.031	.065	-.007	.012	.081	.113	.002	-.016	.890**	.683**	.591**			
15	Impulses (PI)	.282**	.142	.001	.102	-.113	.006	.154	.232*	.033	.058	.648**	.468**	.430**	.434**		
16	Precision Counting (PI)	-0.0109	-.140	.107	-.014	-.050	.030	.067	.088	0.20955	-.139	.470**	.194	.255*	.419**	.435**	
17	Precision Repeating (PI)	0.15701	.028	.102	-.041	-.044	.018	-.035	-.072	-.047	-.062	.649**	.350**	.441**	.592**	.331**	.448**

Note: n = 87, * p < .05, **p < .01

negatively associated with both total BAS score and reward subscale of BAS, while it was positively associated with the impulses subscale of PI (Table 2).

Table 2. *Simple robust linear regression analysis for the fourth session*

Dependent Variable	Predictor	Unstandardized Coefficents (*B*)	Std. Error	*t*	*p*	r^2
Difference from Optimal IRT	Padua Total Score	0.0022	0.0011	2.0422	0.0451	0.0883
CV	Behavioral Inhibition	0.0061	0.0026	2.3334	0.0227	0.0876
Onset of Acquisition	Reward	-0.4704	0.2264	-2.0781	0.0416	0.0934
Onset of Acquisition	Precision (Counting)	0.4058	0.1962	2.0686	0.0425	0.0982
Rapidness of Acquisition	Behavioral Activation	-0.5781	0.2058	-2.8095	0.0065	0.1476
Rapidness of Acquisition	Reward	-1.3515	0.5261	-2.5689	0.0125	0.1236
Rapidness of Acquisition	Impulses	0.7812	0.2014	3.8793	0.0002	0.2407

(ii) Abruptness of acquisition

The mean normalized abruptness index calculated was 0.42 (SEM=0.07, median 0.22, IQR=.40) in the first three sessions while it was 0.77 (SEM=0.15, median 0.06, IQR=1.08) in fourth session. Normalized abruptness index was significantly lower in the first three sessions compared to the fourth session; $t(67)=2.09$, $p=.04$. There was a negative association between fun seeking subscale of BAS scale and normalized abruptness index in the first three sessions (Table 3). There was no significant predictor of abruptness in the fourth session.

c. Steady State Responding:

Trend for slowing or speeding in the steady state IRTs was evaluated. Average slope of the linear regression fits to post-acquisition data points were 0.00 (SEM=0.00, median 0.00, IQR=0.01) in both first three sessions and fourth session suggesting that IRTs remained

Table 3. *Simple robust linear regression analysis for the first three sessions*

Dependent Variable	Predictor	Unstandardized Coefficents (*B*)	Std. Error	*t*	*p*	r^2
PMERR	Behavioral Inhibition	0.0026	0.0011	2.3452	0.0214	0.0699
PMERR	Padua Total Score	0.0005	0.0002	2.5664	0.0120	0.0806
PMERR	Rumination	0.0012	0.0006	2.1715	0.0327	0.0591
PMERR	Impulses	0.0033	0.0010	3.2196	0.0018	0.1305
Expected Reward Rate	Checking	0.0004	0.0001	2.9805	0.0038	0.0956
Perceived Penalty	Behavioral Inhibition	0.0579	0.0254	2.2746	0.0255	0.0629
Perceived Penalty	Barratt	0.0143	0.0065	2.1950	0.0310	0.0587
Perceived Penalty	Non-planning	0.0315	0.0141	2.2233	0.0289	0.0610
Perceived Penalty	Padua Total Score	0.0096	0.0043	2.2208	0.0291	0.0606
Perceived Penalty	Impulses	0.0728	0.0287	2.5402	0.0129	0.0899
Perceived Penalty	Trait Anxiety	0.0258	0.0126	2.0540	0.0431	0.0521
Perceived Penalty	Rumination	0.0272	0.0123	2.2129	0.0296	0.0601
Perceived Penalty	Washing	0.0304	0.0148	2.0548	0.0430	0.0524
Difference from Optimal IRT	Trait Anxiety	0.0043	0.0019	2.3013	0.0239	0.0608
Difference from Optimal IRT	Padua Total Score	0.0017	0.0006	2.6547	0.0095	0.0792
Difference from Optimal IRT	Rumination	0.0046	0.0018	2.5285	0.0133	0.0726
Difference from Optimal IRT	Impulses	0.0106	0.0040	2.6542	0.0095	0.0790
Difference from Optimal IRT	Washing	0.0058	0.0022	2.5871	0.0114	0.0751
Abruptness Index	Fun Seeking	-0.0333	0.0125	-2.6678	0.0092	0.1080
CV	Behavioral Inhibition	0.0056	0.0027	2.0793	0.0406	0.0490
CV	Non-planning	0.0029	0.0015	2.0058	0.0481	0.0457
Empirical IRT	Behavioral Inhibition	0.0175	0.0071	2.4676	0.0156	0.0678
Empirical IRT	Impulses	0.0214	0.0075	2.8666	0.0052	0.0894

steady after the acquisition took place. An exponential-inverse Gaussian mixture distribution function was fit to steady state IRTs of the first three sessions and the fourth session. Bayesian Information Criteria (BIS) was calculated to assess goodness of fit. Average BIS value for the exponential-inverse Gaussian mixture distributions in the first three sessions was 2053 (SEM=22.50) while it was 1170 (SEM=43.80) for the fourth session. When an exponential-Gaussian mixture distribution function was fit to the same dataset instead, the average BIS values increased to 2152 (SEM=24.43) in the first three sessions and to 1224 (SEM=46.14) in the fourth session. Exponential-inverse Gaussian mixture distribution was preferred in both cases because of the lower BIS values.

d. Optimality Analysis:

The optimality analysis of steady state responding in both experiments showed that participants aimed for the optimal IRT that was parameterized by the payoff structure and participants' timing uncertainty. Figure 5A and 5B depict the performance of each participant tested in first three sessions and the fourth session, respectively. These figures show the heat map of the expected reward rates (for normalized DRL schedules) expressed over a parameter space composed of target IRT and the level of timing uncertainty (CV). Ridges of these two "surfaces" are indicated by the black curves, namely the optimal performance curves for the DRL task with two different payoff structures. Optimal performance curves indicate how long participants should aim to wait for (normalized by DRL schedule) before responding again given their level of timing uncertainty and payoff structure.

