DEVELOPMENTAL AND INDIVIDUAL DIFFERENCES IN DECISION-MAKING

A Dissertation

Presented to the Faculty of the Weill Cornell Graduate School

of Medical Sciences

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Johannes Hugo Decker

May 2015

DEVELOPMENTAL AND INDIVIDUAL DIFFERENCES IN DECISION-MAKING

Johannes Hugo Decker, PhD

We are defined by our behavior—how we act and the decisions we make throughout life. The processes underlying decision-making are not fully understood, especially in regards to developmental and clinical populations. Cognitive processes have been proposed to reflect the function of two systems—one fast and automatic, the other slow and deliberative. These dual-systems models fail to acknowledge the complex interconnectivity of neural networks, but have provided a useful framework in psychology for understanding decision-making. An extension of this work uses computational reinforcement-learning algorithms to help characterize potential evaluative processes thought to underlie decision formation. Simple evaluations use error-based feedback from prior responses to track the values of options while increasingly complex evaluations supplement that information to guide goal-directed actions. Such goal-directed decisions are proposed to rely on the ability to form and recruit a cognitive representation of decision relevant information. Relatedly, information received through instruction is thought to have a prolonged biasing effect on habitual learning processes. While the signal underlying simple reinforcement-learning algorithms closely matches neural activity within a well-circumscribed circuit, more complex evaluative processes making up higher-order cognitive models of the world involve a distributed network of brain regions. Given that many of these regions and their connections show dynamic changes across development and

perturbations in clinical populations, there are likely significant differences in the evaluative processes of decision-making among these groups relative to adults.

Chapter One provides an overview of habitual and goal-directed decision-making. Chapter Two tests whether children and adolescents recruit task structure knowledge to make goal-directed decisions to the same extent that adults do. Chapter Three examines how instruction biases learning and decision-making across development. Chapter Four addresses how decisions involving delayed outcomes may be perturbed in anorexia nervosa. Chapter Five summarizes these results and offers a critical assessment of the current state of computational modeling of decision-making in developmental and psychiatric populations. Collectively, in this thesis, I report an initial attempt at using computational modeling as a method for understanding the underlying evaluative processes behind individual and developmental differences in decision-making, and discuss the advantages and disadvantages of this approach.

BIOGRAPHICAL SKETCH

Hugo was born in Hamburg, Germany, and grew up in Menlo Park, California. He studied biology at Stanford University as an undergraduate and had plans of being a pediatrician. During this time he had a variety of research experiences, from x-ray crystallography (Matthews), to genetic engineering (Mendelsohn), and dendritic cell therapy (Sapolsky). As he was uncertain whether to pursue a PhD or MD, his advisors suggested he be a research assistant. He worked on a collaborative project between Peggy MacDonald and Shai Shaham at Rockefeller University, trying to develop a viral model system in *C. elegans*. There, the monthly Shaham, Vosshall, and Bargmann lab meetings led to a growing interest in studying animal behavior. At the Weill-Cornell / Sloan-Kettering / Rockefeller combined MD/PhD program, he did rotations in a human behavioral imaging lab (Casey), computational immunology lab (Altan-Bonnet), and electrophysiology labs (Aksay, Maimon), but ultimately was most interested human behavior. Since joining the Sackler Institute, Hugo has had the great pleasure of working with Catherine Hartley and BJ Casey for the past several years, where he has been studying the development of decision-making.

TO MOLLY

FOR HER LOVE & PERSPECTIVE

TO MAMA AND PAPA

FOR BEING WHO I WANT TO BE WHEN I GROW UP

TO VAL AND MONIKA

FOR VICARIOUS ADVENTURES

ACKNOWLEDGEMENTS

I am immensely grateful for the help that I have received from the following people. I would not have been able to write this dissertation without them.

Catherine Hartley – I cannot thank you enough for being my mentor. The training, guidance, and support that you've given me have made this thesis possible. You have given me so much of your time in, conversation and writing, and I have learned so much from you. I will forever strive to live up to your example.

BJ Casey – Thank you for taking me on as a student. Whether I was rotating, trying to find a project, or working through one, you were immensely supportive. Your high expectations were sometimes daunting, but they helped me become a better researcher. Thank you for being my advocate. Thank you for helping me to see the big picture. And thank you for this experience.

Charles Glatt, Winrich Freiwald, and Jonathan Victor – Thank you for the insights and criticisms you have provided for this thesis.

Bradley Doll, Ross Otto, Nathaniel Daw, Bernd Figner, Joanna Steinglass – For your conversations, writing, editing, and most of all teaching – thank you.

Andrew Drysdale – For being my constant lab companion and sounding board.

Aaron Heller – For your help with statistics and imaging, and helping me find my way at conferences, I thank you.

Sackler Institute – Thank you Alisa Powers, Gloria Pedersen, Danielle Dellarco, Lindsay Hunter, Tracey Shi, Rob Fetcho, Frederico Lourenco, Dave Johnson, Theresa Teslovich, Anfei Li, Michael Dreyfuss, and Ali Cohen. It has been a wonderful experience working with you all.

Thank you to everyone at the WCMC Citigroup Biomedical Imaging Center.

My MD/PhD classmates – for lunches, bike rides, happy hours, dinners, and so on – I thank you.

Financial Support: Medical Scientist Training Program Grant from the National Institute of General Medical Sciences of the National Institutes of Health under award number T32GM007739 to the Weill Cornell/Rockefeller/Sloan-Kettering Tri-Institutional MD-PhD Program

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Developmental and individual differences in decision-making – An introduction

INTRODUCTION

"Behavior has that kind of complexity or intricacy which discourages simple description and in which magical explanatory concepts flourish abundantly. […] Behavior is what an organism is *doing*—or more accurately what it is observed by another organism to be doing. […] It is convenient to speak of this as the action of the organism upon the outside world." – (Skinner 1938)

The understanding of human behavior continues to be a major goal of a diverse set of scientific fields, including economics, sociology, psychology, neuroscience, and computer science. Countless studies have deepened our understanding of behavior and have provided insights into the underlying neural substrates thereof, and yet, how the brain ultimately drives behavior largely remains a mystery. Decision-making is a sub-category of behavior that has received much attention, particularly those decisions motivated by rewards, due to the comparative objectivity with which a decision can be experimentally measured (Edwards 1954). A decision can be defined as follows: an individual is presented with various options, determines the value of those options, takes an action, and evaluates the outcome that is received (Ernst and Paulus 2005). While the presented options, eventual actions, and outcomes are observable, the crucial evaluative step must be inferred from those observations. This evaluation depends on prior information, the observable external environment, and on unobservable internal factors encompassing emotion, mood, attention, memory, impulses, and motivations (Hampshire et al. 2012). Prior information can come from personal experience,

observing others' experiences, and specific rules and instructions that are pertinent to the various goals one wants to achieve. The overarching goal of the study of decision-making is to understand the process by which this set of disparate factors interacts to form an eventual decision. Decisions have broadly been split into two categories, those resembling a habitual action and those that appear more goal-directed. Studying the underlying evaluative processes of these types of decisions may provide greater traction for understanding typical human behavior, and how these processes may go awry in neurological and psychiatric disorders. Clarifying the normal developmental maturation and individual differences in decision-making will hopefully lead to insights that can be applied toward understanding and preventing unhealthy decisions and treating those with difficulty in making appropriate decisions.

Rationale for studying individual differences in decision-making

The ongoing maturation of many behaviors throughout development provides an opportunity for experiences to fine-tune decision-making strategies, but also potentiates poor or risky decisions. The brain of the developing child and adolescent is thought to follow a trajectory passing through multiple sensitive periods that allow age-specific experiences to appropriately influence the formation and strengthening of various neural circuits that are necessary for supporting the increased independence of adulthood (Spear 2000; Munakata, Snyder, and Chatham 2012; Somerville 2013; Galván 2014). However, this increased independence during adolescence can also lead to dangerous situations. For a host of potential reasons, such as an increased tolerance for ambiguous situations (Tymula et al. 2012), heightened sensitivity to rewards (Galvan et al. 2006), an inability to

self-regulate impulses (Heatherton and Wagner 2011), and the heightened influence of peers (Gardner and Steinberg 2005; Steinberg and Monahan 2007; Pfeifer et al. 2011), adolescents often make decisions that can put themselves at risk (Simons-Morton, Lerner, and Singer 2005; Crone et al. 2008; Chein et al. 2012). This increase in risky decision-making likely contributes to the increase in mortality rates seen in adolescence, when individuals are otherwise in prime health (Casey 2014). Indeed, 71% of deaths among persons aged 10-24 in a given year are a result of motor-vehicle crashes, other unintentional injuries, homicide, and suicide (Eaton et al. 2006). In contrast, earlier development of self-control strategies, such as being able to favor distant rewards over immediately rewarding actions, is associated with more rational decision-making and reaching higher levels of academic and social competencies later in life (Mischel, Shoda, and Peake 1988; Casey et al. 2011). Adolescence is also the time period where many psychiatric disorders are prone to develop (P. Cohen, Cohen, and Brook 1993; Wittchen, Nelson, and Lachner 1998; Merikangas et al. 2010; Paus, Keshavan, and Giedd 2008; F. S. Lee et al. 2014), many of which are also associated with an increase in harmful and potentially fatal decisions (Henriksson et al. 1993; Sullivan 1995; Suominen et al. 1998; Sareen et al. 2006).

The widespread impairment of decision-making in numerous psychiatric and neurological disorders has led to a growing medical interest in studying how its underlying processes may be perturbed (Bechara, Damasio, and Damasio 2000). Such decision-making impairments have been proposed to result from distinct types of dysregulation in different cognitive processes (Siegel and Ryan 1989; Barkley et al. 2001). Thus, diminished inhibitory control in ADHD, hypersensitivity to rewards in substance abuse, diminished

reward learning in Parkinson's disease and schizophrenia, and the distorted weight placed on negative outcomes in anxiety and depression are likely due to disruptions in specific subcomponents of decision-making processing (D. Lee 2013). Such dysregulation can have profound consequences. Children with attention-deficit disorder are more likely to progress to problems with substance abuse, potentially resulting from a compounding effect in which early decrements in certain cognitive domains may prevent the appropriate development of other dependent cognitive processes (Nigg and Casey 2005). Poor decision-making has more generally been associated with greater levels of drug use, addiction, and gambling (Brand et al. 2005; Bechara 2005; H. de Wit 2009; Michalczuk et al. 2011). In anorexia nervosa, a defining part of the disease is the abnormal decision to persistently restrict the intake of caloric needs which leads to a significantly low body weight (American-Psychiatric-Association 2013). This preference for delayed outcomes extends beyond decisions about food and body image (Steinglass et al. 2012), suggesting it may underlie a perturbation in a more general underlying evaluative process. Characterizing the evaluative processes that underlie decision-making will lead to a better understanding of both normal and disordered decisions.

Historical account of decision-making

Many different fields of study have provided insight into decision-making processes. Philosophers have debated the concept of free will for millennia (Pereboom 2001). Behaviorists asserted that animal behavior was purely reactive to external stimuli (Thorndike 1898). Economists viewed decisions as rational or irrational and that they revealed an individual's internal preference (Edwards 1954). Psychologists have been focused on determining

how decisions are represented in the mind (Anderson 1990). Cognitive neuroscientists have studied the neural networks that form both an individual's mental state and their eventual decision (Gazzaniga 2004). More recently, the new interdisciplinary field of neuroeconomics has been formed in an effort to bring together these varying approaches that are all studying the same process (Glimcher and Fehr 2013). But first, let us begin with a brief account of the historical study and theory of decision-making.

Early researchers heavily debated the level of "animal intelligence" (Thorndike 1898). Thorndike made repeated observations of learned animal behavior, specifically that of animals escaping an enclosure, and concluded that these complex actions were merely associative, habitual learning. An animal's initial goal to escape was instinctive, which led to more or less random actions, and those actions that led to the animals escape were strongly reinforced. Thus, regardless of the anecdotal evidence suggestive of animal intelligence—his term for what he considered the sloppy scientific methodology of the day—Thorndike believed that all animal action was either instinctive or associative, rather than an "intelligent" evaluation to achieve a specific goal.

Associative learning was the basis of much of modern psychology (Skinner 1938), and from which a careful set of terms were generated to allow for the objective and quantitative description of behavior. A stimulus was defined as an aspect or modification of the environment that affects an individual, and a response was defined as the corresponding action of that individual. Further, a reflex was defined as the learned behavioral association between a stimulus and response. While careful experimentation has extended this vocabulary, many experiments to this day use simple stimulus-

response learning as means of studying behavior and specifically habitual decisions. I will be using the terms habitual and goal-directed decision-making to distinguish between two categories of decisions. Habitual decisions are based on associative, stimulus-response learning whereas goal-directed decisions are based on knowledge of rules, contingencies in the environment, and aspects of outcome not related to reward (Miller and Cohen 2001).

Thorndike's behaviorist account of habitual decision-making dominated for many decades, with no satisfactory evidence surfacing that showed that non-human animals were indeed capable of non-associative, "intelligent", goal-directed decisions. However, through carefully designed and replicable methodology, evidence emerged that rats could indeed make purposeful, goal-directed decisions from information that was not obtained through reward-driven learning (Tolman and Honzik 1930). Tolman's work showed that rats that had been pre-exposed to a maze learned much faster to complete the maze optimally once a reward was introduced relative to a control group that did not have this pre-exposure. This was an example of latent learning, learning of the maze structure that occurred without reinforcement, which allowed rats to make goal-directed decisions at key points in the maze. Many experiments followed that revealed similar types of goal-directed actions were possible, suggesting that animal decisions were not purely habitual. Tolman asked a question that is still being pursued today, 'what are the conditions that favor the learning and utilization of a broad cognitive map versus a narrow map?' Furthermore, he proposed that many psychopathologies resulted from the narrowing of such a cognitive map (Tolman 1948). Since this time, many animal behaviors have been observed that cannot be adequately described by habitual or associative processes. Indeed, the dead-reckoning of foraging

ants, the spatial representation of the digger wasp's nest, and the ability of rats, birds, and fish to represent time, number, and rate all provide evidence that animals are capable of goal-directed behavior (see Gallistel 1989 for a review), and not simply behavior that reflects stimulus-response associations.

Throughout history, behavior has been described as being determined by two separate systems, alluded to throughout this thesis as habitual and goal-directed decisions. From Plato's dichotomy between reason and passion, Freud's Id and Ego, and Thorndike's animal reflexes and intelligent behavior—many accounts of behavioral and cognitive processes have been put into this dual systems framework: heuristic versus analytic (Evans 1984), associative versus rule-based (Sloman 1996), implicit versus verbal (Ashby et al. 1998), hot versus cool (Metcalfe and Mischel 1999), automatic versus controlled (Schneider and Shiffrin 1977; Posner, Nissen, and Ogden 1978), bottom-up versus top-down (Corbetta and Shulman 2002; Posner and Petersen 1990), system I versus system II (Kahneman 2003), and impulsive versus reflective (Strack and Deutsch 2004). The many instantiations of this dual system model present the two systems as acting separately from one another. Where the habitual system is seen as emotional, fast, reflexive, simple, and automatic, the more goal-directed system is seen as reasoned, slow, reflective, complex, and deliberative.

In decision contexts where animals tend to employ habitual learning strategies, humans often take a goal-directed approach. Even for simple procedural learning tasks, in which stimulus-response learning is sufficient, humans often use quite complex strategies (S. C. Hayes et al. 1986). This use of goal-directed strategies is thought to require a cognitive representation of the task at hand, and thus may rely on the ability to verbalize what must be

done. As such, pre-verbal infants and animals often perform similarly to each other but differ remarkably from older children and adults in basic reward learning tasks (Lowe, Beasty, and Bentall 1983). However, it is critical to note that goal-directed decision-making is not necessarily superior to habit, but that either strategy may be appropriate, or superior, in a given environment. Indeed, attempting to use a reasoned approach can actually diminish performance relative to more simple stimulus-response strategies in some tasks (Bocanegra and Hommel 2014). And even complicated decisions are often answered better using non-deliberative approaches (Dijksterhuis et al. 2006). In an elegant example that shows the superiority of habitual learning in certain contexts, pigeons and humans were both trained in a categorization task in which the categories were difficult to verbalize, and whereas all four pigeons could transfer this knowledge to a test phase, only 1 of 10 humans performed above chance (Jitsumori 1993). The pigeons likely learned the categorization through simple associative learning and responded using habitual processes. In contrast, the human participants' reported use of inaccurate verbal rules likely led to their poor performance (Ashby et al. 1998). Thus, recruitment of a verbalized goal-directed strategy can impede habitual learning processes and actions, which athletes are well aware of when they attempt to verbalize their well trained motor actions (Flegal and Anderson 2008). This work suggests that the degree to which instructions elicit a verbal rule or cognitive map may change a normally automatic decision into one that is goal-directed.

The expansive adoption of dual-systems terminology comes from the useful intuition it provides for how humans behave, but there are issues with using such simple toy models. The distinction between these systems is not

always consistent (Sloman 1996), especially the use of “hot and cool” terminology, which has occasionally made interpretation of results difficult (Gladwin and Figner 2014). This confusion is compounded when applying these models across development, when the brain networks proposed to underlie these processes are in a maturational flux. Here, the two systems have been presented as being imbalanced (Casey et al. 2010), where the ontogenetically early maturation of a simpler system relative to the later maturation of a more complex system is manifested in the turmoil of adolescence. This model has been critiqued (Pfeifer and Allen 2012), defended (Strang, Chein, and Steinberg 2013), and reframed to account for the known issues in dividing the brain into two systems (Casey 2014). Neural networks are widely distributed, parallel, and interconnected rather than separate (Mesulam 1990). As such, it is presumed that many evaluative processes occur in tandem across mutually overlapping regions of the brain to eventually result in a single decision. Thus, it is especially critical to consider how these processes interact and inform one another.

Despite their limitations, dual-systems models capture important behavioral distinctions, and have provided an important insight into behavior that has remained valid over time, solidifying the idea that decisions stem from multiple underlying evaluative processes (Sloman 1996). Whereas the goal-directed form of evaluation may involve the recruitment of cognitive representations of external stimuli and internal goals, the habitual system may rely on associative stimulus-response representations. The key point is that these evaluative processes are not separate. A parsimonious and more biologically plausible model suggests that there likely exists as a continuum of how many modalities and cognitive representations can be evoked when

making a specific decision (Kahneman 2003; Marchette, Bakker, and Shelton 2011; Daw et al. 2011; Dolan and Dayan 2013). Adaptive control of behavior in a dynamic environment involves a fluid and contextually sensitive balance between these dissociable learning systems. Whereas habitual behavior allows for the well-honed behavioral routines without forethought or attention, goal-directed behavior enables the flexible adaption of behavior to the dynamic state of our world, and true behavior is a mixture of both. Throughout this thesis, I will use the terms habitual and goal-directed to represent the ends of this continuum. While there is still much to be learned about decision-making, many of the component cognitive processes that support various types of decision-making have been quite well described.

Cognitive processes of goal-directed decision-making

Decision-making depends on various cognitive processes at every stage of a decision—from evaluating potential options, executing an action, and evaluating the outcome (Ernst and Paulus 2005)—including, but not limited to, working memory, planning, and inhibitory control. The set of cognitive processes which support goal-directed behavior has been called executive or cognitive control (Miller and Cohen 2001). Evoking these processes to various extents may change decision behavior. Working memory was initially thought to be a singular system that kept recent information in mind, but is now considered to consist of auditory, visual, and episodic memory components that are coordinated by a central executive (Baddeley and Hitch 1974; Paulesu, Frith, and Frackowiak 1993; Baddeley 2010). Working memory ability is long been known to improve across development (Siegel and Ryan 1989), and may depend on different neural structures across

development (Finn et al. 2010). Decisions have been shown to become less goal-directed and more habitual when this central executive system is taxed with a working-memory task (Otto, Gershman, et al. 2013). Thus, depending on the complexity of the task, working memory capacity may play a crucial role in goal-directed decision-making.

Planning is also crucial for decision-making and relies on proper working memory to keep goal-relevant information in mind. A decision can be planned based both on memories of past events (retrospection) and simulations of potential future outcomes (prospection) (Gilbert and Wilson 2007), both of which rely on a cognitive representation of the world (Gershman, Markman, and Otto 2014). Additionally, the ability to represent and utilize rules relies both on planning and working-memory processes. Rules and explicit information can help form a cognitive representation of the decision problem and must be maintained in working memory at the relevant points of a decision. As may be expected, this ability to hold and use multiple rules in mind increases with age (Diamond, Kirkham, and Amso 2002). The degree to which internal and external rules are followed can influence the types of decision-making strategies that are pursued (Sloman 1996). For example, rules and prior information have the ability to generate confirmation biases (Nickerson 1998), in which new information is weighed subjectively differently to fit with prior beliefs. That is, the normal associative learning thought to underlie habitual action is altered in such a way that it does not accurately track value changes, which has been proposed to occur through the influence of goal-directed representations of the world (Doll, Simon, and Daw 2012).

The ability to inhibit non goal-oriented actions is another crucial aspect of decision-making, and has often been used to probe disorders of impulsivity (Bickel, Odum, and Madden 1999; H. de Wit 2009; Scheres et al. 2010). Multiple factors can lead to impulsive decisions, such as the salience of the inappropriate stimulus (Casey 2005), the inability to keep goal relevant information in mind, or the inability to appropriately apply that information, which can be observed when an individual can verbalize the appropriate action but fails to do so (Strommen 1973). A task commonly used as a measure of impulsivity and self-control is known as delay discounting (Steinberg et al. 2009), which shares some properties with delay of gratification (Mischel, Shoda, and Rodriguez 1989). Discounting tasks require individuals to choose between smaller rewards that are available immediately and larger rewards available after a variable delay. The overall preference for delayed rewards serves as an assessment of self-control and has been considered to reflect a goal-directed decision that requires the cognitive representation of the future (Kurth-Nelson, Bickel, and Redish 2012). Indeed, imagining the future leads to a preference for the larger delayed rewards, potentially because the forced prospective representation of the future simplifies the evaluation of delayed rewards (Benoit, Gilbert, and Burgess 2011; J. Peters and Büchel 2011). Decision tasks can be designed to test each component of a decision, and examining the underlying evaluative process of these steps will help clarify decision-making.

Although perception, emotion, motivation, and long-term memory also support decision-making, the cognitive processes discussed above are critical for the types of decisions that occur in the tasks presented in this thesis. Furthermore, there are well characterized developmental differences observed

in cognitive control (Munakata, Snyder, and Chatham 2012), working memory (Diamond, Kirkham, and Amso 2002; Olesen, Westerberg, and Klingberg 2004), effective use of abstract rules or instruction (Bunge and Zelazo 2006), and response inhibition (Diamond 2006; Somerville and Casey 2010). These processes are proposed to underlie the increased capacity for making goal-directed decisions with age (Munakata, Snyder, and Chatham 2012). Finally, psychiatric diseases are known to have deficits in these cognitive processes as well, which may explain the disrupted decision-making of many individuals with psychiatric conditions in tasks of reward learning and delay discounting (Montague et al. 2012; D. Lee 2013).

