

**Neuroeconomic Studies on Personality and  
Decision-Making**

**A DISSERTATION  
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF MINNESOTA  
BY**

**Daniel Rohan Hawes**

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
Doctor of Philosophy**

**Prof. Terrance Hurley & Prof. Aldo Rustichini**

**July, 2013**

© Daniel Rohan Hawes 2013  
ALL RIGHTS RESERVED

# Acknowledgements

Professor Aldo Rustichini deserves special thanks for being a thoughtful, caring and supportive advisor. His dedication to careful, organized and deliberate research, have been demonstration and inspiration of how to think about scientific theory and how to conduct research. I could not have wished for a better academic mentor.

I am also extremely grateful to the members of my committee Terrance Hurley, Liz Davis and Avner Ben-Ner:

Among other acts of kindness, as Director of Graduate Studies in Applied Economics during my first years in the program, Professor Terry Hurly supported me in deviating from the traditional course work structure of my program. I believe that this put me on a critical path towards pursuing my current research interests. In this context Professor Avner Ben-Ner was similarly influential in supporting my petitions for an approved minor in psychology.

As Director of Graduate Studies in my final years, Professor Liz Davis probably bore the brunt of the pains that accompany my penchant for disorganization, and aversion to filing administrative forms before the very last minute of a looming deadline. Nonetheless, she patiently worked with me, and supported my efforts during multiple fellowship applications and during the final stretches of my dissertation work. For all of this, I am truly thankful.

Collection of the data analyzed in this dissertation was supported by the National Institute of Mental Health (F32 MH077382 to C.G.D.), and the National Science Foundation (DRL 0644131 to J.R.G., SES-0924896 A.R., SES-1156315 A.R. & D.H.)

I further acknowledge the University of Minnesota Doctoral Dissertation Fellowship for financial support.

# Dedication

To you. Literally.

## Abstract

Neural activity causally underlies human cognition and behavior. Investigating the neurobiological principles and computational mechanisms governing brain activity during decision-making provides a way to improve theories of human behavior in the natural as well as social sciences (Glimcher & Rustichini 2004; Rustichini, 2009; Fehr & Rangel, 2009). In this context, the discipline of Neuroeconomics was originally conceived as an endeavor to interrogate neural activity during economic decision-making with the aim of evaluating competing decision theories (Rustichini, 2008; Glimcher, Camerer, Fehr & Poldrack 2009). From this origin, Neuroeconomics has evolved into a full-fledged enterprise of consilience; an attempt to not only test and bridge, but truly unify natural science and social science explanations of human behavior (Wilson, 1998; Glimcher & Rustichini, 2004; Rangel, Camerer & Montague, 2008).

This dissertation binds two neuroeconomic studies of decision-making with an introduction and concluding commentary. The introduction presents a brief introduction to Neuroeconomics, meant to locate both research studies in the existing literature and philosophy of this field. The conclusion provides a brief appraisal of the role of Neuroeconomics in further advancing the kind of research into decision-making reported here.

Both studies in this dissertation comprise investigations of human behavior during experience-based decision-making, with a special focus on the fundamental value computations that underlie such choice behavior.

Study 1 investigates the role of neural reinforcement signals during learning of a strategic decision task from experience.

Study 2 investigates the moderating effect of intelligence on neural reinforcement signals during a sequential binary choice task.

Study 1 is reproduced from (Hawes, Vostroknutov & Rustichini 2013), and study 2 is reproduced from (Hawes, DeYoung, Gray & Rustichini; *under review*).

# Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Dedication</b>	<b>ii</b>
<b>Abstract</b>	<b>iii</b>
<b>List of Tables</b>	<b>vi</b>
<b>List of Figures</b>	<b>viii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Experience and Abstract Reasoning in Learning Backward Induction (Hawes, Vostroknutov &amp; Rustichini, 2012)</b>	<b>5</b>
2.1 Introduction . . . . .	6
2.2 Experimental Paradigm . . . . .	8
2.2.1 Hit-N Game . . . . .	8
2.2.2 The fMRI study . . . . .	10
2.3 Model . . . . .	11
2.3.1 Optimal Information processing . . . . .	14
2.3.2 Model Predictions . . . . .	15
2.4 Results . . . . .	16
2.4.1 Behavioral Results . . . . .	16
2.4.2 The fMRI data . . . . .	25
2.4.3 Prediction error response in the Striatum and Insula . . . . .	26

2.4.4	Prediction error response for Fast Learners . . . . .	28
2.4.5	Experience-based learning and abstract reasoning . . . . .	31
2.5	Conclusions . . . . .	33
2.6	Method and Materials . . . . .	34
2.6.1	MRI data acquisition . . . . .	34
2.6.2	GLM models . . . . .	35
2.6.3	Fast and Slow Learners . . . . .	35
<b>References</b>		<b>37</b>
<b>3</b>	<b>Intelligence Moderates Neural Responses to Monetary Gains and Losses</b> <b>(Hawes, DeYoung, Gray &amp; Rustichini, <i>Under Review</i> )</b>	<b>39</b>
3.1	Introduction . . . . .	40
3.2	Method . . . . .	42
3.2.1	Participants . . . . .	42
3.2.2	fMRI Procedures . . . . .	42
3.2.3	Stimuli and Design . . . . .	43
3.2.4	fMRI Data Analysis . . . . .	45
3.3	Results . . . . .	47
3.3.1	Brain Imaging Results . . . . .	47
3.3.2	Behavioral Results . . . . .	52
3.4	Discussion . . . . .	55
<b>References</b>		<b>61</b>
<b>4</b>	<b>Conclusions</b>	<b>65</b>
<b>References</b>		<b>67</b>

# List of Tables

2.1	<b>Average response Time and Average Correct Rate OLS for both games</b> . . . . .	13
2.2	<b>Response time in G(15, 3): Panel Data analysis</b> . . . . .	22
2.3	<b>Response time in G(17, 4): Panel Data analysis</b> . . . . .	23
2.4	<b>Interaction between Fast/Slow learner, and Early/Late trial on BOLD signal contrast position 11 - position {1, 2} in Striatum.</b>	31
3.1	<b>Regions of interest differentiating between Gain and Loss feedback:</b> . . . . .	57
3.2	<b>Impact of IQ on BOLD signal after losses.</b> Note: The contrast between neural loss and win responses correlates significantly with intelligence. Controlling for working memory and age does not affect the independent effect of intelligence. Loss responses are negative (significant negative intercept term) and higher IQ shifts loss responses closer to zero, i.e. prediction errors after losses decrease with increasing IQ. Results from a robust regression (*p < .05, ** p < .01, *** p < .001) .	57
3.3	<b>Influence of past computer choices on subject choices.</b> Note: Mixed effects logistic panel regression, grouped by subject. N = 3478. <i>H<sub>CPU</sub></i> i-back is a dummy variable, representing whether the computer chose High i periods back. Freq Hcpu is the historic frequency of High choices by the computer, expressed as deviance from 0.5. (*p < .05, ** p < .01, *** p < .001, BIC: Bayesian Information Criterion, LL: Log Likelihood) . . . . .	58



3.4	<b>Influence of conditional transition probabilities on subject behavior.</b> Note: Mixed effects logistic panel regression, grouped by subject. N = 3478. $H_{CPU}$ i-back is a dummy variable, representing whether the computer chose High i periods back. Freq $H_{CPU}$ is the historic frequency of High choices by the computer, expressed as deviance from 0.5. (*p < .05, ** p < .01, *** p < .001, BIC: Bayesian Information Criterion, LL: Log Likelihood) . . . . .	59
3.5	<b>Relation of IQ to conditional probabilities influencing choice behavior.</b> Note: Mixed effects logistic panel regression, grouped by subject. N = 3478. $H_{CPU}$ i-back is a dummy variable, representing whether the computer chose High i periods back. Freq $H_{CPU}$ is the historic frequency of High choices by the computer, expressed as deviance from 0.5. (*p < .05, ** p < .01, *** p < .001, BIC: Bayesian Information Criterion, LL: Log Likelihood) . . . . .	60

# List of Figures

2.1	<b>Board of positions for the game <math>G(15, 3)</math>.</b> This is the board of positions used in the imaging experiment. The superimposed red rectangles indicate the losing positions 3, 7, and 11. The green rectangles indicate the winning positions. In the lower section of the figure two displays indicate the current score of the subject and the (computer) opponent. . . . .	9
2.2	<b>Average response Time and Average Correct Rate.</b> The averages are computed for each subject over the trials for the $G(15, 3)$ (on the left) and $G(17, 4)$ on the right. . . . .	12
2.3	<b>Response time.</b> Average Response time across trials in $G(15, 3)$ and $G(17, 4)$ . The plot shows an unexpectedly long response time for the very first trial, which is driven by subjects' response time at the initial onset of the game (see also figure 7). At the onset of the game subjects appear to require additional time to familiarize themselves with the game environment. Removing the initial position of the initial round produces an increase in response time for $G(15, 3)$ in line with model predictions. . . . .	17
2.4	<b>Error rate by Period.</b> Average error by type for $G(15, 3)$ . . . . .	18
2.5	<b>Last Trial for Error.</b> Whisker plot of trial during which the last error occurred; separated by type. . . . .	19
2.6	<b>Response Times in losing positions, <math>G(15, 3)</math>.</b> For each of the 20 periods in which the game $G(15, 3)$ was played we report the average response time at each of the losing positions, 3, 7, and 11. . . . .	20
2.7	<b>Response Times in winning positions.</b> As in the previous Figure 2.6 we report for each of the 20 periods in which the game $G(15, 3)$ was played the average response time at each of the four winning positions groups. . . . .	21

2.8	<b>Brain activity at the losing position 11 in <math>G(15,3)</math>.</b> Contrast obtained from a GLM with 16 predictors on all 12 subjects. In the GLM we use the same 7 groupings for positions in the game, and differentiate between positions during the first (early) and last (late) 10 trials, for a total of 14 predictors. An additional predictor for computer choices and a constant term describe the full model. The contrast used in the figure shows activation when the current position is 11 during both early and late trials compared to activation at positions $\{1,2\}$ during early and late trials. The map shows activation at a false discovery rate $q < 0.05$ . . . . .	27
2.9	<b>Progression of activation at losing position 11 for Fast learners.</b> GLM and contrasts as for figure 2.8, but limited to Fast Learners. . . . .	28
2.10	<b>Progression of activation in Fast and Slow learners. Panel A:</b> Contrast obtained from a GLM with 16 predictors on 6 subjects classified as Slow Learners. The depicted contrast shows activation at position 11 compared to activation at position 1,2 during late trials. $p < 0.005$ uncorrected, $t > 4.77$ . <b>Panel B:</b> Same model as panel A. The depicted image subtracts the contrast obtained for positions 11 vs $\{1,2\}$ in late periods from the contrast obtained for those positions during early periods. Positive identification of Striatum in this contrast, is driven by a more strongly negative activation at position 11 in late periods for Slow learning subject. <b>Panel C:</b> 12 predictor GLM for 6 subjects classified as Fast Learners. As in panel B, we show the subtraction of the contrast $(11_{\text{early}} - 1,2_{\text{early}}) - (11_{\text{late}} - 1,2)$ . We find activation in Medial Prefrontal Gyrus (MPFG), VLPFC and Striatum. Negative identification in Striatum is driven by a more strongly negative response at position 11 during early trials for Fast Learners. $p < 0.005$ uncorrected for all images depicted here. . . . .	30

2.11	<b>Fast learners at the losing positions.</b>	Contrast for Fast Learners during first 10 periods at $p < 0.005$ . The contrasts used are for each of the three losing position, compared to the predictor given by the game being in position $\{1, 2\}$ . For example the contrast for the position 3 indicates the comparison between position 3 and position $\{1, 2\}$ . The <b>top panel</b> shows the clusters in VLPFC activated for the different contrasts. The <b>lower panel</b> shows the activation for the contrasts $(3, \{1, 2\})$ , $(7, \{1, 2\})$ and $(11, \{1, 2\})$ . Activation in VLPFC is not found in Fast Learners during the last 10 rounds, and Slow Learners show it only during the last 10 rounds for $11-\{1, 2\}$ . Note, from Figure 2.5 that the Fast Learners do not make mistakes past round 10, while Slow Learners commit mistakes even at position $e_{11}$ past round 10. It should be noted here that a direct test of the interaction between subjects' categorization as Fast/Slow Learner and a dummy variable indicating Early/Late trials did not yield a statistically significant effect ( $p=0.158$ two-sided, see table 4). We believe that the failure to identify such an effect at conventional significant level in our data may be due to small sample size, and an insufficiently precise measure of when subjects learn the game. . . . .	32
3.1	<b>Task design:</b>	Subjects engaged in 40 reward relevant trials. At the beginning of these trials a US dollar bill was displayed for 3 seconds, during which subjects pressed one of two buttons indicating their guess of whether a computer generated number would be Low (1-3) or High (4-6). Guesses were followed by two feedback screens, each 1 second in duration. Reward relevant trials were interspersed with 20 reward-neutral trials, which were signaled by a gray rectangle. In this figure (but not in the task itself), a green outline marks the sequence of screens seen during gain trials. A red outline marks the sequence of screens for loss trials, and grey marks reward neutral control trials. Correct/incorrect outcomes were rewarded/punished with \$ US+/-1. . . . .	44

3.2	<b>BOLD response <math>FB^+ &gt; FB^-</math> (<i>Gain &gt; Loss</i>) in the bilateral caudate.</b>	
	BOLD response in caudate after gains is significantly higher than after losses. This results holds for whole-brain analysis as well as a masked regression on the anatomically defined caudate (top left). Event related average BOLD response in caudate replicates the time course identified by Delgado et al. (2000) (top right). Additional areas showing significantly higher activation after gains compared to losses in whole brain analysis include the medial prefrontal cortex and posterior cingulate cortex (bottom panels). . . . .	48
3.3	<b>Individual subject GLM regressors and intelligence.</b>	
	The box plot shows distribution of regression coefficients for Gain, Loss and Control trials extracted from the anatomically defined caudate head ROI. Scatter plots and regression lines depict the relation between IQ and contrasted GLM estimates of BOLD response to feedback for the same region. Potential outliers in the boxplot are depicted as black dots in the scatter plots. Statistical results are unaffected by the removal of these potential outliers. The correlation coefficients between IQ and the contrasts depicted in the 2 <sup>nd</sup> , 3 <sup>rd</sup> , and 4 <sup>th</sup> panel of the figure are $r = .24$ ( $p = .02$ ), $r = 0.01$ ( $p = .93$ ), and $r = -.24$ , ( $p = .02$ ) respectively. . . . .	50

# Chapter 1

## Introduction

Neuroeconomic studies apply the formalism and conceptual structure of mathematical game theory and computational neuroscience to psychological, economic and neurological representations of human behavior. The experimental toolkit of Neuroeconomics draws upon the entire breadth of current neuroimaging techniques, such as functional magnetic resonance imaging, electroencephalography, magnetic encephalography, and transcranial magnetic stimulation (Fehr & Rangel, 2011, Glimcher & Rustichini, 2004). Unlike traditional economic models, which typically constitute "*as if*" descriptions of human behavior, neuroeconomic models aim at "*as is*" representations of human decision-making<sup>1</sup>. This means that traditional economic models aim to describe human choices in terms of mathematical equilibria among the constraints and incentives of a decision situation, without presuming that decision-makers necessarily perform the exact computations which mathematically identify these equilibria (Bernoulli, 1954; Savage, 1954; Koopmans, 1960; Houthakker, 1950; Samuelson, 1938).

In comparison, neuroeconomic models aim at structural description of the decision process itself; i.e. at describing the neurobiological mechanisms underlying basic value computation (e.g. Plassman, O'Doherty & Rangel, 2007), value comparison (e.g. Christopoulos, Tobler, Bossaerts, Dolan & Schultz 2009), and outcome evaluation during simple choices (e.g. San Martin et al, 2010). At the level of theory, Neuroeconomics

---

<sup>1</sup> In this comparison the term traditional economics is limited to descriptive microeconomics, because it is the branch of economics to which Neuroeconomics can be most usefully compared. In this dissertation we do not consider a relation to normative microeconomics, or macroeconomics (normative or descriptive).

additionally aims to provide an organizing framework for how these fundamental value computations link to higher-level behaviors and choice (Dickhaut, Rustichini, Smith, 2009; Krajbich, Camerer, Ledyard & Rangel, 2009). This is what makes Neuroeconomics similar to traditional economics, which similarly provides a formal framework for mapping prespecified utility to complex behaviors (Bernoulli, 1954), and what differentiates Neuroeconomics from Social Cognitive Neuroscience more generally.

The research presented in this dissertation is neuroeconomic in content, method and philosophical approach. While each study stands by itself as an individual experimental investigation of a narrowly defined research question, both studies are connected by a common emphasis on linking fundamental value computations during experience-based decision-making to individual differences in preferences and strategic choice. By doing so, the work in this dissertation also represents an effort of integration between economics and psychology; more specifically the psychology of personality.

Attempts of integrating psychology (and sociology) and economics have a long history in both disciplines (e.g. Simon 1956) and appear to be experiencing a recent revival (Rustichini, DeYoung, Anderson & Burks, 2011; Ferguson, Heckman, Corr, 2011). When considering efforts for integrating psychology and economics, it again makes sense to distinguish between what I have above termed the traditional economic approach versus the neuroeconomic approach:

The traditional economic approach for integrating psychology into economics – at times implicitly, at times explicitly – has chiefly been concerned with the finding that the decision-makers objective reality (i.e. the one modeled by the economist) often systematically deviates from the decision-makers subjective reality; presumably because of psychological biases for what kind of information is attended to, or even how certain types of information are interpreted in context (Strotz, 1955; Kahneman & Tversky, 1979; Bell, 1982; Loomes & Sugden, 1982). Examples of such deviations are numerous, and range from famous violations of expected utility theory during simple binary choice (Allais, 1953; Ellsberg, 1961) to prominent violations of pure self-interest in multi-person decision situations (Loewenstein, Thompson & Bazerman, 1998; Thaler 1988; Henrich et al, 2001). From an economic perspective, Prospect Theory (Kahneman & Tversky, 1979), Cumulative Prospect Theory (Tversky & Kahneman, 1992) theories of Bounded Rationality (Simon, 1982; Gigerenzer 2002), and a body of Social Psychology research

commonly labeled Behavioral Economics (see Ariely & Norton, 2007) can therefore all be understood as attempts at replacing the traditional description of economic man - homo economicus (Mill, 1836) - with alternative agent descriptions of greater psychological validity<sup>2</sup> (e.g. Gilboa, Schmeidler, 1989; Maccheroni, Marinacci & Rustichini, 2006).