We calculated how much participants earned compared to how much they could maximally earn given these endogenous and exogenous parameters. In the first three sessions, participants achieved 95.2% (SEM=.8%, median=98.1%, IQR=6.3%) of the maximum possible expected reward rate for their level of timing uncertainty. In fourth session, this value was 93.6% (SEM=1.1%, median=96.1%, IQR=8.4%). Conservative approach used in the Study 1 was not applicable for this study because the high penalty ensured that responding at the schedule would mean negative reward rates. Briefly, participants nearly maximized their rewards in both conditions and there was no difference between the percentage of maximum possible expected reward rate; $t(67)=1.30$, $p=0.20$. In the first three sessions, there was a significant positive association between the checking subscale of the PI and expected reward rate (Table 3). BIS score, total PI score, impulses and rumination subscales of PI had positive association with the percentage of maximum possible expected reward rate (PMERR) (Table 3, Figure 6). None of the personality scores were associated with the optimality related parameters in the fourth session. When the participants were

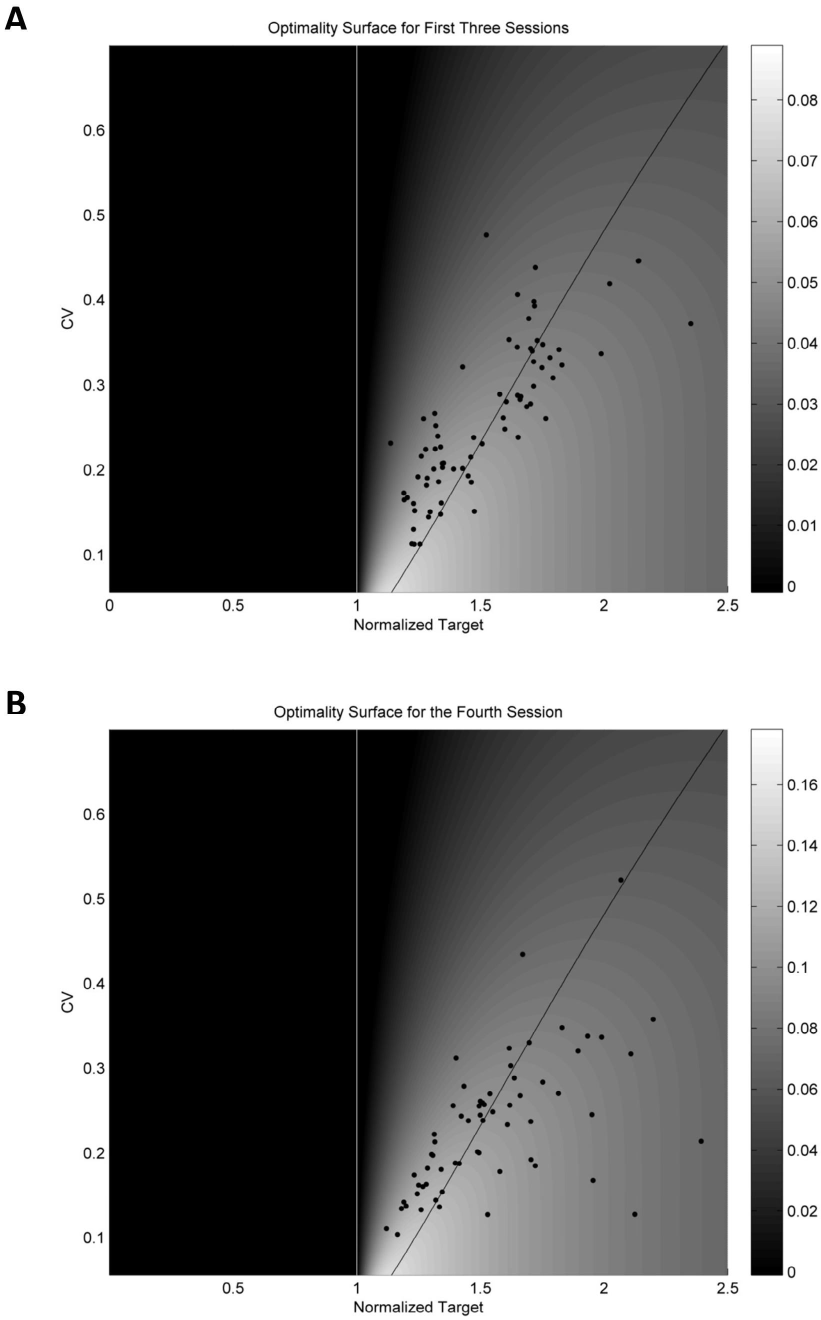

Figure 5: Heat map of expected reward rates for the first three sessions (A) and the fourth session (B) for normalized DRL schedule. Curves denote the ridge of these surfaces indicating the optimal normalized target IRTs for different levels of timing uncertainty. Each data point corresponds to a participant.

grouped according to their bias type, only the participants with speed bias showed significant associations between questionnaire scores and PMERR. Increases in trait anxiety, total PI score, washing, impulses, precision (repeating) subscales of PI produced increases in PMERR in the speed bias group (Table 4, Figure 7).

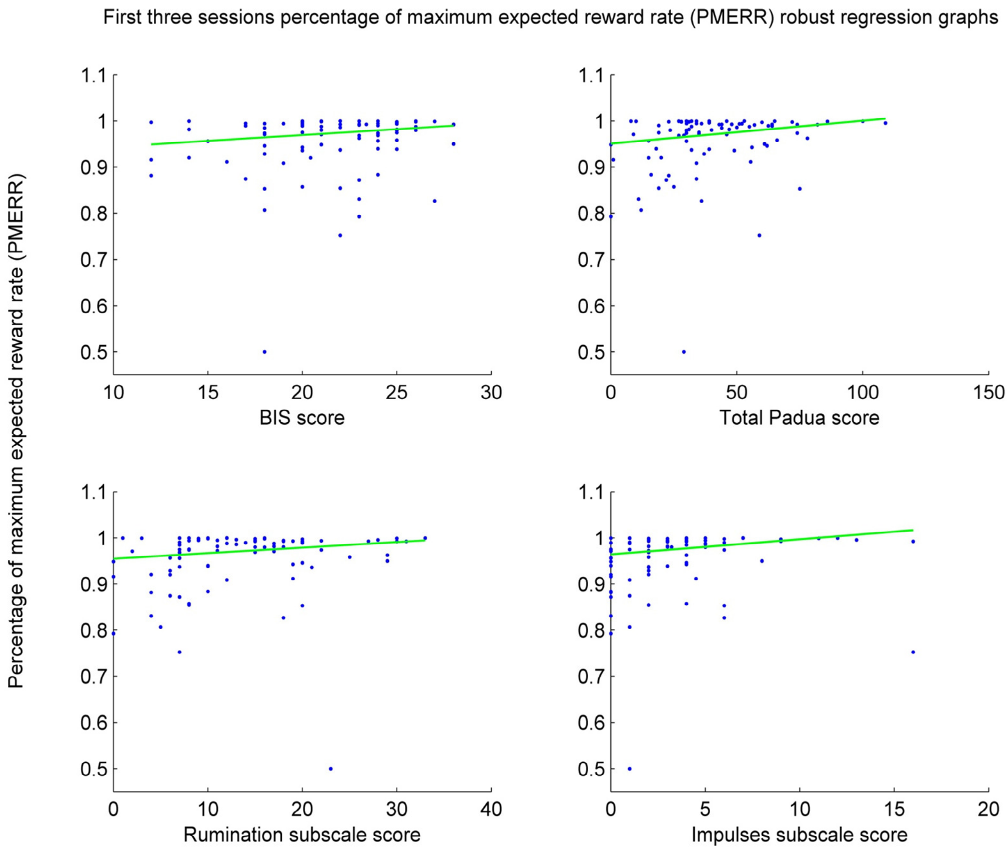

Figure 6: Graphs of significant simple linear robust regression analysis between different personality traits and percentage of maximum expected reward rate (PMERR) in the first three sessions. PMERR is calculated for each participant by finding their expected reward rates using the parameters obtained from their exponential inverse-Gaussian mixture distribution of steady state IRTs and dividing it to the maximum possible expected reward rate at the level of that participants timing uncertainty (CV).