Decision-making neural circuitry

As much of the early work in characterizing the neurocircuitry of decision-making processes was built upon the foundation of a dual-systems model, it is straightforward to characterize the progress made in understanding the underlying neurobiology in those terms. Both habitual and goal-directed action selection strategies rely on large, overlapping, and cooperative neural networks, although they have often been proposed to be supported by separate circuits. The evolution of different brain structures is used as a strong indicator for which networks are required for various cognitive functions. Whereas the subcortical regions known as the basal ganglia have a conserved structure across vertebrate species, there has been tremendous growth and increasing complexity in the prefrontal cortex (Redgrave, Prescott, and Gurney 1999; Grillner and Robertson 2015). It follows that the habitual behaviors that are observed in both lower and more phylogenetically advanced animals are likely supported by subcortical and very basic forebrain regions (Grillner et al.

2008), whereas the increasingly complex goal-directed behaviors of humans likely depend on the expanded prefrontal cortex.

Animal lesion studies, when combined with simple learning tasks, have provided insights into the underlying neural circuitry of decision-making, many of which have been extended to humans with neuroimaging studies. In these experiments, animals are initially trained to perform a specific action through reinforcement and then, by changing the value or contingency of the outcome, their behavior can be assessed to determine whether they are employing a habitual or goal-directed strategy. Early studies of the striatum showed that there was a major division, with the ventral striatum shown to be processing salience and reward information (Kelly, Seviour, and Iversen 1975; Taylor and Robbins 1986; Cardinal et al. 2001), and the dorsal striatum shown to be processing sensorimotor information to help learn and refine motor actions (Graybiel 1995; Voorn et al. 2004). This division was proposed to reflect the learning and action components of a decision, with a shared role in evaluation (Joel, Niv, and Ruppin 2002). Such a division has also been reported in human imaging studies (O'Doherty et al. 2004). The dorsal striatum however can be further subdivided and has been shown to play roles in the cognitive aspects of decision-making in addition to simple motor functions (Balleine, Delgado, and Hikosaka 2007; Bornstein and Daw 2011). Forming stimulus-response associations is the basis of habitual actions, whereas learning about the outcomes of those actions is a hallmark of goal-directed behavior. The dorsolateral striatum has been shown to be necessary for the acquisition and expression of habitual behaviors (Yin, Knowlton, and Balleine 2004; Yin and Knowlton 2006; Yin, Ostlund, and Balleine 2008; Tricomi, Balleine, and O'Doherty 2009), and the dorsomedial striatum has been shown to be a key

node in goal-directed behavior (Yin and Knowlton 2004; Yin et al. 2005; Clarke, Robbins, and Roberts 2008; Brovelli et al. 2011). A rare human case of bilateral lesion to the head of the caudate (dorsomedial striatum) implicates it in the role of numerous cognitive processes, including those thought to underlie goal-directed decision-making (Richfield, Twyman, and Berent 1987). Separately, the hippocampus also projects to the ventral striatum, and is thought to be crucial in the planning aspect of goal-directed decision-making (van der Meer, Kurth-Nelson, and Redish 2012). Taken collectively, the striatum is an important conduit for both habitual and goal-directed decisions.

Another influential contribution to decision-making that must be discussed alongside striatal function is the dopamine neuromodulatory system. Dopamine, through its input on the ventral and dorsal striatum, has long been known to play a role in reward learning (Schultz 1986; Taylor and Robbins 1986) and action selection (Dauer and Przedborski 2003). However, dopaminergic inputs from the ventral tegmental area innervate broad swaths of the cortex as well. Dopamine appears to play a key role in modulating goal-directed behavior (Durstewitz and Seamans 2008; Cools 2011), potentially through its input on the prefrontal cortex as well as in the hippocampus (Goto and Grace 2005). Although the mechanisms through which dopamine modulates goal-directed choice are not well understood, there is evidence that it may stem from its gating of prefrontal and hippocampal information to the striatum, enabling state and other goal-relevant information to inform the learning signal (Goto and Grace 2005). The dopaminergic system undergoes marked developmental changes from childhood to adulthood and this dopaminergic maturation has been proposed to contribute to the maturation of goal-directed behavior in rodents (Naneix et al. 2012). Other neurotransmitter

systems are known to influence multiple cognitive processes that are involved in making decisions (serotonin in impulsivity and emotion, acetylcholine in memory, and norepinephrine in attention and arousal), and therefore it is important to include these neuromodulatory systems in the interpretation of the neural circuits of decision-making (Ernst and Paulus 2005).

The distinction between sub-region roles within the prefrontal cortex is less clear, with many studies highlighting areas that support different aspects of goal-directed decision-making. This is due in part to the wide distribution of cognitive functions among different cortical regions, and the fact that higher cognitive functions are themselves less distinguishable from one another and rely on subcomponents of each other (Manes et al. 2002). An example of this can be seen in the broad cognitive impairments noted in patients with diffuse frontal lobe damage (Russel 1948; Delazer et al. 2007). However, by comparing naturally occurring lesions of the human prefrontal cortex with homologous surgical or ablative lesions in monkeys and rats that result in similar disruptions in cognitive processes related to goal-directed decision-making (e.g. response inhibition or perseverative behavior, temporal ordering of actions, and spatial orientation (B. Kolb 1984)), some separation is possible. Whereas orbital prefrontal (ventral in rats) lesions lead to more problems with response inhibition, dorsolateral prefrontal (medial in rats) lesions lead primarily to problems in ordering and spatial orientation. These regions correspond fairly neatly with recent neuroimaging studies in humans for response inhibition (Aron and Poldrack 2006; Chikazoe et al. 2007), planning and ordering (Tanji and Hoshi 2001; Knutson, Wood, and Grafman 2004), and the use of spatial maps (Hagler and Sereno 2006; Curtis 2006). While the prefrontal cortex has been shown to support many executive functions (Miller

and Cohen 2001), it is becoming increasingly clear that these rely on a broad network of regions outside of the prefrontal cortex, especially the parietal cortex, amygdala, and hippocampus (Corbetta and Shulman 2002; Casey 2014).

The preceding paragraphs point to specific regions involved in various components of the decision-making process, but it is crucial to also consider their interconnectivity, as no region functions in isolation. The cortex has inputs on the striatum in multiple distinct circuits, and after relaying through multiple nuclei of the basal ganglia, these circuits end up looping back to the cortex (Graybiel 1995; Haber and Knutson 2009). The dorsolateral striatum is interconnected with the sensory and motor cortex, the dorsomedial striatum with associative cortex, and the ventral striatum with the prefrontal cortex, amygdala, and hippocampus, but these circuits are not defined by sharp borders and are instead overlapping (Graybiel 1995; van der Meer, Kurth-Nelson, and Redish 2012). Furthermore, the output neurons in these loops project more heavily back to the associative prefrontal cortex than to the motor cortex, suggesting that any separation of decision-making systems in the striatum is likely reintegrated in the prefrontal cortex to assist in the arbitration between multiple evaluative processes (Graybiel 1995). Thus, this corticostriatal circuitry integrates information form a broad network of neural structures that support decision-making (van der Meer, Kurth-Nelson, and Redish 2012).

Decision-making networks in developmental and psychiatric populations

The neural structures that comprise the decision-making network undergo extensive maturational changes across development (M. H. Johnson

2001; Somerville, Jones, and Casey 2010). Neural maturation can be measured using multiple indices, such as local and global structural changes in addition to functional changes. At the local structural level, gray-matter cortical thickness has been shown to increase sharply in infancy and early childhood, peak in late childhood, and decrease across development in a posterior to frontal fashion (Giedd et al. 1999; Gogtay et al. 2004). This change in thickness is thought to reflect the overgrowth and pruning of neuronal synaptic connections (Chechik, Meilijson, and Ruppin 1998), and the posterior to anterior pattern is thought to be related to the gradual improvement of first motor and then more cognitive abilities across development. A recent study focusing on subcortical structures has revealed highly heterochronous development in the striatum, with the caudate tail and putamen showing significant expansion with age and the caudate head showing contraction with age, and that these maturational changes continue into early adolescence (Raznahan et al. 2014). These changes in gray-matter volume suggest that fine-tuning of local circuits continues throughout development. At the global structural level, white-matter volume has been show to increases well into the third decade of life (Imperati et al. 2011), thought to reflect the increasing myelination of axons, which allows for the more efficient communication between regions comprising a neural network. More intact fronto-striatal white matter tracts, as measured by lower levels of radial diffusivity, have been associated with increased levels of cognitive control (Liston et al. 2006). The structural changes that occur across development strongly suggest that the strong integration between different regions of the brain is crucial for the increasing cognitive abilities with increasing age. Structural changes, however, do not speak to the changes in

network function that are thought to underlie developmental differences in behavior.

Functional imaging studies, in contrast, allow for the BOLD (blood-oxygen-level dependent) signal of every region of the brain to be directly compared with behavior as it occurs, allowing for a richer characterization of which networks may play a role in decision-making across development (Luna and Sweeney 2004; Kotsoni, Byrd, and Casey 2006). Supporting the structural observations, functional imaging studies have found that greater cognitive control is associated with an increased recruitment of the frontostriatal network (Dalley, Everitt, and Robbins 2011), and that the deficits in cognitive control of children and adolescents are associated with lower BOLD signal in these regions (Rubia et al. 2006; Rubia et al. 2007; Somerville, Hare, and Casey 2011). While better performance in tasks gauging executive functioning is generally associated with increased BOLD signal (Crone and Dahl 2012), some studies report heightened responses in children with equivalent performance to adults (Pfeifer and Allen 2012). This discrepancy likely reflects the difference between the tendency to recruit a cognitive process and the efficiency with which it is done, highlighting the care needed in interpreting developmental differences (Strang, Chein, and Steinberg 2013). In addition to the differential recruitment of the same network, some studies suggest that children and adults recruit different networks for the same task (Thomas et al. 2004; Finn et al. 2010). These studies provide evidence that decision-making networks are recruited differentially across development, which could reflect a difference in the underlying evaluative processes themselves, or how they are being recruited to inform choices across the lifespan.

The diverse symptomatology of various psychiatric disorders has led to the fractionated study of their underlying neurobiological substrates; however by focusing on the overlapping symptom of disrupted decision-making, insights from one disorder may be informative for the others (Sharp, Monterosso, and Montague 2012). Additionally, it is critical to understand the development of decision-making and its neural circuitry as most psychiatric diseases appear in adolescence (Casey, Oliveri, and Insel 2014). Disruptions in the neuromodulatory systems play a role in many psychiatric disorders—dopamine in substance abuse, attention deficit-hyperactivity disorder, and schizophrenia; serotonin in depression, anxiety, and anorexia nervosa—which are known to be important for decision-making (Montague et al. 2012). Dopamine is crucial for both habitual and goal-directed actions and serotonin has been associated with impulsivity and self-control (D. Lee, Seo, and Jung 2012). Currently, a trend in psychiatry has been to focus on the changes in neural circuitry that are at the root of various psychiatric diseases (van der Meer, Kurth-Nelson, and Redish 2012). An important characteristic of substance use disorders is an underlying change in brain circuitry that may persist beyond detoxification (American Psychiatric Association 2013), which carries the strong implication that there is a disruption in the system underlying reward learning and habit formation (Huys et al. 2014). In anorexia nervosa, both disease specific and neutral tasks have been used to reveal differential neural signaling from healthy controls in many regions of the decision-making network, including the anterior cingulate (Zastrow et al. 2009), head of the caudate (Rothemund et al. 2011), and dorsolateral prefrontal and parietal cortices (van Kuyck et al. 2009). The broad array of cognitive processes and associated neuronal changes observed in various psychiatric disorders makes

it difficult to generate a clear network model that underlies the observed cognitive changes in these disorders. By focusing on perturbed decision-making that is devoid of disease specific stimuli, it may be revealed that different aspects of a common evaluative process are affected in different psychiatric disorders.

Computational approaches towards understanding decision-making

Computational models have been used to formalize the distinct evaluative processes postulated by dual systems theories of decision-making. Marr proposed that to understand an informational processing system—vision in his example, but the proposal is equally valid for decision making—it must be characterized at three levels: computational theory, representation and algorithm, and hardware implementation (Marr 1982). The computational theory captures the goal of the process and the logic of the strategy behind its implementation. In psychology a decision has been defined as observing and evaluating a set of options, and taking an action to obtain an outcome – with habitual and goal-directed responses serving as two strategies to do so. The representation and algorithm is a set of mathematical rules that allows input (external stimuli and internal states) to be transformed into output (action). The hardware implementation is the brain, but determining the mechanism of this implementation and which networks are involved in various cognitive processes is a pursuit that will long be followed in neuroscience. Research focusing on any of these levels of understanding a cognitive process will inform the others.

While it is agreed that there is likely a continuum of decision-making between purely habitual choices and those that are goal-directed (Dolan and

Dayan 2013), these two extremes of decision behavior have proven quite easy to represent as mathematical models. Two classes of reinforcement-learning algorithms that are formalized in computational theory capture the key properties of habitual and goal-directed learning and have been proposed to approximate their underlying neural computations (Daw, Niv, and Dayan 2005). “Model-free” learning recruits trial and error feedback to efficiently update a cached action value associated with a stimulus. In contrast, “model-based” learning algorithms select actions via a flexible but computationally demanding process of searching a cognitive model or “map” of potential state transitions and outcomes. Rescorla and Wagner described an early model-free learning algorithm, based on a reward prediction error signal that encodes the discrepancy between expected and experienced outcomes (A. R. Wagner and Rescorla 1972). This algorithm has been modified for various types of related learning (Sutton and Barto 1998), but each version relies on a simple evaluation system that uses error-based feedback from prior responses to track the value of various options. In contrast, a model-based algorithm uses a representation of all possible action outcome possibilities and calculates which action is likely to lead to the highest reward given the current information. Furthermore, a behavioral strategy that favors a mixture of the habitual and goal-directed approaches can be implemented computationally as a weighted average of these two separate reinforcement-learning algorithms (Daw et al. 2011).

Model-free and model-based learning are proposed to recruit distinct underlying neural processes (Daw, Niv, and Dayan 2005). Dopaminergic input to the striatum, carrying a signal that resembles a computational reward prediction error, is thought to support such a model-free learning process

(Schultz, Dayan, and Montague 1997; Pagnoni et al. 2002; McClure, Berns, and Montague 2003; O'Doherty et al. 2003). Developmentally, the striatal prediction error signals thought to underpin model-free learning appear to be relatively mature from childhood onwards (Galvan et al. 2006; J. R. Cohen et al. 2010; van den Bos et al. 2012). Model-based learning also engages striatal prediction error signals (Daw et al. 2011), but these signals are thought to integrate information about states and outcomes that stem from a more extensive network of brain regions. Both the hippocampus and the prefrontal cortex, regions with strong connectivity to the striatum (Pennartz et al. 2011), are proposed to play a role in the representation and search of a cognitive model of the task (Hassabis et al. 2007; Pfeiffer and Foster 2013; Wilson et al. 2014). Model-based learning also depends on the dorsolateral prefrontal cortex (Smittenaar et al. 2013), which may reflect the engagement of working memory and cognitive control processes (Miller and Cohen 2001; Otto, Gershman, et al. 2013). Studies that examine the development of these circuits and their relationship to the underlying evaluative processes of decision-making may serve as a way to understand suboptimal decision-making across development and in various psychiatric populations (Stephan and Mathys 2014). Extending computational algorithms that were developed to study adult decision-making such that they characterize the decisions of developmental and psychiatric populations is one approach to studying these evaluative processes.

While the model-free and model-based algorithms include both the learning and action components of a decision in their algorithms, alternative models for the action selection algorithm can be used independently to describe decision-making in tasks where no learning occurs. This is the

approach that is used in tasks known as delay (or probability) discounting, in which a choice must be made between two differently valued options that are available after different delays (Laibson 1997). This type of choice has also been suggested to reflect goal-directed decision-making, as evaluating the delayed reward can be considered as a search through a cognitive representation of the future (Kurth-Nelson, Bickel, and Redish 2012). Therefore, a computational model that only consists of action-selection phase can still be used to probe the underlying process of the decision-making that supports goal-directed behavior.

The Current Thesis

In this thesis, the goal is to further the understanding of the evaluative processes underlying decision-making across development and within anorexia nervosa. We explored how children, adolescents, and adults make decisions when different levels of task information, such as task structure or advice, could be used to form an evaluation, and thereby sought to understand the reliance on simpler, habitual versus more complex, goal-directed strategies of decision-making across development. Additionally, by comparing how healthy individuals and those with anorexia nervosa evaluated decisions involving delayed outcomes while they underwent a functional neuroimaging scan, we probed whether these groups recruit different neural circuits, or the same circuits to differing extents.

The thesis has been structured as follows: **Chapter 1, "Individual and developmental differences in decision-making – An introduction,"** provided the basic framework by which psychologists and cognitive neuroscientists have studied decision-making, habitual and goal-directed, two

strategies at extreme ends of a decision-making continuum. Furthermore, developmental and psychiatric disruptions in cognitive processes and neural circuitry were highlighted to offer an indication that the underlying evaluative processes may be different form those seen in healthy adults. Finally, it summarized the computational approach to understanding decision-making, and offered a rationale for studying the development of goal-directed strategies. **Chapter 2, "Model-based behavior emerges across development,"** examines the developmental recruitment of task structure knowledge in making goal-directed decisions. **Chapter 3, "Experiential learning outweighs instruction prior to adulthood,"** (Decker, Lourenco, Doll, & Hartley. 2015. Cognitive Affective Behavioral Neuroscience), tests for developmental differences in how instruction biases learning and decision-making. **Chapter 4, "On weight and waiting: delay discounting in anorexia nervosa pretreatment and posttreatment,"** (Decker, Figner, Steinglass. 2015. *Biological Psychiatry*), utilizes functional magnetic resonance imaging, alongside a delay discounting decision-task task in order to explore differences in evaluative processes between healthy controls and individuals with anorexia nervosa (in both an underweight and weight-restored state). **Chapter 5, "Individual and developmental differences in decision-making: Conclusions and implications,"** summarizes the results of the current thesis. Furthermore, it offers a critical assessment of the utility of using computational modeling to understand individual and developmental differences in decision-making, and discusses future directions to increase the relevance of this approach.

Chapter 2: Model-based behavior emerges across development

We are defined by our behavior—how we act and the decisions we make throughout life. Psychological theories distinguish between two broad types of decisions, goal-directed and habitual. Mirroring this distinction, the field of reinforcement learning defines “model-based” and “model-free” classes of algorithms that capture key aspects of these two forms of action. Model-based learning generates and searches a cognitive map of potential paths and outcomes, enabling flexible behavioral adaptation to a dynamic environment. Model-free learning incrementally updates and stores a cached action value or policy associated with a stimulus, allowing the execution of well-honed behavioral routines without forethought or attention. While adults have been shown to rely on a mixture of these two strategies, the developmental trajectory of action selection strategies has not yet been examined. In this study, we adapted a two-stage Markov reinforcement-learning task for use across development, which allowed us to estimate model-based and model-free contributions to choice behavior in each age group. Whereas a model free strategy was evident in the choices across all age groups, model-based influence on choice only emerged in adolescence and continued to increase with age.

INTRODUCTION

Learning to select actions that will yield the best outcomes is a lifelong challenge. From even very young ages, children demonstrate competence in making many simple value-based decisions (Jacobs and Klaczynski 2002). However, there is also abundant evidence that the choices of children and adolescents differ in important qualitative ways from those of adults. Younger individuals tend to persist in performing actions that were previously rewarded

even after those outcomes were no longer valued relative to other potential actions (Piaget 1954; Klossek, Russell, and Dickinson 2008). Children and adolescents often make seemingly shortsighted choices that prioritize immediate gains over longer-term rewards (Mischel, Shoda, and Rodriguez 1989; Green, Fry, and Myerson 1994). Such choices have been proposed to reflect a regulatory failure, in which insufficient executive control leads to the performance of a pre-potent action or the prioritization of a hedonically alluring outcome over a more valuable alternative (Posner and Rothbart 2000). Indeed, studies of cognitive development have clearly demonstrated that executive functions such as inhibitory control and working memory improve as one matures into adulthood, conferring increased ability to withhold suboptimal actions and maintain goal-relevant information (Diamond 2006). However, an alternative but not necessarily contradictory account is that this normal cognitive maturation may alter the learning processes through which actions are evaluated, yielding developmental differences in the estimates of which actions are best.

Several prominent theoretical models of decision making distinguish two types of evaluative processes that can inform one's choices (Sloman 1996; Kahneman 2003). A slower and more deliberative process compares possible actions and their likely consequences to identify the action most likely to obtain the current goal. In contrast, a more rapid and automatic process links previously rewarded actions to the cues and contexts with which they were associated, enabling reflexive repetition of actions that were successful in the past. A large psychological and neuroscientific literature provides support for such a distinction, suggesting that "goal-directed" and "habitual" forms of action evaluation yield distinct behavioral tendencies and have largely

dissociable neural substrates (Dolan and Dayan 2013). Two classes of algorithms formalized in computational theory capture the key properties of goal-directed and habitual learning and have been proposed to approximate their underlying neural computations (Daw, Niv, and Dayan 2005). “Model-based” learning algorithms select actions via a flexible but computationally demanding process of searching a cognitive model or “map” of potential state transitions and outcomes. In contrast, “model-free” learning recruits trial and error feedback to efficiently update a cached action value associated with a stimulus. Adaptive control of behavior in a dynamic environment involves a fluid and contextually sensitive balance between these dissociable learning systems. Whereas model-free learning allows us to carry out well-honed behavioral routines without forethought or attention, model-based learning enables us to flexibly adapt our behavior to the dynamic state of our world.

In adulthood, these two learning systems are proposed to operate in parallel, competing for control over behavior (Dickinson 1985). Previous studies have indeed demonstrated that adults exhibit a mixture of both model-free and model-based learning strategies during decision-making tasks (Daw et al. 2011), and that reliance upon a given strategy is sensitive to the cognitive or affective demands placed on the individual (Otto, Gershman, et al. 2013; Otto, Raio, et al. 2013; Dias-ferreira et al. 2009). However, to date, there has been little study of when these action selection strategies typically develop, and how their relative recruitment changes as individuals mature from childhood to adulthood. In this study, we sought to characterize maturational changes in model-based and model-free learning using a two-stage reinforcement-learning task that can dissociate the extent to which each learning system informs an individual’s choices. Model-free and model-based

algorithms make different predictions of behavior in the task. A model-free chooser relies only on reward information. In contrast, a model-based chooser also takes into account the probabilistic state transition structure in the task. We examined the extent to which children, adolescents, and adults exhibited the behavioral signatures of these two forms of learning. Whereas a model free strategy was evident in the choices across all age groups, model-based influence on choice only emerged in adolescence and continued to increase with age. Together, these results suggest that the ability to recruit model-based evaluative processes to inform one's choices increases with age, highlighting a critical component of the gradual emergence of goal-directed behavior.