More recent research related to the integration of economic decision theory and personality theory from psychology has primarily focused on personality traits to

a) explain behavioral heterogeneity between actors (Borghans, Duckworth, Heckman & ter Weel, 2008; Almlund Duckworth, Heckman & Kautz, 2013, Denissen & Penke, 2008),

b) within-actor changes of revealed preferences and personality across tasks (e.g. Roberts & Jackson, 2008 & Dingemans, 2010), and

c) observed changes of actors over the life-span (personality as investment), (e.g. Heckman, 2007; Cunha, Heckman & Schennach, 2010).

These attempts have been primarily motivated by the ability of psychological traits to predict economic outcomes (e.g. Roberts, Kuncel, Shiner, Caspi & Goldberg, 2007; Borghans, Duckworth, Heckman & ter Weel, 2008; Coates, Gurnell & Rustichini, 2009), and appear to be ushering in a shift away from the traditional economic actor - composed of predefined parameters for risk sensitivity and attitude towards temporally delayed rewards - towards a more complex actor definition comprising multiple parameters reflecting basic psychological traits (Rustichini, 2009).

In regards to integration with psychology, the neuroeconomic approach differs from traditional economics mostly in terms of its level of analysis. Like traditional economics, Neuroeconomics considers the pervasive role of psychological personality during decision making, and attempts to account for personality-based individual differences in its description of the decision-maker (Rustichini, 2009). However, the neuroeconomic revision of the two-dimensional decision maker appears to also entail a fractioning of the decision-maker into multiple systems (e.g. Sanfey, Loewenstein, McClure & Cohen 2006; van den Bos & McClure, 2013). Hence, instead of equipping the decision-maker with a psychologically informed utility function (as would be the traditional approach,

---

<sup>2</sup> From a psychological perspective this research is probably better understood as an attempt to provide behaviorist psychology with a formal framework for mapping external context to behavior (Prospect Theory and Bounded Rationality), or as providing formal notation for social psychological theories (Behavioral Economics).



Neuroeconomic research appears to be moving in the direction of describing the individual in terms of an emergent entity of (two or more?) potentially competing/potentially cooperating systems; each with its own utility function and psychological properties<sup>3</sup> .

The neurobiological system under consideration in this dissertation is referred to as the persons dopaminergic reward system. It consists of the ventral tegmental area, nucleus accumbens and connected parts of the prefrontal cortex. This reward system has been identified as critical to various forms of reinforcement learning, which appears to constitute a fundamental and pervasive process of how the brain assigns value to choice alternatives from experience (Dayan & Daw, 2008). Activity in the dopaminergic system during reinforcement learning and cognitive decision tasks have been linked to individual differences in intelligence (Previc, 1999; Schlagenhaut et al, 2012). Intelligence, or cognitive ability, occupies a special position in the description of economic behavior: In various applied and experimental studies it has been demonstrated that cognitive skills significantly affect important life-time outcomes such as health, divorce, education, income, but also correlate with systematic differences in attitude to risk, temporal delay, and strategic reasoning in competitive games (Shamosh & Gray, 2008; Burks, Carpenter, Goette, Rustichini, 2009; Shamosh et al 2008). In other words, intelligence links dimensions of decision-making (risk and discounting) traditionally thought to be independent (Rustichini, 2009). Because this link extends beyond general improvements of information processing along different phases of the decision process, and across different domains of decision making (Burks, Carpenter, Goette, Rustichini, 2009), Neuroeconomics should be concerned with understanding these simultaneous correlations between intelligence and preferences at the level of the decision makers fundamental value computations<sup>4</sup> . The studies presented below should be read with the above context in mind, and viewed as inceptive contributions towards this greater theoretical objective.

---

<sup>3</sup> This kind of fractioning of the decision-maker is of course not unique to Neuroeconomics (see for example Fudenberg Levine, 2006 for a dual self model of self-control), but neuroeconomic investigations of psychological personality appear more closely aligned with these kind of representations of the decision-maker, than is the case for traditional models that consider personality. Philosophically, the multiple-systems approach to decision making has strong roots in neuroscience and psychology (e.g. Schneider & Shiffrin, 1977).

<sup>4</sup> A more detailed discussion of the role of intelligence is included in study 2.

## Chapter 2

# Experience and Abstract Reasoning in Learning Backward Induction (Hawes, Vostroknutov & Rustichini, 2012)

Backward induction is a benchmark of game theoretic rationality, yet surprisingly little is known as to how humans discover and initially learn to apply this abstract solution concept in experimental settings. We use behavioral and functional magnetic resonance imaging (fMRI) data to study the way in which subjects playing in a sequential game of perfect information learn the optimal backward induction strategy for the game. Experimental data from our two studies support two main findings:

First, subjects converge to a common process of recursive inference similar to the backward induction procedure for solving the game. The process is recursive because earlier insights and conclusions are used as inputs in later steps of the inference. This process is matched by a similar pattern in brain activation, which also proceeds backward, following the prediction error: brain activity initially codes the responses to losses in final positions; in later trials this activity shifts to the starting position.

Second, the learning process is not exclusively cognitive, but instead combines experience based learning and abstract reasoning. Critical experiences leading to the adoption of an improved solution strategy appear to be stimulated by brain activity in the reward system. This indicates that the negative affect induced by initial failures facilitates the switch to a different method of solving the problem. Abstract reasoning is combined with this response, and is expressed by activation in the ventrolateral prefrontal cortex. Differences in brain activation match differences in performance between subjects who show different learning speeds.

## 2.1 Introduction

Backward induction (BI) is a recursive algorithm, wherein inferences regarding a decision problem made at an earlier stage are applied to the process of deriving yet further inferences on the problem. As a mathematical construction, backward induction constitutes a benchmark of game theoretic rationality that prescribes the behavior of rational players in finite sequential games of perfect information. In game theory, the concept is at the basis of abstract theorems; most notably Zermelo's theorem (1908) on the existence of equilibria in pure strategies for those games, or Selten's theorem (1967, 1974) characterizing Sub-game perfect equilibria.

In contrast to its applicability in mathematical proofs, backward induction has at

times been considered inapposite as a descriptive account of the cognitive processes operating in human subjects during the sort of strategic interactions that game theorists would conceptually represent as sequential games (Johnson, 2002). In part, backward inductive reasoning has been considered an unlikely description of the human thought process during such game situations, because it requires the cognitive enactment of a solution concept that is considerably complex and ostensibly unnatural. However, we find evidence that subjects playing the sequential game of perfect information in our experiment have a common pattern of learning the optimal solution, and that although individuals may differ in their speeds for producing this pattern, the path is common, and reproduces the steps of the backward induction algorithm.

To argue effectively this conclusion, we first review what the backward induction algorithm prescribes. We illustrate backward induction in finite sequential games of perfect information. These are games in which players alternate in actions, know and remember precisely the choices made by other players in previous stages of the game, and know exactly the payoff structure for all players involved. A strategy for a player is a rule assigning a move at every decision point. For these games, backward induction prescribes the following procedure to construct a strategy for every player: At the very last stage of the sequential game, when the final player makes the last choice of the game, she should move to maximize her payoff from the choice made at this stage. We can call this step the last stage. Since the game ends at the last stage, and because payoffs are known, rationality prescribes the outcome of the last player's choice. In other words, all players, including the last player herself, can unambiguously determine what would constitute the payoff maximizing choice given the options available at the last stage. The second to last player should anticipate the unambiguous criteria according to which the last player will choose, and conclude that his second to last choice will ultimately yield the payoffs induced by the optimal move of the last player. After this has been established, the original game is effectively replaced by a game with shorter length where the last move is eliminated and the payoff at the last stage is defined to be the payoff following the optimal choice of the last player. Iterating this process until the initial decision point is reached produces a strategy for every player.

Backward inductive reasoning relies in an essential way on the human ability for recursive thought, which itself has been posited as a prerequisite for quintessential human

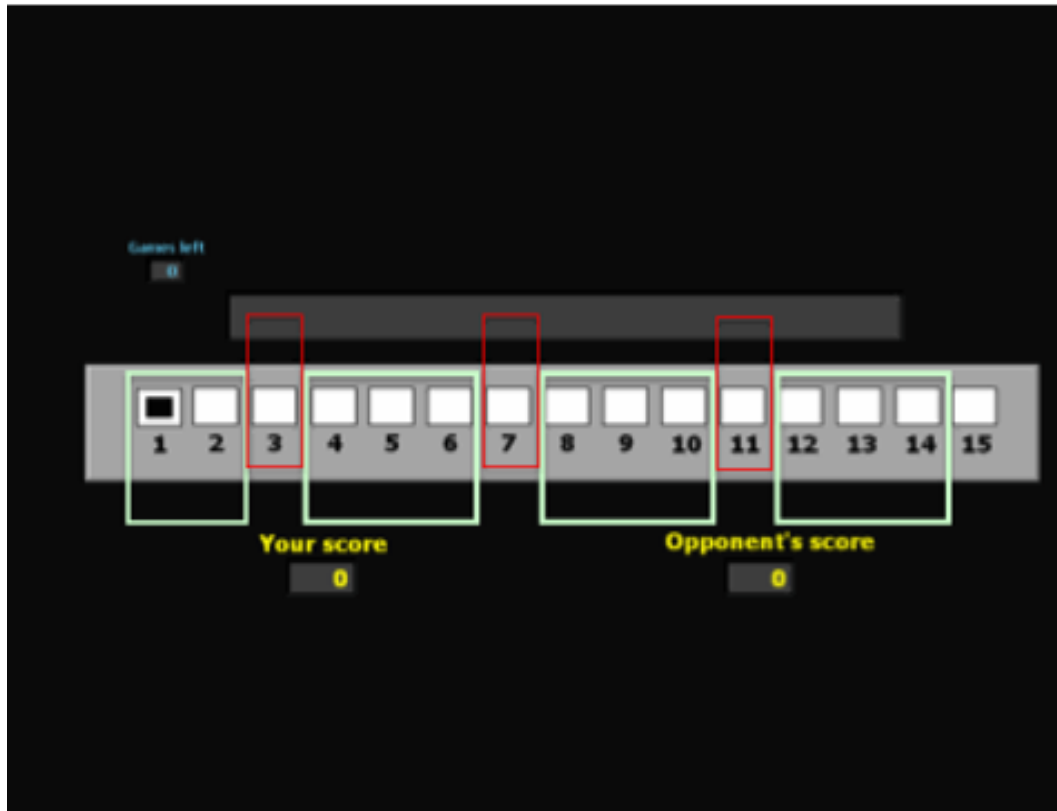
achievements such as language acquisition and basic numeracy (Hauser, 2002). Hence, to the extent that this innate human ability to think recursively manifests itself in strategic games, there may indeed exist a link between fundamental cognitive processes in humans and the abstract game theoretic concept of backward induction. Following this conjecture, we conduct two studies of a particular strategic social interaction that facilitates recursive learning. These studies were designed to address two fundamental questions: First, what are the neural correlates of recursive learning in the strategic environment, and second, how do the cognitive processes involved in recursive learning connect to abstract backward inductive reasoning.

## **2.2 Experimental Paradigm**

### **2.2.1 Hit-N Game**

The finite sequential game which subjects play in our experiment is played by two parties on a virtual playing board, and is the same as used in (Gneezy, 2010). The board used to display the game in the imaging study is presented in Figure 2.1.

Figure 2.1: **Board of positions for the game  $G(15, 3)$ .** This is the board of positions used in the imaging experiment. The superimposed red rectangles indicate the losing positions 3, 7, and 11. The green rectangles indicate the winning positions. In the lower section of the figure two displays indicate the current score of the subject and the (computer) opponent.



For the basic variant of the Hit-N game used in this experiment, the first player to move is allowed to move a single common playing piece on the board, and she is allowed to move it only forward, by 1, 2, or 3 positions; no more or no less. The move then goes to the second player, who is allowed the same action of moving the figure 1, 2, or 3 positions forward. From thereon the opportunity to move according to the 1-2-or-3-only rule alternates between the two players. The player who reaches the final position (15 in experiment 1) first wins that game. We refer to this game as  $G(15, 3)$ . A second game in our experiment involves the game  $G(17, 4)$  which is played on a virtual playing board of length 17, and allows players to move 1, 2, 3 or 4 positions forward.

We apply backward induction reasoning to this game to derive the optimal strategy: Players moving in position 12, 13 or 14 can win by reaching position 15 immediately. It follows that players moving at 11 have lost, since they can only move to 12, 13 or 14, where the opponent, as we have just seen, wins. Players can now replace the original game with the shorter game where the first player to reach position 11 wins: a move is optimal in the original game if and only if it is optimal in the reduced game. The same argument, repeated, shows that the player who first gets to position 7 wins; after which it can be concluded that the first to reach 3 wins. In summary, all positions different from 3, 7 and 11 are winning positions, because from there the player who is moving can reach either position 3, 7 or 11, and win: she just has to be sure to move there. On the other hand, positions 3, 7 and 11 are losing positions, and there is not much that the player moving there can do but hope for an error of the opponent. The argument we have just presented is the BI solution to  $G(15, 3)$ . A similar argument shows that the losing positions in  $G(17, 4)$  are  $\{2\}, \{7\}, \{12\}$ , and the groups of winning positions are  $\{1\}, \{3, 4, 5, 6\}, \{8, 9, 10, 11\}, \{13, 14, 15, 16\}$ .

### **The Behavioral study**

We use data from (Gneezy, 2010) as a behavioral sample, and focus here on error rate, response time, and their relation. A total of 72 subjects competed in 20 trials of  $G(15, 3)$ , and 52 out of the 72 subjects played an additional 10 trials of  $G(17, 4)$ . The incentive structure for  $G(15, 3)$  promised \$5 for winning more than 5 trials over the 20 game period, and \$20 for winning more than 11 trials. For  $G(17, 4)$  subjects were promised \$10 for winning more than 5 games.

#### **2.2.2 The fMRI study**

A total of 12 subjects participated in the MRI study. They played first 20 trials of  $G(15, 3)$ , then 20 trials of  $G(17, 4)$  against a computer. The game, incentives and instructions include three modifications to those used in the behavioral study. First, subjects are informed that they are playing a computer, programmed to win and subject to small errors. Also subjects play 20 trials of  $G(17, 4)$  (compared to 10 trials in study 1). Finally subjects were allowed 10 seconds to make a choice on each of their turns.

Data were collected at the Center for Magnetic Resonance Research (CMRR) at University of Minnesota using a 3T Siemens Trio scanner. Both studies were approved by the Institutional Review Board (IRB) at the University of Minnesota. Subjects in both studies signed an informed consent form after they were given the instructions.

## 2.3 Model

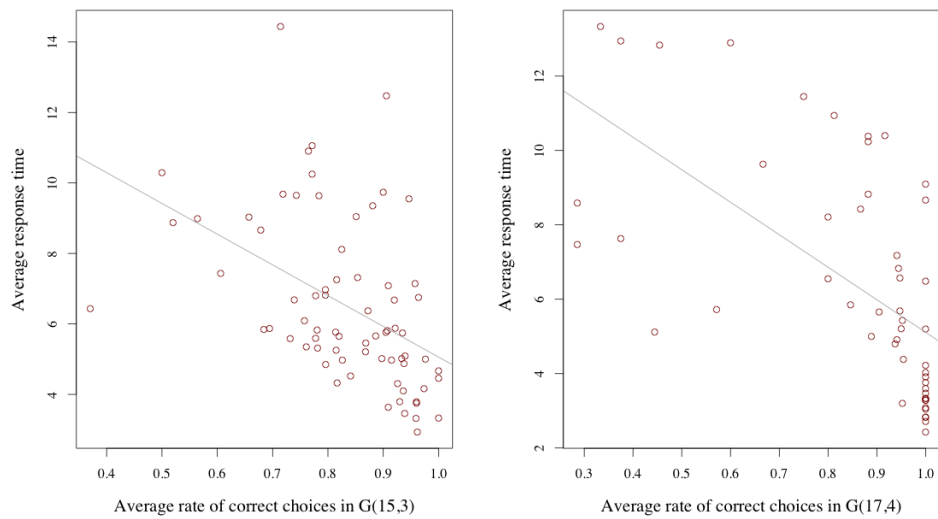
To motivate the need for a theoretical model and the structure we are going to use, we begin by considering the relation between two key observable variables, response time and error rate. The response time is the length of the time interval between the moment in which the move of the opponent (another player or the computer) is observed and the moment in which the subject makes his next move. To define the error rate, we focus on  $G(15, 3)$  and note that at every winning position one, and only one, of the possible moves is correct, and the other two are incorrect. An error is the choice of the wrong move, and the error rate is the frequency of this event, conditional on the position being a winning position for the subject (these are the only positions at which an error is possible). The correct response rate is the difference from 1 of the error rate.

How are these two variables related? It may be reasonable to assume that, everything else being equal, a longer response time is associated with a higher correct response rate. This for example would be the case if the response time were varied exogenously, since by thinking about the problem for a longer time the subject would be more likely to achieve a richer understanding of what constitutes a good move. We point out that the condition of everything else being equal is crucial for this assertion. Considering now, that the length of the response time is not exogenous, but is decided upon by the subject who is reasoning about the decision, the relationship between response time and error rate may, however, be a different; indeed reversed: Since the reasoning activity can be assumed - in some measure - costly, a decision maker may compare and trade-off the estimated returns and costs from the reasoning activity. If the returns are estimated to be low, he may prefer to discontinue the process. If they are high, he might continue. Consider also, ability as an individual characteristic: An individual with lower cognitive skills may find the returns to his reasoning unsatisfactory, stop early, and be more likely to make the wrong choice. Similarly, a subject who has not acquired a basic familiarity



with the game may conclude very little from his examination, stop cognitively engaging, and commit errors at a high rate. Both cognitive ability and problem familiarity are subsumed under the concept of ability. Considering response time as a choice variable together with differences in ability, the average relation at the individual level between response time and correct response rate may therefore be negative. In our data we find this to be the case. Figure 2.2 illustrates this point.