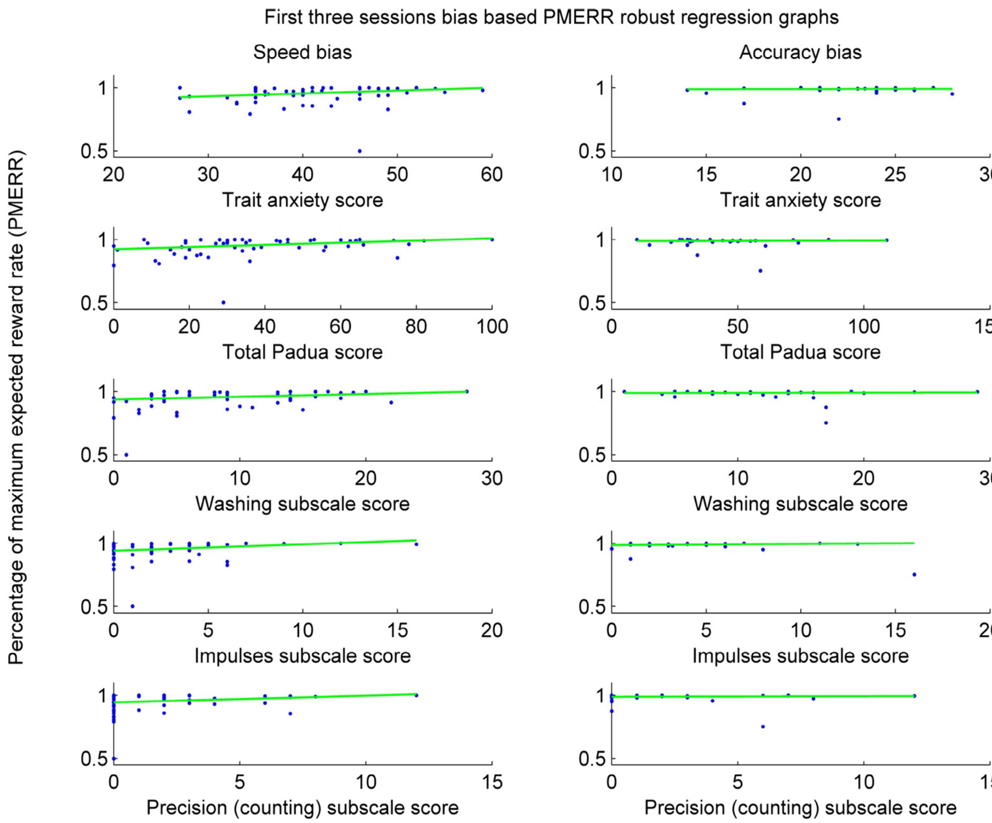

Figure 7: Graphs of significant simple linear robust regression analysis between different personality traits and percentage of maximum expected reward rate (PMERR) grouped by participants' bias type in the first three sessions. Speed bias is defined as having lower average steady state IRT than the optimal IRT, while accuracy bias is defined as having higher average steady state IRT than the optimal IRT. Only the relationships of the speed bias group are significant.

Difference from optimal IRT variable was positively associated with trait anxiety, total PI score and rumination, washing and impulses subscales of PI in the first three sessions (Table 3, Figure 8). While only total PI score was positively associated with the difference metric in the fourth session (Table 2, Figure 9). Perceived penalty was positively associated with BIS, BIS-11, non-planning subscale of BIS-11, total PI score, PI subscales of impulses, rumination washing and trait anxiety in the first three sessions (Table 3, Figure 10). There were no significant associations of perceived penalty with questionnaires in the fourth session.

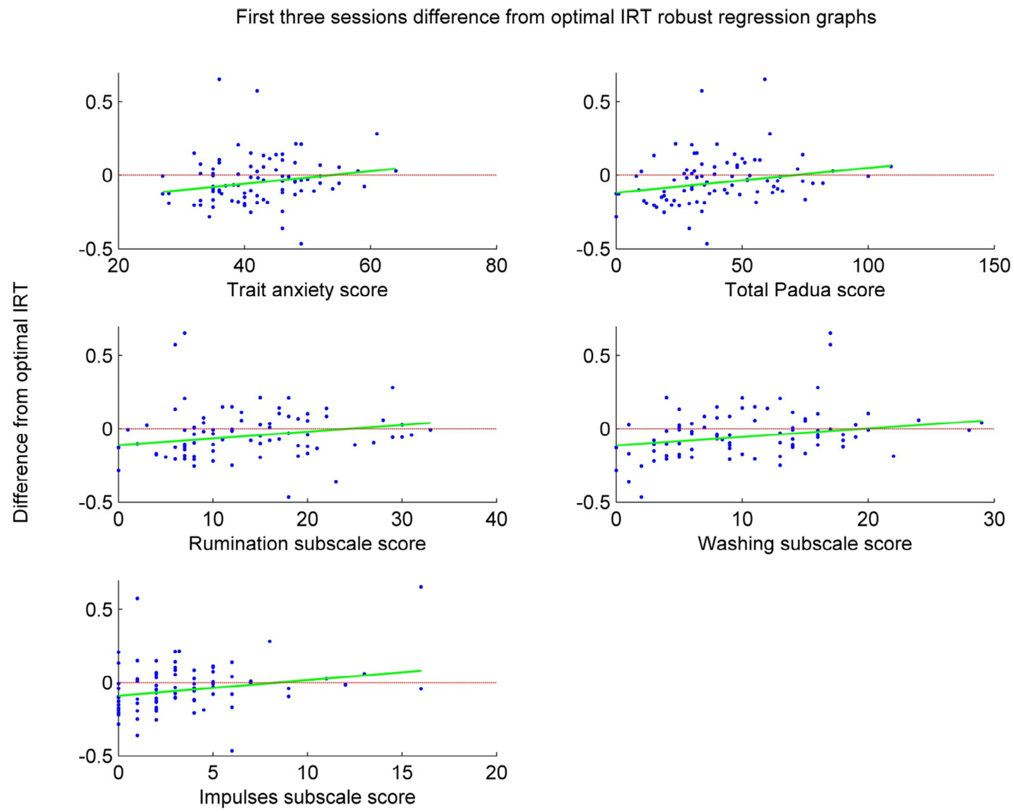

Figure 8: Graphs of significant simple linear robust regression analysis between the personality traits and difference from optimal IRT variable in the first three sessions. Difference variable is calculated by subtracting the normalized IRT from the normalized average steady state IRT. It provides a measure of temporal distance from the optimal IRT, positive values indicating waiting longer than the optimal while negative values indicate waiting shorter from the optimal IRT. Red horizontal line signifies the zero, the value which would make participants optimal.