METHODS

Participants

A total of 82 individuals (30 children, 28 adolescents, and 22 adults) participated in the task. Using a behavioral exclusion criteria discussed below, the final sample included 59 subjects: 20 children (11 females; M = 9.80; SD = 1.54, 8-12 years), 20 adolescents (12 females; M = 15.35; SD = 1.39; 13-17 years), and 19 adults (11 females; M = 21.63; SD=2.03; 18-25 years). All participants were recruited from the New York metropolitan area and provided written informed consent according to the procedures of the Weill Cornell Medical College Institutional Review Board. All participants were compensated $30 regardless of their performance.

Spaceship-task

This task was adapted from the Daw et al. (2011) two-step task to use child-friendly stimuli and timing, and a child-friendly narrative. Participants

were instructed they would be traveling through space to collect space-treasure (Figure 2.1A). Participants chose between two spaceships in the first stage, which took them to one of two planets, and then chose between two aliens (a different pair on each planet). Each spaceship traveled more frequently (70% versus 30%) to one planet than the other throughout the entire experiment. Choosing an alien revealed whether the participant was rewarded (a picture of 'space treasure') or not (an empty circle of the same size). Each alien was rewarded according to a slowly drifting probability (bounded between 0.2 and 0.8). This required participants to explore throughout the experiment. Participants had 3 seconds to choose the left or right stimuli (press 1 or 0), their choice was animated for 1 second, and they saw crossed out stimuli for 4 seconds if they failed to do so. The inter-trial interval was 1 second.

Prior to the task, participants were verbally instructed and experientially trained on the task. Participants were instructed that each alien had a mine, and that how often an alien would bring up treasure changed slowly depending on how easy it was to dig. A set of 10 forced choices for one alien with a good mine, rewarded 7/10 times, was used as an example of probabilistic feedback. 20 choices between 2 aliens, rewarded 20% and 80% of the time, were used as an example of good and bad aliens (all participants correctly identified the 'good alien'). And 20 full trials were used as an example of the probabilistic nature of transitioning to different second-stage states, and participants were told that the planet the spaceship goes to most of the time would stay the same for the whole game. Finally, participants were reminded that that the alien mines were independent, and that there were not special patterns of choices that would lead to rewards. The spaceships, planets, and

aliens used in the tutorial were different from those in the task. The full game consisted of 200 trials in four blocks separated by breaks.

Exclusion Criteria

One child and one adult were excluded for inattentiveness during the task. One adult was excluded for complete choice invariance throughout the task. A fundamental assumption in the task is that participants are trying to obtain rewards. We used two reward-sensitivity based criteria to exclude subjects whose behavior was not consistent with such a goal. First, for trials following common transitions, the proportion of trials in which participants repeated a first stage choice had to be at least 0.1 greater when rewarded than unrewarded. This restriction requires that participants appear to be pursuing reward, but is unbiased as to whether they do so via a model-free or model-based strategy. Second, when participant arrived at a second stage state in which the were rewarded on the previous trial, they were required to choose the same previously rewarded stimulus at least 55% of the time. This removed participants whose behavior would not be well explained by our analysis, such as choosing randomly or using a non-reward based strategy.

Behavioral Analysis

The group logistic regression analysis has been described previously (Daw 2011; Otto, Gershman, et al. 2013). Briefly, a generalized linear mixed-effects regression analysis of group behavior data was performed using the lme4 package for the R-statistics language. 1st-stage choice (stay/switch from previous trial) was modeled by independent predictors of previous reward (reward/no reward), previous transition (rare/common), age (z-score

transformed), and all two-way and three-way interactions as fixed-effects, as well as per-participant random adjustment to the fixed intercept ('random intercept'), and per-participant adjustment to previous outcome, transition, and outcome-by-transition interaction terms ('random slopes'). The terms of interest are for a main effect of reward (model-free term), a reward-by-transition type interaction effect (model-based term), and the reward-by-age and reward-by-transition-by-age interaction effects. The first 9 trials for every subject were removed, as were trials in which an individual failed to make a first or second stage choice (median: child=3.5, adolescent=0.5, adult=0). This approach enabled separate estimation of the evidence for intact habitual and goal-directed learning strategies across development. Additionally, this analysis was performed for each age group separately, removing the age and age-interaction terms. Response time data for the second-stage actions was analyzed similarly using a linear mixed-effects analysis with current transition and age as independent predictors. Finally, the relationship of how model-based a participant was, determined by the individual random-effects estimates of the reward-by-transition interaction, with age and response time differences was examined through correlational analyses.

Additionally, the reinforcement learning computational learning model consisted of a weighted combination of a model-free SARSA(λ) (Sutton and Barto 1998) temporal difference algorithm that incrementally updates a fixed value for the first-stage choice based on reward history and a model-based "tree-search" reinforcement learning algorithm (explicit computation of Bellman's equation (Sutton and Barto 1998)), which represents all possible choices and associated outcomes from which to choose. In addition to the w

parameter, the algorithm includes parameters for learning rate (α), eligibility trace (λ).

The hybrid model consists of model-free and model-based algorithms, both of which estimate a state-action value functions Q(s,a) that map each state-action pair to its expected future value. On a given trial, the first and second stages, actions, and rewards are denoted as $s_{1,t}$, $s_{2,t}$, $a_{1,t}$, $a_{2,t}$, $r_{1,t}$ (always zero), and $r_{2,t}$. The model-free temporal difference algorithm updates the state action values according to the following formula:

$$Q_{TD}(s_{i,t}, a_{i,t}) = Q_{TD}(s_{i,t-1}, a_{i,t-1}) + \alpha\delta_{i,t}$$

where

$$\delta_{i,t} = r_{i,t} + Q_{TD}(s_{i+1,t}, a_{i+1,t}) - Q_{TD}(s_{i,t-1}, a_{i,t-1})$$

As such, δ is the reward-prediction error (RPE). At stage one, reward $r = 0$ and the RPE is driven by the estimate of the second stage action value. At the second stage, $r = 0$ or 1, and there is no third stage that drives additional value. The eligibility trace λ, which is only carried over across stages for one tiral, is used to further update the first-stage action by the second-stage RPE according to:

$$Q_{TD}(s_{1,t}, a_{1,t}) = Q_{TD}(s_{1,t}, a_{1,t}) + \alpha\lambda\delta_{2,t}$$

For the model-based algorithm, the first state learning function differed from the model-free algorithm in that it took into account the 70/30-transition probability structure and computed cumulative state-action values from all possible outcomes. The second-stage action value estimate is the same across the model-free and model-based algorithms. As such, the model-based algorithm updates the first stage action values according to the following formula:

$$Q_{MB}(s_{1,t}, a_{j,t}) = P(s_B|s_A, a_j) \max_{a\in\{a_A,a_B\}} Q_{TD}(s_B, a) + P(s_C|s_A, a_j) \max_{a\in\{a_A,a_B\}} Q_{TD}(s_c, a)$$

These action value estimates are combined by a simple linear combination using a weighting factor (w), ranging from 0 (no model-based component) to 1 (only model-based).

$$Q_{net}(s_A, a_j) = wQ_{MB}(s_A, a_j) + (1-w)Q_{TD}(s_A, a_j)$$

Finally, we estimate the probability of a choice using the softmax formula with Q_{net}:

$$P(a_{i,t} = a|s_{i,t}) = \frac{\exp(\beta[Q_{net}(s_{i,t}, a) + p \cdot \text{rep}(a)])}{\exp(\beta[Q_{net}(s_{i,t}, a_A) + p \cdot \text{rep}(a_A)]) + \exp(\beta[Q_{net}(s_{i,t}, a_A) + p \cdot \text{rep}(a_A)])}$$

Here, inverse temperature (β) determines how deterministic the participant is. The rep(a) function is set to 1 if the participant repeats the first stage choice, and 0 if not repeated or for a second stage choice. Perseveration (*p*) captures how likely the participant is to repeat the previous choice (p>0) or switch away (p<0). This model then fit 5 separate parameters, $\alpha, \lambda, w, \beta,$ and p. These models were fit by maximum a posteriori estimation (Daw 2011; den Ouden et al. 2013), taking the beta distributions beta(1.1, 1.1) as a prior for $\alpha, \lambda,$ and w parameters and gamma(3,1) for β as a priors. The parameter estimates were chosen to ensure smooth parameter boundaries and be uninformative over the previously observed ranges of parameter estimates in adults on this task (Daw et al. 2011). This nests with pure model-free (w=0, with or without fixed λ) and pure model-based (w=1, λ arbitrary), which were compared by AIC (Akaike 1974). As the relative weighting of model-based over model-free learning, w, is not normally distributed (ranging

from 0 to 1), we performed age correlation analyses on the log-transformed value of w.

Results

We modified an established two-stage reinforcement-learning task (Daw et al. 2011; Otto, Gershman, et al. 2013) to be maximally engaging for a broad developmental cohort. Participants performed 200 trials in the task, each of which consisted of two choices. In the first-stage state, participants made a choice between two stimuli (spaceships), each of which made probabilistic transitions to each of two second-stage states (the red or purple planets). For example, choosing the blue spaceship led to the red planet with 70% probability (the common transition), and led to the purple planet with 30% probability (the rare transition). On each planet, participants made a second choice between two stimuli (aliens), each of which were associated with an independent probability of reward (space treasure), which changed slowly according to a Gaussian random walk across trials (Figure 2.1A). These shifting reward probabilities encouraged participants to explore different choices throughout the task in order to maximize rewards.

Critically, this task structure enables the dissociation of model-free and model-based learning strategies in the task. Whereas a model-based chooser uses a cognitive model of the transitions and outcomes in the task to select actions, a model-free chooser simply repeats previously rewarded actions (Figure 2.1B). Thus, how a previous trial influences the first-stage choice on the next trial depends on the participant's learning strategy. For example, consider a trial in which a participant chooses the blue spaceship, makes a rare transition to the purple planet, chooses an alien, and is rewarded.

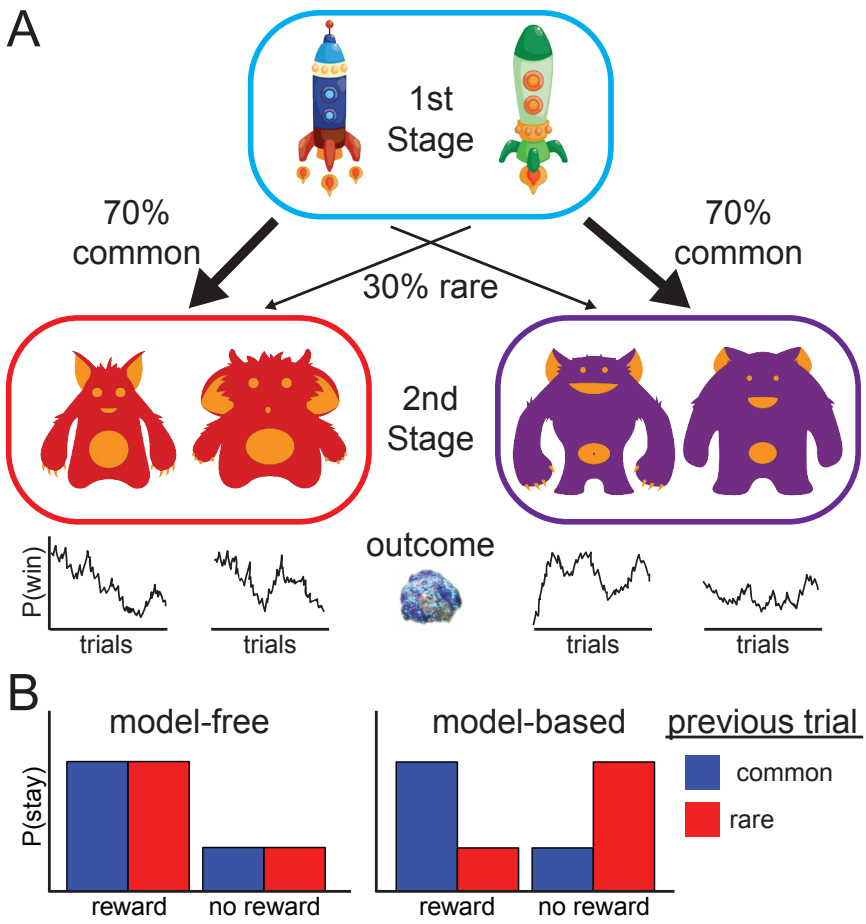

Figure 2.1 Two-stage spaceship task design and idealized behavior. (A) For each trial, participants made a choice between two stimuli (spaceships) in the first-stage state and made probabilistic transitions to each of two second-stage states (the red or purple planet). This transition probability was fixed such that each first-stage stimulus led preferentially to a particular second-stage state (70% versus 30%). Participants similarly chose between two stimuli (aliens) in either second-stage state and rewarded (space treasure) according to a slowly drifting probability (bounded between 0.2 and 0.8). B) Model-free and model-based strategies make different predictions about first-stage choices for trials following a rare transition. Whereas a model-free chooser (left) is more likely to stay with their first-stage choice when rewarded and switch when unrewarded, a model-based chooser (right) is more likely to switch when rewarded and stay when unrewarded.

A model-free learner is likely to repeat the previous first-stage choice (blue spaceship) regardless of the transition that led to the reward. In contrast, a

model-based chooser—taking into account the state transition structure—is likely to switch to the green spaceship in order to increase the likelihood of returning to that rewarded state. Below we assess the recruitment of these two learning strategies at different stages of development by examining trial-by-trial switching or staying as a function of the transition type (common or rare) and outcome (reward or no reward) on the previous trial.

First, examining the behavior visually provides an initial indication of whether the participants are following a more model-free or model-based pattern of behavior (Figure 2.2). The proportion of trials for which a participant stayed with their previous first-stage choice serves as an overall measure of their reward seeking behavior. A stay, or repetition, of a first-stage choice suggests a belief that that spaceship is more likely to lead to a highly rewarded alien. Children were more likely to choose the same spaceship following a reward, regardless of the previous transition, exhibiting a typical model-free pattern of behavior. In contrast, adolescents and adults showed a more complex pattern of behavior. After a rare transition, they were less likely to stay following a reward and more likely to switch following no reward relative to when the previous trial had a common transition. There was still an overall greater tendency to stay following rewarded trials than unrewarded trials suggesting a mixture of both model-free and model-based strategies were being used in the adolescents and adults.

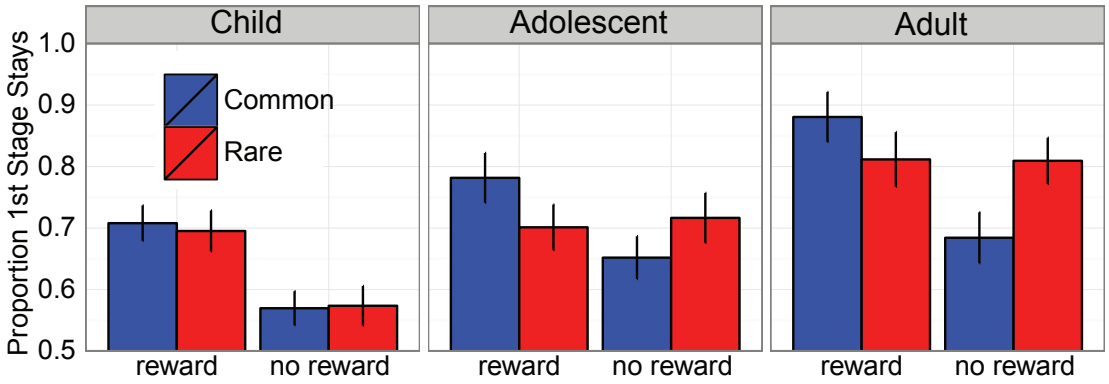

Figure 2.2 Proportion of 1st-stage choices that were repeated by trial type. Children were more likely to repeat their first stage choice after a reward, regardless of whether the previous transition was common or rare, reflecting a model-based strategy. Adolescents and adults exhibited a mix of both model-free and model-based strategies, taking into account both the previous reward and transition.

Generalized mixed-effects regression

To analyze the age-related differences in the recruitment of these learning systems, we used logistic regression to test for the influence of a given trial's transition type (common or rare) and reward (reward or no reward) on the subsequent first-stage choice (stay or switch) (Figure 2.1B). A purely model-free strategy disregards transition information and relies solely on previous reward to determine whether to repeat the first-stage action or not (a main effect of reward). In contrast, a model-based strategy incorporates both transition and reward information (a reward-by-transition interaction effect). In order to characterize the developmental changes in these learning strategies, we included age and age-interactions in this analysis (Table 2.1). The behavioral signatures of both model-free and model-based learning were evident in the full cohort of participants, who showed both a significant main effect of reward ($p<1e-8$) and a reward-by-transition interaction ($p<1e-6$). However, only the model-based learning signature exhibited a significant

increase with age (reward-by-transition-by-age interaction p<0.001; Figure 2.3). Notably, the regression line has its intercept near the childhood age range, highlighting the lack of model-based behavior in the child age group.

Table 2.1 Result of the mixed-effects logistic regression for the full sample. Examining the effect that age, previous reward, and previous transition type have on first-stage choice repetition.

Predictor	Estimate	X^2 (df=1)	p-value
Intercept	1.20	64.22	<1e-14
Reward	0.37	34.28	<1e-8
Transition	0.05	2.64	0.13
Age	0.43	13.55	0.0003
Reward by Transition	0.26	28.47	<1e-6
Reward by Age	0.10	3.14	0.124
Transition by Age	0.01	0.37	0.77
Reward by Transition by Age	0.17	14.37	0.0004

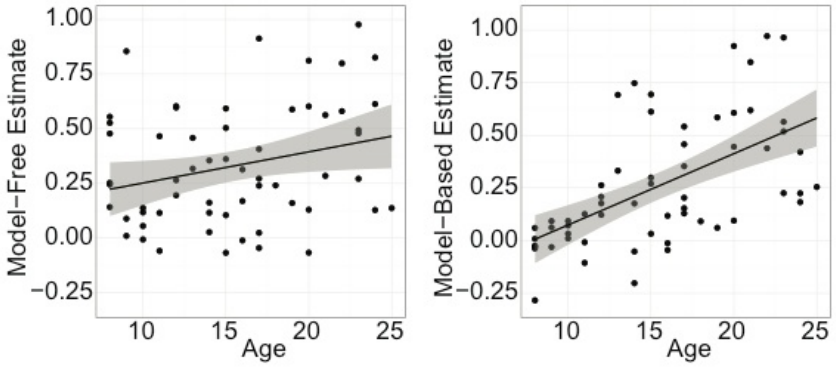

Figure 2.3 Individual estimates of model free and model-based behavior as a function of age. The model-free signature of behavior (main effect of reward) was above zero and did not increase significantly with age. The model-based signature of behavior (reward-by-transition interaction) was not present in children but increased significantly with age.

These results replicate previous findings that adults exhibit a mixture of both model-free and model-based strategies (Daw et al. 2011; Otto, Gershman, et al. 2013), and extends this result to adolescents. Additionally, while

perseveration—the tendency to repeat actions regardless of outcome or transition—was significantly positive ($p < 1e\text{-}14$), this tendency increased with age (main effect of age $p<0.001$). Collectively, these results suggest that whereas the behavioral signature of model-free learning is evident from childhood through adulthood, the recruitment of model-based learning increases as one develops from childhood into adulthood.

This regression analysis was repeated within each age group, removing the age terms, to determine whether this mixture of model-free and model-based strategies was present in all age groups (Table 2.2). Children ($p<0.001$), adolescents ($p<0.01$), and adults ($p<1e\text{-}5$) all show a main effect of reward, whereas only adolescents ($p<0.002$) and adults ($p<0.0001$), but not children ($p=0.65$), show a reward-by-transition interaction effect. Thus the mixture of strategies is evident only in adolescents and adults, while the choices of children tend to rely principally on the model-free strategy.

Table 2.2 Result of the mixed-effects logistic regression for each age group. Examining the effect of previous reward and transition type on first-stage choice repetition.

Group	Predictor	Estimate	X^2 (df=1)	p-value
Child	Intercept	0.61	20.45	<1e-5
N=20	Reward	0.30	11.79	0.0006
	Transition	0.01	0.07	0.79
	Reward by Transition	0.02	0.21	0.65
Adolescent	Intercept	1.19	17.28	<0.0001
N=20	Reward	0.22	7.39	0.0066
	Transition	0.09	2.34	0.13
	Reward by Transition	0.35	10.00	0.0016
Adult	Intercept	1.85	26.32	<1e-6
N=19	Reward	0.56	20.43	<1e-5
	Transition	0.07	5.06	0.024
	Reward by Transition	0.49	15.64	<0.0001

To determine whether these age group differences might reflect a lack of understanding of the transition structure of the task, we asked participants to recall the state transition structure at the end of the task, “Which spaceship traveled to the red planet most of the time.” (2 children, 3 adolescents, and 2 adults were not asked about their explicit recognition of the transition structure). There was no difference by age group (X-squared = 0.6701, df = 2, p-value = 0.7153, 14 children, 14 adolescents, and 15 adults answered correctly), suggesting that participants from each age group had explicit awareness of the transition structure. Next, we examined participant response times (RT) to the second stage choice (choosing an alien) as a function of transition type. If participants were not aware of the transition structure, we would expect no differences in RT at the second stage as a function of transition type (common or rare). A linear mixed-effects model revealed that participants were slower after rare transitions than common transition (main effect of transition type: 42 ms, X^2=38.6, p<1e-7), there was a general quickening of responses with age (main effect of age: -78 ms, X^2=9.2, p=0.0036), but there was no transition type by age interaction (p=0.40) (Figure 2.4). This analysis provides evidence that children are processing the transition structure of the task, even though this information does not appear to directly inform their choices. Finally, we examined whether these RT differences were associated with model-based behavior (Figure 2.5), as has recently been shown in adults (Deserno et al. 2015). Within adults (r=0.578, p=0.0095) and adolescents (r=0.51, p=0.022), individuals with longer response times after rare relative to common transitions were more model-based. In other words, the degree to which these participants incorporated transition information into their first stage choices was associated with how

much they slowed their responses for rare versus common trials. This relationship was not seen in children (r=0.084, p=0.72), further suggesting that children did not integrate this information into their choices.