Figure 2.2: **Average response Time and Average Correct Rate.** The averages are computed for each subject over the trials for the  $G(15, 3)$  (on the left) and  $G(17, 4)$  on the right.



The simple regression in Table 2.1 of the correct response rate on the individual average response time confirms the negative relation, again in both games.

Table 2.1: **Average response Time and Average Correct Rate OLS for both games**

	Error15 b/se	Error17 b/se
Avg. Correct G(15, 3)	-8.581*** (1.766)	
Avg. Correct G(17,4)		-8.633*** (1.797)
Constant	10.993*** (0.952)	11.010*** (1.017)
r2	0.252	0.316
N	72	52

Given the observed relation it appears particularly prudent that we consider a model in which response time is endogenously determined, and that reflects the notion that subjects choose to think about a problem, decide whether to stop thinking, and only then select a move.

### 2.3.1 Optimal Information processing

In our experiment, at each turn, a player observes the position in the game, considers a set of potential cues and insights, and tries to identify the best move at the current position. At any point in time before choosing a move, he can terminate the process and then make a move determined by the conclusions reached up to this point. If he does not terminate the process, he has to decide the intensity of the effort devoted to the decision. The quality of his decision will then depend on his ability to reason about the game as well as his effort in doing so. We consider ability as an individual characteristic of the player, and this may describe both a player's natural, general skills, as well as her acquired understanding of the game. We also consider effort as a choice variable. Ultimately, both effort and ability contribute positively to the agent's problem solving success.

We model the above process as an optimal information acquisition problem to be solved in the time interval before the move. In the model, the subject has to choose an action, and has beliefs over which of the feasible actions (for example, the set  $\{1, 2, 3\}$  in  $G(15, 3)$ ) is currently the best. In every instant during this process the agent can observe an informative signal on what the best action is, update her belief, and decide whether to continue the information acquisition process or to stop and choose what at the current belief is the optimal action. The model outlined above constitutes a general inter-temporal decision problem which can be formulated as a dynamic programming problem with an action set that consists of the agent's effort and the decision to continue or stop processing information about the game. The state space of the problem is the set of beliefs over the action set, assigning to each action the probability that it is the best action. Information acquired in every instant is a partially informative signal on the true state; that is, on which among the feasible actions is the optimal one.

### 2.3.2 Model Predictions

It is clear that if ability is so low that any processed signal is entirely non-informative, the optimal time spent should be zero, and that correct response rates in this case will consequently be low. This is likely to occur in the early stages of the game, when subjects are just beginning to familiarize with the task, and lack even the basic insights to make even minor headway into the problem. At this stage we should observe a short response time and a high error rate. The effect should also be more pronounced at the difficult positions, those further from the end: this is because reasoning about the best move can only produce useful insights when the individual has some idea of what happens in later stages of the game, at positions closer to the end. In the initial rounds this understanding of the game at later stages is lacking, and the subject may prefer to discontinue the reasoning soon because it is not producing any useful insights.

At the opposite extreme, if ability is so large that the signal is completely informative, only a short time will be necessary while still leading to a high correct response rate. This is likely to occur of course at the late stages of the game, when a subject has an overall understanding of the optimal strategy. It is also likely to occur at the final positions, where very simple reasoning can provide the conclusion.

Between these two extremes, where signal is partly informative, the optimal policy will prescribe a positive response time. Overall the relation between ability and response time is non-monotonic: likely to be increasing for low values of ability, and decreasing for higher values.

A specific conclusion of the model is that the response time at a position is not necessarily monotonically increasing or decreasing with experience, but might instead be first increasing and then decreasing. At the early stages, low experience, which corresponds to low ability, induces an early stopping of the reasoning process (the information acquisition in our model), a short response time and a high error rate. At intermediate stages, as the subject acquires some basic understanding of the game, reasoning becomes more informative, hence stopping is postponed. Finally, in later periods the response time declines as subjects simply implement a solution algorithm which they now understand.

We will see that subjects' behavior broadly matches these predictions, and provide the conceptual framework for the analysis of the imaging data.

## 2.4 Results

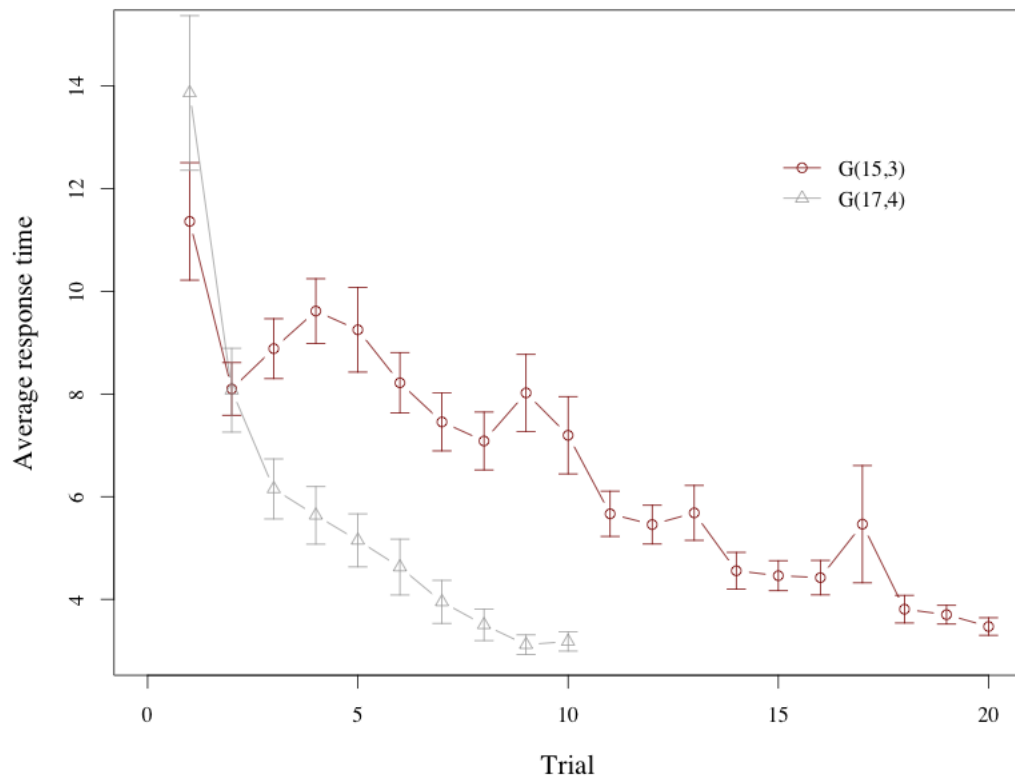
### 2.4.1 Behavioral Results

We review the basic behavioral results presented in (Gneezy, 2010) to prepare for the analysis of the imaging data. To analyze error rate, we define a subject's error  $j$  as a subject's failure to move the marker to  $j$ , whenever this is possible and moving to  $j$  is part of the winning strategy. In  $G(15, 3)$  the possible errors of interest are failures to move the marker to any of the positions 3, 7, 11 or 15 whenever this would be possible. The error rate at  $j$ ,  $e_j$ , is the fraction of times the error is made over the times the subject could avoid the error. For example  $e_3$  is calculated as the number of times the subject had the opportunity of moving her opponent to position 3, yet failed to do so, divided by the times the subject held the move at position 1 or 2 in the game. The average error rate is the number of errors made at a winning position divided by the number of times the subject was in a winning position.

Response times for subjects show a marked decline across trials: see Figure 2.3, with subjects requiring more than 8 seconds on average to make a choice during the first three periods of the game, but not even half of that during the last 3 periods.

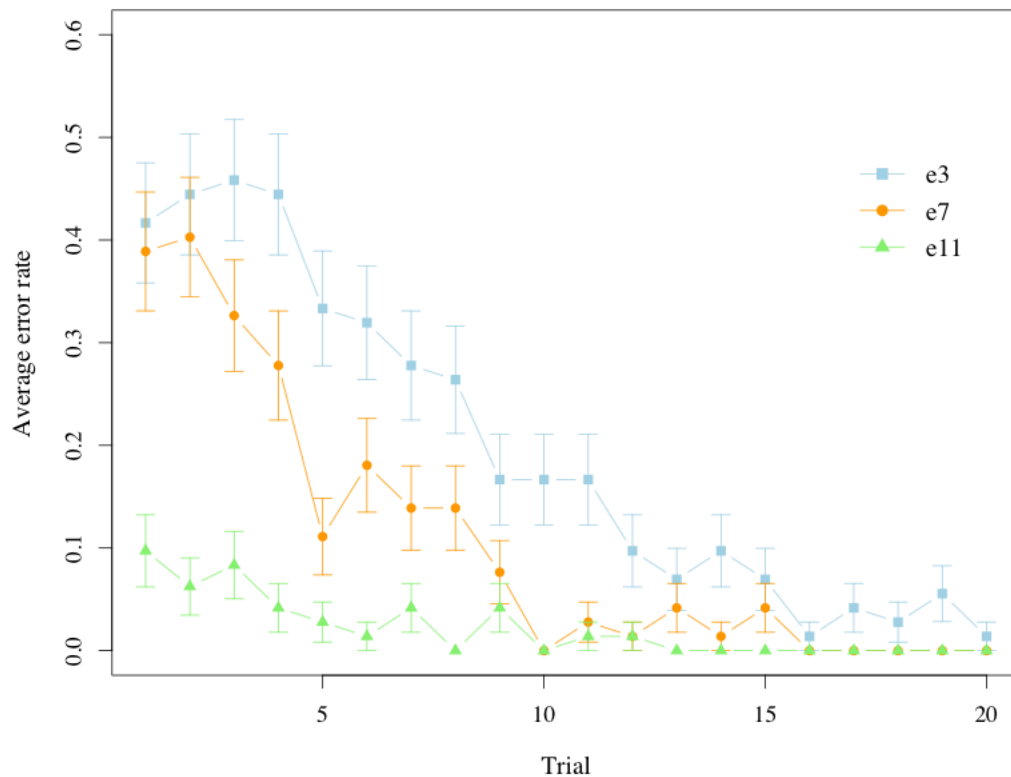
There is a substantial difference in the evolution of the Response Time in the two games. Consider first the Game  $G(15, 3)$ . Note that the first trial has a very special role, since it is the one where subjects get acquainted with the task, and the rules of the game. If we ignore the first trial we see that the response time increases from the second to the fourth trial, and then declines, as the model predicts.

Figure 2.3: **Response time.** Average Response time across trials in  $G(15, 3)$  and  $G(17, 4)$ . The plot shows an unexpectedly long response time for the very first trial, which is driven by subjects' response time at the initial onset of the game (see also figure 7). At the onset of the game subjects appear to require additional time to familiarize themselves with the game environment. Removing the initial position of the initial round produces an increase in response time for  $G(15, 3)$  in line with model predictions.



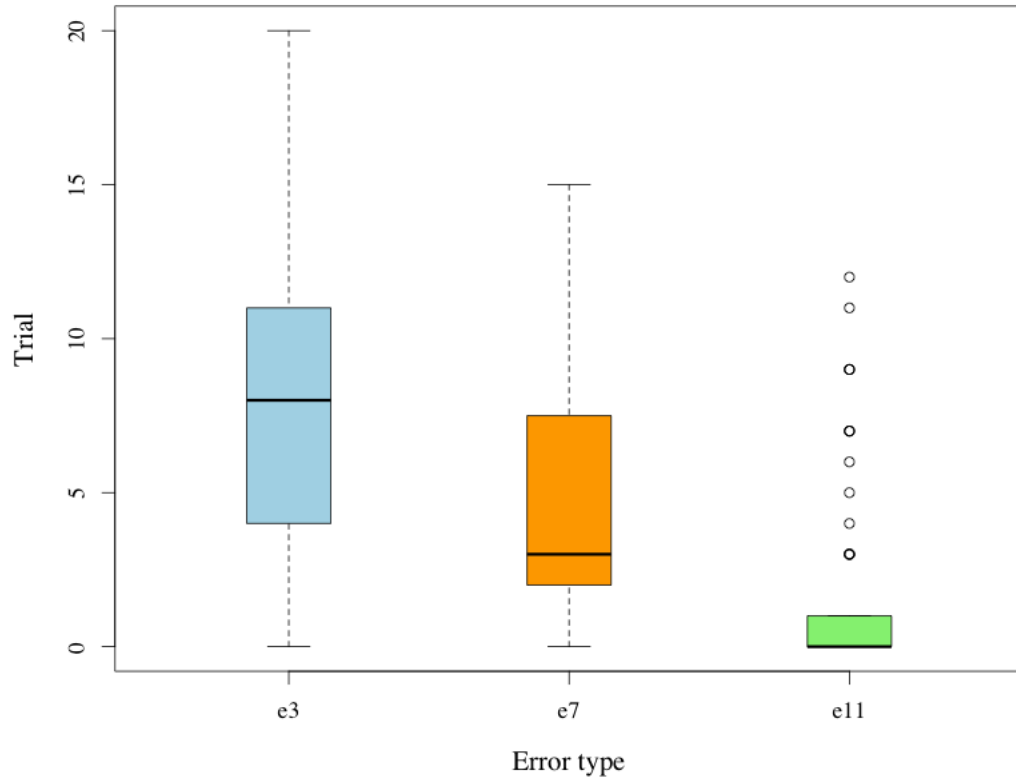
For the first trial of  $G(15, 3)$  the error rate is 0.38, which is significantly lower than the average error rate that would be expected if choices were made randomly. Across 20 trials of study 1, the error rate steadily declines until it almost reaches zero: see Figure 2.4.

Figure 2.4: **Error rate by Period.** Average error by type for  $G(15, 3)$ .



The four possible errors in  $G(15, 3)$  occur at significantly different rates. No subject deviates from the winning strategy choice at the final 3 positions ( $e_{15} = 0$ ). Error rates and average period marking the last occurrence of a particular error are lower for positions closer to the game's end ( $e_3 \geq e_7 \geq e_{11}$ ): see Figure 2.5.

Figure 2.5: **Last Trial for Error.** Whisker plot of trial during which the last error occurred; separated by type.



Each of the differences between  $e_3$ ,  $e_7$ , and  $e_{11}$  is statistically significant ( $p < 0.01$ ), and the pattern suggests that subjects indeed learn to identify losing positions in a sequential manner that begins from the game's final positions. These observations indicate that subjects progress through a sequence of minor realizations towards becoming proficient in the Hit game. The above trends for  $G(15, 3)$  replicate in  $G(17, 4)$ . For both games we observe lower error rates at later positions, and an overall decrease of error rates over repeated trials. Average response times decline across trials in both games.

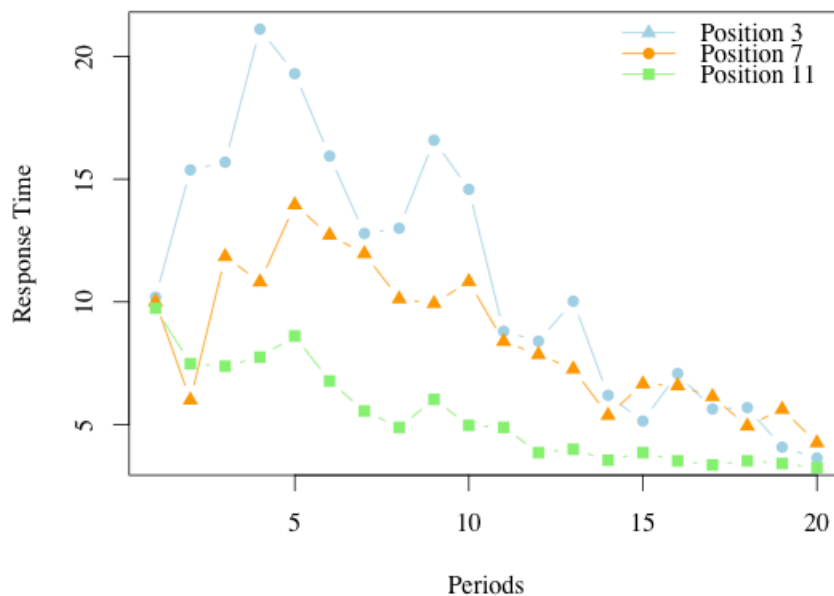
Subjects make significantly fewer mistakes in  $G(17, 4)$  than in  $G(15, 3)$  indicating



that subjects transfer some of their acquired skill to the new game. Observing however, that only 20 out of 72 subjects manage to commit zero errors in  $G(17, 4)$ , it is likely, that most subjects have not fully developed the explicit BI solution to the sequential game after 20 trials of  $G(15, 3)$ .

Figure 2.6 illustrates the average response time in the losing positions, for each of the periods.

Figure 2.6: **Response Times in losing positions,  $G(15, 3)$ .** For each of the 20 periods in which the game  $G(15, 3)$  was played we report the average response time at each of the losing positions, 3, 7, and 11.