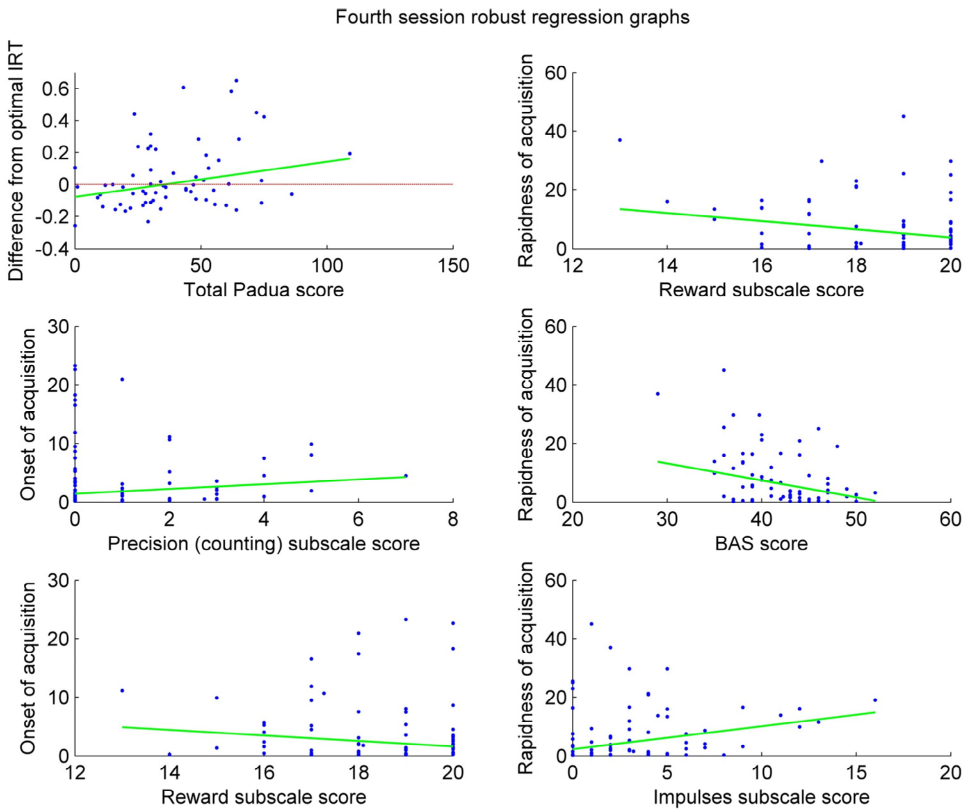

Figure 9: Graphs of significant simple linear robust regression analysis between the personality traits and various task parameters in the fourth session. Onset of acquisition signifies the duration between the start of the session and the start of dynamic range. Rapidness of acquisition on the other hand signifies the duration between the start of the session and end of the dynamic range. Red horizontal line signifies the zero, the value which would make participants optimal.

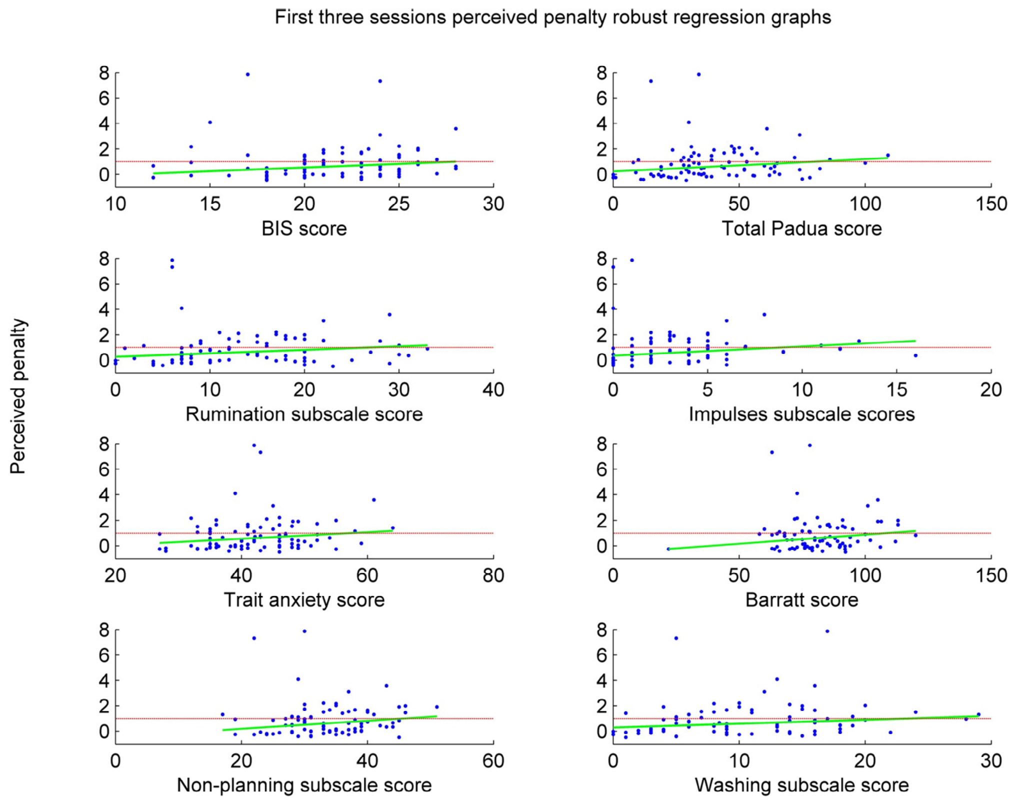

Figure 10: Graphs of significant simple linear robust regression analysis between personality traits and perceived penalty in the first three sessions. Perceived penalty is calculated by finding the penalty/reward ratio that would make participants steady state IRT equal to the optimal IRT of their timing uncertainty (CV) level. Red horizontal line signifies 1, the point where perceived penalty is equal to reward.