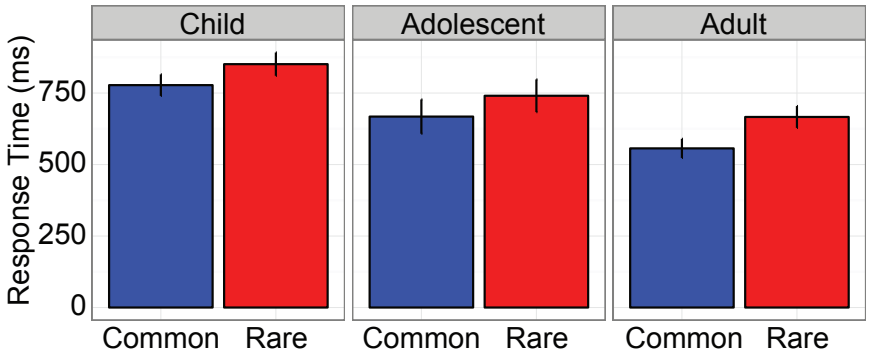

Figure 2.4 Second stage response time as a function of transition type. Children, adolescents, and adults all responded more slowly following a rare transition than after a common transition.

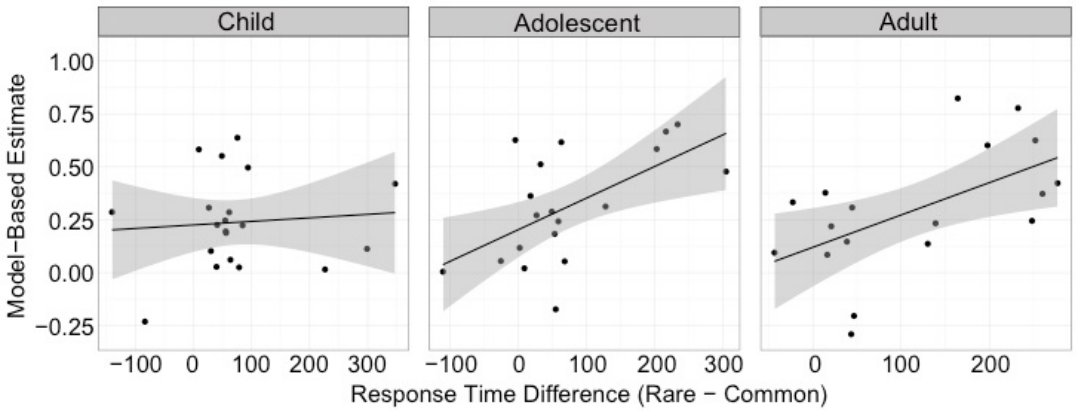

Figure 2.5 The relationship of model-based estimates to the difference in response times between rare and common transitions, by age group. In adults and adolescents, those participants that have longer response latencies for rare relative to common transitions are those that show more model-based behavior. This relationship is not present in the child group, suggesting that any awareness in transition differences is not incorporated in their decisions.

Hybrid model-free and model-based algorithm

Motivated by these results, we used a hybrid reinforcement-learning algorithm that combines model-free and model-based learning systems to determine the degree to which participants deploy one system over the other. The model-free system uses a simple temporal-difference reinforcement algorithm to cache the value of various actions, whereas the model-based system uses the transition structure to form evaluations prospectively. This approach provides a few advantages over the group regression analysis. We considered these models separately, and combined using a weight parameter, w, which captures the degree to which an individual favors a model-based or model-free strategy. This hybrid model fit significantly better than chance, at $p<0.05$ by likelihood ratio testing for 16/20 children, 19/20 adolescents, and all 19 adults. We estimated the parameters individually for each participant by maximum a posteriori likelihood (Table 2.3, Figure 2.6). By non-parametric Kruskal-Wallis testing, the only reinforcement-learning parameter that differed across age groups was the inverse temperature, β (X2 = 10.1, p = 0.0064), suggesting that the different age groups were differently deterministic in their choices. (learning rate, α, p = 0.19; eligibility trace, λ, p = 0.90; perseveration, p, p = 0.14; hybrid weight, w, p = 0.12). As we had a special interest in the hybrid parameter, w, we did a post-hoc exploratory analysis. As suggested by the data, we examined the correlation between age-squared and $\log(w)$ (r = -0.40, p = 0.0015; this remains significant with non-transformed w, p = 0.0057; Figure 2.7).

Table 2.3 Parameter estimates, mean and quartiles across subject in each age group

Age Group	Quartile	β	α	λ	p	w	LL	p-R2
Children	25%	1.549	0.177	0.431	0.175	0.103	-265.341	0.024
Children	50%	1.946	0.395	0.701	0.222	0.176	-253.862	0.070
Children	75%	2.482	0.703	0.861	0.338	0.625	-233.567	0.144
Adolescents	25%	1.591	0.318	0.286	0.225	0.322	-256.231	0.072
Adolescents	50%	1.999	0.617	0.687	0.372	0.542	-238.928	0.132
Adolescents	75%	3.834	0.799	0.943	0.498	0.683	-166.069	0.356
Adults	25%	2.186	0.427	0.357	0.225	0.064	-223.214	0.191
Adults	50%	3.837	0.601	0.720	0.372	0.434	-172.925	0.376
Adults	75%	4.943	0.703	0.909	0.498	0.648	-140.461	0.493
All	25%	1.642	0.299	0.327	0.176	0.120	-256.457	0.064
All	50%	2.229	0.555	0.717	0.317	0.434	-233.643	0.157
All	75%	3.681	0.724	0.922	0.455	0.648	-171.059	0.361

β inverse temperature; α learning rate; λ eligibility trace; p perseveration; w model-based weight; LL log-likelihood; p-R2 pseudo-R^2

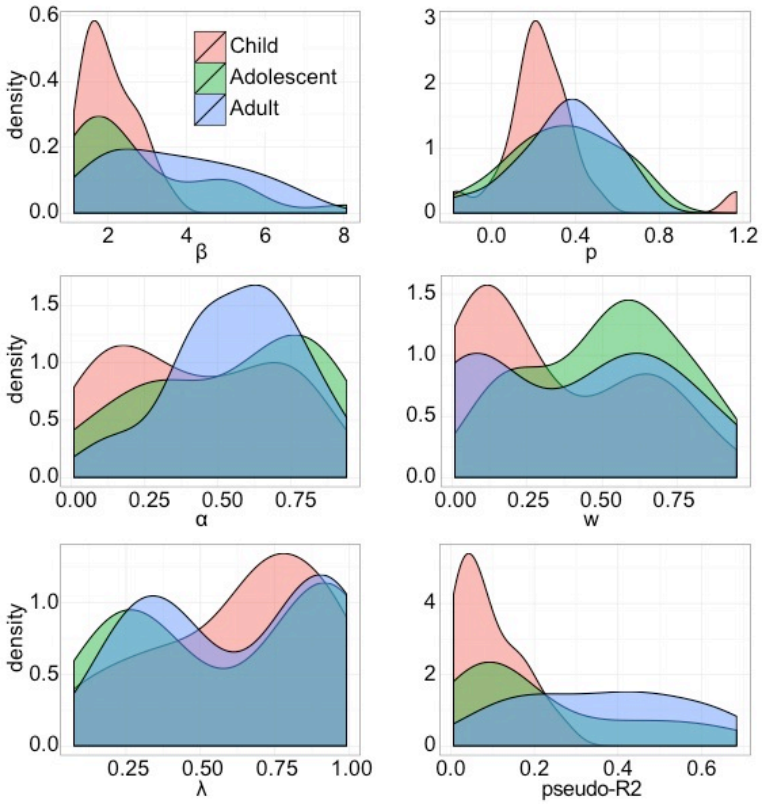

Figure 2.6 Density plots of parameter estimates. Inverse temperature, β; learning rate, α; eligibility trace, λ; perseveration, p; hybrid weight, w.

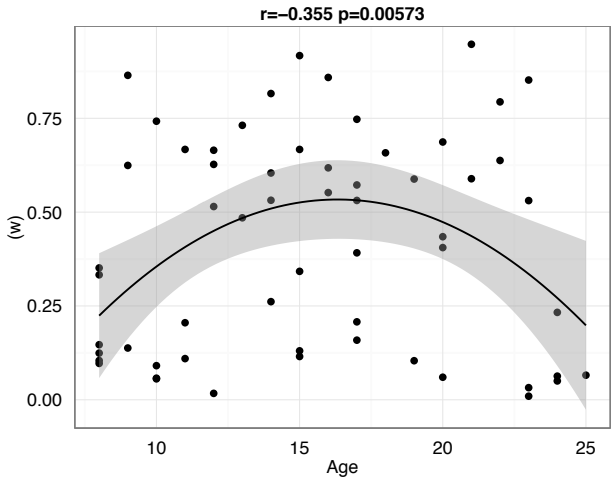

Figure 2.7 The relationship between age and the log-transformed model-based parameter. The adolescent group showed the most prominent mixture of strategies whereas the child and adult group have a higher proportion of members that show a purely model-free strategy.

Next, we performed classical model comparisons to determine whether the parameters chosen in the hybrid model were justified by the data, relative to four simpler models and the null model (Table 2.4). We tested the special cases where the eligibility trace, λ, was set to 0 or 1, and the full model-free, w = 0, and model-based, $w = 0$. While the hybrid model performed well for most participants, often a simpler model was sufficient for a handful of participants. 4 children, 13 adolescents, and 9 adults were best explained by a version of the hybrid model (including the λ = 0/1 hybrid models). 8 children, 2 adolescents, and 5 adults were best fit by the model-free temporal difference model. 5 children, 4 adolescents, and 5 adults were best fit by the model-based model. This shows that at the population level there definitely does seem to be a mix of strategies, but with quite a bit of individual and developmental variability.

Table 2.4 Model Comparisons between Hybrid Model and Nested Models

Age Group	Model	-LL	Number Favoring Hybrid	Aggregate LRT Favoring Hybrid	Median AIC	Total Number Favored
Children N=20	Hybrid	4918	-	-	517.7	0
	$\lambda = 0$	4975	10	$X^2_{20} = 115$ p < 5e-15	520.3	1
	$\lambda = 1$	4926	0	$X^2_{20} = 16$ p = 0.71	515.9	3
	$w = 0$	4929	1	$X^2_{20} = 23$ p = 0.31	517.2	8
	$w = 1$	4987	7	$X^2_{40} = 139$ p < 1e-12	519.7	5
	Null	5402	16	$X^2_{100} = 969$ p < 1e-100	544.8	3
Adolescents N=20	Hybrid	4264	-	-	487.9	0
	$\lambda = 0$	4308	8	$X^2_{20} = 87$ p < 5e-10	489.5	2
	$\lambda = 1$	4271	2	$X^2_{20} = 13$ p = 0.86	488.2	11
	$w = 0$	4307	7	$X^2_{20} = 85$ p < 1e-9	488.1	2
	$w = 1$	4320	8	$X^2_{40} = 110$ p < 5e-8	489.4	4
	Null	5476	19	$X^2_{100} = 2423$ p < 1e-100	553.1	1
Adults N=19	Hybrid	3355	-	-	355.9	1
	$\lambda = 0$	3434	9	$X^2_{19} = 158$ p < 5e-23	354.8	2
	$\lambda = 1$	3367	3	$X^2_{19} = 24$ p = 0.19	354.3	6
	$w = 0$	3430	9	$X^2_{19} = 151$ p < 5e-22	390.2	5
	$w = 1$	3456	10	$X^2_{38} = 202$ p < 5e-24	352.9	5
	Null	5257	19	$X^2_{95} = 3804$ p < 1e-100	554.5	0

DISCUSSION

In this study, we examined developmental changes in the recruitment of model-free and model-based evaluation systems in a sequential decision-making task. We found that children, adolescents, and adults all exhibited the behavioral signature of model-free learning, showing a tendency to repeat an initial choice that eventually led to reward. In contrast, there was no evidence

of model-based learning in children's choices. Although participants of all ages were able to distinguish common from rare transitions, the model-based ability to recruit this task transition knowledge to inform one's choices only emerged in adolescence and continued to strengthen into adulthood.

Model-based behavior stems from the ability to form a cognitive model of the environment, and use it to inform goal-directed choices. This capacity involves multiple component processes, including the recruitment of executive functions, such as working memory and cognitive control (Daw and Shohamy 2008). Introducing concurrent working memory load during the two-stage task renders choices more model-free (Otto, Gershman, et al. 2013), providing evidence of the role of working memory in model-based behavior. Similarly, individuals who exhibit greater cognitive control on independent tasks are also more model-based (Otto et al. 2015). Multiple components of executive function including working memory (Diamond, Kirkham, and Amso 2002; Olesen, Westerberg, and Klingberg 2004), cognitive control (Munakata, Snyder, and Chatham 2012), effective use of abstract rules or instruction (Bunge and Zelazo 2006; Decker et al. 2015), and response inhibition (Diamond 2006) exhibit a protracted maturational trajectory, and are proposed to underlie an increased capacity for goal-directed behavior with age (Munakata, Snyder, and Chatham 2012). The gradual developmental emergence of many of these cognitive processes likely contributes to our observed increase in model-based choice with age.

A striking result in our study is that, although children's first stage actions were model-free, there was both implicit and explicit evidence that they had formed a cognitive model of the task structure. Children, like adolescents and adults, were able to verbally report the common transition after completing

the task. Moreover, analysis of response time data showed that participants of all ages exhibited behavioral slowing following rare transitions, reflecting implicit knowledge of the transition structure. A previous study in adults found that that the magnitude of this behavioral slowing effect predicted greater model-based choice (Deserno et al. 2015). However, in our study, only adolescents and adults showed this correlation. Thus, although children exhibit knowledge of the transition structure, they do not recruit this knowledge prospectively in their subsequent choices at the first stage. This result accords with a dissociation commonly observed in many studies of cognitive development between the age at which knowledge of a task is present, and when it is behaviorally evident in task performance (Zelazo, Frye, and Rapus 1996). The emergent ability to recruit transition knowledge in one's choice strategies may reflect a developmental shift from reactive engagement of cognitive control following surprising transitions, to proactive cognitive control engaged at the first stage choice (Munakata, Snyder, and Chatham 2012; Botvinick et al. 2001).

Model-based learning algorithms reproduce several defining features of goal directed behavior (Daw, Niv, and Dayan 2005; Sutton and Barto 1998). Goal-directed behavior is distinguished from habitual behavior by two key properties: sensitivity to changes in the contingency between an action and its outcome, and sensitivity to changes in the value of the outcome itself. Perseveration in either condition reveals an action to be under habitual control (Dickinson 1985; Balleine and O'Doherty 2009; Yin and Knowlton 2006). Model-free learners do not recruit representations of state transitions or specific outcomes that are necessary to inform goal-directed behavior. In several canonical assays of cognitive development, younger children

perseverate in performing previously rewarded actions following a change in contingency; however, this behavior disappears at a slightly later developmental stage. For example, in Piaget's A not B task, after repeating a reinforced action several times (e.g. reach left to obtain a hidden toy), babies aged 10 months or younger are impaired in a critical test trial where they must perform a new action (reach right), but at 12 months old this is no longer seen (Piaget 1954). In other tasks of increased complexity, this developmental emergence of sensitivity to changes in contingency can be observed at later ages (Zelazo, Frye, and Rapus 1996; Gerstadt, Hong, and Diamond 1994; Kirkham, Cruess, and Diamond 2003). Similarly, sensitivity to outcome devaluation has been reported to emerge across development (Klossek, Russell, and Dickinson 2008). A parsimonious account for these developmental changes in behavior may be a transition from a model-free to a model-based process of evaluating actions. This transition may be a general characteristic of cognitive development that occurs at different ages for tasks of varying complexity, as the capacity to form and recruit a model of the task improves.

Studies in adults suggest that model-free and model-based learning recruit distinct underlying neural processes (Daw, Niv, and Dayan 2005). The model free algorithm relies on error-driven feedback to compute actions values. Dopaminergic input to the striatum, carrying a signal that resembles a computational reward prediction error, is thought to support such a model-free learning process (Schultz, Dayan, and Montague 1997; Pagnoni et al. 2002; McClure, Berns, and Montague 2003; O'Doherty et al. 2003). In addition to the central role of the ventral striatum in estimating model-free values, the dorsolateral striatum is critical for model-free action selection (Yin, Knowlton,

and Balleine 2004; Tricomi, Balleine, and O'Doherty 2009). Developmentally, the striatal prediction error signals thought to underpin model-free learning appear to be relatively mature from childhood onwards (Galvan et al. 2006; J. R. Cohen et al. 2010; van den Bos et al. 2012). These neural data are consistent with our present result that model-free choice behavior was evident in children, adolescents, and adults.

Model-based learning also engages striatal prediction error signals (Daw et al. 2011), but these signals are thought to integrate information about states and outcomes that stem from a more extensive network of brain regions. Both the hippocampus and the prefrontal cortex, regions with strong connectivity to the striatum (Pennartz et al. 2011), are proposed to play a role in the representation and search of a cognitive model of the task (Hassabis et al. 2007; Pfeiffer and Foster 2013; Wilson et al. 2014). The hippocampus supports the learning of sequential relationships that constitute state transition information (Turk-Browne et al. 2010; DuBrow and Davachi 2014) and orbital and medial prefrontal regions are implicated in learning associations between actions and their specific outcomes (Corbit and Balleine 2003; Valentin, Dickinson, and O'Doherty 2007; Schoenbaum, Saddoris, and Stalnaker 2007). Model-based learning also depends on the dorsolateral prefrontal cortex (Smittenaar et al. 2013), which may reflect the engagement of working memory and cognitive control processes (Miller and Cohen 2001). Additionally, the dorsomedial striatum plays a central role in the selection and performance of model-based actions (Yin et al. 2005; Yin, Ostlund, and Balleine 2008; Bornstein and Daw 2011; McNamee et al. 2015). While the present behavioral study cannot speak to underlying neural substrates, this literature suggests that the developmental emergence of model-based

learning may reflect the burgeoning integration of a prefrontal-hippocampal-striatal circuit that recruits learned information about states and outcomes to take goal-directed action (Shohamy and Turk-Browne 2013). Corticostriatal connectivity exhibits a protracted maturational trajectory from childhood through adulthood, paralleled by increased efficacy of cognitive functions that depend on this circuitry (Liston et al. 2006; Imperati et al. 2011). Although studies suggest early maturation of hippocampal dependent learning processes (Amso and Davidow 2012), developmental changes in hippocampal-striatal connectivity have not been well-characterized to date. Future studies examining the development of these circuits and their relationship to behavior may elucidate the neurocircuitry underlying the increase in model-based learning with age.

Dopamine appears to play a key role in modulating the recruitment of model-based learning (Durstewitz and Seamans 2008; Cools 2011), potentially through its influence on the prefrontal cortex as well as the hippocampus (Arnsten 2009; Shohamy and Adcock 2010). Manipulations or measures that predict increased or decreased dopamine signaling have been shown to yield a corresponding influence on model based choice (Wunderlich, Smittenaar, and Dolan 2012; S. De Wit et al. 2012; Deserno et al. 2015). Although the mechanisms through which dopamine modulates model-based choice are not well understood, there is evidence that it may stem from its gating of prefrontal and hippocampal information to the striatum, enabling state and other goal-relevant information to inform the learning signal (Goto and Grace 2005).The dopaminergic system undergoes marked changes across the lifespan (Wahlstrom et al. 2010), which have been proposed to

contribute to the emergence of goal-directed behavior in adolescence (Naneix et al. 2012), and its decline in aging (Eppinger et al. 2013; Worthy et al. 2014).

The balance between model-based and model-free learning is proposed to have important implications for real world decision-making (Lucantonio, Caprioli, and Schoenbaum 2014). Impulsive decision-making is often defined as a propensity to prioritize immediate rewards or hedonic experiences, over a more highly valued long-term goal (Mischel, Shoda, and Rodriguez 1989). Model-free learners may be particularly susceptible to such impulsive choices, as they may be prone to repeat an action that previously yielded a small immediate reward and fail to prospectively consider the long-term opportunity cost of such actions (Kurth-Nelson, Bickel, and Redish 2012; Story et al. 2014). Furthermore, insensitivity to the devaluation of a reinforcer may lead a model-free learner to perseverate in taking previously rewarded actions when they are no longer beneficial. Experimentally, children and adolescents have been found to exhibit greater impulsivity and perseveration in their choices than adults (Klossek, Russell, and Dickinson 2008; Mischel, Shoda, and Rodriguez 1989). Such developmental differences in decision-making strategies can have important real-world repercussions. This may be particularly true during adolescence, when increased exploration and autonomy confers greater opportunity to make new choices, with less parental protection from their consequences. Indeed, the greatest perils of adolescence are those associated with poor decision-making (e.g. reckless driving, unprotected sex, suicide) (Eaton et al. 2006), underscoring the importance of understanding how decision-making changes across development. The developmental emergence of model-based learning observed in the present study represents an expansion in the repertoire of evaluative processes that

are available to inform one's actions. This increasing ability to incorporate a model of the complex and changing environment into one's evaluations may promote the maturation of goal-directed decision-making from childhood to adulthood.

LIMITATIONS

In this section, some of the limitations of using computational model to describe behavior, in general, and as it applies specifically to the results of this study, are discussed. The first recognition that must be made is that there is no model can fully capture behavior. However, that does not mean modeling is a worthless endeavor. Rather, it means that the interpretations must be measured so as not to overstate the findings. These issues, and others, are discussed below, along with some thoughts as to how these issues can be partially reconciled.

Reinforcement learning algorithm limitations

Care needs to be taken when interpreting the parameter estimates and overall goodness of fits when modeling behavior (Nassar and Gold 2013; Shteingart and Loewenstein 2014). An assumption in all the previous analyses and interpretations of the data in this chapter is that the models accurately reflect the underlying neural process determining the behavior. But, how well does the model fit the data, or is there a better model to describe the data? This is a difficult question to answer (Daw 2011). The first approach is to determine how well the model fits the data, often by reporting negative log likelihood estimates, or corrected versions such as the AIC and BIC that account for the number of parameters included. However, these scores are

somewhat difficult to understand in isolation, and are only relevant to a model on the same set of data. Across studies that have different number of trials, or within the same study if missed trials are thrown out (as is the case in our study), these numbers do not serve as an intuitive representation of how well the data fits. Reporting a pseudo-r^2 statistic can provide the same intuition as the normal r^2 statistic. This provides a number between 0 and 1 that can serve as a goodness of fit, relative to the null or random model. The formula used here is $1 - L/R$ (Camerer and Hua Ho 1999), where L is the log likelihood under the fit model, and R is the log likelihood under chance $(log(.5) \cdot N)$ and N is the number of trials. All but 3 children and 1 adolescent were fit better to some model than the random model, suggesting the parameterization of the models does a decent job of explaining participant's behavior. However, there are also concerns about using pseudo-r^2 measures as an indicator of model fit, and other approaches are likely better for assessing the appropriateness of a model (Lukacs, Burnham, and Anderson 2010).