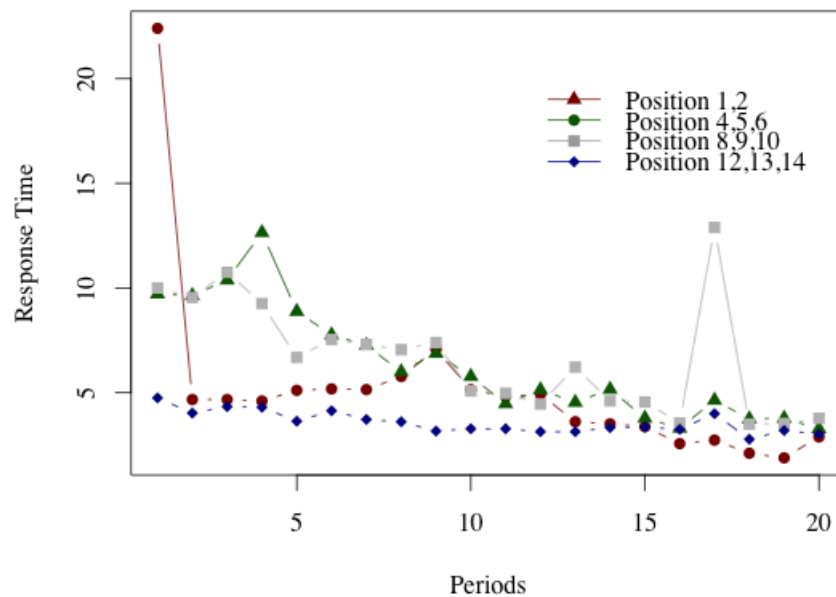


For position 11, the losing position which is closest to the end, the highest response time occurs in the first period, and declines in the periods thereafter. The peak for position 7 is reached at period 4, and that for position 3 is reached at period 5. As the model predicts, the response time is non monotonic over the periods. For example the response time in period 7 is low at the initial stages, when subjects typically have

a limited understanding of the game, but increases as the insight that the position 11 is a losing position is acquired and becomes available in the analysis of what to do at position 7. In later periods the response time at position 7 declines.

A similar relation can be seen in Figure 2.7, which illustrates the average response time at winning positions for  $G(15, 3)$ .

Figure 2.7: **Response Times in winning positions.** As in the previous Figure 2.6 we report for each of the 20 periods in which the game  $G(15, 3)$  was played the average response time at each of the four winning positions groups.



In this case too, the peak for the middle positions (winning positions  $\{4, 5, 6\}$  and  $\{8, 9, 10\}$ ) is reached after an initial low value. The peak is reached at period 4 for  $\{4, 5, 6\}$  and at period 3 for  $\{8, 9, 10\}$ . The response time at the very first positions  $\{1, 2\}$  increases slowly; the maximum is reached at period 8, after an initial spike in period 1 which is likely to be due to the fact that the very first instance of position 1 is also the subjects' very first encounter with the game. The response time at the easy

positions  $\{12, 13, 14\}$  monotonically declines after the initial period.

The figures we have seen present instructive average values over individuals' response times. A more accurate description is provided by the panel data regressions in Table 2.2 for  $G(15, 3)$  game and Table 2.3  $G(17, 4)$  for which the dependent variable is the response time and the time variable for the panel is the index of the period. The independent variables are dummy variables corresponding to the groups of positions. They are indexed in increasing order according to their position on the board, left to right. For example, the first group of winning positions (*Win Pos 1*) in game  $G(15, 3)$  indicates the set of positions  $\{1, 2\}$ . The second group of losing positions for  $G(17, 4)$  indicates the position 7. In both regressions the variable dropped is the final group of winning positions, that is  $\{10, 11, 12\}$  for  $G(15, 3)$  and  $\{13, 14, 15, 16\}$  for  $G(17, 4)$ .

Table 2.2: **Response time in G(15, 3): Panel Data analysis**

	RT151	RT152	RT153
	b/se	b/se	b/se
Win Pos 1	1.472*** (0.447)	1.293*** (0.432)	1.299*** (0.432)
Win Pos 2	2.979*** (0.446)	2.690*** (0.432)	2.679*** (0.432)
Win Pos 3	3.095*** (0.456)	3.009*** (0.442)	3.005*** (0.442)
Losing Pos 1	4.223*** (0.499)	4.820*** (0.484)	4.849*** (0.484)
Losing Pos 2	6.726*** (0.474)	6.946*** (0.459)	6.969*** (0.459)
Losing Pos 3	1.280*** (0.466)	1.299*** (0.451)	1.306*** (0.451)
period		-0.381*** (0.021)	-0.561*** (0.088)
period square			0.009** (0.004)
Constant	3.800*** (0.396)	7.801*** (0.447)	8.460*** (0.544)
r2			
N	5044	5044	5044

Table 2.3: Response time in G(17, 4): Panel Data analysis

	RT171	RT172	RT173
	b/se	b/se	b/se
Win Pos 1	2.972*** (0.920)	2.957*** (0.886)	2.958*** (0.874)
Win Pos 2	3.307*** (0.904)	3.140*** (0.870)	3.066*** (0.859)
Win Pos 3	2.711*** (0.913)	2.653*** (0.879)	2.637*** (0.868)
Losing Pos 1	2.866*** (0.984)	3.327*** (0.948)	3.420*** (0.936)
Losing Pos 2	4.078*** (0.943)	4.205*** (0.908)	4.262*** (0.896)
Losing Pos 3	0.802 (0.926)	0.773 (0.891)	0.774 (0.880)
period		-0.964*** (0.091)	-3.110*** (0.359)
period square			0.204*** (0.033)
Constant	3.866*** (0.694)	8.429*** (0.797)	12.565*** (1.033)
r2			
N	1442	1442	1442

The constant value is similar in both games, and around 4s. The main effect of learning the game is estimated by the variables *period* and *period*<sup>2</sup>, indicating a significant and fast (particularly in the game  $G(17, 4)$ ) decline over time. The other variables confirm what we have seen in the aggregate analysis of the figures. Most notably, the increase in response time at losing positions is significantly higher than the one induced by winning positions; making more likely the conjecture that subjects carry over into the analysis of positions further from the end, insights they have obtained from the losing position 11, and possibly search for equivalent insight among positions earlier in the game.

## 2.4.2 The fMRI data

### Expected activation patterns and regions of interest

On the basis of the model and the analysis of the behavioral data we can formulate hypotheses to be tested in the study of the imaging data.

Learning of the method of backward induction should begin with the negative affective response experienced with moving at position 11, and realizing that the game is lost at that point. This experience should involve the reward system, particularly the Striatum (Schultz, 1997). We explore this hypothesis in section 2.4.3.

The predicted striatal response should be stronger, and occur earlier with subjects for whom behavioral evidence indicates that they possess a better understanding of the optimal strategy. We explore this hypothesis in section 2.4.4.

Further, the analysis of behavioral data has shown longer response times at the losing positions of game  $G(15, 3)$ . The brain activation at these three positions should be similar, but should occur at different points in time during the experimental session. Brain activation should involve both areas associated with reward system and areas involved in abstract reasoning. We test this hypothesis in section 2.4.5 (see in particular in figure 2.11).

One of our main assertions is, that the affective response induced by the understanding that the game is lost at position 11 should occur together with activation of frontal areas involved in planning, particularly VLPFC (Crescentini, 2011). This hypothesis is also examined in section 2.4.5

In what follows we present results obtained from an event-related random effects general linear model (rfxGLM) with 16 predictors. Predictors are dummy variables indicating the 7 sets of positions for  $G(15, 3)$  over the first 10 trials (Early) and the last 10 trials (Late). A dummy variable indicating the computer's turn, and a constant term complete the model. The omitted variable corresponds to a resting period between trials. Unless explicitly stated, all results reported here are significant at an uncorrected threshold of  $p \leq 0.005$ ;  $t(12) \geq 3.49$  for the full sample, or with  $t(5) \geq 4.77$  when split into Fast and Slow Learners. Fast Learners are defined as the 6 subjects with the lowest average error rate over both games. These are incidentally also the 6 subjects with the most wins in  $G(15, 3)$ . Correspondingly, Slow Learners are the 6 subjects with the

highest average error rates.

The model and observed behavior suggests that subjects become proficient at the Hit-15 game via a sequence of insights pertaining to their experience at losing positions; the generic manifestation of which is the avoidance of the losing position at 11, followed by avoidance of position 7, and for some subjects avoidance of position 3. These adaptations, which are likely accompanied by (conscious) realization of these positions as losing positions happen at dramatically varying rates between subjects, and have critical relation to models of prediction error processing and temporal difference learning (see e.g. (Schultz, 1997), or (Daw, 2010)). According to models of prediction error-based learning, unexpected occurrences of losing positions should be accompanied by corresponding BOLD signal change in areas involved with prediction error (PE) tracking, such as the Striatum (Schultz, 1997) and Insula (Preuschoff, 2008). We expect to see these PE responses whenever subjects first realize that a given position is a losing position, and also when subjects are unexpectedly placed onto an already identified losing position; both of which necessitate a yet incomplete understanding of the game, when played against a reasonably proficient opponent such as the computer program used for this study. This expectation follows, because prediction error responses should become less pronounced as subjects gain greater insight into the game as a consequence of their increased ability to accurately predict the games outcome. Hence, once the game's losing positions have been identified, finding oneself at a subsequent losing position becomes almost perfectly predictable at earlier stages, wherefore prediction errors should eventually approach zero.

### 2.4.3 Prediction error response in the Striatum and Insula

All subjects in the fMRI sample learn to identify position 11 as a losing position at some point during the game. In agreement with the idea that the identification of position 11 as a losing position induces an activation in the reward system, we find significant differences in striatal activation for subjects considering a move at losing position 11 compared to when considering a move at winning position  $\{1, 2\}$ . The difference in activation is in the direction of a negative prediction error, and an illustration is provided in Figure 2.8. (See also Supplementary Material for time course graphs of BOLD activation)

Figure 2.8: **Brain activity at the losing position 11 in  $G(15, 3)$ .** Contrast obtained from a GLM with 16 predictors on all 12 subjects. In the GLM we use the same 7 groupings for positions in the game, and differentiate between positions during the first (early) and last (late) 10 trials, for a total of 14 predictors. An additional predictor for computer choices and a constant term describe the full model. The contrast used in the figure shows activation when the current position is 11 during both early and late trials compared to activation at positions  $\{1, 2\}$  during early and late trials. The map shows activation at a false discovery rate  $q < 0.05$ .

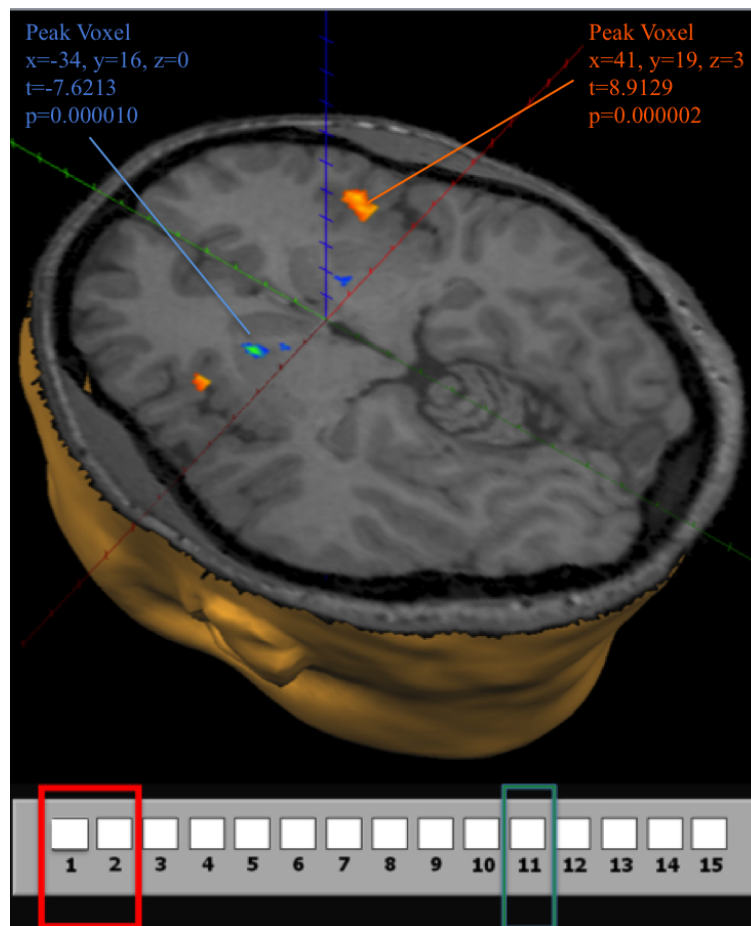


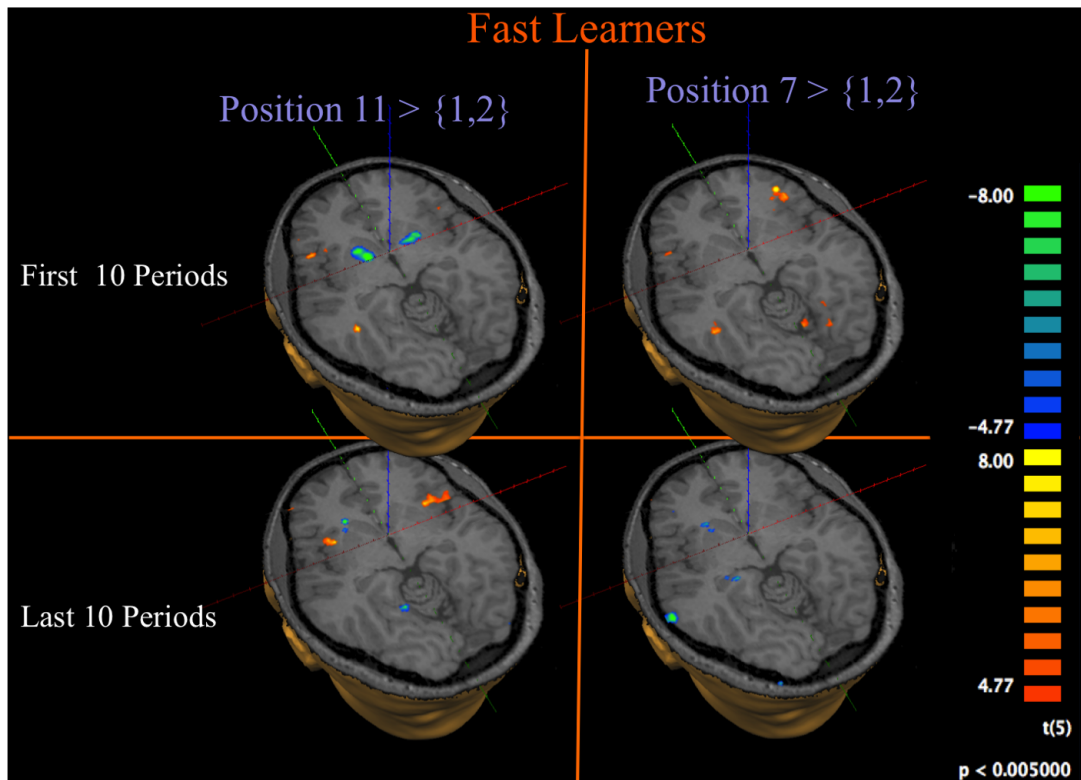


Figure 2.8 also shows significant positive activation of the left and right Insula at coordinates (41, 19, 3), as subjects perceive the near inevitability of losing the game at position 11. This activation is consistent with the Insula's involvement in processing negative affect, and its role in signaling negative prediction errors (Seymour, 2004)

#### 2.4.4 Prediction error response for Fast Learners

Given our main interest in the neural signature of the sequential, recursive way in which subjects learn the solution to the Hit-N game, we concentrate in Figure 2.9 on the Fast Learners; those subjects who actually manage to quickly reduce the amount of errors they make in the game.

Figure 2.9: **Progression of activation at losing position 11 for Fast learners.** GLM and contrasts as for figure 2.8, but limited to Fast Learners.



The left panel of Figure 2.9 contrasts activation at position 11 to activation at

position  $\{1, 2\}$  for Fast Learners. Consistent with the role of the Striatum in signalling prediction errors, we find that subjects show a strong initial negative response in the Striatum at losing position 11 during the first 10 rounds, which diminishes or disappears during the last 10 rounds. Panel C in Figure 2.10 shows that this change of Striatal activity for Fast Learners is statistically significant at an uncorrected threshold of  $p \leq 0.005$ , ( $t(5) \geq 4.77$ ). At the same threshold, we observe significant activity in the Insula during both time periods.

Our analysis also shows strong activity in the Insula at position 7 compared to  $\{1, 2\}$  during early trials, and eventually activity in the Striatum at position 7 during late trials; indicating a shift of the prediction error from position 11 to position 7; the sequence - as we have already shown - in which subjects learn the losing positions.

Figure 2.10: **Progression of activation in Fast and Slow learners.** **Panel A:** Contrast obtained from a GLM with 16 predictors on 6 subjects classified as Slow Learners. The depicted contrast shows activation at position 11 compared to activation at position 1,2 during late trials.  $p < 0.005$  uncorrected,  $t > 4.77$ . **Panel B:** Same model as panel A. The depicted image subtracts the contrast obtained for positions 11 vs {1,2} in late periods from the contrast obtained for those positions during early periods. Positive identification of Striatum in this contrast, is driven by a more strongly negative activation at position 11 in late periods for Slow learning subject. **Panel C:** 12 predictor GLM for 6 subjects classified as Fast Learners. As in panel B, we show the subtraction of the contrast (11early-1,2early) - (11late-1,2). We find activation in Medial Prefrontal Gyrus (MPFG), VLPFC and Striatum. Negative identification in Striatum is driven by a more strongly negative response at position 11 during early trials for Fast Learners.  $p < 0.005$  uncorrected for all images depicted here.