The empirical normalized IRTs of the participants in the first three sessions and the fourth session was compared. Normalized empirical IRTs in the fourth session were significantly longer than the normalized IRTs in the first three sessions; $t(67)=2.09$, $p=0.04$. There was also a positive association between the behavioral inhibition, impulses subscale of the PI and normalized IRTs for the first three sessions (Table 3). There was no significant predictor of IRTs in the fourth session.

Coefficient of variations obtained in the first three sessions were significantly higher than the CVs obtained in the fourth session; $t(67)=2.46$, $p=0.02$. CV's calculated from the first three sessions had positive association with BIS and non-planning subscale of BIS-11 (Table 3). But for the fourth session CV's, only the positive association with behavioral inhibition was significant (Table 3). The relationship between these variables and CV was mirrored by the optimal IRTs, and maximum possible expected reward rate, given that Optimal IRTs and maximum possible expected reward rate depend on CV.

3. Discussion

The second study revealed that most of the participants were able to adapt to unarticulated lowering of the schedule in the fourth session, and after the acquisition, they were close to optimality in their steady state. Nevertheless, six participants were not able to adapt to the new schedule and continued to respond similar to the first three sessions. The second major finding of the second study was the connection between personality traits and parameters of the DRL task. Personality traits predicted the acquisition and steady state performance both in the first three sessions and the fourth session. Predictors of acquisition parameters were mostly distinct from predictors of steady state performance in both first three sessions and the fourth session. BAS and its subscales' scores were predictive only of acquisition parameters while BIS, BIS-11 and trait anxiety scores were only predictive of steady state performance. Padua and its subscales were predictive of both acquisition and the steady state performance.

The acquisition and steady state performance in the fourth session appears to be consistent with the expectations. Participants were able to reduce their average IRTs to maximize their reward rates based on the new schedule. Evidence of adaptation to the new schedule indicates that participants had exploratory responses, since only way for them to discover the schedule change was to respond below the original schedule. Manipulation in the fourth session can be considered as a alternative form of exploration/exploitation trade-off tasks (Cohen, McClure, & Yu, 2007) in the domain of temporal decision-making. Different from these tasks (such as n-armed bandit task), in DRL, participants have a continuous range of durations that they can target to earn reward rather than a fixed amount of options (n). Within these options, every payoff environment has single target duration that is optimal for that participant given their timing uncertainty; therefore the exploration of the target duration space has great importance. Participants who stick to their previous IRTs after the halving of the schedule would be exploiting the environment and earn reward rates similar to the previous sessions. But the participants who shorten their IRTs (within a certain range) would be exploring the changing environment which would increase their reward rate. These findings supplement the task acquisition findings of the Study 1, showing that the acquisition of the task contingency can occur in steady state, and expand the time-based decision making literature. Future studies can build upon this finding by introducing the halving of the schedule randomly, similar to the manipulation done in the beat-the-clock task (Simen et al., 2011).

The steady state performance parameters of timing uncertainty (CV) and empirical IRT were differentially modulated by personality questionnaires in both first three and the fourth session (Table 2 and 3). Complex non-linear relation between the optimality and these parameters make interpretations based solely on these two parameters abstruse. To ease this problem in translation, relations between task parameters and personality questionnaires will be explained through difference from optimal IRT, perceived penalty and percentage of maximum expected reward rate (PMERR).

Difference from optimality parameter is calculated by subtracting normalized empirical IRT from the normalized optimal IRT. By subtracting the optimal IRT that is calculated by using timing uncertainty (CV), the new parameter becomes sensitive to possible interactions between the CV and empirical IRT. Positive values in the parameter indicate that the participant is waiting longer then they should to be optimal, while negative values indicate they are waiting shorter than they should. As expected, trait anxiety, total PI scores, and PI subscales of rumination, impulses, washing were positively associated with the difference parameter in the first three sessions. Only the total PI score had a positive association with the difference parameter in the fourth session. These associations are most likely to be caused by the punishment/harm avoidance tendencies that increase with anxiety and OCD symptoms. These sensitivities to negative stimuli would push participants wait longer to avoid penalty that occurs when they have responses earlier then the schedule.

While the difference metric is an important tool to distinguish the average wait durations with respect to optimal IRT, it doesn't necessarily reflect the optimality. There is a non-linear relation between temporal distance to optimal duration and proximity to optimality in terms of reward rates..This effect caused by the shape of the exponential inverse Gaussian distribution and the schedule of the DRL can be clearly seen in the optimality surface (5A and 5B). Participants with higher scores on BIS, total PI, and rumination and impulses subscale of PI were found to be closer to optimal compared to participants who scored lower in the first three sessions. Behavioral inhibition system that is measured by the BIS scale is theorized to underlay sensitivity to punishment and it is activated when conflicting concurrent goals occur (Gray & McNaughton, 2003). Optimal solution in the DRL task relies on the balancing two conflicting goals, waiting longer to avoid penalty and decreasing wait time to increase reward rate. Based on theoretical definition of BIS, better performance in the DRL task by high BIS participants is not surprising. On the other hand, it is harder to track the role of PI and its subscales on optimal performance. Rumination was shown to be related to anxiety and neuroticism when it was measured by another scale (Muris, Roelofs, Rassin, Franken, & Mayer, 2005) and PI (Oppen et al., 1995). Anxiety and neuroticism in turn are closely related to the BIS both conceptually and statistically (Gray & McNaughton, 2003;

Jorm et al., 1998). The findings of Study 2 also corroborate these findings, showing that rumination subscale has the strongest correlation with trait anxiety and BIS compared the other PI subscales (Table 1). The effect of BIS and rumination scores on PMERR is most likely caused by the same mechanism. Unfortunately it is not possible to use stepwise methods with robust regression to investigate this relationship more rigorously. Unlike the rumination subscale, impulses subscale neither share the same theoretical underpinnings with BIS nor correlated with it (Table 1). The items of this subscale measure obsessional impulses to harm oneself and others (Oppen et al., 1995) which at face value, shouldn't be related to DRL performance. Alternatively, the effect of this subscale can be explained by net increase in OCD symptoms rather than the specifics of the symptoms itself.

The separation of the optimality analysis based on bias type provided a more detailed look into the effects of personality scores. When the robust regression analysis is done separately for the speed and accuracy bias groups, the effect of BIS and rumination on the optimal performance disappeared in both. Most likely cause for this was the decrease in the degrees of freedom. Effect of PI and Impulses scores remained significant for the speed bias group despite the loss of degrees of freedom but disappeared for the accuracy bias group (Table 4). Based on this interaction, it can be postulated that the increases in optimality with these personality trait scores is caused by the increases in the difference parameter (Table 3). When the difference parameter increases, participants with negative scores (speed bias) move closer to 0 (optimality) while participants with positive difference scores (accuracy bias) move away from it. Trait anxiety and washing were also positively associated with PMERR in the speed bias group, again caused by their tendency to increase the average IRT with respect to optimal IRTs of that participant.