Another way to examine whether the model fits the data well is by simulating new data with the estimated parameters and comparing performance. Before doing so, let us discuss what we might predict from the model-based, model-free, and mixed strategies. Model-based strategies allow for more information to be considered and a complex set of choices to be evaluated, and therefore are often seen as a superior behavioral strategy to model-free decision-making. There are, of course, many instances when a model-free strategy is more appropriate, which reflects the often-observed phenomena that initially goal-directed actions become habitual (Grafton, Hazeltine, and Ivry 1995; Dickinson 1994), both with positive and negative behaviors (Vanderschuren and Everitt 2004). In our task, we can use the

proportion of trials that were rewarded as a measure of performance, excluding the first 9 trials and any trial the participant missed (Range: 0.44-0.63; M: 0.53; SD: 0.04). Correlations with various study parameters are shown in Table 2.5. While a few parameters are somewhat predictive of how well participants performed, none did as well as age (Figure 2.8), suggesting that it was some combination of these parameters that is most predictive of performance.

Table 2.5 Relationship between proportion of trials that were rewarded and various study parameters.

	β	α	λ	p	w	p-R2	Int	Rew	Reward-Transition	RT	Age
r	0.21	0.11	0.02	0.03	0.15	0.25	0.18	0.17	0.23	0.22	0.26
p-value	0.11	0.43	0.90	0.83	0.27	0.052	0.18	0.20	0.077	0.095	0.05

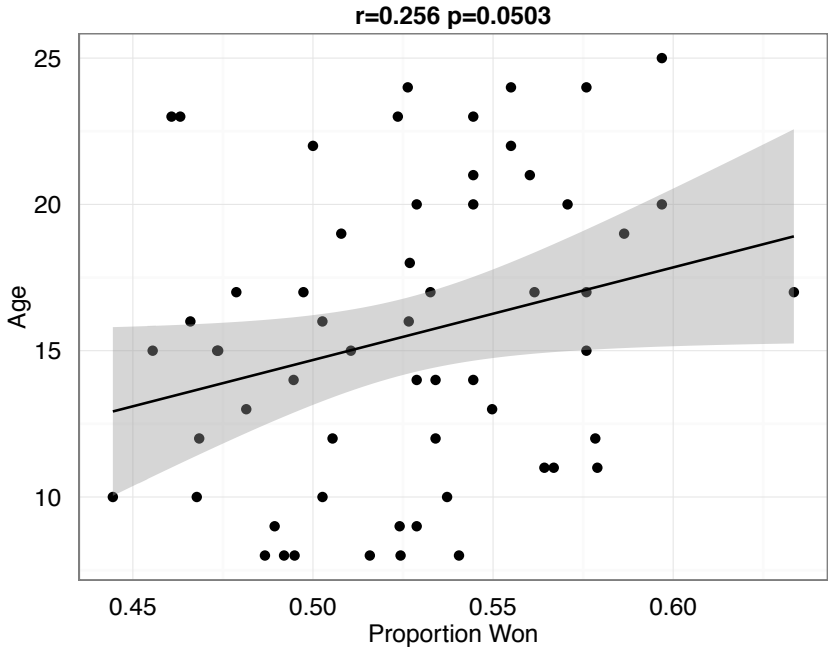

Figure 2.8 Relationship between proportion of trials rewarded and age. There is a non-significant increase in proportion of trials that were rewarded with age.

Testing for an optimal strategy using simulation

This raises the question as to what might be the optimal strategy in this task. What decision-making strategy will lead to the highest amount of reward? The design of the task, with the transition structure and probabilistic nature of obtaining rewards, likely makes it difficult for a participant to determine the optimal strategy. As both the model-based and model-free approaches are valid strategies that can lead to similar levels of total reward, our assumption is that the task reflects differences in a tendency to choose a strategy, rather than an ability to find and use the optimal strategy. To probe this question further, we ran a number of simulations of the reinforcement-learning algorithm while varying a set of parameters to determine which set would generally lead to the highest amount of reward. We were especially interested if there would be a noticeable difference in model-free, model-based, or mixed strategies. We ran 21 sets of simulations to determine if any specification led to higher rewards: one with $\beta = 0$ to signify a random chooser, one with the parameters from the participant that had the highest proportion of rewarded trials, one with the parameters from the participant that had the lowest proportion of rewarded trials, and then a systematic variation of three of the five parameters that varied the most between the age groups ($\beta, p,$ and w (Figure 2.9). Each simulation was run 100 times, and the mean and standard deviation of the proportion of rewarded trials was reported. As can be seen, the random choice had the lowest average score, and while there was a slight increase in number of rewards with higher β, the differences were minimal and only trending towards significance when taking 20 samples each from the lowest and highest scoring simulations (t=-2.02, p=0.051). An additional simulation was run with $\beta = 25$, and performance did not improve further; suggesting

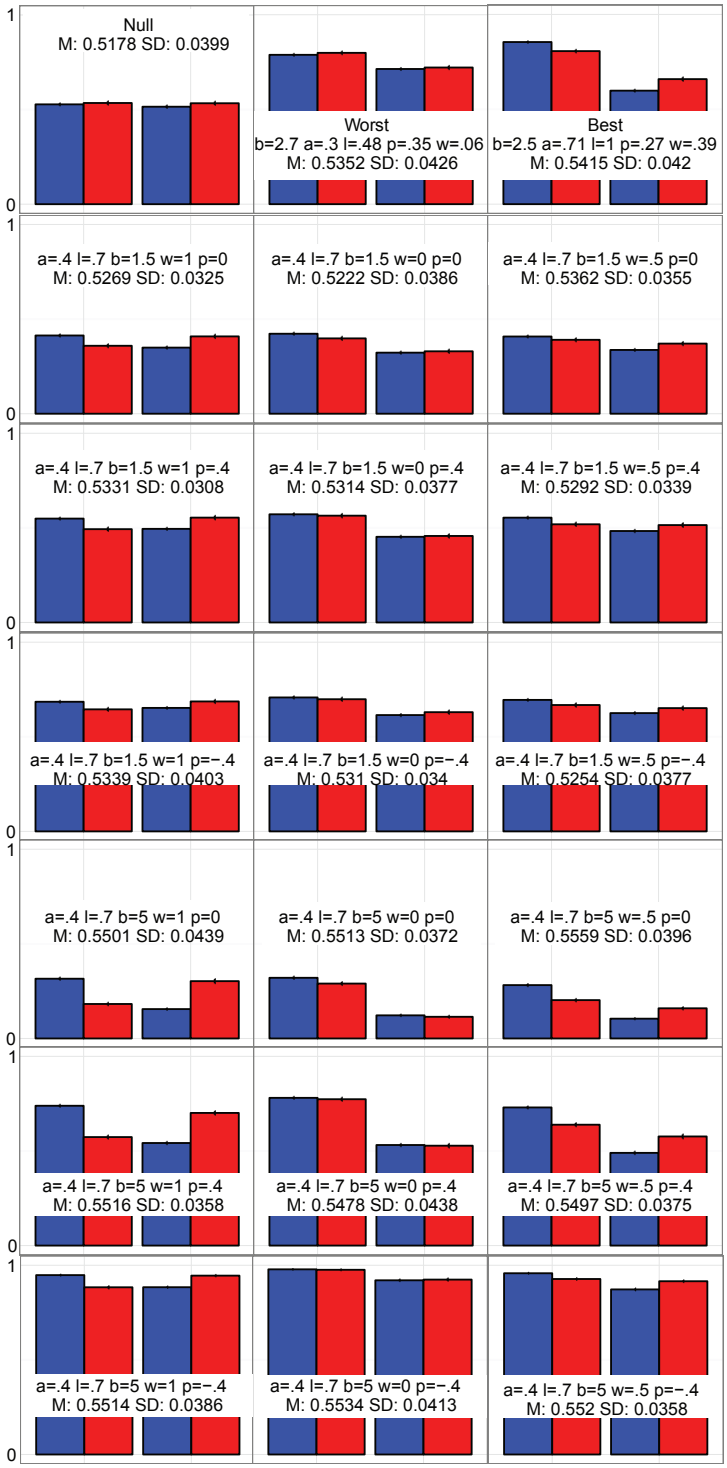

Figure 2.9 Simulated Choices (100x). The columns show a model-based, model-free, and mixed strategy from left to right. Parameterization had little influence on reward total.

performance of the model may be at asymptote. Similarly, simulations using the parameters of the highest and lowest rewarded participants did not, on average, achieve the same proportion of trials that were rewarded, and were just outside the 95% confidence interval. The worst proportion was 0.444 (lower 95% bound at that parameterization was 0.45), and the best proportion was 0.633 (upper 95% bound at that parameterization was 0.626). While it is possible that this high performing participant used a superior strategy than some form of the hybrid algorithm discussed above, the higher score may also have been due to chance. All other participant scores fell within the 95% confidence interval produced by the simulations. Choosing a model-free, model-based, or mixed strategy had little influence on performance, therefore, a tendency to use one strategy can be viewed as an individual preference, rather than uncovering an optimal strategy.

Further modeling considerations

While it is clear that the reinforcement models fit individual behavior better than the null (or random choice) model for most participants, and there is no prevailing optimal strategy, there are still other reasons why the retrieved parameters may be a poor indicator of participant choices. For example, unaccounted factors may lead to a biasing of the estimates in the model. Two additional approaches are commonly used to determine the appropriateness of using a specific model of participant behavior. The first is to analyze the data using different analytical methods and determine whether it provides a similar interpretation. Another approach is to analyze a subset of the trials for which there is reason to believe that there may be a breakdown in the assumption in the choice of parameterization (Nassar and Gold 2013). For

example, if a task has different phases or blocks of trials, then using fixed parameter estimates to account for every trial is likely inappropriate, and using adjustable or multiple parameters might be warranted. Our task design did not include subsets of trials that would be categorically different from one another, as the reward probabilities drifted slowly. As such, this second approach is not available in this study.

We did however use the first approach of using a different analytic technique. The data were presented graphically, through group logistic regression, and individual reinforcement learning model fits. The graphical representation allows for a visual comparison of the age group patterns of behavior to one another and the pure model-free and model-based strategies. While there is some agreement between these approaches, namely that a mixture of strategies is used in adolescents and adults, the regression and algorithmic approach results are somewhat difficult to reconcile in children. The regression analysis suggests that children show model-free behavior, but the reinforcement-learning approach suggests their mixture of model-free and model-based strategies does not differ from adolescents or adults. This may be due to the inappropriateness of fitting a complex model to children when a simpler one (w=0) fits similarly well. Another possibility is that children do use a model-based strategy, but that it is different from the one that has been shown to fit adult behavior. There are no suggestions from the behavior that might signal that an alternative model is being used, but this possibility should be kept in mind when interpreting the results. Many studies in adults under various conditions have suggested that a hybrid of model-free and model-based selection strategies is an appropriate depiction of adult behavior (Daw et al. 2011; Otto, Raio, et al. 2013; Otto, Gershman, et al. 2013; Skatova,

Chan, and Daw 2013; Otto et al. 2015), however that does not guarantee that it is appropriate in children. Indeed, the developmental differences on which the hypothesis that model-based behavior would emerge across development could also predict that the type of model-based strategy simply changes across development. Separately, the peak in the hybrid weight parameter seen in adolescence may tell us something about the development of the model-based strategy that the regression analysis does not. Perhaps adolescents, once capable of recruiting the model-based strategy tend to preferentially do so, whereas adults may be better at determining when a model-free approach might suffice, or default to using one or the other.

The regression analysis assumes that each decision was based on the previous trial type (reward and transition). We likened a participant that was solely sensitive to reward to be showing model-free behavior, whereas one that incorporated reward and transition information to be model-based. However this characterization is only loosely related to the model-free and model-based algorithms. Only considering the previous trial precludes the possibility that participants are basing their decisions on multiple, or a pattern of, previous trials. Methodologically, using information from the past two trials would generate 16 different trial types and three trials back would generate 64 trial types, requiring 19 and 83 independent variables and making this analysis intractable for the current study. The tutorial preceding the task explicitly informed participants that no pattern of choices would lead to more treasure and that rewards only depended on the slowly drifting probability at the second stage. There are, of course, still valid behavioral reasons why an individual might base a decision on more than the previous trial. An individual could, for example, decide to switch due to a model-free or model-based strategy, but

withhold that switch for one or more trials, which would be completely missed by our analysis approach. However, the qualitative effects of a model-free and model-based strategy are considered to be well characterized in adults when using the single previous trial (Otto, Raio, et al. 2013).

Another potential strategy participants could use is hypothesis testing. That is, a good strategy may be to test a specific spaceship and alien, then switch and test another, and then exploit the best combination. As was alluded to earlier, the task structure makes this strategy somewhat difficult due to the constant slow drift in reward probabilities and fixed transition probabilities makes clear testing difficult. This strategy may, however, be reflected in an individual's baseline tendency to repeat the previous trial, captured in the regression analysis by the intercept term and by reinforcement-learning algorithm by the perseveration parameter. While this tendency to repeat does obscure the pure model-free or model-based behavior, the trials for which participants do switch their first-stage choice reveals a pattern of behavior that can be still be defined as either reward (model-free) or reward-and-transition (model-based) sensitive.

The reinforcement-learning fits do take into account the entire history of an individual's prior actions and rewards at each subsequent decision. That is, a reward prediction error is calculated after the second stage, which updates both second-stage action values and first-stage model-free action values. We make the assumption that participants base their action selection on the model-free and model-based estimates of value that were determined by the reward history they experienced (as well as the perseveration and inverse-temperature parameters discussed in the methods). However, determining whether the parameters that go into the model reflect the true neural

computations sufficiently well is difficult, requiring a measured interpretation of those parameters.

Deterministic choices and perseveration

The two analytical approaches revealed an aspect of participant behavior that is not well described by either model-free or model-based directed action. Descriptively, model-free and model-based choice strategies do not account for the systematic tendency to repeat a previous choice regardless of the previous outcome or transition information (Lau and Glimcher 2005; Schönberg et al. 2007 - supplement). This base tendency to repeat a previous decision can be observable in the Figure 2.2 as an average shift away from the 0.5 probability of repeating a choice, in the regression analysis as the intercept term (Table 2.1), and in the reinforcement analysis as the perseverance parameter (p) (Table 2.3). Positive p values indicate a higher likelihood of repeating a choice, and negative values a higher likelihood of switching. A purely model-free or model-based learner should have no tendency to repeat or switch their choice (p = 0), and should determine their choice solely according to the values that each algorithm has calculated. This type of pure behavior is not seen, which is why the perseveration parameter and intercept terms were included in our analytical approaches. Perseveration is not involved in producing the model-free or model-based learning signal (the reward prediction error). Rather, it is part of the action selection strategy and should be included in the descriptive models of habitual and model-based decision-making. A modeling choice may be done when considering where to place the p parameter: in our formulation it is multiplied by the β term, but it

could alternatively be added. This parameterization makes slightly different assumptions about behavior. In our formulation there is the assumption that the baseline tendency to stay is also somewhat non-deterministic, whereas the other assumes it is. Our formulation requires the combined term $p * \beta$ to be used to correspond to the intercept term of the regression model (Table 2.1), and this product correlates with age as well (r=0.34, p=0.0081). Perseveration is an important aspect of decision-making that should be considered developmentally.

Relatedly, there is another aspect of the action selection stage of decision-making (as opposed to the evaluation stage) that is not part of the descriptive characterization of model-free and model-based decision-making – how deterministic an individual is when making a choice. In other words, how likely is it that an individual will choose the higher valued option given the difference that was calculated between them? In the group regression analysis this cannot be determined on a per-participant basis, but is reflected at the group level to some extent by how much variance remains in the model. The residual error, or variance, in the data that is not described by the model can then be ascribed to noisy, uncertain, or non-deterministic behavior, or because a wrong or simplified model is being utilized. Which of these factors, and to what degree, drive the unexplained variance cannot be determined. However, in the individual reinforcement learning analysis of behavior, a parameter (β, or the softmax inverse-temperature parameter) is included that estimates the extent to which a participant's behavior is deterministic (Schönberg et al. 2007), separate from the unexplained variance. Alternatively this can be considered as a parameter that captures how noisy a participant's decisions are. A β = 0 means that all choices will have the same probability of being

chosen regardless of their difference in value and the decision will be random, whereas a large positive number (approaching ∞) means that any difference in value will have a probability approaching 1 of choosing the higher valued action (Sutton and Barto 1998). Here too a developmental increase in β is seen (r=0.40, p=0.0015), suggesting that participants become more deterministic with age, which should be included in a descriptive characterization of action selection strategies.

Let us first assume that the models being used are correct (or are sufficiently accurate representation of behavior), and that children are not basing their decisions on a wholly different structure of computational models, so as to consider these developmental differences. Though the inverse-temperature and perseveration parameters are not presented in the descriptive characterization of model-free and model-based behavior, there are interesting hypotheses as to where these developmental differences may originate. The inverse-temperature parameter reflects differences in the randomness, or noise, of behavior, but it does not suggest the underlying cause of this randomness (Nassar and Gold 2013). It could simply be due to the task complexity; participants with poor attention or an inability to keep the previous trial's information in working memory would have some non-deterministic behavior. Alternatively, this random behavior could be purposeful, which has been proposed to be an important strategy for learning (Daw et al. 2006; Luksys, Gerstner, and Sandi 2009), with suggestions that lower deterministic behavior may be predictive of the risky or exploratory behavior seen in adolescence (Tymula et al. 2012). Together, the perseveration and inverse temperature parameters can reflect a tendency to exploit a known option, or explore unknown options (Daw et al. 2006;

Schönberg et al. 2007; Lau and Glimcher 2005). An individual that has a tendency to repeat a previous action and does not show much choice variability is exploiting an action for which the outcome value is known, whereas as a person that often switches or is quite variable in their behavior might be exploring the other options in the hope of finding a more beneficial outcome. The task in our study was not appropriately designed to determine whether the developmental differences seen in the perseverance and inverse temperature parameters are related to differences in exploratory or exploitative behavior, as a repetition after a common transition is different from a repetition after a rare transition. Regardless, these aspects of behavior seem to be present across development and should fall under the description of model-free or model-based behavior.

How does imaging constrain interpretation of the change in model-based behavior across development?

In the experiment described in Chapter 2, there was evidence of a gradual developmental increase in goal-directed (model-based) behavior, whereas the level of habitual (model-free) responding was fairly consistent across all age groups. There are multiple possible causes leading to this developmental difference in behavior, which were discussed in the previous limitations section. Neuroimaging, additionally, can be used to help constrain these interpretations. Recent efforts to extend animal studies of the neural substrates of habitual and goal-directed action into humans have largely suggested that homologous networks are involved (Balleine and O'Doherty 2009). Lesions in the infralimbic cortex (medial prefrontal and orbital prefrontal cortex) and nucleus accumbens (ventral striatum), dorsolateral striatum

(putamen), and dorsomedial striatum (anterior caudate) have revealed their roles in reward evaluation, habitual action selection, and goal-directed action selection respectively; animals: (Schultz, Dayan, and Montague 1997; Ostlund and Balleine 2007; Yin, Knowlton, and Balleine 2004; Yin et al. 2005), and by imaging for humans: (McClure, Berns, and Montague 2003; Schoenbaum, Saddoris, and Stalnaker 2007; Tricomi, Balleine, and O'Doherty 2009; Valentin, Dickinson, and O'Doherty 2007). Furthermore, many regions are crucial to cognitive processes involved in goal-directed behavior – including the dorsolateral prefrontal and parietal regions for working-memory (J. D. Cohen et al. 1997; Otto, Gershman, et al. 2013; McNamee et al. 2015), the dorsal anterior cingulate cortex for conflict resolution (Botvinick et al. 2001; Davidson et al. 2004), the hippocampus for mapping sequences or events to goals (Bornstein and Daw 2012; Pfeiffer and Foster 2013; Pennartz et al. 2011), and an as yet poorly defined role for the amygdala (Schoenbaum, Saddoris, and Stalnaker 2007). Collectively, this provides many regions that could be examined for the corresponding developmental differences in recruitment of model-based strategies.

Adults that performed a task similar to ours exhibited neural signals in the ventral striatum that tracked both the purported model-free and model-based reward prediction errors determined by the reinforcement learning algorithm (Daw et al. 2011). Next, I will discuss analytic approaches building from this finding that can be used to probe the developmental differences observed in behavior in our study. The clearest behavioral result was a differential recruitment of the model-based strategy with age, which is reflected in the proportion of first-stage choices that were repeated as a function of trial type (previous reward and transition). To examine this effect neurally, we will

look for brain regions that correlate with a set of time series that match first stage choices (stay or switch), previous outcome (reward or no reward), previous transition (common or rare), and their interactions, and we will specifically examine the reward-by-choice (model-free like decisions) and reward-by-transition-by-choice (model-based like decision) effects. Unlike the computational-neural approach discussed below, these neural signals will capture habitual and goal-directed behavior without assuming a specific learning model. As this first stage reflects a decision point, rather than the reward prediction error, we expect to find the largest developmental difference in the putamen and anterior caudate. We expect that all age groups would show a larger signal in the putamen for reward-stay than reward-switch, and that there would be no age dependent effect. Additionally, we would expect to find an age related effect in the anterior caudate for the reward-by-transition-by-choice interaction, specifically in the contrast of [reward-common-stay + no reward-rare-stay] – [reward-common-switch + no reward-rare-switch]. Such a result would provide evidence that while all age groups are using some model-free, habitual strategy to base their decision, there is a developmental increase in the amount of model-based, goal-directed strategy being used.

A second examination would use the reinforcement-learning parameters to look at the difference in reward-prediction-error signaling in the ventral striatum. This will be also prove to be somewhat difficult to interpret, as these error signals are highly correlated, and both have been found to correlate with striatal activity (Daw et al. 2011). Also, it is important to note the individual differences we see in our participants, such that some children show model-based behavior and some adults show model-free behavior, without much mixture. As such, if we use a mixed reward prediction error using each

individual's hybrid weight parameter, there may be no difference seen. However, if we separately examine a purely model-free, model-based, and half mixed prediction error signal, we predict that these will differ as a function of age and the individual hybrid weight parameter. Preliminary imaging analyses in 5 participants (2 children, 1 adolescent, 2 adults) shows that the model-free prediction error signal is present in the ventral striatum, suggesting that this analysis is warranted.