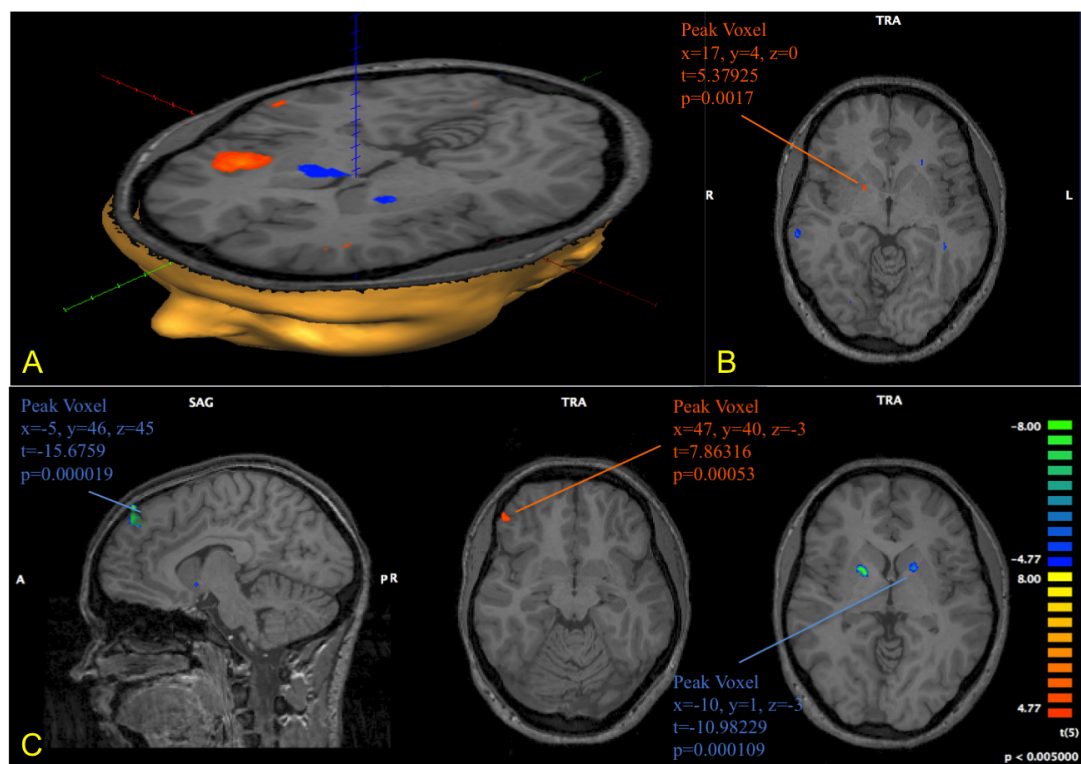


Figure 2.10 provides support to the above observations by overlaying the contrasts of early and late activity at position 11 (compared to  $\{1, 2\}$ ) for both Fast and Slow learners. Slow learners exhibit detectable striatal activation in direction of a prediction error only during the last 10 trials; consistent with the observation that these subjects learn the game according to the same general pattern, but at a slower pace, than subjects classified as Fast Learners. However, the direct test of the effects of Early/Late periods, Fast/Slow learners, and the interaction term of these classifications, shown in table 2.4, did not identify a statistically significant effect for the interaction ( $p=0.158$ ).

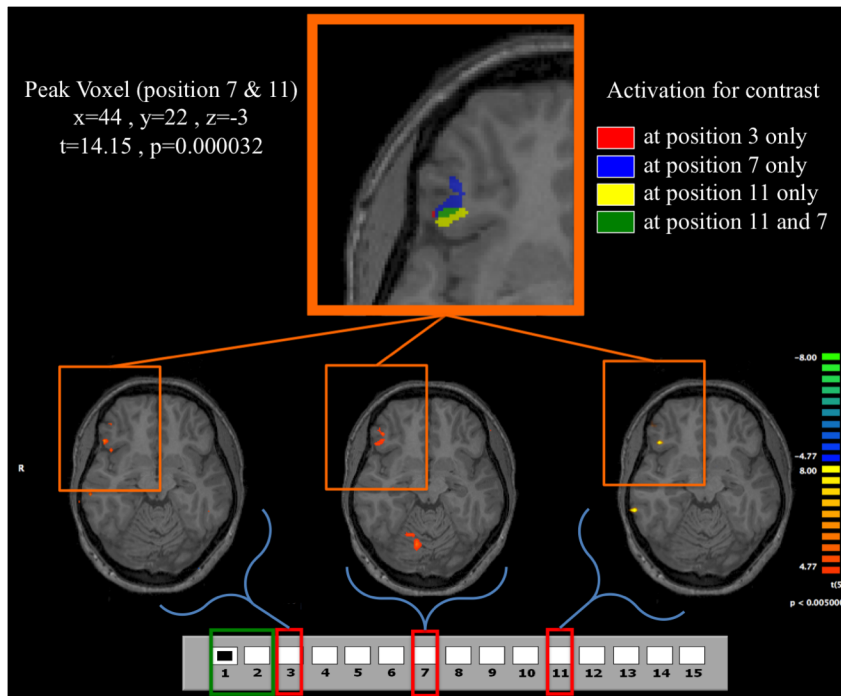
Table 2.4: **Interaction between Fast/Slow learner, and Early/Late trial on BOLD signal contrast position 11 - position  $\{1, 2\}$  in Striatum.**

	Perc. BOLD b/se
Dummy for Fast Learners	-0.09366 (0.08354)
Dummy for First 10 Periods	0.15197 (0.32534)
Interaction Term	0.17345 (0.11814)
Constant	1.02843*** (0.23005)
r <sup>2</sup>	0.2144
N	24

#### 2.4.5 Experience-based learning and abstract reasoning

The center image of panel C in Figure 2.10 identifies a cluster of voxels in the ventro-lateral prefrontal cortex (VLPFC) (47, 40, 3) with strong positive activation during the early losing position for Fast Learners. In what follows, we investigate the activity in this area when subjects are in losing positions for the game  $G(15, 3)$ .

Figure 2.11: **Fast learners at the losing positions.** Contrast for Fast Learners during first 10 periods at  $p < 0.005$ . The contrasts used are for each of the three losing position, compared to the predictor given by the game being in position  $\{1, 2\}$ . For example the contrast for the position 3 indicates the comparison between position 3 and position  $\{1, 2\}$ . The **top panel** shows the clusters in VLPFC activated for the different contrasts. The **lower panel** shows the activation for the contrasts  $(3, \{1, 2\})$ ,  $(7, \{1, 2\})$  and  $(11, \{1, 2\})$ . Activation in VLPFC is not found in Fast Learners during the last 10 rounds, and Slow Learners show it only during the last 10 rounds for  $11-\{1, 2\}$ . Note, from Figure 2.5 that the Fast Learners do not make mistakes past round 10, while Slow Learners commit mistakes even at position  $e_{11}$  past round 10. It should be noted here that a direct test of the interaction between subjects' categorization as Fast/Slow Learner and a dummy variable indicating Early/Late trials did not yield a statistically significant effect ( $p=0.158$  two-sided, see table 4). We believe that the failure to identify such an effect at conventional significant level in our data may be due to small sample size, and an insufficiently precise measure of when subjects learn the game.



Our analysis of losing positions, illustrated in Figure 2.11, shows statistically significant increases of activation in the VLPFC at all of the losing positions during  $G(15, 3)$ . Given this region's association with tasks requiring spatial imagery in deductive reasoning (see (Knauff, 2002) or (Crescentini, 2011)) the observation of higher activity during losing positions is of particular interest, as it indicates the special contribution that the experience of a losing position seems to make towards subject's progress in learning the game.

Figure 2.11 shows overlapping regions of activation for all losing positions experienced by Fast Learners that is most pronounced at position 11, and least pronounced at position 3; once again highlighting the critical nature of the initial losing position 11 for subject's learning experience with the game.

## 2.5 Conclusions

We have explored how subjects learn to play the Hit-N game, and how this process converges for all subjects to learning the optimal strategy with the method of backward induction. We found strong evidence for a sequential learning process in which subjects learn the losing positions at the game's end first. We showed that the behavioral characteristics (in error rate and response time) of this sequential learning process are consistent with a basic search model in which subjects choose an optimal search effort conditional on their ability and associated search costs.

We have also shown a neural pattern of activation in the brain's reward system, including the Insula and Striatum, that mirrors the behaviorally implied pattern of subjects learning to identify losing positions from the game's end. In particular, we find that the rate at which subjects learn to identify losing positions is also reflected by a differential onset of prediction error response between Fast and Slow Learners. A critical finding of our study is the implication of the prefrontal cortex in subject's progression towards finding the solution to the Hit-N game. Here we find that activity in VLPFC is higher at losing positions than at corresponding winning positions. Taken together, these findings point towards a cognitive process in which the affective experience of a losing position feeds critically into the subject's abstract cognitive engagement with the task.

While most of our discussion concentrated on subject’s success in recursively learning to identify losing positions in the Hit-N game, it is clear that such a process - although enabling subjects to master any length Hit-N game - is not equivalent to an abstract, explicit understanding of the BI solution to the game; one which could be transferred instantaneously to other similar games, such as  $G(17, 4)$ . We see then, in both of our studies, that most subjects, despite quickly becoming highly proficient in  $G(15, 3)$ , fail to instantaneously achieve proficiency in  $G(17, 4)$ . Instead, subjects require an abbreviated learning period also for the second game. What seems remarkable about the transition of behavior from  $G(15, 3)$  to  $G(17, 4)$  is that subjects, even without ostensibly having explicit knowledge of the BI solution at the time they begin  $G(17, 4)$ , nonetheless commit fewer errors, and require a shorter learning phase for the theoretically more difficult second game. This observation provides strong indication that the recursive learning algorithm that enables learning of  $G(15, 3)$  is also a contributor to the development of a precursory understanding of the game’s abstract solution. One implication of this finding is that complex cognitive insights, such as understanding that backward inductive reasoning provides a solution to the general Hit-N game, can arise from the interaction of experience-based reward system responses and abstract reasoning within a relatively simple model. The fact that an experience-based understanding derived from playing  $G(15, 3)$  is effective in improving subject’s performance in  $G(17, 4)$  suggests that at least some higher-order cognition and insights might be motivated and prepared by joint activity in the brain’s reward system and prefrontal cortex.

## 2.6 Method and Materials

### 2.6.1 MRI data acquisition

High resolution anatomical images were acquired first, using a Siemens *t1*-weighted 3d flash 1mm sequence. Then, functional images were acquired using echo planar imaging with Repetition Time (TR) 2000ms, Echo Time (TE) 23ms, flip angle 90 degrees,  $64 \times 64$  matrix, 38 slices per scan, axial slices 3mm thick with no gap. The voxel size was  $3 \times 3 \times 3$  mm.

The data were then preprocessed and analyzed using Brain Voyager QX 2.1. The anatomical images were transformed into Talaraich space in 2 steps: first the cerebrum

was rotated into anterior commissure - posterior commissure (AC-PC) plane using trilinear transformation, second we identified 8 reference points (AC, PC, and 6 boundary points) to fit the cerebrum into the Talaraich template using trilinear transformation. We preprocessed functional data by performing slice scan time correction, 3D movement correction relative to the first volume using trilinear estimation and interpolation, removal of linear trend together with low frequency non-linear trends using a high-pass filter. Next, we co-registered functional with anatomical data to obtain Talaraich referenced voxel time courses, to which we applied spatial smoothing using a Gaussian filter of 7 mm.

### 2.6.2 GLM models

fMRI analysis was performed in BrainVoyagerQX version 2.1. Contrasts obtained for  $G(15, 3)$  are based on the results of an event-related general linear model with random effects using 16 predictors. 7 predictors signify the period in which a subject contemplates any of the positions  $\{1, 2\}$ ,  $\{3\}$ ,  $\{4, 5, 6\}$ ,  $\{7\}$ ,  $\{8, 9, 10\}$ ,  $\{11\}$ ,  $\{12, 13, 14\}$  during the first 10 trials of  $G(15, 3)$ . Another 7 predictors signify the same position during the last 10 trials. An additional predictor for times in which the computer is moving and an intercept term describe the model. Contrasts obtained for  $G(17, 4)$  are based on the results of an event-related general linear model with random effects using 16 predictors. 7 predictors signify the period in which a subject contemplates any of the positions  $\{1\}$ ,  $\{2\}$ ,  $\{3, 4, 5, 6\}$ ,  $\{7\}$ ,  $\{9, 10, 11\}$ ,  $\{12\}$ ,  $\{13, 14, 15, 16\}$  during the first 10 trials of  $G(17, 4)$ . Another 7 predictors signify the same position during the last 10 trials. An additional predictor for times in which the computer is moving and an intercept term describe the model.

### 2.6.3 Fast and Slow Learners

The fMRI study consists of 12 subjects. For analysis comparing Fast and Slow Learners in  $G(15, 3)$ , subjects were split into groups according to their overall error rate (a subject is slow if the error rate is larger than 40 per cent), which also constitutes a splitting according to Wins in  $G(15, 3)$  (a subject is slow if the number of wins in that game is less than five). Both are median values, but they are also values at which there is a



large change of performance.

# References

- [1] Aymard, S and Serra, D, (2001) "Do individuals use backward induction in dynamic optimization problems? An experimental investigation," *Economic Letters*, 73, 287-292.
- [2] Bouton, C, (1901-1902) "Nim, a Game with a Complete Mathematical Theory," *The Annals of Mathematics*, 3, 1-4, 35-39
- [3] Crescentini, C and Seyed-Allaei, S and De Pisapia, N and Jovicich, J and Amati, D and Shallice, T, (2011), "Mechanisms of Rule Acquisition and Rule Following in Inductive Reasoning", *Journal of Neuroscience*, 31, 21, 7763-7774.
- [4] Daw, ND and Gershman, SJ, Seymour, B, Dayan, P and Dolan, RJ, (2010), "Model-Based Influences on Humans' Choices and Striatal Prediction Errors", *Neuron*, 69, 6, 1204-1215.
- [5] Dufwenberg, M, Sundaram, R and Butler, D, (2009) "Epiphany in the Game of 21", manuscript.
- [6] Fey M, McKelvey R, and Palfrey T, (1996) "An experimental study of Constant-Sum Centipede Games," *International Journal of Game Theory*, 25, 269-287.
- [7] Fischbacher, U (2007): "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 10, 2, 171-178.
- [8] Fudenberg, D and Tirole, J, (1991) *Game Theory*, The MIT Press, Cambridge, Massachusetts.

- [9] Gneezy, U Rustichini, A and Vostroknutov, A (2010) "Experience and insight in the race game, *Journal of Economic Behavior and Organization* 75, 2, 144-155.
- [10] Hauser, D., Chomsky, N and Fitch, WT (2002) "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?", *Science* 22, 298, 5598, 1569-1579.
- [11] Johnson, EJ and Camerer, C and Sen, S (2002) "Detecting failures of backward induction: Monitoring information search in sequential bargaining", *Journal of Economic Theory*, 104, 1, May 2002, 16-47.
- [12] Knauff, M and Mulack, T and Kassubek, J and Salih, HR (2002), "Spatial imagery in deductive reasoning: a functional MRI study", *Cognitive Brain*
- [13] Preuschoff, K and Quartz, S R and Bossaerts, P, (2008), "Human Insula Activation Reflects Risk Prediction Errors As Well As Risk" , *Journal of Neuroscience*, 28, 11, 2745-2752.
- [14] Schultz, W, Dayan, P, Montague PR,(1997), "A Neural Substrate of Prediction and Reward", *Science* 275, 1593 - 1598.
- [15] Selten, R (1965) "Spieltheoretische behandlung eines oligopolmodells mit nachfragertragheit," *Zeitschrift fur die Gesamte Staatswissenschaft*, 12, 301-324.
- [16] Selten, R and Stoecker, R (1986) "End behavior in sequences of finite prisoner's dilemma supergames," *Journal of Economic Behavior and Organization*, 7, 47-70.
- [17] Seymour, B, O'Doherty, JP, Dayan, P, Koltzenburg, M, Jones, AK, Dolan, RJ, Friston, KJ, and Frackowiack, RS, (2004) "Temporal difference models describe higher-order learning in humans" *Nature* , 429, 664-667.
- [18] Zermelo, E (1912) "Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels," *Proceedings of the Fifth International Congress of Mathematicians*, Vol. II, Cambridge, 501.

## Chapter 3

# Intelligence Moderates Neural Responses to Monetary Gains and Losses (Hawes, DeYoung, Gray & Rustichini, *Under Review* )

The nature of the relation between cognitive skills and economic preferences is a key question in psychology and economics. Examining subject's intelligence (IQ) and responses to probabilistic feedback during a simple decision task allows us to investigate neural correlates of cognitive ability at the foundational level of gain/loss processing. Our neuroimaging results for 94 subjects show that typical declines in striatal BOLD signal after monetary punishment are significantly less pronounced for subjects with higher IQ. This finding strongly implicates IQ in the ex-post processing of decision outcomes, thereby opening up the hitherto unconsidered possibility that cognitive mechanisms underlying outcome evaluation may be central to how intelligence influences preferences and decision, especially attitudes towards risk. We further investigate the role of IQ for outcome evaluation on subject's behavior on our decision task to demonstrate a correlation between IQ and the extent to which past decision outcomes influence future

choices. Specifically we find that larger IQ predicts behavior to be more strongly correlated with an extended period of previously experienced decision outcomes, while lower IQ predicts behavior to be correlated exclusively to the most recent decision outcomes. In addition to showing the moderating effect of intelligence on neural responses to gains and losses, our findings illustrate the existence of a link between intelligence and choice behavior that extends beyond the ex-ante comparison of decision options to include ex-post evaluation of decision outcomes. Importantly, this identified link suggests that observed correlations between intelligence and preferences may be rooted in a unitary process of how decision outcomes are experienced.

### 3.1 Introduction

The ability to maintain and manipulate mental models of the environment - including abstract representations of the relationships between actions and outcomes, and the contingencies on which they depend - is critical to most instances of learning, planning, and goal directed decision-making (Johnson-Laird, 2010; Li, Delgado, & Phelps, 2011; Solway & Botvinick, 2012). During experience-based decision making, humans rely on mental models not only during the process of making a choice, but also while evaluating feedback and interpreting newly observed relations between actions and outcomes following that choice (Hampton, Bossaerts, & O'Doherty, 2006; Rangel, Camerer, & Montague, 2008; Glimcher & Rustichini, 2004; Liljeholm & O'Doherty, 2012; Daw, Gershman, Seymour, Dayan, & Dolan, 2011). Stable systematic differences in the mental models that decision makers entertain may therefore have a two-fold influence on decision making. First, a direct influence on which actions and action-outcome contingencies are considered and compared during a choice task; and second, a more indirect influence on the evaluation of histories of observed rewards following decisions: Identical histories of feedback will be experienced differently by evaluating decision makers if they evaluate it using different mental models. This difference in experience is likely, eventually, to affect behavior, especially during trial-and-error decision making. In addition, domain general stable influences of mental models on feedback evaluation and the experience of decision outcomes are likely to eventually manifest in stable differences in preferences.