Perceived penalty metric is an alternative way to measure relative sensitivity to punishment compared to reward. It is calculated by finding the ratio of penalty to reward that would make the average IRT of the participant equal to the optimal solution. Increases in trait anxiety, BIS, BIS-11 ,non-planning subscale of BIS-11 , total PI, rumination, washing, impulses subscales of PI predicted the increases in perceived penalty measure. BIS score's relationship with the perceived penalty is in line with the theoretical background of the scale (Corr, 2004; Gray & McNaughton, 2003) as a punishment sensitivity measure. Participants with both high BIS (Boksem et al., 2006; De Pascalis et al., 2010) and high trait anxiety (Lerner & Keltner, 2000) were shown to be more responsive to negative stimuli, supporting the findings of this study While the relation of perceived penalty with Padua scale and its subscales has theoretical backing (Fullana et al., 2004); same can't be said for BIS-11 and its non-planning subscale. BIS-11 and its subscale's effect on perceived penalty might be caused by their high correlation with trait anxiety, BIS and rumination scores in this study (Table 1).

In addition to the optimality in the first three sessions; personality questionnaire scores predicted the acquisition performance in both first three sessions and the fourth session. For acquisition of the first three sessions, strongest predictor of abruptness index was fun seeking subscale of the BAS scale while the onset or rapidness of acquisition was not affected. Increases in fun seeking decreased the abruptness index, which meant less time spent to reach steady state after the learning has started. In the fourth session increases in BAS score and reward subscale decreased the time spent to reach steady state while increase in impulses subscale increased the duration. Onset of acquisition was negatively associated with reward subscale while it was positively associated with the precision (counting) subscale of the PI. Apparent distinction between first three and the fourth sessions in acquisition patterns might be caused by the distinct teaching signals. Initial acquisition depends on decreasing the number of negative feedback and increasing the number of positive feedback. But the acquisition of fourth session requires abandoning the established responding pattern to establish a more rewarding one. Learning in the fourth session does not depend on decreasing the negative stimuli, but only increasing the reward rate (within a certain range of IRTs). Participants that could not adapt to the schedule change continue to earn reward at the same rate of their previous sessions. Therefore differentiation between the reward subscale of the BAS that measures reward sensitivity and the acquisition in the fourth session is expected. This finding is also in line with the literature; compared to the other subscales of BAS, fun seeking subscale has stronger correlation to the novelty seeking measure while reward subscale has stronger correlation with reward dependence measure of the Tridimensional Personality Questionnaire (Carver & White, 1994). The positive association between the impulses and precision (counting) subscales of PI and acquisition is expected. OCD patients are shown to be more prone to habitual behavior, less adaptive to the rule changes in the tasks and more harm avoidant (Evans et al., 2004; Gillan et al., 2011; Remijnse et al., 2006). While these findings should translate into longer or impaired acquisition in fourth session, there isn't any evidence that specifically points to these subscales.

CONCLUSION

In conclusion, Study 1 established that humans are capable of close to optimal performance in non-instructed temporal decision task, when their premature responses are penalized. Response-outcome contingencies of the task were enough for participants to reach close to the optimal performance, experience with the duration and instructions were not necessary. These findings support the idea that non-instructed DRL can be used with humans, providing healthier comparison of human and animal performance by eliminating the confounding factors created by instructions and prior experience with the schedule duration. Study 2 replicates the findings of the first study with a much larger sample and shows that most of the participants are capable of adjusting to the unannounced lowering of the target duration in the fourth session in order to maximize their reward rate. In addition to that, findings of the second study indicate that the personality traits measured by Trait anxiety scale, BIS/BAS Scale, Barratt Impulsiveness Scale and Padua Inventory is capable of predicting the parameters of acquisition and steady state performance in DRL. These two studies together emphasize the importance of DRL's efficacy in characterizing the relevant decision-making concepts with its parameters and its generalizability to animal studies.

REFERENCES

Balcı, F., Freestone, D., & Gallistel, C. R. (2009). Risk assessment in man and mouse. *Proceedings of the National Academy of Sciences, 106*(7), 2459-2463.

Balcı, F., Freestone, D., Simen, P., Desouza, L., Cohen, J. D., & Holmes, P. (2011). Optimal temporal risk assessment. *Front Integr Neurosci, 5*.

Balcı, F., Meck, W. H., Moore, H., & Brunner, D. (2009). Timing deficits in aging and neuropathology. In J. L. Bizon & A. G. Woods (Eds.), *Animal models of human cognitive aging* (pp. 1-41). Totowa, NJ: Humana Press.

Bannerman, D., Yee, B., Good, M., Heupel, M., Iversen, S., & Rawlins, J. (1999). Double dissociation of function within the hippocampus: a comparison of dorsal, ventral, and complete hippocampal cytotoxic lesions. *Behav Neurosci, 113*(6), 1170-1188.

Baunez, C., & Gubellini, P. (2010). Effects of GPi and STN inactivation on physiological, motor, cognitive and motivational processes in animal models of Parkinson's disease. *Prog Brain Res, 183*, 235-258.

Beaver, J. D., Lawrence, A. D., Van Ditzhuijzen, J., Davis, M. H., Calder, A. J., & Woods, A. (2006). Individual differences in reward drive predict neural responses to images of food. *Journal of Neuroscience, 26*(19), 5160-5166. doi: 10.1523/JNEUROSCI.0350-06.2006

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and economic behavior, 52*(2), 336-372.

Bejerot, S., Schlette, P., Ekselius, L., Adolfsson, R., & Knorring, L. v. (1998). Personality disorders and relationship to personality dimensions measured by the Temperament and Character Inventory in patients with obsessive‐compulsive disorder. *Acta Psychiatr Scand, 98*(3), 243-249.

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

Blum, K., Oscar-Berman, M., Barh, D., Giordano, J., & Gold, M. (2013). Dopamine Genetics and Function in Food and Substance Abuse. *Journal of Genetic Syndromes & Gene therapy*.

Boksem, M. A., Tops, M., Wester, A. E., Meijman, T. F., & Lorist, M. M. (2006). Error-related ERP components and individual differences in punishment and reward sensitivity. *Brain Res, 1101*(1), 92-101.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision, 10*(4), 433-436.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience, 6*(10), 755-765.

Camara, E., Rodriguez-Fornells, A., & Münte, T. F. (2010). Microstructural brain differences predict functional hemodynamic responses in a reward processing task. *Journal of Neuroscience, 30*(34), 11398-11402. doi: 10.1523/JNEUROSCI.0111-10.2010

Carter, R. M., MacInnes, J. J., Huettel, S. A., & Adcock, R. A. (2009). Activation in the VTA and nucleus accumbens increases in anticipation of both gains and losses. *Frontiers in Behavioral Neuroscience, 3*(AUG). doi: 10.3389/neuro.08.021.2009

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of personality and social psychology, 67*(2), 319.