An additional analysis will probe the behavioral result that suggests that all participants, including children, were aware of the rare transitions, according to their response time slowing after a rare transition. This analysis will be similar to the reward prediction analysis, except that it will follow a state prediction error signal, which is only produced by a model-based evaluative process. Here, we predict that a neural signal in the hippocampus will correlate with this state prediction error, and that it will vary as a function of how much the participant slowed, on average their response to a rare trial.

Additionally, any regions found in the above analyses will be used as seeds for functional and structural connectivity, which may provide some insight into the individual and developmental differences seen in behavior. Namely, we expect to see decreased corticostriatal connectivity in those participants with the weakest recruitment of the model-based strategy. Examining hippocampal-striatal and hippocampal-cortical connections may also show differences by how model-based an individual is, reflecting the difference between implicit awareness of the transition structure and recruitment thereof. Taken together, these analytic approaches would provide strong evidence that using model-free and model-based algorithms is an

appropriate representation of habitual and goal-directed behavior across development.

There are of course concerns with computational modeling in neuroimaging, similar to those discussed in regards to behavior earlier. A specific worry is that neural signals could correlate strongly with trial-by-trial values that were determined from a computational model of behavior, and this would be used as a strong indicator that the computational model was somehow “correct”. However, it is important to remember that differing parameterization will provide highly correlated reward prediction error signals, even if there was a systematic bias in those parameters (Daw 2011; Wilson and Niv 2013). Thus, strong correlation would be suggestive that a process similar to the one parameterized by our approach is occurring in the brain, it does not indicate proof. However, the approach using first stage choices is a non-modeling approach to get at the same question. Regardless of these potential limitations, imaging will still serve to shine light on some of the outstanding questions brought up in the behavioral results.

Chapter 3: Experiential learning outweighs instruction prior to adulthood*

Throughout our lives, we face the important task of distinguishing rewarding actions from those that are best avoided. Importantly, there are multiple means by which we acquire this information. Through trial and error, we use experiential feedback to evaluate our actions. We also learn which actions are advantageous through explicit instruction from others. Here, we examined whether the influence of these two forms of learning on choice changes across development by placing instruction and experience in competition in a probabilistic learning task. Whereas inaccurate instruction markedly biased adults' estimation of a stimulus' value, children and adolescents were better able to objectively estimate stimulus values through experience. Instructional control of learning is thought to recruit prefrontal-striatal brain circuitry, which continues to mature into adulthood. Our behavioral data suggest that this protracted neurocognitive maturation may cause the motivated actions of children and adolescents to be less influenced by explicit instruction than those of adults. This absence of confirmation bias in children and adolescents represents a paradoxical developmental advantage of youth over adults in the unbiased evaluation of actions through positive and negative experience.

* Decker, J. H., Lourenco, F. S., Doll, B. B. & Hartley, C. A. Experiential reward learning outweighs instruction prior to adulthood. *Cogn. Affect. Behav. Neurosci.* (2015). doi:10.3758/s13415-014-0332-5

INTRODUCTION

Learning to obtain rewards and avoid punishment is critical for the survival of all organisms. An approach to this challenge that is employed across species is trial and error-based learning. By aggregating positive and negative feedback stemming from our previous actions, we are able to estimate how beneficial a given action might be in the future. Although such feedback-driven learning is effective, the need to learn about the consequences of our actions through direct experience can be inefficient at best, and dangerous when the potential outcomes are extremely negative.

Recruiting a sophisticated capacity for symbolic communication, humans regularly circumvent these shortcomings of experiential learning by conveying the value of an action through rules, advice, or other forms of explicit instruction. By selecting actions based on instruction, a learner is able to benefit from the prior experience and knowledge of others. The utility of transmitting information through instruction is particularly evident in the context of development. Children and adolescents receive a great deal of instructed information from parents, teachers, and public policy campaigns that seek to educate and protect them, as well as from their peers. An assumption inherent in providing such guidance is that instruction can direct children and adolescents' behavior as effectively as, or better than, their own experiential learning. To date, there have been few studies directly examining whether the efficacy of learning from instructions versus experience changes across development. However, our understanding of the cognitive processes and neural circuits implicated in such learning, and their maturational trajectories, suggests that there may be qualitative changes in the recruitment of instructed versus experiential learning across development.

Previous research has demonstrated that providing adults with instruction or advice induces a behavioral “confirmation bias”, in which recommended actions are valued more highly than those learned solely through experience, even when those recommendations turn out to be inaccurate (Biele, Rieskamp, and Gonzalez 2009; Biele et al. 2011; Doll, Hutchison, and Frank 2011; Doll et al. 2009). This instructional biasing of experiential learning is thought to stem from the influence of the prefrontal cortex, implicated in rule-guided behavior (Bunge and Zelazo 2006; Miller and Cohen 2001), on feedback-based evaluative processes in the striatum (McClure, Berns, and Montague 2003; O’Doherty et al. 2003; Pagnoni et al. 2002). This process has been modeled computationally as an instruction-consistent distortion of error-driven reinforcement learning signals (Doll et al. 2009). Developmentally, the striatal signals implicated in feedback-based reward learning appear to be relatively mature in children and adolescents (J. R. Cohen et al. 2010; Galvan et al. 2006; van den Bos et al. 2012). In contrast, connectivity between the prefrontal cortex and the striatum exhibits marked structural changes from childhood through adulthood (Liston et al. 2006; Imperati et al. 2011). Consistent with the proposal that these connectivity changes reflect fine-tuning of the information exchange between these regions (Somerville and Casey 2010), cognitive functions that depend on the integrity of frontostriatal pathways typically show continued maturation into adulthood (Liston et al. 2006; Rubia et al. 2006; Somerville and Casey 2010) . This neural model suggests that the biasing influence of explicit instruction on value-based choices might be diminished in children and adolescents, predisposing them to exhibit greater reliance on experiential learning.

In the present behavioral study, we tested this hypothesis by having children, adolescents, and adults complete a probabilistic reward-learning task consisting of a learning phase immediately followed by a test phase. In the learning phase, three pairs of stimuli were presented, and participants could learn experientially, through trial and error, which stimulus within each pair was most likely to yield reward. Importantly, participants were given inaccurate instruction that a lower-valued stimulus within one pair was likely to be rewarding. Participants could discover that this information was inaccurate through the subsequent positive or negative feedback following each choice. During the test phase, participants were presented with all possible pairings of the six stimuli from the learning phase, and attempted to select the higher valued option, receiving no feedback. By comparing their performance for instructed and uninstructed stimuli of equal value, we could determine the extent to which false instruction biased their experiential learning of the true stimulus value, providing a quantitative measure of the influence of instruction on experiential learning. Previous studies in adults have demonstrated that inaccurate instruction strongly biases experiential value learning (Doll, Hutchison, and Frank 2011; Doll et al. 2009; Staudinger and Büchel 2013). We hypothesized that children and adolescents would be less susceptible to this bias, instead relying predominantly upon their own experience to guide their choices.

METHODS

Participants

Participants were recruited through community-based events (e.g. street fairs) and flyers posted within institutions in the New York City

metropolitan area. All participants (or parents of minors) were screened by phone prior to participation to ensure that the participant had no history of diagnosed neurological or psychiatric disorders, was not taking medication, and was typically developing cognitively and behaviorally (based on self or parental report). We also ensured that all participants were not color blind. All participants provided written consent to participate and were paid for their participation. Participants were debriefed following the experiment about the misleading nature of the instructions.

A total of 87 (51 female) paid volunteers completed the study and were included in the analyses: 30 children (18 females, aged 6-12 years, M = 9.5, SD = 1.8), 31 adolescents (15 females, aged 13-17 years, M = 14.8, SD = 1.5) and 26 adults (17 females, aged 18-34 years, M = 23.0, SD = 4.3). Previous studies (Biele, Rieskamp, and Gonzalez 2009; Doll et al. 2009) have reported large instruction bias effect sizes in adults (d = 0.9 and d = 1.0–1.3, respectively). As we considered the possibility that children or adolescents might show a smaller effect, we targeted a sample size of 25 participants per group, which would enable us to detect a significant effect of at least 0.6 in each age group with 80% power (alpha of .05, two-tailed). Additional participants were recruited to ensure adequate power in the event of subject attrition, particularly in the child and adolescent groups.

Behavioral Paradigm

Participants completed an instructed probabilistic selection task (Doll et al. 2009) that was adapted for use across development, which consisted of a learning phase immediately followed by a test phase. Participants were told that their task was to feed a hungry mouse by helping him find the cheese

hidden behind one of two mouseholes. During learning, participants saw one of three stimulus pairs on each trial, referred to here as AB, CD and EF, which consisted of uniquely colored mouseholes (Figure 3.1). These stimuli were chosen to make them easily distinguishable and to be as engaging as possible for our younger participants.

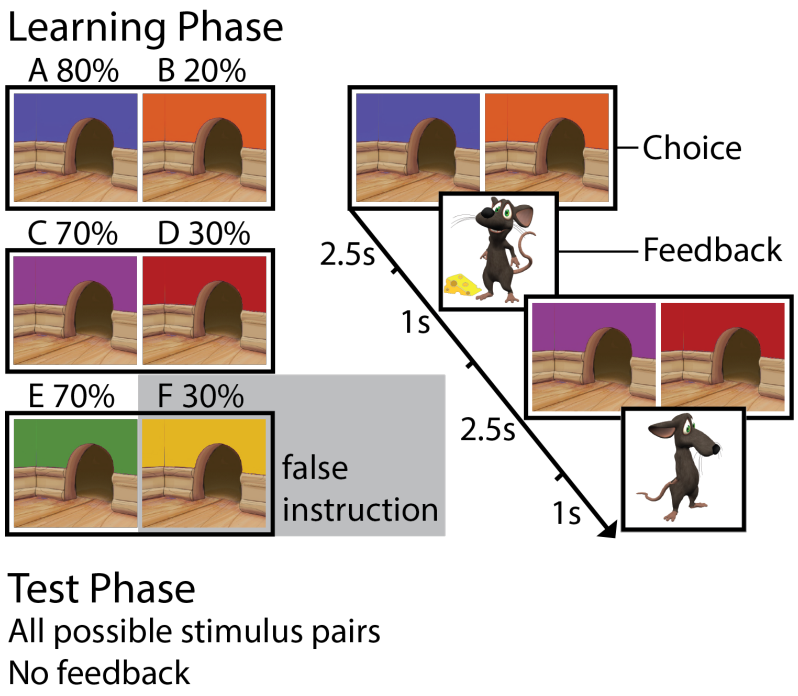

Figure 3.1 Probabilistic learning paradigm. The learning phase consisted of 180 choices between 6 probabilistically reinforced stimuli presented in 3 pairs. Participants were falsely instructed that one stimulus had a high likelihood of being rewarded when in actuality it did not. Positive or negative feedback was given following each trial. The test phase, consisting of all 15 possible stimulus pairs with no feedback, enabled assessment of the extent to which learned stimulus values were biased by instruction.

Participants were given positive or negative feedback (a happy mouse with cheese, or a sad mouse) after each choice during the learning phase, indicating whether they made a “correct” or an “incorrect” choice. Although participants did not receive monetary rewards, previous studies suggest that

purely cognitive feedback in learning tasks recruit similar underlying neurocircuitry to reward-based reinforcement learning (Rodriguez, Aron, and Poldrack 2006; Daniel and Pollmann 2010; van den Bos et al. 2012). Though both stimuli in each pair were occasionally correct or incorrect, each pair had an optimal choice. Stimuli were probabilistically reinforced; for the AB pair, choosing “A” resulted in positive feedback for 80% of the trials, whereas “B” led to positive feedback for 20% of the trials. The other two pairs (CD and EF) had reward contingencies of 70%(C/E) and 30%(D/F), but participants were given inaccurate instruction about stimulus F (verbatim: “We’ll get you started with a hint – this mousehole has a good chance of containing cheese”). This instruction was provided in textual format on the screen, accompanied by an image of the recommended mousehole. Thus, the instruction did not have a clear social source and was not directly associated with the experimenter, or any specific individual. Before starting, participants completed a brief quiz on the task instructions during which they were prompted to recall the recommended stimulus and were again visually reminded of this instruction. The participants saw each stimulus pair 60 times, pseudo-randomized in 10-trial blocks, with side of presentation counterbalanced for each participant. Participants had 2.5 seconds to choose a stimulus and received feedback for 1 second.

Before the test phase, participants were told they would now be tested on what they had just learned. Participants were presented all 15 possible stimulus pairings (3 original, 12 novel), but were not given feedback after making a choice. For each pair, they were asked to “choose the mousehole that feels more correct based on what you’ve learned; if you’re not sure which one to pick, go with your gut feeling”. Participants saw each pair 6 times,

randomly intermixed in the 90 test trials. There was a blank screen between trials (150ms duration), and no time limit when making a choice.

Data Analysis

Learning phase choice behavior data were analyzed using a generalized linear mixed-effects model using the lme4 package for the R-statistics language (Bates et al. 2014). Optimal choice (i.e., choosing the higher probability option) was modeled with independent predictors of age group (factors: children, adolescents, adults), pair (factors: AB, CD, EF), trial (1:180, z-normalized), and all 2-way and 3-way interactions. We used a maximal random-effects structure (Barr et al. 2013), including a per-participant adjustment to the intercept ("random intercepts"), as well as per-participant adjustments to the pair, trial, and pair-by-trial interaction terms ("random slopes"). In addition, we included all possible random correlation terms among the random effects. P-values and 95% confidence intervals of the log-odds were determined through bootstrapping with 1000 simulations using the bootMer function as implemented in the lme4 package, and p-values for analyses of variance were determined using likelihood-ratio-tests as implemented in the mixed function of the afex package. Data are presented visually in Figure 3.2 using mean percent choice for each pair by age group in 10-trial blocks.

We assumed that any biasing influence of instruction on experiential learning would be revealed by a tendency to make instruction-consistent choices in the test phase, an effect previously observed when participants are instructed either that a suboptimal stimulus is good, or an optimal stimulus is bad (Doll et al. 2009). We examined participants' test phase choices for pairs

that included the equally valued but differentially instructed 30% stimuli (D and F) to determine the extent to which learned stimulus values were biased by instruction. We first assessed whether participants chose in accordance with the instruction for the equally valued pair of 30% stimuli (DF pair; 30:30 instructed; Figure 3.3A). We then compared performance for a set of pairs from the test phase to generate an instruction-bias score: AD (80:30), AF (80:30 instructed), DB (30:20), and FB (30 instructed:20, see Figure 3.3B). These were chosen because any difference in performance — measured as the proportion of choices of the optimal, higher probability option — between the two 80:30 pairs or 30:20 pairs, is likely due to instruction as they are otherwise identical. The bias score is the mean of two difference scores: the difference between AD and AF performance, and the difference between FB and DB performance, each of which can vary between -1 and +1. Positive numbers indicate an instruction-consistent bias while negative numbers indicate an instruction-inconsistent bias (i.e., participants chose against instruction), and values close to zero indicate no instruction bias. The choice behavior data for these pairs were analyzed similarly to the training phase analysis except the independent variables were age group (children, adolescents, adults), pair (factors: easy 80:30, hard 30:20), and instruction (factors: instructed, uninstructed). Post-hoc testing was performed using t-tests of the instruction-bias score mentioned above. To establish that all age groups exhibited above chance experiential learning, we performed an additional analysis testing performance on all uninstructed pairs (A, B, C, D combinations) with age group as the independent variable.

Response-time data from each phase of the task were analyzed separately using a linear mixed-effects model. Models were constructed as

before, with response time as the dependent variable and choice and its interactions added as additional independent variables. P-values were determined using conditional F-tests with Kenward-Roger correction of degrees-of-freedom, as implemented in the ANOVA function (with Type III F-tests) from the car package.

We used reinforcement-learning models to attempt to characterize how participants integrated the positive and negative feedback received during the learning phase. We used participants' test phase choices as indication of their learned stimulus values (Doll, Hutchison, and Frank 2011; Doll et al. 2009; M. J. Frank, Seeberger, and O'reilly 2004; M. J. Frank and O'Reilly 2006; M. J. Frank et al. 2007) and fit two reinforcement-learning models to these test phase choices by maximum a posteriori estimation (Daw et al. 2011; den Ouden et al. 2013) using Matlab optimization toolbox (MATLAB 2012). For each model, we estimated the parameters that best captured how learning phase feedback could be integrated to yield the choices observed in the test phase. The first model was a standard reinforcement-learning model, in which prediction errors (δ) are used to update the values (Q) associated with each stimulus. Feedback that is better than expected yields a positive prediction error and feedback is worse than expected yields a negative prediction error (δ). The learning rate parameter (α) determines the extent to which these prediction errors are incorporated into the updated stimulus value. This learning algorithm has been widely used to model an experiential trial-and error-based learning process (Bayer and Glimcher 2005; Pessiglione et al. 2006; Watkins and Dayan 1992). Specifically, we updated stimulus values (Q) on each trial according to the following model:

$$Q_s(t+1) = Q_s(t) + \alpha\delta(t) \qquad \text{(for all stimuli A,B,C,D,E,F)}$$

where $\delta(t) = r(t) - Q_s(t)$ is the difference between the outcome at time t (1 reward, 0 no reward) and the current expected stimulus value.

The second reinforcement learning model includes an additional bias parameter (α_I) that alters the integration of feedback following choices of the instructed stimulus, enabling instruction-consistent feedback to be amplified (multiplying positive prediction errors, $\delta_{+,}$ which were set to zero on negative prediction error trials, by the bias parameter), and instruction-inconsistent feedback to be diminished (dividing negative prediction errors, $\delta_{-,}$ which were set to zero on positive prediction error trials, by the bias parameter) (Doll et al. 2009; Doll, Hutchison, and Frank 2011). For the instruction bias reinforcement-learning model, stimulus values were updated as follows:

$Q_s(t+1) = Q_s(t) + \alpha\delta(t)$ (for uninstructed stimuli A,B,C,D,E)

$Q_s(t+1) = Q_s(t) + \alpha_I\delta_+(t) + \delta_-(t)/\alpha_I$ (for instructed stimulus F)

For both models, the final stimulus values were then fit to participants' test-phase choices, with each trial modeled using the softmax choice rule:

$$P_{s1}(t) = \frac{1}{1 + e^{-\beta(Q1-Q2)}}$$

Where the inverse temperature parameter (β) describes how deterministic an individual's choices are given the difference in Q-values. Parameter estimates were compared at the group level using non-parametric tests. Model fits were compared to one another using the Akaike Information Criterion (Akaike 1974).

We took the Beta(1.1, 1.1) distribution as a prior for the learning rate parameter (α), and the Gamma(1.2,5) distribution as a prior for both the bias (α_I) and inverse temperature (β) parameters. These priors were chosen to be uninformative over the previously observed ranges of parameter estimates in

similar tasks and to ensure smooth parameter boundaries (Daw et al. 2011; Daw 2011).

RESULTS

Learning phase

During the learning phase (Figure 3.2), there was a significant difference in performance by age group (X^2 = 6.96, df = 2, p = 0.031). Children performed significantly worse than adults (log-odds difference = -0.547, 95% confidence interval (CI) (-1.018, -0.161), p = 0.015), and children showed a trend toward worse performance than adolescents (log-odds difference = -0.357, CI (-0.721, 0.082), p = 0.095), but adolescents did not differ from adults (log-odds difference = -0.190, CI (-0.627, 0.198), p = 0.40). Performance also differed significantly depending on the stimulus pair (X^2 = 29.7, df = 2, p < 0.0001). Performance for the easy, uninstructed pair (AB 80/20) was significantly better than the falsely instructed pair (EF 70/30) (log-odds difference = 1.046, CI (0.721, 1.387), p < 0.005), and marginally better than the hard, uninstructed pair (CD 70/30) (log-odds difference = 0.183, CI (-0.001, 0.338), p = 0.055). Performance for the CD pair (70/30) was also significantly better compared to the instructed EF pair (70/30; log-odds difference = 0.863, CI (0.573, 1.182), p < 0.005). There was also a significant linear improvement in performance across the learning phase (X^2 = 36.1, df = 1, p < 0.0001, log-odds estimate = 0.297, CI (0.202, 0.378)).

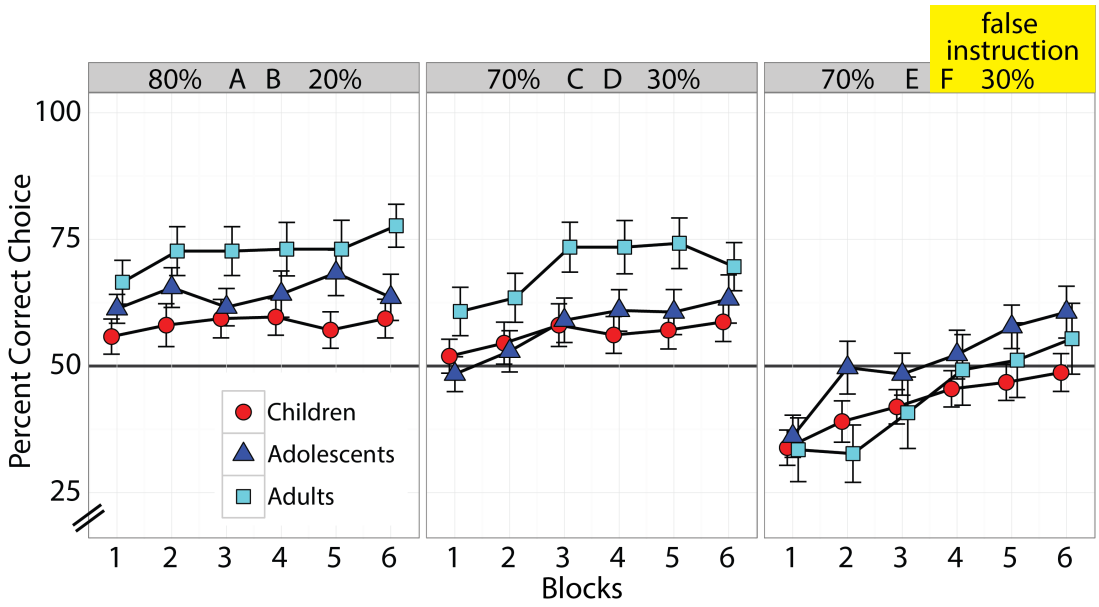

Figure 3.2 Learning phase performance. Adults (cyan squares) perform better than children (red circles) and marginally better than adolescents (blue triangles) for the two uninstructed pairs (AB, CD). All groups initially adhere to the false instruction (F) and gradually learn through experience to select the higher valued alternative (E), with adolescents showing the fastest improvement. Error bars are SEM.