Furthermore, cognitive defaults applied to such domains (e.g., attitudes toward uncertainty and temporal delay) may be influenced by how persons with different cognitive models of these outcomes experience the resolution of uncertainty, or the delivery of a delayed reward.

One likely source of systematic individual differences in mental models is *intelligence*, which has been described as the cognitive ability to manage complexity (Gottfredson, 1997). Links between model complexity and intelligence could help to explain why standard measures of IQ predict systematic differences in decision-making, including variation in preferences for risk and temporally delayed rewards (Dohmen, Falk, Huffman, & Sunde, 2010; Shamosh & Gray, 2008; Shamosh et al., 2008; Rustichini, 2009). In particular, higher intelligence has been associated with reduced aversion to beneficial risk taking, as well as to reduced discounting of delayed rewards (Dohmen et al., 2010; Rustichini, 2009; Burks, J, Goette, & Rustichini, 2009; Benjamin, Brown, & Shapiro, 2012). Because greater intelligence is likely to be associated with the ability to maintain mental models of increasing complexity, one possible explanation for the correlation of intelligence with the two theoretically separate domains of decision making (risk and temporal delay) may be rooted in systematic differences in how subjects with different levels of intelligence experience probabilistic rewards. Thus, an important necessary first step towards fully identifying the role of intelligence for preferences and decision making is to identify the extent to which intelligence relates to *ex-post* evaluation of decision outcomes. At the foundational level this entails investigating the relation between intelligence and randomly obtained monetary gains and losses.

In this research we establish a link between IQ and monetary gain/loss processing by evaluating subjects' behavior and neural responses during a simple decision task, very similar to a previously examined paradigm used by Delgado et al (Delgado, Nystrom & Fissel; 2000). In our decision task, participants guessed whether a computer-generated number would be high or low, and received monetary gains and punishments depending on the correctness of these guesses. By experimentally manipulating subjects' performance on this task to be pseudorandom and fixed, our task design eliminated the opportunity for subjects to experience different performance histories, and thus minimized any potential concern that between-subject variance in choice behavior was driven by differences in the history of obtained rewards. Instead, remaining individual differences

in responses to gains and losses during our task are restricted to intrinsic differences in preferences for reward/punishment, or to individual differences in how the experience of reward/punishment is interpreted in relation to the decision-task.

## 3.2 Method

### 3.2.1 Participants

We collected data from 100, male, right handed-subjects. Subjects were administered the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), which provides an estimate of full-scale IQ using 4 sub-tests (Vocabulary, Similarities, Block Design, and Matrix Reasoning). Average IQ was 122.9 (min: 95.5, max: 148.0, SD: 11.6). Subjects further completed a battery of questionnaires and cognitive tasks that included an n-back working memory task. During the working memory task, subjects viewed a series of words and indicated whether each word matched the one appearing three previously. Correct and incorrect responses were then combined into an indicator of working memory. Median age for our subjects was 22 (min: 18, max: 38). Data for 6 subjects were discarded because of excessive head-motion in the scanner, leaving 94 participants in our analysis.

### 3.2.2 fMRI Procedures

Subjects performed the experimental tasks reported here, and three additional unrelated tasks for a scanning time of 1.25 hours. Imaging data were collected using a 3-Tesla Siemens Trio scanner at the Yale Magnetic Resonance Research Center. For each participant, a high-resolution T 1-weighted anatomical image (MPRAGE, time repetition[TR] = 2500 ms; time echo [TE] = 3.34 ms; inversion time = 1100 ms; flip angle =  $7^\circ$ ; slices = 256, voxel size =  $1 \times 1 \times 1$ mm) and 180 contiguous functional volumes (gradient-echo EPI sequence; TR = 2000ms; TE = 25ms; field of view [FOV] = 240cm; flip angle =  $80^\circ$ ; voxel size =  $3.75 \times 3.75 \times 4$  mm ) were acquired. Participants viewed stimuli projected onto a screen through a mirror mounted on the head coil. Responses were made using fiber-optic response buttons, using the fingers of the right hand. Stimuli were presented in PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993).

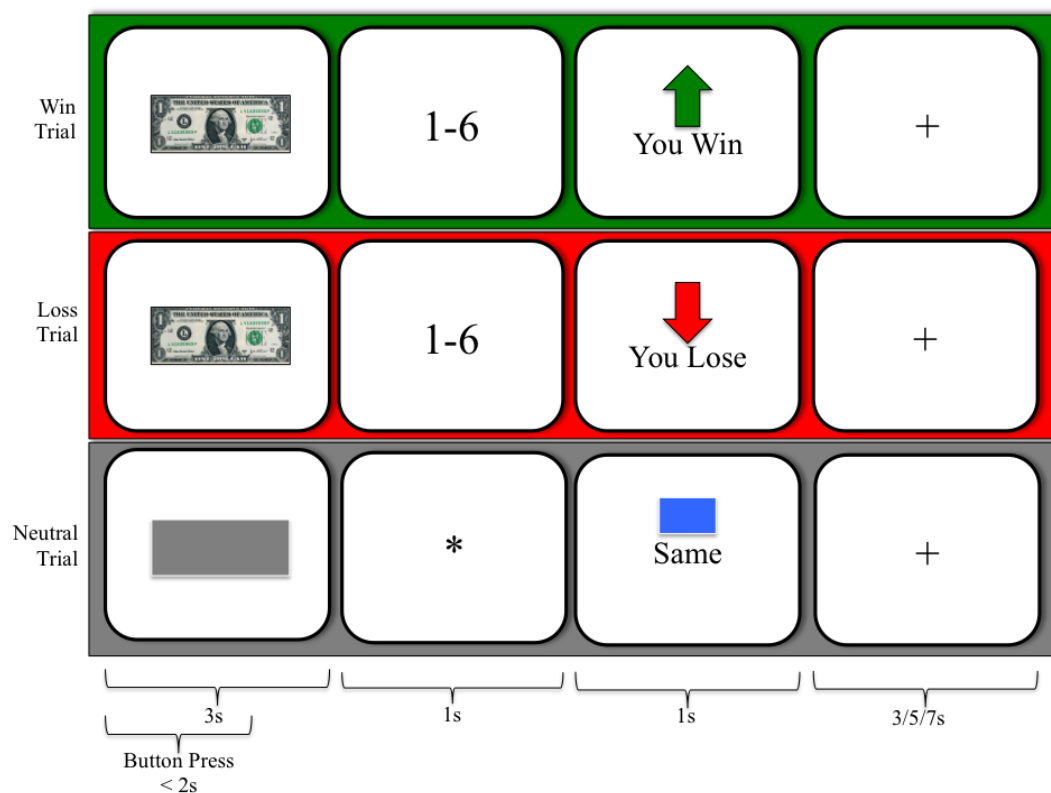
### 3.2.3 Stimuli and Design

Subjects were instructed to guess whether an upcoming computer generated number would be either Low (in the set  $\{1,2,3\}$ ) or High ( $\{4,5,6\}$ ). Subjects were rewarded with \$ 1 (USD) for each correct guess, and received a punishment of \$ -1 for each incorrect guess. The guessing game was interspersed with 20 reward-neutral control trials, during which subjects were also instructed to press a button, but for which they received neither feedback, nor monetary rewards/punishments. During the guessing game, subjects saw a sequence of stimuli as shown in figure 1. During reward-relevant trials, subjects first saw a 1 dollar bill displayed on the screen for 3 seconds. During this time subjects indicated their guess regarding the upcoming number. After these 3 seconds, subjects first saw a computer generated number for 1 second, and then, depending on trial type, a green upward arrow, or a red downward arrow containing the words you win or you lose for another 1 second. Control trials started with a 3 second display of a gray rectangle of the same size as the dollar bill, followed by an asterisk for 1 second, and a blue rectangle containing the word same for another 1 second.

Unbeknownst to subjects, the computer's number generating process was fixed so as to confront each subject with the exact same pseudo-randomized sequence of outcomes, irrespective of the subjects choice behavior (i.e., a high or low number was selected on each trial based on the fixed sequence and the subject's guess). Hence, each subject saw the same sequence of gain and loss trials (20 gains and 20 losses in total) in the same order, interspersed with 20 reward-neutral control trials during which no outcome was revealed. All trials were separated by fixation periods of 3, 5 or 7 seconds duration (jittered). The task design is illustrated in Figure 3.1.



Figure 3.1: **Task design:** Subjects engaged in 40 reward relevant trials. At the beginning of these trials a US dollar bill was displayed for 3 seconds, during which subjects pressed one of two buttons indicating their guess of whether a computer generated number would be Low (1-3) or High (4-6). Guesses were followed by two feedback screens, each 1 second in duration. Reward relevant trials were interspersed with 20 reward-neutral trials, which were signaled by a gray rectangle. In this figure (but not in the task itself), a green outline marks the sequence of screens seen during gain trials. A red outline marks the sequence of screens for loss trials, and grey marks reward neutral control trials. Correct/incorrect outcomes were rewarded/punished with \$ US+/-1.



### 3.2.4 fMRI Data Analysis

All data were preprocessed and analyzed using BrainVoyager QX 2.1 (Brain Innovation, Maastricht, The Netherlands). T1-weighted anatomical images were transformed into Talairach space (Talairach & Tournoux, 1988) by manually identifying the anterior and posterior commissure and then applying BrainVoyager default settings for spatial normalization. Functional data were preprocessed by applying slice time correction (using cubic spline interpolation), alignment of slices (using cubic spline interpolation to the first non-discarded scan time within a scan run), motion correction (using trilinear interpolation), spatial smoothing (using a 4 mm Gaussian kernel), linear trend removal, and temporal high-pass filtering (using BrainVoyager default settings). Subjects were eliminated if motion correction indicated deviations in the estimated center of mass greater than 3 mm, leading to the elimination of 6 subjects. Functional data was then co-registered to T1 anatomical images to create 4d data (anatomical 3D + time course) for GLM analysis, using BrainVoyagers default settings.

For statistical analysis, we computed a general linear model on the percentage BOLD signal normalized time course of 94 subjects. Our model contained one predictor for the onset of payoff-relevant trials ( $R^{Relevant}$ ), one predictor for the onset of payoff-neutral trials ( $R^{Neutral}$ ), and one predictor each for the feedback period of gain ( $FB^+$ ), loss ( $FB^-$ ) and control ( $FB^0$ ) trials. Each of these five predictors was specified as a zero-one variable (box-car), 2 seconds in duration when equal to one, which was further convolved with a double gamma function estimate of the hemodynamic response, using BrainVoyager default settings. The 2s predictors indicating feedback covered the time period during which subjects saw the computer generated number as well as the arrows/rectangle. Because the maximum time taken by subjects to enter their guesses was just less than 2 seconds, we set the predictor indicating onset of payoff relevant and payoff neutral trials to cover exactly 2 seconds from trial onset, thus increasing comparability of our predictors by having them all be of equal length. The GLM for our analysis was therefore:

$$BOLD = a + b \times R^{Relevant} + c \times R^{Neutral} + d \times FB^+ + e \times FB^- + f \times FB^0 + error$$

Regions of interest were identified according to t-tests performed on the whole brain contrast  $FB^+ - FB^-$ . A Bonferroni corrected p-value threshold of .01 was used as

a cut-off for identifying significant clusters. Table 3.1 lists all clusters with significant activation for this contrast.

We extracted regression weights for the above GLM separately for each subject to correlate with IQ scores. Subject-specific regression weights were obtained by performing the above GLM on voxels falling into an anatomical template of the caudate nucleus. Coordinates for the template are listed in Table 3.1. Importantly, we performed individual differences analysis on regressors obtained from performing our GLM on this a priori anatomically defined region of interest (ROI), rather than from the functionally identified region showing the strongest contrast between gain and loss feedback. This choice was made because regions that are most relevant for individual differences are often not the same as those showing the strongest main contrast, precisely because individual differences reduce the strength of the main effect. However, the anatomical region we used falls entirely into the functionally identified regions listed in Table 3.1. Furthermore, all analysis reported here produce results that are statistically significant at  $p < .05$  also when functionally identified ROIs are used.

We performed exploratory analysis on 3 selected functionally identified ROIs which showed significant activation for gains compared to losses. These regions comprised the medial prefrontal cortex (0, 44, 1) and posterior cingulate cortex- two regions shown to correlate with the subjective utility of monetary outcomes of risky choices (Wu, Delgado, & Maloney, 2011), and also the right inferior/middle frontal gyrus (-42, 44, 9), which has been associated with neural representations of loss aversion (Tom, Fox, Trepel & Poldrack, 2007).

Total brain size was calculated in Freesurfer using the `asegstats2table` command in its default settings (Fischl & Dale, 2000).

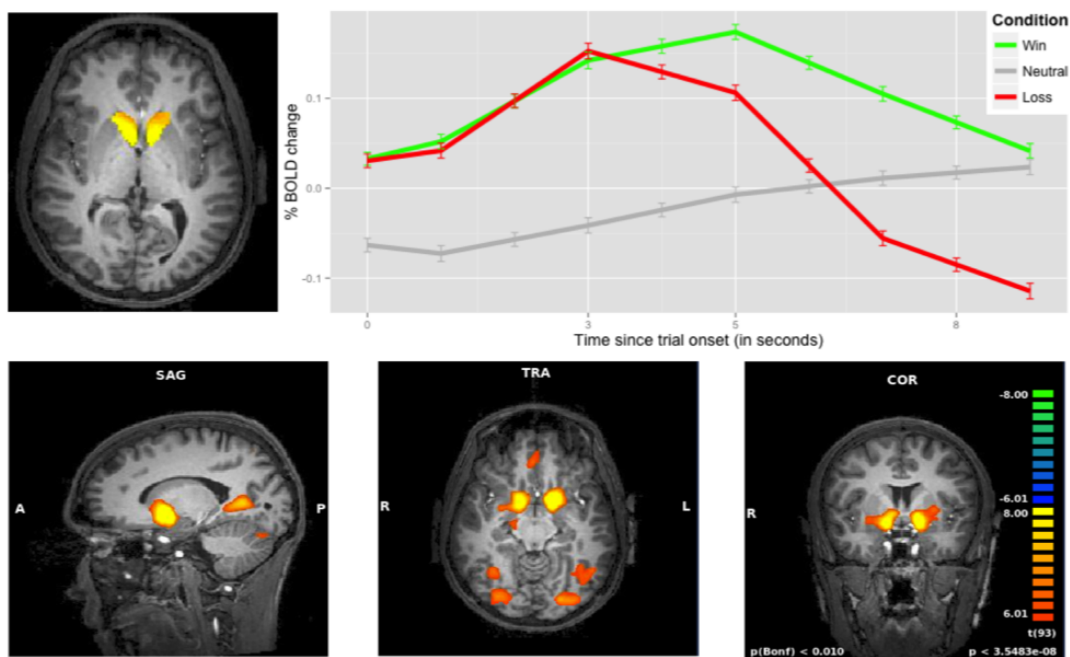
Regressions in Tables 3.2, 3.3, 3.4, and 3.5 were performed using the statistical programming software R, and the general linear modeling package `lme4` (Bates, Maechler & Bolker, 2012). Regressions of behavioral data were computed using mixed effects linear logistic regression; coefficients were conditioned at the subject level where appropriate (Tables 1 and 2). The regression reported in Table 3.2 was performed using robust regression to control for potential influence of outliers. Robust regression, ordinary least square regression, and removal of potential outliers all produced results with a statistically significant main effect of IQ as reported in the results section.

## 3.3 Results

### 3.3.1 Brain Imaging Results

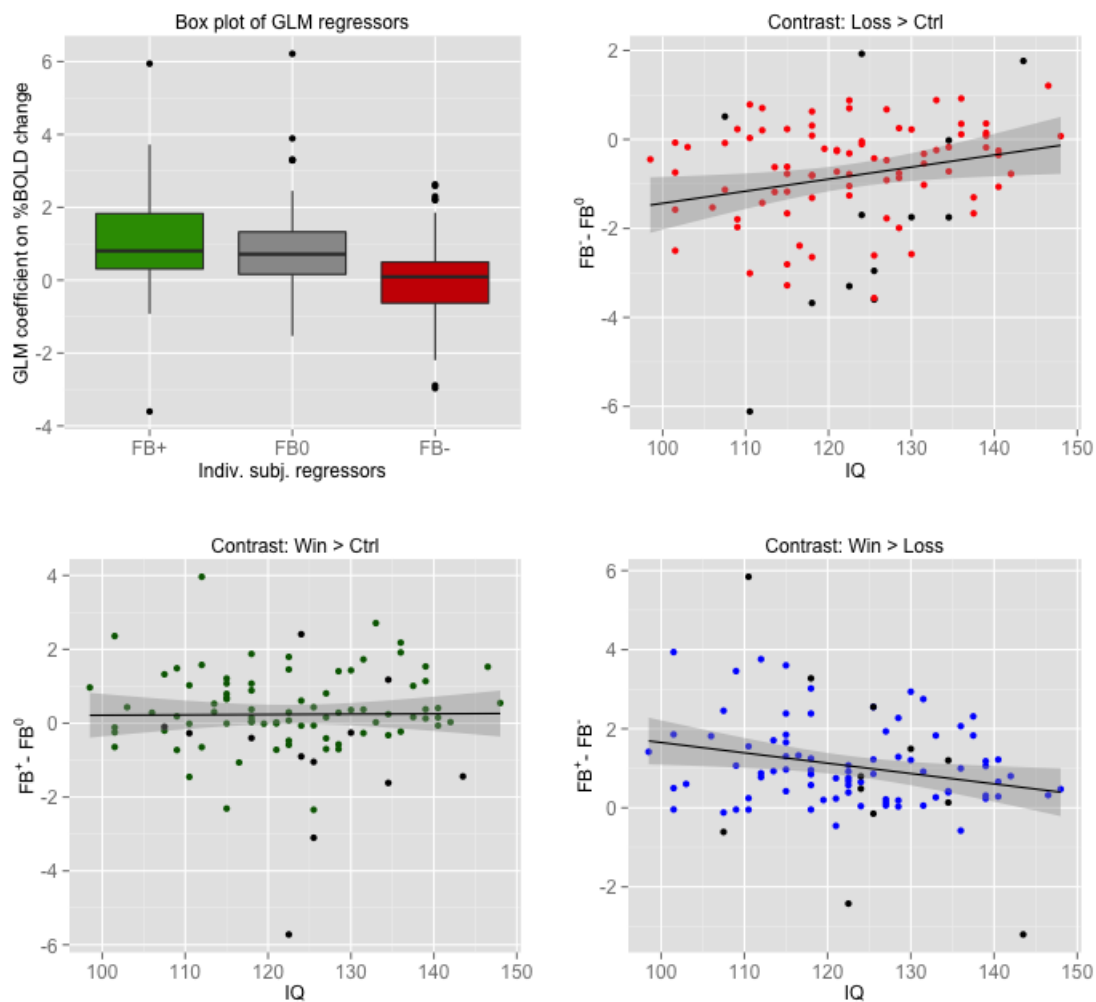
Averaged across all subjects, ignoring effects of IQ, neuroimaging results for our large sample replicate and extend findings reported by Delgado et al (2000) on a sample of 9 subjects: We find anticipatory increases in BOLD response in the caudate at the onset of reward relevant trials. This BOLD response remains elevated after revelation of a gain, or decreases steeply below baseline in the case of losses. Figure 3.2 shows this pattern of BOLD response for the masked caudate area (barycenter at (12, 9, 0), Talairach).