Cho, Y. H., & Jeantet, Y. (2010). Differential involvement of prefrontal cortex, striatum, and hippocampus in DRL performance in mice. *Neurobiology of learning and memory, 93*(1), 85-91.

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I Stay or Should I Go? How the Human Brain Manages the Trade-off between Exploitation and Exploration, 933.

Corr, P. J. (2004). Reinforcement sensitivity theory and personality. *Neuroscience And Biobehavioral Reviews, 28*(3), 317-332.

Çoşkun, F., Sayali Ungerer, C., Emine, G., & Balci, F. (In press 2014). *Optimal Time Discrimination*.
Costa, V. C. I., Bueno, J. L. O., & Xavier, G. F. (2005). Dentate gyrus-selective colchicine lesion and performance in temporal and spatial tasks. *Behav Brain Res, 160*(2), 286-303.
Dalley, J. W., Cardinal, R. N., & Robbins, T. W. (2004). Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. *Neuroscience & Biobehavioral Reviews, 28*(7), 771-784.
De Pascalis, V., Varriale, V., & D'Antuono, L. (2010). Event-related components of the punishment and reward sensitivity. *Clinical Neurophysiology, 121*(1), 60-76. doi: http://dx.doi.org/10.1016/j.clinph.2009.10.004
Endrass, T., Kloft, L., Kaufmann, C., & Kathmann, N. (2011). Approach and avoidance learning in obsessive–compulsive disorder. *Depression and anxiety, 28*(2), 166-172.
Evans, D. W., Lewis, M. D., & Iobst, E. (2004). The role of the orbitofrontal cortex in normally developing compulsive-like behaviors and obsessive–compulsive disorder. *Brain and cognition, 55*(1), 220-234.
Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion, 7*(2), 336.
Frank, M. J., & Hutchison, K. (2009). Genetic contributions to avoidance-based decisions: striatal D2 receptor polymorphisms. *Neuroscience, 164*(1), 131-140.
Fullana, M. À., Mataix‐Cols, D., Trujillo, J. L., Caseras, X., Serrano, F., Alonso, P., . . . Torrubia, R. (2004). Personality characteristics in obsessive‐compulsive disorder and individuals with subclinical obsessive‐compulsive problems. *British Journal of Clinical Psychology, 43*(4), 387-398.
Gable, P. A., & Harmon-Jones, E. (2013). Trait behavioral approach sensitivity (BAS) relates to early (<150 ms) electrocortical responses to appetitive stimuli. *Soc Cogn Affect Neurosci, 8*(7), 795-798. doi: 10.1093/scan/nss072
Gallistel, C. R., Fairhurst, S., & Balsam, P. (2004). The learning curve: implications of a quantitative analysis. *Proc Natl Acad Sci U S A, 101*(36), 13124-13131.
Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107*(2), 289-344.
Gehring, W. J., Himle, J., & Nisenson, L. G. (2000). Action-monitoring dysfunction in obsessive-compulsive disorder. *Psychol Sci, 11*(1), 1-6.
Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*(3), 279-325. doi: 10.1037/0033-295X.84.3.279
Gillan, C. M., Papmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A., Robbins, T. W., & de Wit, S. (2011). Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *American Journal of Psychiatry, 168*(7), 718-726.
Gizer, I. R., Ficks, C., & Waldman, I. D. (2009). Candidate gene studies of ADHD: a meta-analytic review. *Hum Genet, 126*(1), 51-90.
Gray, J. A. (1981). A critique of Eysenck's theory of personality *A model for personality* (pp. 246-276): Springer.
Gray, J. A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. Oxford: Oxford University Press.
Gray, J. A., & McNaughton, N. (2003). *The neuropsychology of anxiety: An enquiry into the function of the septo-hippocampal system*. New York: Oxford University Press.
Güleç, H., Tamam, L., Güleç, M. Y., Turhan, M., Karakuş, G., Zengin, M., & Stanford, M. S. (2008). Psychometric properties of the Turkish version of the Barratt Impulsiveness Scale-11. *Klinik Psikofarmakoloji Bülteni, 18*(4), 251-258.
Hahn, T., Dresler, T., Ehlis, A. C., Plichta, M. M., Heinzel, S., Polak, T., . . . Jakob, P. M. (2009). Neural response to reward anticipation is modulated by Gray's impulsivity. *Neuroimage, 46*(4), 1148-1153. doi: 10.1016/j.neuroimage.2009.03.038

Holmes, A. J., Bogdan, R., & Pizzagalli, D. A. (2010). Serotonin transporter genotype and action monitoring dysfunction: a possible substrate underlying increased vulnerability to depression. *Neuropsychopharmacology, 35*(5), 1186-1197.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature neuroscience, 13*(8), 1020-1026.

Jorm, A. F., Christensen, H., Henderson, A. S., Jacomb, P. A., Körten, A. E., & Rodgers, B. (1998). Using the BIS/BAS scales to measure behavioural inhibition and behavioural activation: Factor structure, validity and norms in a large community sample. *Personality and Individual Differences, 26*(1), 49-58.

Karg, K., Burmeister, M., Shedden, K., & Sen, S. (2011). The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic moderation. *Arch Gen Psychiatry, 68*(5), 444.

Kheifets, A., & Gallistel, C. (2012). Mice take calculated risks. *Proceedings of the National Academy of Sciences, 109*(22), 8776-8779.

Kusunoki, K., Sato, T., Taga, C., Yoshida, Y., Komori, K., Narita, T., . . . Ozaki, N. (2000). Low novelty‐seeking differentiates obsessive‐compulsive disorder from major depression. *Acta Psychiatr Scand, 101*(5), 403-405.

Landy, M. S., Trommershäuser, J., & Daw, N. D. (2012). Dynamic Estimation of Task-Relevant Variance in Movement under Risk. *The Journal of Neuroscience, 32*(37), 12702-12711. doi: 10.1523/JNEUROSCI.6160-11.2012

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition & Emotion, 14*(4), 473-493.

Lyoo, I. K., Lee, D. W., Kim, Y. S., Kong, S. W., & Kwon, J. S. (2001). Patterns of temperament and character in subjects with obsessive-compulsive disorder. *The Journal of clinical psychiatry, 62*(8), 637-641.

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal “time cells” bridge the gap in memory for discontiguous events. *Neuron, 71*(4), 737-749.

Maner, J. K., Richey, J. A., Cromer, K., Mallott, M., Lejuez, C. W., Joiner, T. E., & Schmidt, N. B. (2007). Dispositional anxiety and risk-avoidant decision-making. *Personality and Individual Differences, 42*(4), 665-675.

Maner, J. K., & Schmidt, N. B. (2006). The role of risk avoidance in anxiety. *Behavior Therapy, 37*(2), 181-189.