The linear improvement in performance across the learning phase differed by age group (X^2 = 7.45, df = 2, p = 0.024). Children showed slower improvement in performance than adults (log-odds difference = -0.283, CI (-0.502, -0.066), p < 0.005), marginally slower improvement than adolescents (log-odds difference = -0.183, CI (-0.363, 0.037), p = 0.080), and there was no difference between adolescents and adults (log-odds difference = -0.100, CI (-0.309, 0.109), p = 0.344). The linear improvement in performance also differed by stimulus pair (X^2 = 2.53, df = 2, p = 0.009). There was a slower improvement in performance for both the easy uninstructed pair (AB) compared to the falsely instructed pair (EF) (log-odds difference = -0.265, CI (-0.397, -0.135), p = 0.010), and the hard uninstructed pair (CD) compared to the EF pair (log-odds difference = -0.169, CI (-0.300, -0.034), p=0.030),

suggesting that performance quickly reached asymptote in the uninstructed pairs. However, there was no difference between the two uninstructed pairs (log-odds difference = -0.096, CI (-0.208, 0.016), p = 0.110). There was a marginally significant age group by stimulus pair interaction effect (X^2 = 9.33, df = 4, p = 0.053). Whereas adults performed better than children and adolescents on uninstructed pair choices (AB/CD), their performance decreased for the instructed pair (EF), particularly when compared to adolescents (Fig. 2). Finally, we found no evidence of a pair-by-age-by-trial interaction (X^2 = 2.89, df = 4, p = 0.576).

Test phase

Replicating previous results (Doll et al. 2009), during the test phase, adults showed a bias towards the instructed stimulus (F) when it was part of the equally valued DF stimulus pair (30:30 instructed; t = 5.98, df = 25, p < 0.0001, mean = 81.4%, CI (70.6, 92.2)). However, children only showed a marginal effect (t = 1.98, df = 29, p = 0.0573, mean = 60.8%, CI (49.6, 71.9)) and adolescents showed no effect of instruction (t = 0.66, df = 30, p = 0.515, mean = 53.8%, CI (42.1, 65.4)) (Figure 3.3*A*). Children (t = -2.64, df = 54.0, p = 0.011, percent difference = -20.7, CI (-35.8, -5.5)), and adolescents (t = -3.56, df = 55.0, p = 0.0008, percent difference = -27.6, CI (-43.2, -12.1)) chose the instructed stimulus significantly less than adults, and there was no difference in preference between children and adolescents (t = 0.92, df = 59.0, p = 0.36, percent difference = 7.0%, CI (-8.8, 22.8)). This provided an initial indication of a persistent instruction bias in adults that is absent in children and adolescents.

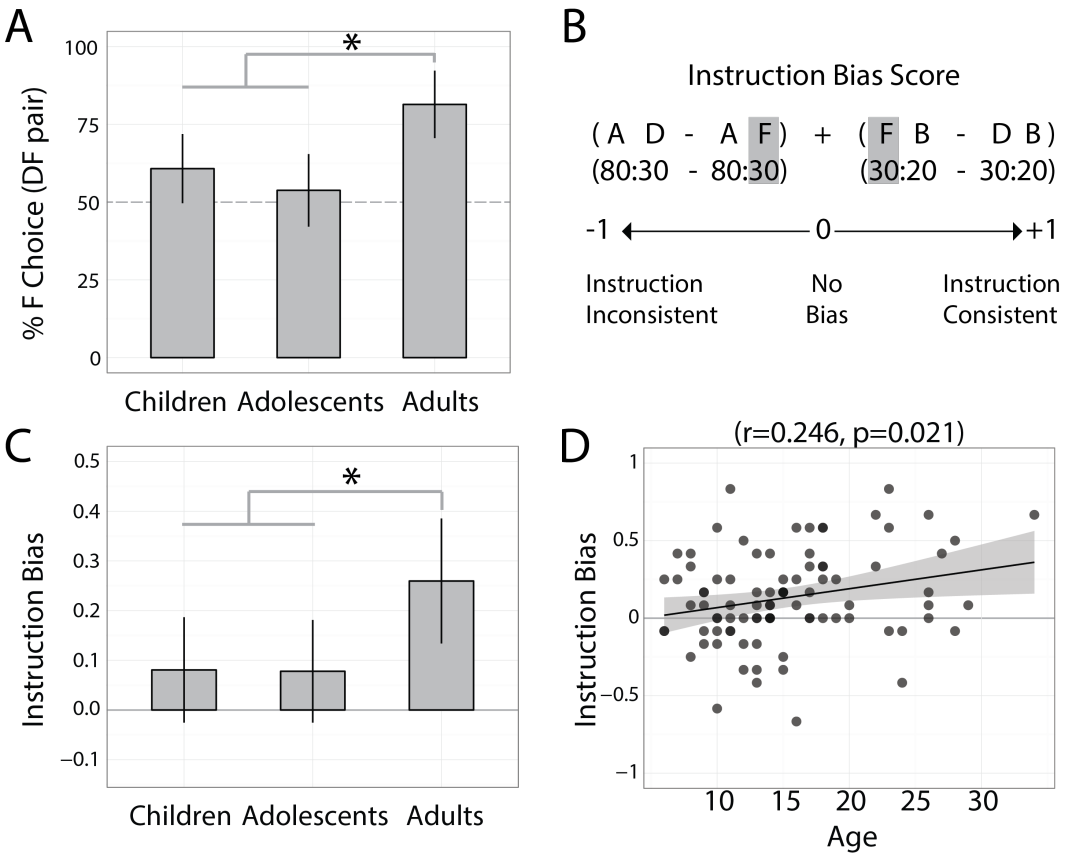

Figure 3.3 Test phase performance. a) Percentage of F choices when seeing the DF pair (30:30 instructed). b) The instruction bias is the average of two difference scores between stimulus pairs of equal probability that are differentially instructed. c) Adults show a significantly larger instruction bias than both children and adolescents. d) There is a significant increase in instruction bias with age (darker circles indicate two data points). Error bars (a, c) are 95% confidence intervals; shading (d) is SEM.

The computed instruction bias score (Figure 3.3*B*) assesses choice preference across a broader set of pairs that are equally valued, but differentially instructed. Again, we found that adults were more biased than children (t = 2.24, df = 51.8, p = 0.029, instruction bias difference = 0.179, CI (0.018, 0.340)), and adolescents (t = 2.29, df = 51.0, p = 0.026, instruction bias difference = 0.182, CI (0.026, 0.341)), and that children were not differently biased from adolescents (t = 0.002, df = 58.7, p = 0.998, instruction bias difference = 0.003, CI (-.284, .295); Figure 3.3*C*). This pattern of age-

group differences was also present for each individual subcomponent of the bias score (80/30 and 30/20). The effect of age group remained significant when gender was included as a predictor of instruction bias, with only adults showing a significant bias (p=0.0005). There was also a significant gender interaction for the adult group (p=0.0133), but not children (p=0.40) or adolescents (p=0.25). This effect was due to adult females having a larger instruction bias than adult males (t = 2.57, df = 24, p = 0.017, instruction bias difference = 0.298, CI (0.059, 0.537)). We also found the instruction bias to increase linearly with age (r=0.246, p = 0.021; Figure 3.3*D*).

We examined the choices for the pairs that make up the instruction bias (Figure 3.3*B*) – pairs AD 80/30, AF 80/30 instructed, DB 30/20, and FB 30/20 instructed – using a mixed-effects regression with pair difficulty, instruction, age group, and all interactions as independent variables. As expected, we found a significant pair-by-instruction interaction effect (i.e., an instruction bias) on performance (X^2 = 23.19, df = 1, p < 0.0001, log-odds estimate = -0.510, CI (-0.729, -0.302)), indicating that the false instruction impaired performance for the otherwise easy pair (80/30), and improved performance for the otherwise hard pair (30/20). Additionally, this instruction bias effect differed across age groups (X^2 = 11.59, df = 2, p = 0.0031), mirroring the effects seen in the instruction bias score. Children and adolescents were equally unaffected by instruction (log-odds difference = 0.024, CI (-0.428, 0.471), p = 0.945), whereas adults were significantly more affected by instruction than children (log-odds difference = -0.793, CI (-1.311, -0.278), p < 0.002) and adolescents (log-odds difference = -0.769, CI (-1.325, -0.267), p = 0.008). We found no other significant effects of pair or instruction (both p > 0.6).

There was an overall difference in age group optimal choice for these pairs (X^2 = 16.58, df = 2, p = 0.0003), reflecting age-related differences in overall probabilistic learning, independent of instruction. Children (log-odds difference = -1.310, CI (-1.963, -0.665), p < 0.002) and adolescents (log-odds difference = -0.868, CI (-1.486, -0.275), p = 0.006) chose less optimally than adults, but not differently from one another (log-odds difference = -0.443, CI (-1.024, 0.115), p = 0.108).

Performance on all uninstructed pairs (A, B, C, and D combinations) showed that all age groups performed better than chance (children: log-odds estimate = 0.436, CI (0.128, 0.744), p = 0.006; adolescents: log-odds estimate = 0.674, CI (0.367, 0.981), p < 0.0001; adults: log-odds estimate = 1.28, CI (0.932, 1.621), p < 0.0001). Children (log-odds difference = -0.840, CI (-1.302, -0.378), p = 0.0004) and adolescents (log-odds difference = -0.603, CI (-1.063, -0.143), p = 0.010) performed less well than adults and were not significantly different from each other (log-odds difference = -0.238, CI (-0.672, 0.197), p = 0.284), demonstrating age differences in experiential learning, similar to those evident in learning phase uninstructed choices.

Both during the learning and test phases, response times were unrelated to choice, instruction, age group, or any of their interactions. Response times decreased significantly over the learning phase (response time effect = -51.4 ms, SEM = 7.75 ms, X^2 = 35.58, df = 1, p < 0.0001), with no difference by age group. During the test phase response times were significantly longer for the hard pairs (30:20) than easy pairs (80:20), regardless of instruction (response time difference = 244.8 ms, SEM = 38.5ms, $F_{1,93.58}$ = 9.58, p = 0.0026). There were no other significant effects on test phase response times.

Reinforcement Learning (RL)

To explore the process by which instruction might influence the integration of feedback during the learning phase, we fit participants' test-phase choices using a standard and modified instruction bias RL model. The standard RL model describes a feedback-driven learning process that has been proposed to underlie experiential reward learning (Bayer and Glimcher 2005; Pessiglione et al. 2006). The instruction bias RL model adds a bias parameter that amplifies the influence of instruction-consistent outcomes and diminishes instruction-inconsistent feedback, yielding an instruction bias (Doll et al. 2009). The standard RL model is equivalent to an instance of the instruction bias RL model in which the bias parameter is set to 1 (i.e. no bias). Model comparison based on the median AIC values for each age group (see Table 3.1) indicated that the choices of children and adolescents were better fit by the standard RL model, whereas adults were better fit by the modified model that includes an instruction bias parameter. This result suggests that children and adolescents recruited an undistorted feedback-based integration process during the learning phase, while adults biased the integration of feedback for the instructed stimulus, altering the weighting of positive and negative outcomes in an instruction-consistent manner.

These age differences in choice consistency parallel the differences in test phase performance observed for the experientially learned stimuli in the regression results. In the instruction bias RL model, bias parameter estimates (α_I) exhibited the same pattern of age group differences observed for both the instruction bias score and the regression analysis (H = 14.25, p = 0.0009). Adults' bias parameter estimates were significantly higher than both children's

(W = 156.5, p = 0.0002) and adolescents' (W = 237, p = 0.021), with no difference between children and adolescents (W = 430.5, p = 0.62).

Table 3.1. Reinforcement learning model parameter fits. α – learning rate. β – softmax inverse temperature parameter, α_I – bias parameter, AIC - Akaike Information Criterion. The non-parametric Kruskal-Wallis (KW) and Mann-Whitney-Wilcoxon (MWW) tests were used to compare group parameter estimates. Standard (RL) and modified (Mod RL) reinforcement-learning models.

		Median			KW	MWW		
Model	Param	Child	Adolescent	Adult	Chi-sq (df=2)	Child vs. Adolescent	Child vs. Adult	Adolescent vs. Adult
RL	α	0.429	0.083	0.036	H=12.87 p=0.002	W=579 p=0.10	W=621 p=0.0002	W=504 p=0.11
	β	0.80	1.25	3.89	H=16.26 p=0.0003	W=361 p=0.14	W=150 p<0.0001	W=238 p=0.009
	AIC	131.2	125.2	117.2				
Mod RL	α	0.289	0.046	0.054	H=8.82 p=0.013	W=622 p=0.023	W=565 p=0.004	W=396 p=0.92
	β	1.32	2.89	4.55	H=18.71 p<0.0001	W=316 p=0.031	W=127 p<0.0001	W=262 p=0.024
	α_I	1.42	1.44	3.31	H=14.25 p=0.0009	W=430.5 p=0.62	W=156.5 p=0.0002	W=237 p=0.021
	AIC	135.8	128.3	100.0				

Collectively, these modeling results suggest qualitative differences as a function of age in the manner in which instruction influenced experiential feedback-based learning. Whereas instruction biased the integration of feedback during learning for adults, both children and adolescents were less influenced by instruction, integrating feedback in a relatively unbiased manner.

DISCUSSION

In this study, we examined whether the influence of instruction on experiential learning changes across development. Despite age differences in performance, children, adolescents, and adults were all able to recruit

experiential feedback to learn to preferentially select the higher-valued stimulus of each uninstructed pair. In all age groups, choices for the instructed pair were initially biased toward the inaccurately recommended stimulus and gradually (rapidly in adolescents) shifted toward the higher-valued alternative stimulus as participants received continued negative feedback. However, performance during the test phase suggested marked qualitative differences across development in how instruction influenced the processing of this experiential feedback. Consistent with previous findings (Doll et al. 2009; Staudinger and Büchel 2013), we found that adults showed a strong instruction-consistent bias, suggesting that inaccurate instruction distorted their feedback-based value learning. In contrast, both children and adolescents showed a minimal influence of instruction on test phase performance, suggesting that they integrated positive and negative feedback more objectively during the learning phase in order to estimate the value of the instructed stimulus. These data suggest that when explicit instruction or advice conflicts with experiential feedback about the value of an action, children and adolescents, unlike adults, weight their own experience more heavily.

Our analyses of instruction bias focused on decisions made during the test phase, in which participants' novel choices revealed the value estimated for each stimulus through the integration of learning phase feedback. In contrast, choices during the learning phase can reflect potentially heterogeneous evaluation strategies adopted by participants (e.g. hypothesis testing across multiple trials, (M. J. Frank et al. 2007; Doll, Hutchison, and Frank 2011)), which may obscure current stimulus value estimates. Past studies employing variants of this task suggest that test phase choices may provide the most reliable indication of learning and are selectively sensitive to

various pharmacological, genetic, psychological, and neurological factors thought to alter the incremental experiential learning process (M. J. Frank, Seeberger, and O'reilly 2004; M. J. Frank and O'Reilly 2006; M. J. Frank et al. 2007; Doll et al. 2009; Doll, Hutchison, and Frank 2011). In our study, age group differences in the influence of instruction were not evident in choices made during the learning phase. However, we saw robust evidence of instruction bias in the test phase choices of adults, but not children and adolescents. Our reinforcement-learning analyses establish a link between feedback received during the learning phase and test phase decisions by formalizing potential underlying processes for learning stimulus values. Crucially, reinforcement-learning parameters for the initial learning phase were fit to the test phase choices that reflect final learned stimulus values. These analyses suggest that whereas experienced outcomes during the learning phase were objectively weighted in children's and adolescents' value estimates, adults biased the weighting of outcomes for the instructed stimulus to be more consistent with the explicit instruction they had received.

Experiential learning is thought to depend critically on dopaminergic prediction errors, through which the striatum can learn the value of an action (McClure, Berns, and Montague 2003; O'Doherty et al. 2003; Pagnoni et al. 2002; Schultz, Dayan, and Montague 1997). Explicit instruction is proposed to bias this striatal learning process through the top-down influence of the prefrontal cortex (Biele, Rieskamp, and Gonzalez 2009; Biele et al. 2011; Doll et al. 2009; Doll, Hutchison, and Frank 2011; Li, Delgado, and Phelps 2011; Staudinger and Büchel 2013), which enables task-relevant rules and instructions to influence goal-directed behavior (Miller and Cohen 2001). A theoretical model supported by our reinforcement learning analyses (Doll et al.

2009) posits that the prefrontal cortex amplifies the effect of instruction-consistent outcomes and diminishes the influence of instruction-inconsistent outcomes on striatal learned values. This produces a behavioral “confirmation bias”, through which recommended actions are more highly valued than those learned solely through experience, even when the recommendation is inaccurate. Previous studies examining the instructional control of experiential value learning in adults largely support this model, demonstrating both the hypothesized alteration of striatal feedback-driven error signals (Biele et al. 2011), and a correlation between instruction-guided choice outcomes and prefrontal cortex activation (Li, Delgado, and Phelps 2011). Collectively, this evidence suggests that functional interaction between the prefrontal cortex and the striatum may mediate the instructional biasing of learning that we observed in our adult participants.

By extension, the relative absence of instructional influence on experiential learning in children and adolescents may stem from reduced functional efficacy of prefrontal-striatal pathways prior to adulthood. Functional imaging studies have observed intact striatal prediction error signals from childhood onwards (Galvan et al. 2006; van den Bos et al. 2012), consistent with evidence of feedback-based experiential reward learning across development (J. R. Cohen et al. 2010; van den Bos et al. 2009; van den Bos et al. 2012; S. Peters et al. 2014). In contrast, both structural and functional connectivity between the prefrontal cortex and the striatum exhibit marked changes from childhood through adulthood (Liston et al. 2006; Imperati et al. 2011). Cognitive functions that depend on the integrity of this neural pathway typically show continued maturation into adulthood (Liston et al. 2006; Rubia et al. 2006; Somerville and Casey 2010), suggesting that developmental

changes in frontostriatal connectivity may facilitate information exchange between these regions. Based on the neuroscientific model of instructional control of learning in adulthood, we hypothesize that the prolonged maturation of prefrontal-striatal connectivity underlies the resistance of children and adolescents to the biasing effects of inaccurate instruction in our task. Our present study focused solely on behavior. However, we expect that a functional imaging study of our task might find that adults exhibit an instruction-consistent bias in striatal prediction error signals for choices of the instructed stimulus during the learning phase, with positive signals amplified and negative signals diminished relative to those for the equally valued uninstructed stimulus. We expect that this biased signaling in adults would be accompanied by greater PFC-striatal connectivity following instructed versus uninstructed choices outcomes. In contrast, we expect that children and adolescents would show no such differences in prediction error signals for instructed versus uninstructed stimuli. These specific hypotheses about the potential neural substrates of our behavioral results could be tested directly in a subsequent developmental neuroimaging study.

This maturational increase in instructional influence on learning observed in this study concurs with a broader literature suggesting a gradual developmental emergence of cognitive control (Bunge and Zelazo 2006; Diamond 2006; Munakata, Snyder, and Chatham 2012). A primary challenge of cognitive development is to acquire knowledge across a variety of stimulus domains about the nature of the environment, which is accomplished in large part through inductive statistical learning. Such experientially acquired knowledge may be more flexibly applied and easily generalized than explicit rule-based learning, a principle that has long been recognized in pedagogical

theory (R. L. Hayes 1993; D. A. Kolb 1984). Implicit learning processes typically recruit evolutionarily conserved subcortical regions including the basal ganglia (Bischoff-Grethe et al. 2004; Rauch et al. 1997). Such learning is evident early in development (Amso and Davidow 2012; Saffran, Aslin, and Newport 1996; Kirkham, Slemmer, and Johnson 2002), and may continue to improve into adulthood (Thomas et al. 2004). Although reduced prefrontal control is often portrayed as a developmental handicap, it may confer distinct advantages by enabling implicit learning to occur unhindered (Thompson-Schill, Ramscar, and Chrysikou 2009). Providing instruction, whether false or veridical, has been shown to interfere with multiple forms of implicit experiential learning, reducing task performance relative to when no instruction is given (Reber 1989). Increased sensitivity to underlying patterns in the reinforcement of actions may facilitate children and adolescents' acquisition of language, social conventions, and other complex behaviors.

An important consideration not addressed in this study is whether the social source of instruction might modulate its influence. In the present study, the instruction provided to participants was simply presented on the screen, lacking any specific social origin. In contrast, real world advice often comes from peers (friends, classmates, colleagues) or authority figures (parents, teachers, bosses), which may yield different effects on behavior than a printed message. The source of advice may be a particularly important factor during adolescence—a period of increasing independence and heightened sensitivity to peers (Chein et al. 2012; Galván 2014; Gardner and Steinberg 2005; Jones et al. 2014; Steinberg and Monahan 2007). Moreover, the influence of instruction has been shown to depend on the perceived expertise of the advisor (Meshi et al. 2012), and peers and authority figures may be viewed as

experts in different behavioral domains at different developmental stages. Thus, advice from different social sources may vary in salience across both age groups and decision contexts. Future work might explore whether manipulating the social source of instruction alters the developmental differences in instruction bias reported here.

Both parents and policy-makers commonly rely upon rules and instruction to deter children and adolescents from actions that carry potentially harmful consequences. In particular, increased independence during adolescence often presents opportunities to experiment with behaviors (e.g. sex, drug experimentation, reckless driving) that frequently yield positive social or hedonic outcomes, but can have rare yet serious negative effects. Positive experienced outcomes may come to predominate in adolescents' risk estimates (Reyna and Farley 2006). In our study, participants of all ages initially adhered to instruction, consistent with other evidence that the actions and decisions of adolescents can be influenced by advice (Engelmann et al. 2012). However, when the feedback they received provided evidence contradictory to their prior instruction, both children and adolescents, but not adults, showed greater reliance upon their own experience. Public policy campaigns attempting to deter adolescents from risky behavior through explicit guidance or information have had limited efficacy (Ennett et al. 1994; Trenholm et al. 2007). The present results suggest a cognitive mechanism underlying such resistance to instruction. This finding highlights the importance of research aimed at identifying effective ways for both parents and public health campaigns to advise adolescents as they navigate real-world risky behavioral domains.

In summary, by placing instruction and experience in competition, we show here that the relative weighting of these two sources of information shifts over the course of development. Consistent with the protracted maturation of the circuitry implicated in instructional control of learning, children and adolescents showed less influence of instruction on choice than adults. Whereas instruction alters the processing of experiential feedback in adults, our results suggest that children and adolescents remain attuned to the true reward contingencies in their environment, enabling experience to prevail in directing their actions. Many aspects of cognition (e.g. working memory (Crone et al. 2006), attentional control (Rueda, Posner, and Rothbart 2005), and executive function (Diamond 2006) improve as individuals mature from childhood through adulthood, typically conferring advantages for adults in learning and decision-making. Similarly, the effective recruitment of instruction to guide one's actions may generally be advantageous, allowing an individual to benefit from the knowledge and prior experience of others. However, our results suggest that this ability may also come at the cost of introducing pronounced bias in the processing of experiential feedback. The absence of confirmation bias in children and adolescents observed in this study represents a paradoxical developmental advantage of youth over adults in the unbiased evaluation of actions through positive and negative experience.