Figure 3.2: **BOLD response  $FB^+ > FB^-$  ( $Gain > Loss$ ) in the bilateral caudate.** BOLD response in caudate after gains is significantly higher than after losses. This results holds for whole-brain analysis as well as a masked regression on the anatomically defined caudate (top left). Event related average BOLD response in caudate replicates the time course identified by Delgado et al. (2000) (top right). Additional areas showing significantly higher activation after gains compared to losses in whole brain analysis include the medial prefrontal cortex and posterior cingulate cortex (bottom panels).



We obtained predictors of percentage BOLD change following rewarding ( $FB^+$ ) and punishing ( $FB^-$ ) feedback by applying the event-related general linear model described in the methods section. The brain regions showing significantly more activation after gains than after losses are shown in Figure 3.2 and listed in Table 3.1. No brain areas showed significantly larger activation after losses compared to gains. We focused our investigation of the role of intelligence during feedback processing on the caudate (12, 9, 0) and the medial prefrontal cortex (0, 44, 1), because of their joint implication for reward/punishment processing, reinforcement learning and decision making in previous studies with similar task design (e.g. Delgado, Nystrom & Fissel, 2000., Li & Daw, 2011, Van denBos, Cohen, Kahnt & Crone, 2012). Use of a priori defined regions of interest (ROI) for the investigation of individual differences is preferable to investigation of ROIs showing the greatest gain/loss contrast, because important influences of individual differences may often imply small main (i.e. group average) effects. Alternatively, identifying ROIs based on the identified presence of individual difference effects is undesirable for well-documented reasons (Vul, Harris, Winkielman & Pashlert, 2009). We extracted, individually for each subject, the regression coefficients for BOLD responses after gains, losses, and control sequences for our regions of interest. The panels of Figure 3.3 illustrate correlations of the gain and loss responses with intelligence in the Caudate.

Figure 3.3: **Individual subject GLM regressors and intelligence.** The box plot shows distribution of regression coefficients for Gain, Loss and Control trials extracted from the anatomically defined caudate head ROI. Scatter plots and regression lines depict the relation between IQ and contrasted GLM estimates of BOLD response to feedback for the same region. Potential outliers in the boxplot are depicted as black dots in the scatter plots. Statistical results are unaffected by the removal of these potential outliers. The correlation coefficients between IQ and the contrasts depicted in the 2<sup>nd</sup>, 3<sup>rd</sup>, and 4<sup>th</sup> panel of the figure are  $r = .24$  ( $p = .02$ ),  $r = 0.01$  ( $p = .93$ ), and  $r = -.24$ , ( $p = .02$ ) respectively.



Notably, the correlation between increasing intelligence and neural responses to outcomes is significant and positive for losses ( $\beta_{reg} = 0.31$ ,  $p = .028$ ), but essentially flat for gains ( $\beta_{reg} = -0.01$ ,  $p = .952$ ). This finding is more strikingly illustrated by the contrast between gain and loss responses, which decreases significantly with higher intelligence scores ( $\beta_{reg} = -0.31$ ,  $p = .014$ ), indicating smaller decline of BOLD response after losses for higher IQ. Table 3.2 shows the effect of intelligence on the neural responses to losses when controlling for working memory and brain size using robust methods to attenuate possible influences of potential outliers. Using robust regression (as done for the results in Table 3.2), or conservatively removing potential outliers identified in the boxplot of Figure 3.3 does not meaningfully affect the regression in Table 3.2 and retains the statistically significant effect of IQ.

We performed the same analysis on BOLD responses after gains and losses for the medial prefrontal cortex (mPFC) (0, 44, 1), and in a more exploratory fashion- on the posterior cingulate cortex (pCC) (-1,-32,-28) and left inferior/middle frontal gyrus (IMFG) (-42, 44, 9). mPFC and PCC are believed to show BOLD signal activation in proportion to the experienced utility from outcomes of risky choices (Wu, Delgado & Maloney, 2011), and activation in the IMFG has been associated with behavioral loss aversion (Tom, Fox, Trepel & Poldrack, 2007). The conjecture that the relation between subjects risk preferences and intelligence is shaped by the experience of outcomes to risky choices implies a relation between intelligence and responses to gains compared to losses (i.e. the contrast Win - Loss) in brain regions whose activation correlates with subjective utility, hence investigation of these regions was also of importance to our study: We find a correlation between intelligence and the contrast of gain and loss responses in the mPFC:  $r = -.208$ ,  $p = .045$ , and in the IMFG ( $r = -.246$ ,  $p = .017$ ). No correlation was found for the posterior cingulate cortex ( $r = -.111$ ,  $p = .286$ ).

The above findings provide strong evidence for a link between intelligence and expect processing, or experience, of monetary losses. Given the prominent roles of the caudate and the medial prefrontal cortex during reinforcement learning, and our tasks conceptual identity with standard reinforcement learning tasks, our results provide the necessary basis for investigating the hypothesis that outcomes of probabilistic events have differential impact on negative and positive reinforcement signals for subjects who



differ in IQ. Support for this hypothesis would signify an important step towards explaining – at the level of theoretical neuroscience – how individual differences in intelligence relate to long-term differences in attitudes to risk: The implied mechanism of this hypothesis states that a functional role of intelligence during reinforcement systematically leads to divergent experiences of punishment during probabilistic decision making, with higher intelligence correlating with reduced (i.e. less negative) prediction errors after negative outcomes. One probable functional account of how the above described moderating effect of IQ may arise during reinforcement, may be due to an effect of IQ on the complexity of the mental model a subject applies to a feedback learning task. We explore this possibility in relation to the behavioral data of our task.

### 3.3.2 Behavioral Results

Mixed-effects logistic panel regression (Table 3.3), grouped by subject, shows that subjects' guesses in any given trial were influenced primarily by the computer choice observed in the most recent two trials, with subjects more often choosing the option not recently selected by the computer. The outcome (i.e. gain or loss) of the previous trials did not exhibit a significant effect on subjects' choices. Thus, we do not find evidence of a pervasive use of a win-stay/lose-shift heuristic. Response times did not substantially differ with respect to subject IQ ( $r = -.14$ ,  $p = .16$ ). Additionally, the observed overall frequency of High choices by the computer does not influence subjects' behavior after controlling for the computer's choice one and two trials back (Table 3.3, column 3), indicating that subjects modified their choices with respect to past observations of computer behavior in a manner that extends beyond responding to simple frequency. Stated more affirmatively, subjects appeared engaged, consciously or unconsciously, in an attempt to learn and exploit perceived patterns in the reward generating process by modeling recently observed computer choices. Our subsequent analysis shows that this behavioral process is systematically moderated by intelligence.

To aid interpretation of our results we consider a decision and feedback-learning process during which more intelligent subjects construct richer and more complex mental models of the computer's choice generating process. Such a process could be formally expressed within the framework of standard model-based reinforcement-learning (Li et al., 2011; Liljeholm & O'Doherty, 2012; Daw et al., 2011; den Ouden, Friston, Daw,

McIntosh, & Stephan, 2009; Glscher, Daw, Dayan, & O’Doherty, 2010), during which decision-makers up-date their beliefs regarding state-transitions and state-action reward values via prediction error assessment. State-action rewards describe the value of reward that is associated with taking a particular action in a particular state. State transitions describe the probability with which states succeed one another. In our task state-action reward values are deterministically fixed and known to each subject, while state transitions (i.e. whether the computer generated number will be high or low) are unknown and have to be learned. Consequentially, the prediction error a subject experiences in our task depends solely on the subject’s beliefs regarding the process that determines the computer generated number, and how he estimates this process from the feedback in the task.

Given the described behavioral findings, we operationalize the idea that subjects with higher IQ use richer mental reinforcement models as the hypothesis that more intelligent subjects integrate a longer history of past observations into their estimation and belief-updating procedure. To investigate this claim, we first considered the observed trial-by-trial frequencies of observed computer choices for each subject. In particular we considered three frequencies:  $F_t^0$ , the unconditional frequency of observed computer choices at trial  $t$ ,  $F_t^1$  the frequency of observed computer choices up to trial  $t$  conditional on the computers choice in the previous trial, and  $F_t^2$ , the computer choices conditional on the computers choices two trials back<sup>1</sup>. Based on each of these frequencies, we calculated the trial-by-trial expected value of guessing High for each subject and entered these values into a mixed-effects logistic panel regression, grouped by subject, in order to assess the impact of these conditional frequencies on subject choice (Table 3.4). As already demonstrated by the results in Table 3.3, expected values based on the unconditional frequency of computer choices,  $EV^0$ , were unrelated to subject choices. However, expected values based on conditional frequencies relating to one and two trials back,  $EV^1$  and  $EV^2$  respectively, significantly affected subjects guessing behavior. For subjects at the lower range of intelligence for our sample, conditional frequencies one trial back positively predicted guessing behavior, whereas subjects in the higher range of intelligence displayed an increasing effect of the events two trials back (Note that, in our sample, lower IQ individuals have an average level of IQ with respect to the overall

---

<sup>1</sup> Considering no more than two previous trials is justified in light of our analysis described above.

population.).

Because our experiment was not designed to differentiate between competing possible models of how subjects use past information to determine future choices, our analysis remains restricted to showing that intelligence predicts the extent to which information further back in time predicts subjects behavior. To illustrate this clearly without assuming any specific functional form for information integration, Table 3.5 combines information from one period and two periods back into a single regressor  $EV^{1+2}$ , consisting of the un-weighted average of  $F_t^1$  and  $F_t^2$ . In this analysis, controlling for  $F_t^1$  does not eliminate the hypothesized effect that the influence of the composite information on subjects guesses significantly increases with IQ. Hence, we find that subjects with higher IQs were influenced by events one and two periods back, whereas subjects with lower IQs were chiefly responding to events only one period back.

At this point it is important to point out that by considering more information in our task, high IQ subjects are likely to more quickly learn the true structure of the process assigning rewards, despite the fact that rewards are not actually contingent on previous trial outcomes in the experiment. For an easy intuition of why this is the case, one may consider the simple case in which subjects formulate reward expectancies by averaging reward experiences over histories of observations that differ in length. In this environment higher IQ subjects who consider longer histories will make estimates that are on average closer to the true mean reward value (zero in our task) than will subjects with lower IQ, whose trial-by-trial estimates will on average exhibit larger deviations from the true mean (in particular, under the proposition that subjects will choose the action that maximizes expected reward, this deviation will lead to an overestimate of the expected reward for each trial). This arithmetic relation between belief updating based on differentially rich histories of past observations and the expected precision of reward predictions according to such beliefs may provide the basis for the already described differences in neural responses after losses: If loss responses in the caudate are interpreted as prediction errors, then subjects with higher IQ - whose predictions more closely resemble equal probability for each outcome - would experience smaller prediction errors after losses, a consequence of formulating beliefs with regards to richer histories of past observations. However, this interpretation should be viewed as an exploratory account of a possible mechanism underlying the more strongly supported neural evidence

reported in this manuscript, and will require further research for substantiation.

### 3.4 Discussion

By providing neural and behavioral support for a functional role of intelligence during gain and loss processing, the results reported here bring us closer to understanding how intelligence may influence preferences over uncertain outcomes. Our main result demonstrates that IQ moderates BOLD responses to losses. This finding, in combination with our behavioral results seems compatible with the hypothesis that more intelligent subjects evaluate random rewards and punishments using mental models of greater complexity, and that this influence of IQ manifests as differences in the experienced reinforcement-relevant prediction errors when outcomes are revealed. It is clear that the precise mechanics of this account will need further investigation and description within a formal model.

Our findings relate directly to previously identified correlations of decision behavior with increasing intelligence, in particular increased willingness to take risks and reduced temporal discounting. Our results offer the possibility that these joint correlations might be founded in systematic differences in the way rewards and punishments are experienced ex-post by decision makers with different cognitive ability. This effect of intelligence would be in addition to, but distinct from, its role during the ex-ante evaluation of decision options.

Extending the results of our behavioral analysis, our results recommend investigating the functional role of IQ during reinforcement learning, with respect to the hypothesis that higher intelligence correlates with the ability to maintain richer mental models of feedback contingencies, one result of which may be generally reduced negative prediction errors after probabilistic events, and consequently reduced aversion to potential risks. Current research in theoretical neuroscience demonstrates that the predictive power of reinforcement learning algorithms applied to human behavior and neural activity can be improved by considering the effect that experimentally induced increases in cognitive load or limitations in working memory have on a decision-makers consideration of state-transitions and success in learning action-outcome relations (Otto, Gershman, Markman,

& Daw, 2012; Collins & Frank, 2012). Hence, our results add to an emerging literature suggesting the potential of improving reinforcement learning models in theoretical neuroscience by considering the impact of stable individual differences in cognitive ability on feedback learning (van den Bos, Cohen, Kahnt & Crone, 2012; Otto, Gershman, Markman, & Daw, 2012; Collins & Frank, 2012). Our results for medial prefrontal cortex and left inferior frontal gyrus further suggest that investigation into the relation between intelligence and experienced utility may benefit particularly from analyzing the functional integration of utility coding prefrontal areas with subcortical regions during reinforcement learning<sup>2</sup> .

Despite explicit references to reinforcement mechanisms in this manuscript, our main conclusions do not depend on this particular mechanistic interpretation of our results, and instead may remain focused on the positive finding that intelligence moderates responses to monetary losses. This finding, by itself, demonstrates that previously observed decision-correlates of intelligence may stem from differences in how options are perceived and outcomes experienced, rather than following exclusively from differences in the acuity with which options are computationally compared during the process of making decisions. This possibility had not been previously demonstrated.

Finally, in a separate view, our results provide the neurophysiological basis for a potential bridge between theoretical neuroscience and psychological research on intelligence: Although it has been shown that general intelligence is predicted by subjects' efficiency during associative learning (Kaufman, DeYoung, Gray, & Brown, 2009), the mechanism via which associative learning contributes to general intelligence has remained essentially unexplored. Further investigation of a moderating effect of IQ on model complexity may contribute toward providing a functional, neurophysiological, account of how IQ influences associative learning.

---

<sup>2</sup> Notably, a link between developmental changes in reinforcement learning and striatal-medial prefrontal cortex connectivity has been identified in at least one previous study (van den Bos, Cohen, Kahnt, & Crone, 2012). Additionally, neural responses in medial orbitofrontal cortex and dorsomedial striatum have been shown to co-vary as a function of causal contingency (Tanaka, Balleine, & O'Doherty, 2008).