McNaughton, N. (2006). The role of the subiculum within the behavioural inhibition system. *Behav Brain Res*(2), 232. doi: 10.1016/j.bbr.2006.05.037

Meck, W. H. (1988). Hippocampal function is required for feedback control of an internal clock's criterion. *Behav Neurosci, 102*(1), 54-60.

Meck, W. H. (2005). Neuropsychology of timing and time perception. *Brain and cognition, 58*(1), 1-8.

Meck, W. H., Church, R. M., & Olton, D. S. (1984). Hippocampus, time, and memory. *Behav Neurosci, 98*(1), 3-22. doi: 10.1037/0735-7044.98.1.3

Miu, A. C., Heilman, R. M., & Houser, D. (2008). Anxiety impairs decision-making: psychophysiological evidence from an Iowa Gambling Task. *Biol Psychol, 77*(3), 353-358.

Mueller, E. M., Nguyen, J., Ray, W. J., & Borkovec, T. D. (2010). Future-oriented decision-making in Generalized Anxiety Disorder is evident across different versions of the Iowa Gambling Task. *J Behav Ther Exp Psychiatry, 41*(2), 165-171.

Muris, P., Roelofs, J., Rassin, E., Franken, I., & Mayer, B. (2005). Mediating effects of rumination and worry on the links between neuroticism, anxiety and depression. *Personality and Individual Differences, 39*(6), 1105-1111. doi: http://dx.doi.org/10.1016/j.paid.2005.04.005

Nijs, I. M., Franken, I. H., & Smulders, F. T. (2007). BIS/BAS sensitivity and the p300 event-related brain potential. *Journal of Psychophysiology, 21*(2), 83.

O'Donell, P., Greene, J., Pabello, N., Lewis, B. L., & Grace, A. A. (1999). Modulation of cell firing in the nucleus accumbens. *Annals of the New York academy of sciences, 877*(1), 157-175. doi: 10.1111/j.1749-6632.1999.tb09267.x

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

Oppen, P. v., Heokstra, R. J., & Emmelkamp, P. M. G. (1995). The structure of obsessive-compulsive symptoms. *Behav Res Ther*(1), 15.

Pajkossy, P., Dezső, L., & Zoltay Paprika, Z. (2009). The opposite effect of trait and state anxiety on Iowa Gambling Task. *Learning & Perception, 1*(2), 279-295.

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of clinical psychology, 51*(6), 768-774.

Pellegrino, L. J., & Clapp, D. F. (1971). Limbic lesions and externally cued DRL performance. *Physiology & Behavior, 7*(6), 863-868.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision, 10*(4), 437-442.

Pennartz, C., Ito, R., Verschure, P., Battaglia, F., & Robbins, T. (2011). The hippocampal–striatal axis in learning, prediction and goal-directed behavior. *Trends Neurosci, 34*(10), 548-559.

Pfohl, B., Black, D., Noyes Jr, R., Kelley, M., & Blum, N. (1990). A test of the tridimensional personality theory: association with diagnosis and platelet imipramine binding in obsessive-compulsive disorder. *Biological psychiatry, 28*(1), 41-46.

Pothuizen, H. H., Jongen - Rêlo, A. L., Feldon, J., & Yee, B. K. (2005). Double dissociation of the effects of selective nucleus accumbens core and shell lesions on impulsive - choice behaviour and salience learning in rats. *European Journal of Neuroscience, 22*(10), 2605-2616. doi: 10.1111/j.1460-9568.2005.04388.x

Remijnse, P. L., Nielen, M. M., van Balkom, A. J., Cath, D. C., van Oppen, P., Uylings, H. B., & Veltman, D. J. (2006). Reduced orbitofrontal-striatal activity on a reversal learning task in obsessive-compulsive disorder. *Arch Gen Psychiatry, 63*(11), 1225-1236.

Richter, M. A., Summerfeldt, L. J., Joffe, R. T., & Swinson, R. P. (1996). The Tridimensional Personality Questionnaire in obsessive-compulsive disorder. *Psychiatry Research, 65*(3), 185-188.

Sanavio, E. (1988). Obsessions and compulsions: the Padua Inventory. *Behav Res Ther, 26*(2), 169-177.

Savitsky, K., Medvec, V. H., Charlton, A. E., & Gilovich, T. (1998). " What, Me Worry?": Arousal, Misattribution, and the Effect of Temporal Distance on Confidence. *Personality and Social Psychology Bulletin, 24*(5), 529-536.

Schneider, T. (2001). Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *Journal of Climate, 14*(5), 853-871.

Shepperd, J. A., Grace, J., Cole, L. J., & Klein, C. (2005). Anxiety and outcome predictions. *Personality and Social Psychology Bulletin, 31*(2), 267-275.

Simen, P., Balcı, F., Cohen, J. D., & Holmes, P. (2011). A model of interval timing by neural integration. *The Journal of Neuroscience, 31*(25), 9238-9253. doi: 10.1523/JNEUROSCI.3121-10.2011

Şişman, S. (2012). Davranışsal İnhibisyon Sistemi/Davranışsal Aktivasyon Sistemi Ölçeği'nin Türkçeye Uyarlanması: Geçerlik ve Güvenirlik Çalışması. *Psikoloji Çalışmaları Dergisi, 32*(2), 1-22.

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

Stanford, M. S., Mathias, C. W., Dougherty, D. M., Lake, S. L., Anderson, N. E., & Patton, J. H. (2009). Fifty years of the Barratt Impulsiveness Scale: An update and review. *Personality and Individual Differences, 47*(5), 385-395. doi: http://dx.doi.org/10.1016/j.paid.2009.04.008

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior research methods, 37*(3), 498-505.

Uslaner, J. M., & Robinson, T. E. (2006). Subthalamic nucleus lesions increase impulsive action and decrease impulsive choice− mediation by enhanced incentive motivation? *European Journal of Neuroscience, 24*(8), 2345-2354. doi: 10.1111/j.1460-9568.2006.05117.x
Wearden, J. H., & Culpin, V. (1998). Exploring scalar timing theory with human subjects. In V. DeKeyser, G. d'Ydewalle & A. Vandierendonck (Eds.), *Time and the dynamic control of behavior* (pp. 33-49). Gottingen: Hogrefe and Huber.
Werner, N. S., Duschek, S., & Schandry, R. (2009). Relationships between affective states and decision-making. *International Journal of Psychophysiology, 74*(3), 259-265.
Wheeler, M. G. (2009). *GABAergic Transmission in the Nucleus Accumbens-Ventral Pallidum Path and DRL Task Efficiency in Rats.* (Master's thesis), Emory University.
Yin, B., & Troger, A. B. (2011). Exploring the 4th dimension: hippocampus, time, and memory revisited. *Front Integr Neurosci, 5*, 36. doi: 10.3389/fnint.2011.00036
Young, B., & McNaughton, N. (2000). Common firing patterns of hippocampal cells in a differential reinforcement of low rates of response schedule. *The Journal of Neuroscience, 20*(18), 7043-7051.