Table 3.1: **Regions of interest differentiating between Gain and Loss feedback:**

ROI ( $FB^+ > FB^-$ )	Peak X	Peak Y	Peak Z	Nr. of Voxels	Avg. t-stat.	Avg. p-value
Medial prefrontal G. (mPFC)	1	49	3	4160	7.98	<.001
Posterior Cingulate C. (pCC)	-1	-32	28	9,15	<.001	
Right Caudate	-1	6	-3	2873	9,91	<.001
Left Caudate	13	3	-5	2740	11.33	<.001
Parahippocampal G.	17	-15	-16	4227	7.15	<.001
Left medial Frontal G.(lmFC)	-42	44	9	1050	6.50	<.001
Anatomically defined Caudate	8	7	-4	5613	6.86	<.001

Table 3.2: **Impact of IQ on BOLD signal after losses.** Note: The contrast between neural loss and win responses correlates significantly with intelligence. Controlling for working memory and age does not affect the independent effect of intelligence. Loss responses are negative (significant negative intercept term) and higher IQ shifts loss responses closer to zero, i.e. prediction errors after losses decrease with increasing IQ. Results from a robust regression (\*p < .05, \*\* p < .01, \*\*\* p < .001)

	Model 1 (b/se)	Model 2 (b/se)	Model 3 (b/se)
Dependent variable: Caudate BOLD $FB^- > FB^0$			
Intercept	-4.26 *** (1.5)	-5.25 *** (2.05)	-7.69 ** (2.3)
IQ	0.03 * (0.01)	0.03 * (0.01)	0.03 ** (0.01)
Head Size ( $mm^3$ )		0.004 (0.14)	0.009 (0.14)
Working Memory (dprime)		0.12 (0.15)	0.21 (0.16)
Adj-R2	.06	.05	.08
p-value	.03	.07	.03

Table 3.3: **Influence of past computer choices on subject choices.** Note: Mixed effects logistic panel regression, grouped by subject.  $N = 3478$ .  $H_{CPU}$  i-back is a dummy variable, representing whether the computer chose High i periods back. Freq Hcpu is the historic frequency of High choices by the computer, expressed as deviance from 0.5. (\* $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ , BIC: Bayesian Information Criterion, LL: Log Likelihood)

	Model 1 (b/se)	Model 2 (b/se)	Model 3 (b/se)
Dependent variable: Subject's choice of $H$			
Intercept	0.091 (0.051)	0.453 (0.084)	0.368 (0.07)
Freq. of $H_{CPU}$	-1.136 (0.333)	-0.091 (0.374)	-0.23 (0.360)
$H_{CPU}$ 1-back		-0.210 (0.073)	-0.273 (0.087)
$H_{CPU}$ 2-back		-0.44 (0.073)	-0.373 (0.086)
$H_{CPU}$ 3-back		-0.110 (0.073)	
$H_{CPU}$ 1-back $\times$ Won		-0.168 (0.10)	
$H_{CPU}$ 1-back $\times$ Lost		-0.12 (0.09)	
BIC	4802	4784	4790
LL	-2389	-2368	-2367

Table 3.4: **Influence of conditional transition probabilities on subject behavior.** Note: Mixed effects logistic panel regression, grouped by subject.  $N = 3478$ .  $H_{CPU}$  i-back is a dummy variable, representing whether the computer chose High i periods back.  $Freq H_{CPU}$  is the historic frequency of High choices by the computer, expressed as deviance from 0.5. (\* $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ , BIC: Bayesian Information Criterion, LL: Log Likelihood)

	Model 1 (b/se)		Model 2 (b/se)	
Dependent variable: Subject's choice of $H$				
Intercept	0.102 (0.05)	.	0.06 (0.05)	
EV $m_0(H)$	-1.039 (2.36)			
EV $m_1(H)$	2.66 (1.22)	*	2.45 (1.06)	*
EV $m_2(H)$	-1.87 (1.00)	*	2.00 (0.93)	*
EV $m_0(H) \times IQ$	0.001 (0.02)			
EV $m_1(H) \times IQ$	-0.02 (0.01)	*	-0.02 (0.01)	*
EV $m_2(H) \times IQ$	0.02 (0.01)	*	0.02 (0.01)	*
BIC	4828		4828	
LL	-2382		-2389	



Table 3.5: **Relation of IQ to conditional probabilities influencing choice behavior.** Note: Mixed effects logistic panel regression, grouped by subject. N = 3478.  $H_{CPU}$  i-back is a dummy variable, representing whether the computer chose High  $i$  periods back. Freq  $H_{CPU}$  is the historic frequency of High choices by the computer, expressed as deviance from 0.5. (\*p < .05, \*\* p < .01, \*\*\* p < .001, BIC: Bayesian Information Criterion, LL: Log Likelihood)

	Model 1 (b/se)	
Dependent variable: Subject's choice of $H$		
Intercept	0.066 (0.05)	
EV $m_1(H)$	4.45 (1.59)	**
EV $m_{1+2}(H)$	-4.00 (1.59)	*
EV $m_1(H) \times IQ$	-0.04 (0.01)	**
EV $m_{1+2}(H) \times IQ$	0.03 (0.015)	*
BIC	4828	
LL	-2389	

# References

- [1] Bates, D., Maechler, M. & Bolker (2012). lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0. <http://CRAN.R-project.org/package=lme4>
- [2] Benjamin, D.J., Brown, S. & Shapiro, J. (2012). Who is behavioral? cognitive ability and anomalous preferences. *Journal of the European Economic Association*, (forthcoming).
- [3] Burks, S., J, C., Goette, L. & Rustichini, A. (2009).. Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Science*, 106(7746).
- [4] Collins, A. G. E. & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? a behavioral, computational, and neurogenetic analysis. *Eur J Neurosci*, 35(7), 1024-35. doi: 10.1111/j.1460-9568.2011.07980.x
- [5] Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P. & Dolan, R. J. (2011, Mar). Model-based influences on humans choices and striatal prediction errors. *Neuron*, 69(6), 1204-15. doi: 10.1016/j.neuron.2011.02.027
- [6] Delgado, M.R., Nystrom, L. & Fissel, C. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology*, 84(6), 3072-3077.

- [7] den Ouden, H. E. M., Friston, K. J., Daw, N. D., McIntosh, A. R. & Stephan, K. E. (2009, May). A dual role for prediction error in associative learning. *Cereb Cortex*, 19(5), 1175-85. doi: 10.1093/cercor/bhn161
- [8] Dohmen, T.J., Falk, A., Huffman, D. & Sunde, U. (2010). A risk aversion and impatience related to cognitive ability? *American Economic Review*, 100, 1238-1260.
- [9] Fischl, B. & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 11050-11055.
- [10] Glimcher, J., Daw, N., Dayan, P. & O'Doherty, J.P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, 66(4), 585-95. doi: 10.1016/j.neuron.2010.04.016
- [11] Glimcher, P.W. & Rustichini, A. (2004). Neuroeconomics: the consilience of brain and decision. *Science*, 306(5695), 447-52. doi: 10.1126/science.1102566
- [12] Gottfredson, L.S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24, 791-32.
- [13] Hampton, A.N., Bossaerts, P. & O'Doherty, J.P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci*, 26(32), 8360-7. doi: 10.1523/JNEUROSCI.1010-06.2006
- [14] Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proc Natl Acad Sci U S A*, 107(43), 18243-50. doi: 10.1073/pnas.1012933107
- [15] Kaufman, S. B., DeYoung, C. G., Gray, J. R., & Brown, J. (2009). Associative learning predicts intelligence above and beyond working memory and processing speed. *Intelligence*, 37, 374-382.
- [16] Li, J., & Daw, N. D. (2011). Signals in human striatum are appropriate for policy update rather than value prediction. *The Journal of Neuroscience*, 31(14), 5504-5511.

- [17] Li, J., Delgado, M. R. & Phelps, E. A. (2011, Jan). How instructed knowledge modulates the neural systems of reward learning. *Proc Natl Acad Sci U S A*, 108(1), 55-60. doi: 10.1073/pnas.1014938108
- [18] Liljeholm, M. & O'Doherty, J. P. (2012). Contributions of the striatum to learning, motivation, and performance: an associative account. *Trends in Cognitive Sciences*, 16(9), 467-475.
- [19] Otto, A. R., Gershman, S. J., Markman, A. B. & Daw, N. D. (2012). The curse of planning: Dissecting multiple reinforcement learning systems by taxing the central executive. *Psychological Science*, (in press).
- [20] Rangel, A., Camerer, C., & Montague, P. R. (2008,). A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci*, 9(7), 545-56. doi: 10.1038/nrn2357
- [21] Rustichini, A. (2009, Dec). Neuroeconomics: what have we found, and what should we search for. *Curr Opin Neurobiol*, 19(6), 672-7. doi: 10.1016/j.conb.2009.09.012
- [22] Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275, 1593-1599.
- [23] Shamosh, N. A., DeYoung, C. G., Green, A. E., Reis, D. L., Johnson, M. R., Conway, A. R. , . . . Gray, J. R. (2008). Individual differences in delay discounting: relation to intelligence, working memory, and anterior prefrontal cortex. *Psychological Science* 19(9), 904-11. doi: 10.1111/j.1467-9280.2008.02175.x
- [24] Shamosh, N.A. & Gray, J.R. (2008). Delay discounting and intelligence: A meta-analysis. *Intelligence*, 36, 289-305.
- [25] Solway, A. & Botvinick, M.M. (2012). Goal-directed decision making as probabilistic inference: a computational framework and potential neural correlates. *Psychol Rev*, 119(1), 120-54. doi: 10.1037/a0026435
- [26] Sutton, R. S. & Barto, A. G. (1998). *Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning)*. A Bradford Book.

- [27] Talairach, J. & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. Thieme, New York.
- [28] Tanaka, S. C., Balleine, B. W. & O'Doherty, J. P. (2008). Calculating consequences: Brain systems that encode the causal effects of actions. *J Neurosci*, 28(26), 6750-5. doi:10.1523/JNEUROSCI.1808-08.2008
- [29] Tom, S.M., Fox, C.R., Trepel, C. & Poldrack, R. A. (2007) The Neural Basis of Loss Aversion in Decision-Making Under Risk. *Science*, 315(5811), 515-518. doi:10.1126/science.1134239
- [30] Van denBos, W., Cohen, M.X., Kahnt, T. & Crone, E.A. (2012). Striatum-medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. *Cereb Cortex*, 22(6), 1247-55. doi: 10.1093/cercor/bhr198
- [31] Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on psychological science*, 4(3), 274-290. Wechsler, D. (1999). Wechsler abbreviated scale of intelligence manual. San Antonio, TX.
- [32] Wu, S., Delgado, M.R. & Maloney, L. (2011). The neural correlates of subjective utility of monetary outcome and probability weight in economic and in motor decision under risk. *J Neurosci*, 31(24), 8822-31.

## Chapter 4

# Conclusions

The first study in this dissertation identifies a role of experience-based reinforcement signals for abstract reasoning during complex strategic choice, and discusses implications for strategic learning in general. The study also identifies behavioral correlates of strategically allocated cognitive effort during the decision task: Subjects adjust the time spent reasoning about which choice to make, in relation to the expected return of increased reasoning. Hence, the study also hints at an underlying principally economic<sup>1</sup> process during which cognitive effort is allocated as a scarce resource during decision-making. This aspect of the first study links it firmly to the second study during which a link between intelligence (in economic terms a measure of intrinsic individual differences in the ability to allocate cognitive resources) and differences in reinforcement signals after gains and losses is established.

The findings of the second study were considered as a possible mechanism underlying simultaneous correlations between intelligence and risk attitudes and intelligence and attitudes to temporal delay.

Since each study includes its own discussion section, the overall conclusion to this dissertation is used more as an opportunity to extrapolate more speculatively from the findings presented here to one possible further direction of this research for which the neuroeconomic approach might seem especially relevant: Both studies identify a role for cognitive ability during decision-making. Taken together they seem to suggest that the

---

<sup>1</sup> Principally economic, because it deals with the allocation of a scarce resource within the framework of maximizing a decision objective.

process via which people allocate cognitive resources during decision-making is

- a) principally economic in nature,
- b) can impact immediate decision and learning trajectories, and
- c) may underlie long-term preferences and decision tendencies.

Understanding and explicitly modeling decisions as emerging from a trade-off between costly cognitive effort and the anticipated resulting improvement of decision outcomes seems a critical future direction for cognitive neuroscience as well as psychology. Neuroeconomics, because of its intellectual history in economics, appears particularly suited for advancing this line of research. For one, observed cognitive performance, is not solely determined by cognitive ability. Instead, the decision-maker's incentives for improving performance given a particular cognitive strategy, and/or for identifying superior cognitive strategies (possibly in an attempt to reduce long-term cognitive effort) co-determine behavioral outcomes. The relation between experimentally observed cognitive effort and underlying cognitive ability (and individual differences in general) is therefore one of equilibrium, and one which can benefit from the type of modeling that is historically associated with economics; especially modeling aimed at overcoming identification problems when linking observed outcomes to latent traits. (Borghans, Golsteyn, Heckman & Humphries, 2011; Heckman, Stixrud & Urzua, 2006). This research tradition may be carried over by Neuroeconomics into current investigations of the brain and how we make decisions.

# References

- [1] Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'cole Amricaine. *Econometrica: Journal of the Econometric Society*, 503-546.
- [2] Almlund, M., Duckworth, A.L., Heckman, J. and T. Kautz (2011). *Personality Psychology and Economics*. IZA Discussion Paper No. 5500, February 2011.
- [3] Ariely, D., & Norton, M. I. (2007). Psychology and Experimental Economics A Gap in Abstraction. *Current Directions in Psychological Science*, 16(6), 336-339.
- [4] Bernoulli, D., Jan, N., & Risk, O. F. (2007). Exposition of a New Theory on the Measurement of Risk, 22(1), 2336.
- [5] Borghans, L., Duckworth, A. L., Heckman, J. J., & Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4), 972-1059.
- [6] Borghans, L., Golsteyn, B. H., Heckman, J., & Humphries, J. E. (2011). Identification problems in personality psychology. *Personality and individual differences*, 51(3), 315-320.
- [7] Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences*, 106(19), 7745-7750.
- [8] Christopoulos, G.I., Tobler, P.N., Bossaerts, P., Dolan, R.J. & Schultz, W. (2009) Neural correlates of value, risk, and risk aversion contributing to decision making under risk. *J. Neurosci.*, 29, 1257412583.



- [9] Coates, J. M., Gurnell, M., & Rustichini, A. (2009). Second-to-fourth digit ratio predicts success among high-frequency financial traders. *Proceedings of the National Academy of Sciences*, 106(2), 623-628.
- [10] Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3), 883-931.
- [11] Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience*, 8(4), 429-453.
- [12] Denissen, J. J., & Penke, L. (2008). Motivational individual reaction norms underlying the Five-Factor model of personality: First steps towards a theory-based conceptual framework. *Journal of Research in Personality*, 42(5), 1285-1302.
- [13] Dickhaut, J., Rustichini, A., Smith, V., "A neuroeconomic theory of the decision process." *PNAS* 2009, 106(52). pp. 22145-22150.
- [14] Fehr, E., & Rangel, A. (2011). Neuroeconomic Foundations of Economic ChoiceRecent Advances. *Journal of Economic Perspectives*, 25(4), 330. doi:10.1257/jep.25.4.3
- [15] Ferguson, E., Heckman, J. J.& Corr, P. (2011). Personality and economics: Overview and proposed framework. *Personality and Individual Differences*, 51(3), 201209. doi:10.1016/j.paid.2011.03.030
- [16] Fudenberg, D., & Levine, D. K. (2006). A dual-self model of impulse control. *The American Economic Review*, 1449-1476.
- [17] Gigerenzer, G. (2002). *Bounded rationality: The adaptive toolbox*. the MIT Press.
- [18] Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of mathematical economics*, 18(2), 141-153.
- [19] Glimcher, P. W., & Rustichini, A. (2004). Neuroeconomics: the confluence of brain and decision. *Science (New York, N.Y.)*, 306(5695), 44752. doi:10.1126/science.1102566
- [20] Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior (No. w12006). National Bureau of Economic Research.

- [21] Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences*, 104(33), 13250-13255.
- [22] Heckman, J.J., 2011. "Integrating Personality Psychology into Economics," NBER Working Papers 17378, National Bureau of Economic Research, Inc.
- [23] Henrich, Joseph, et al. "In search of homo economicus: behavioral experiments in 15 small-scale societies." *The American Economic Review* 91.2 (2001): 73-78.
- [24] Houthakker, H. S. (1950), Revealed preference and the utility function, *Economica*, May,159-174.
- [25] Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291
- [26] Koopmans, T. C. (1960) Stationary ordinal utility and impatience. *Econometrica*, 28, 287-309.
- [27] Krajbich, I., Camerer, C., Ledyard, J., & Rangel, A. (2009). Using neural measures of economic value to solve the public goods free-rider problem. *Science (New York, N.Y.)*, 326(5952), 5969. doi:10.1126/science.1177302
- [28] Loewenstein, G. F., Thompson, L., & Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social psychology*, 57(3), 426-441.
- [29] Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, 92(368), 805-824. Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*, 643-669.
- [30] Maccheroni, F., Marinacci, M., & Rustichini, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6), 1447-1498.
- [31] Mill, John Stuart. "On the Definition of Political Economy, and on the Method of Investigation Proper to It," *London and Westminster Review*, October 1836. *Essays*

on Some Unsettled Questions of Political Economy, 2nd ed. London: Longmans, Green, Reader & Dyer, 1874, essay 5, paragraphs 38 and 48.

- [32] Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27(37), 99848. doi:10.1523/JNEUROSCI.2131-07.2007
- [33] Previc, F. H. (1999). Dopamine and the origins of human intelligence. *Brain and cognition*, 41(3), 299-350.
- [34] Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2(4), 313-345.
- [35] Roberts, B. W., & Jackson, J. J. (2008). Sociogenomic personality psychology. *Journal of personality*, 76(6), 1523-1544.
- [36] Rustichini, A. (2009). Neuroeconomics: what have we found, and what should we search for. *Current opinion in neurobiology*, 19(6), 672-677.
- [37] Rustichini, A., DeYoung, C. G., Anderson, J., & Burks, S. (2011) Toward the Integration of Personality Theory and Decision Theory. Working Paper
- [38] Samuelson, P. (1938), A Note on the Pure Theory of Consumer Behavior. *Economica* 5(17): 61-71
- [39] San Martn, R., Manes, F., Hurtado, E., Isla, P., & Ibaez, A. (2010). Size and probability of rewards modulate the feedback error-related negativity associated with wins but not losses in a monetarily rewarded gambling task. *NeuroImage*, 51(3), 1194204. doi:10.1016/j.neuroimage.2010.03.031
- [40] Sanfey, A. G., Loewenstein, G., McClure, S. M., & Cohen, J. D. (2006). Neuroeconomics: cross-currents in research on decision-making. *Trends in cognitive sciences*, 10(3), 108-116.

- [41] Savage, L.J. 1954. *The Foundations of Statistics*. New York: John Wiley.
- [42] Schlagenhaut, F., Rapp, M. A., Huys, Q. J., Beck, A., Wstenberg, T., Deserno, L., & Heinz, A. (2012). Ventral striatal prediction error signaling is associated with dopamine synthesis capacity and fluid intelligence. *Human brain mapping*.
- [43] Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological review*, 84(1), 1-66
- [44] Shamosh, N. A., & Gray, J. R. (2008). Delay discounting and intelligence: A meta-analysis. *Intelligence*, 36(4), 289-305.
- [45] Shamosh, N. A., DeYoung, C. G., Green, A. E., Reis, D. L., Johnson, M. R., Conway, A. R., Enge, R.W, Braver, T.S. & Gray, J. R. (2008). Individual differences in delay discounting relation to intelligence, working memory, and anterior prefrontal cortex. *Psychological Science*, 19(9), 904-911.
- [46] Simon, H.A. (1959). *Theories of Decision-Making in Economics and Behavioral Science*. *The American Economic Review*, Vol. 49, No. 3, pp. 253-283
- [47] Simon, H. A. (1982). *Models of bounded rationality: Empirically grounded economic reason* (Vol. 3). MIT Press (MA).
- [48] Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3), 165-180.
- [49] Thaler, R. H. (1988). Anomalies: The ultimatum game. *The Journal of Economic Perspectives*, 2(4), 195-206.
- [50] Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4), 297-323.
- [51] van den Bos, W., & McClure, S. M. (2013). Towards a general model of temporal discounting. *Journal of the experimental analysis of behavior*, 99(1), 58-73.