LIMITATIONS

Reinforcement learning algorithm limitations

This study shares many of the same limitations as those discussed in the limitations section of Chapter 2, so the discussion will be focused primarily on aspects and considerations that are specific to the analysis approach taken

in this chapter. In addition to the typical concerns that come with modeling behavior using trial-by-trial estimates of a learning process, this experiment makes a critical change to the modeling procedure. Whereas reinforcement-learning algorithms generally model learning and action-selection for every trial throughout a task, here, the learning and action-selection portions of the algorithm are separated into different phases (Doll et al. 2009). More specifically, only the learning phase is used to calculate the prediction errors that train the learning portion of the algorithm, and only the test phase choices are used to calculate the action-selection probabilities. The reasons for doing so are as follows: first, the study is not interested in determining which action-selection strategies are being used by participants during the learning phase, but only how instruction influences that learning. As such, regardless of whether participants use atypical action-selection strategies during the learning phase, such as choosing randomly or hypothesis testing, no attempt is made to model those decisions. Second, there is no feedback during the test phase and so no additional learning can occur. It follows that participants should thus only choose the stimuli they think would reward them on a given trial, rather than some strategy to help them learn. Therefore, we make the assumption that the final derived stimulus values from the learning phase are transferred over to the test phase decisions. While this has the nice property of not needing to estimate the true choice strategy of the learning phase, it also offers no chance to interpret the decision strategy during the learning phase.

Previous studies have attempted other modeling approaches to explain the decision behavior of adults. The bias (the approach used in our analysis), override, strong prior, and insight models each have a corresponding neuropsychological interpretation (Doll et al. 2009). The bias model assumes

that reinforcement learning of the non-instructed stimuli corresponds to a plain SARSA algorithm (Sutton and Barto 1998), and that learning about the instructed stimuli includes a scaling factor. This scaling factor is the bias parameter, and simply amplifies instruction-consistent prediction errors and dampens instruction-inconsistent prediction errors. This provided a good account of the learning phase and test phase in prior work. The override model assumes that learning proceeds according to reinforcement learning for all stimuli, but that there is an additional scaling factor during decisions involving the instructed stimulus. While this account was qualitatively similar to the bias model, the override model fit the behavior less well, especially during the test-phase. The strong prior model assumes that instruction leads to a high initial estimated value of the instructed stimulus and lowers the learning rate for that specific stimulus, and this model was shown to fit both learning and test phases poorly. Finally, the insight model assumes that a participant, after experiencing instruction-inconsistent outcomes, has the insight that the instruction was incorrect, and will make a sharp switch to favoring the uninstructed paired stimulus. While this account captured the learning phase data fairly well, it failed to account for the bias seen in the adult test phase choices. Given these results, we selected to focus our attention on the bias model, but it is possible that one of these other models may be better than the standard reinforcement-learning model currently being used.

A more parsimonious model would likely incorporate all of these approaches. As participants of all ages initially followed the instruction, adolescents sharply shifted their learning phase behavior, and adults showed the instruction bias behavior, a single model would need to be designed that allowed any of these aspects to drive behavior. The model could, for example,

keep the initial values of each of the six stimuli as free parameters, and allow for the values in a pair to be swapped on any given trial. However, we expect that the unmodified reinforcement-learning algorithm would continue to fit child and adolescent behavior as well as any other, when correcting for free parameters.

Another issue in the modeling approach is that it does not take into consideration that the values of the learning phase stimuli within a pair are yoked. While the point is not belabored during the instructions, it is stated that one stimulus in each pair is better than the other, and becomes fairly obvious quite early during the learning phase, and this yoking is the basis behind the insight model. The algorithm could model this information by having a prediction error update two values, including the uninstructed stimulus that was paired with the instructed stimulus. If the value estimates of the paired stimuli are yoked, it would suggest that there may also be an indirect instruction bias, going in the opposite direction of the main result. However, a separate scaling parameter is likely unnecessary for this indirectly instructed stimulus, as there is no behavioral evidence in the test phase showing that subject representations of these values were tied to one another.

What is the optimal strategy?

Unlike the two-stage task discussed in Chapter 2, there is a clear optimal strategy in this task, and no age group seems to adopt it. The optimal strategy is to always choose the stimulus with the higher likelihood of yielding a reward. There were 6 participants (3 adolescents and 3 adults) who showed perfectly optimal strategy for all three pairs by the end of the learning phase, and 4 additional (3 adolescents and 1 adult) that showed close to optimal. Yet,

as a whole, participants seem to asymptote at around the reinforcement probability, known as probability matching behavior (Stanovich 2003; Doll et al. 2009). This behavior is readily observable with probabilistic reinforcers, and is known as the Gambler's Fallacy (Tversky and Kahneman 1971). The fallacy is as follows: even knowing the probability of an event, like a coin toss, people generally assume outcomes to be self-correcting, such that if multiple heads come up in a row, multiple tails must follow to bring the average back to 0.5. In our task, this might manifest in an individual predicting that on a given trial, the reward might appear with the lower probability stimulus. However, it has been shown that with large financial incentives, meaningful and regular feedback, and extensive training, participants can eventually adopt an optimal strategy (Shanks, Tunney, and McCarthy 2002). As none of these are the case in our study, save feedback, we are not surprised by this result in the learning phase, and do not think that it informs the interpretation of the test phase behavior.

Finally, the poor model fits in this study need to be discussed. Unlike the studies in Chapter 2 and Chapter 4, the proposed models fit many participants' data no better than a chance model. And while each age group seemed to fit the model overall, group differences were present, and the poor individual-level fits suggest that there may be some serious problems either with the parameterization of the models, or with the design of the task. The test phase, with its 12 novel pairs, is difficult, and it is quite easy to make internally inconsistent decisions. For example, if someone were to indicate that A > B, B > C, and C > A, it would be impossible to estimate their difference in relative values and a fitting function would predict that they are all equal. Thus, because of the relatively poor fits, we do not rest our interpretations on the

modeling results, and instead use the bias score and regression analysis to make our claim.

Limitations to the bias score and regression analysis

The bias score and DF-preference (the relative preference between the two equally valued but differentially instructed 30% reinforced stimuli) are not especially robust as they are based on choice proportions, but we explored the possibility of including more trials to increase this robustness. The bias score is the average of two proportions, whereas the DF-preference is based on only one. It may be valid, however, to increase the number of measurements that go into the bias score, if it is appropriate to include the stimulus that was paired with the instructed stimulus as an indirectly instructed stimulus. To determine this, we examined the CE-preference (equally valued, differentially instructed 70% reinforced stimuli), and the indirect instruction bias score, which is calculated equivalently to the normal bias score but with E and C stimulus swapped for F and D. There was no indication in either measure for a bias in any age group (all p-values > 0.15). As such, it was determined that these trials should not be incorporated as measures of the instruction bias. However, in an effort to examine the consistency in the instruction bias score, we plotted the range of the two proportions that form the instruction bias score (Figure 3.4). The general low within-participant variability suggests that, despite the low number of measurements, the instruction bias measure is internally consistent.

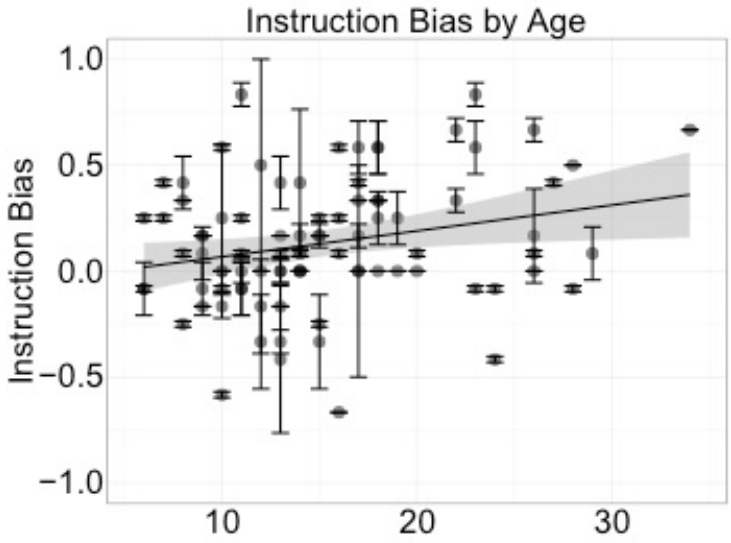

Figure 3.4 Variance of instruction bias measurement. Most participants are consistent in both measures of the instruction bias score. Darker circles indicate two participants are overlapping.

The regression analysis is a more robust approach, as it uses all 24 test phase trials that make up the instruction bias score in its prediction of test phase choice. However, this number of trials is still fairly small to estimate 4 random effects for each participant and 7 fixed-effects overall; thus, the returned statistics are likely not very trustworthy. Thus, rather than relying simply on likelihood-ratio-tests, as we did in the other chapters, here we employed parametric bootstrapping. That the results hold with this more stringent test provides added support that the observed bias effect in adults, which was undetectable in children and adolescents, was valid.

Another concern is that participants were not learning to the same extent during the learning phase. This is especially concerning for the instructed pair for which most participants were choosing the optimal choice only 50% of the time. However, just because they were choosing this option, does not guarantee they were not learning, and when looking at overall

performance during the test phase, all age groups were performing above chance. Nevertheless, other versions of this study have used a ‘to criterion’ learning phase, rather than a fixed number of trials (Doll et al. 2014). That way, once a pair was learned (measured as choosing in proportion with the reward probability), it was not overly trained for the participants who learned it. This also runs into the problem of different training amounts, but is a good way of guaranteeing that learning occurred. However, if the hypothesis is that children may have learned about the instructed stimulus less than adults, then it might be expected that they would be more likely to be biased, which is not what we see. This is something to consider in future implementations of the task.

Chapter 4: On weight and waiting: delay discounting in anorexia nervosa pre-treatment and post-treatment*

Individuals with anorexia nervosa (AN) override the drive to eat, forgoing immediate rewards in favor of longer-term goals. We examined delay discounting and its neural correlates in AN before and after treatment to test a potential mechanism of illness persistence. Inpatients with AN (n=59) and healthy controls (HC, n=39) performed a delay discounting task at two timepoints. A subset (n=30 AN, n=22 HC) participated in fMRI scanning during the task. The task consisted of a range of monetary choices with variable delay times, yielding individual discount rates—the rate by which money loses value over time. Before treatment, the AN group showed a preference for delayed over earlier rewards (i.e., less steep discount rates) compared with HC; after weight restoration, AN did not differ from HC. Underweight AN showed slower response times for earlier versus delayed choices; this reversed with treatment. Underweight AN showed abnormal neural activity in striatum and dorsal anterior cingulate; normalization of behavior was associated with increased activation in reward regions (striatum and dorsal anterior cingulate) and decision-making regions (dorsolateral prefrontal cortex and parietal cortex). The undernourished state of AN may amplify the tendency to forgo immediate rewards in favor of longer-term goals. The results suggest that behavior that looks phenotypically like “excessive self-control” does not correspond with enhanced prefrontal recruitment. Rather, the results point to alterations in cingulo-striatal circuitry that offer insights on the potential role of perturbed decision-making neural systems in the perpetuation of AN.

* Decker, J. H., Figner, B. & Steinglass, J. E. On Weight and Waiting: Delay Discounting in Anorexia Nervosa Pretreatment and Posttreatment. *Biol. Psychiatry* 1–9 (2015). doi:10.1016/j.biopsych.2014.12.016

INTRODUCTION

Anorexia nervosa (AN) is a serious disorder with a mortality rate six times as high as expected among young women (Arcelus and Mitchell 2011). Despite ongoing research, the neurobiology of AN remains poorly understood. One defining characteristic of AN is the ability to resist the drive to eat. Individuals with AN demonstrate a capacity to forgo receipt of food reward—to override biological hunger cues and "postpone" eating. This common feature of AN is examined in the current study using a behavioral task that quantifies delay discounting, a measure of one aspect of self-control.

The clinical phenomena seen in AN have been understood as manifestations of excessive self-control, dating back to early descriptions by Hilde Bruch of "iron determination" (Bruch 1978). In a recent study using a monetary delay discounting task, we measured preferences between smaller-but-immediate rewards versus larger-but-delayed rewards, providing an estimate for individual discount rates—reflecting how rapidly a reward loses subjective value as a function of how long one must wait to receive it (Green, Fry, and Myerson 1994). Individuals with AN had discount rates that were significantly less steep than their healthy peers (i.e., they preferred larger-but-delayed rewards) (Steinglass et al. 2012), suggesting greater "self-control" or "patience." This result was provocative in part because it is uncommon to find heightened self-control in this task in a psychiatric population, and because discount rates have been shown to have behavioral correlates and thus ecological validity (Sharp, Monterosso, and Montague 2012). Steeper discounting (i.e., less patience) is associated with poorer self-control as evidenced by increased tendencies toward impulse shopping and gambling (Chabris et al. 2008), and lower achievement later in life (Mischel, Shoda, and

Rodriguez 1989; Ayduk et al. 2000). These established links between discount rate and behavior have implications for AN, where the core disturbances are maladaptive eating behaviors.

Measuring self-control in a monetary reward paradigm has advantages in AN, as the reward value of food is uncertain (G. K. W. Frank et al. 2012). Disrupted decision-making around money is not part of the AN diagnosis, therefore, the presence of an abnormality in delay discounting of monetary reward may indicate an attribute that extends beyond eating-related abnormalities and that can provide clues to underlying neurobiology. In AN, the tendency to choose delayed monetary rewards suggests a disposition that may contribute to persistent maladaptive eating choices. Perhaps brain function is altered in AN in a way that makes it easier to resist the temptation of short-term reward—resisting both monetary smaller-sooner rewards and food rewards such that the waiting for future weight loss is less of a burden. Paradigms with non-food rewards have previously shown reward processing abnormalities supporting this avenue of investigation in AN (Keating et al. 2012).

Neurobiologically, fronto-striatal reward and fronto-parietal control networks are implicated in delay discounting, including the medial prefrontal cortex, ventral striatum, dorsolateral prefrontal cortex (dlPFC) and inferior parietal lobule (McClure et al. 2004; Kable and Glimcher 2007; Ballard and Knutson 2009; Carter, Meyer, and Huettel 2010). The increased patience in delay discounting seen in AN raises the question as to whether individuals with AN may exhibit functional abnormalities in these decision-making systems. Existing data in healthy adults suggest that choosing the larger-later rewards is associated with activity in the dlPFC (McClure et al. 2004; Figner et

al. 2010), a region identified with self-control processes (Miller and Cohen 2001; Hare, Camerer, and Rangel 2009). Neuroimaging studies in AN have pointed to potential abnormalities in regions relevant for delay discounting (striatum and dorsal anterior cingulate) (Keating et al. 2012; G. K. Frank et al. 2005; Bischoff-Grethe et al. 2013; A. Wagner et al. 2007), yet this hypothesis has not been investigated.

Clarification of the neurocognitive underpinnings of AN is critical for developing new treatment targets for this potentially severe illness. In this study, we aimed to examine delay discounting behavior before and after treatment, along with the associated neural systems. We hypothesized that individuals with AN would show less steep discounting both before and after treatment, as compared with healthy peers, suggesting a possible underlying trait consistent with excessive self-control and that this would be associated with greater activity in the dlPFC.

METHODS

Participants

Participants were individuals with AN presenting to the Columbia Center for Eating Disorders/ New York State Psychiatric Institute (NYSPI) and healthy controls (HC)(Table 1). Eligible patients were between 16 and 45 years old, met DSM-5 (American-Psychiatric-Association 2013) criteria for AN, restricting (AN-R) or binge-purge (AN-BP) subtype, and were receiving inpatient treatment. Individuals were excluded if they had an estimated IQ less than 80, history of a neurological, bipolar, or psychotic disorder, substance abuse in the last 6 months, or if they were pregnant. Anxiety or depressive disorders, which

commonly co-occur, were not an exclusion when AN was the primary diagnosis (Hudson et al. 2007).

A subset of these individuals were recruited to participate during an fMRI scan (n=48) if they were 16-25 years old, female, with no contraindication to MRI, and not taking psychotropic medication. Medications are not routinely used for AN on the inpatient unit, due to lack of evidence of utility (Attia et al. 1998).

HC were matched for age, sex, and ethnicity and were included if they had no current or past psychiatric illness, no significant medical illness, no psychotropic medications, and a BMI in the normal range (18-25 kg/m^2). This study was approved by the NYSPI Institutional Review Board, and after complete description of the study to the participants, written informed consent was obtained.

Procedures

Height and weight were measured on a beam balance scale (Detecto). Participants were administered the Eating Disorders Examination semi-structured interview (Z. Cooper and Fairburn 1987), and the Wechsler Test of Adult Reading (Wechsler 2001). Testing occurred twice. Individuals with AN were tested within 1 week of hospital admission (Session 1), and after weight restoration to a BMI of 19.5 kg/m^2(Session 2). Time between sessions was group-matched.

Delay Discounting Task

Participants made binary choices between amounts of money available at various delays: smaller-sooner (SS) and larger-later (LL), adapted from

(McClure et al. 2004) . The SS options were available “today” (NOW) or “in 2 weeks” (NOT-NOW). The delay to the LL option was either 2 or 4 weeks after the SS option. The relative difference in dollar amounts between the SS and LL options (i.e., (LL-SS)/SS) was 1, 3, 5, 10, 15, 20, 25, 35, or 50%. The SS amounts ranged from $15 to $85 dollars.

This factorial design results in 36 trials—2 (NOW or NOT-NOW) by 2 (2 or 4 week delay) by 9 (relative percentage difference). Two sets of trials were used during the fMRI scan. These sets were duplicated, with one duplicate presenting the LL option as the default (ACCELERATE) and the other presenting the SS as the default (DELAY), for a possible total of 144 trials. Participants outside the scanner used only one duplicate set, for 72 trials. The order and frame (ACCELERATE or DELAY) was counterbalanced within and between participants. For half of the participants, ACCELERATE trials were paired with amazon gift cards and DELAY trials with cash, also counterbalanced between participants.

In the scanner, participants had 10 seconds to indicate their choice, and received feedback for 2 seconds, indicating that their choice was recorded. During feedback, the triangle below the chosen option turned green while the triangle below the alternative option disappeared. Feedback was followed by a variable inter-trial-interval ranging from 7 to 8 seconds. The task was presented in 4 runs of 4 minutes each, with a brief break between runs to allow participants to rest. Runs were presented in one of two counterbalanced orders, either (A = Acceleration; D = Delay) A-D-D-A or D-A-A-D. Outside of the scanner there was no time limit.

Participants were instructed that there were no right or wrong answers and to choose the option they truly preferred because at the end of the

experiment they would be paid according to their choice on one of the trials. After finishing participation, one trial was selected by a random number generator, and the participant was paid according to their preference on that trial (e.g., if they had selected an SS of $24 today over an LL of $36 in 4 weeks, they received $24 that day).

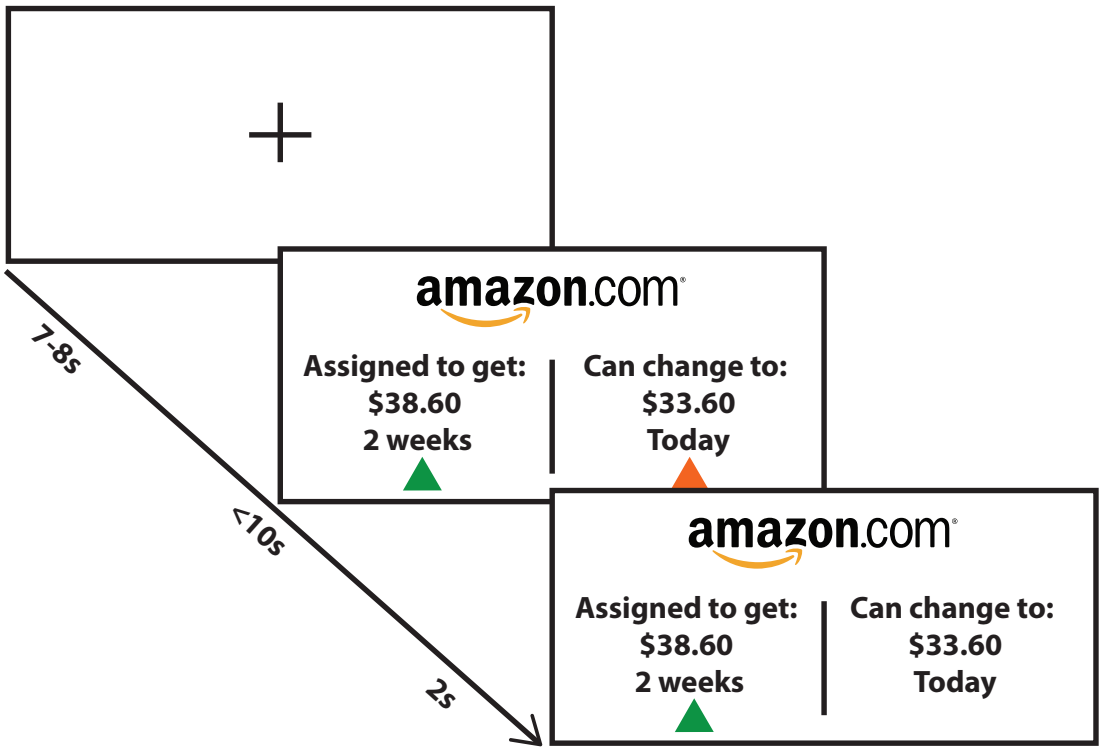

Figure 4.1 Delay discounting task design. Individuals were presented with a choice between a smaller amount of money available sooner (SS) and a larger amount available later (LL). Amounts ranged from $15 to $85 and time of delivery for SS choices was either Now or in 2 weeks, and the time of delivery for LL was 2 or 4 weeks after the SS. Outside the scanner, there was no time limit for responding. In the fMRI version, there was a fixation cross between trials. All task parameters (i.e. monetary values, time differences) were the same inside and outside the scanner.

fMRI Data Acquisition

Imaging was performed on a 1.5T Philips Intera scanner, with an 8-channel head coil. High-resolution T1-weighted anatomical images were

acquired using an SPGR sequence (TR=25sec, TE=3.7ms, angle=30°, FOV=256mm, 256x204 matrix, 128 slices, voxel size 1x1x1mm). The task was performed during 4 functional runs using an EPI sequence (TR=2000ms, TE=40ms, FOV=192mm, 64x63 matrix, 33 axial slices, voxel size 3x3x4mm, 180 TRs). Trials advanced with participants' selections, and were not synchronized with TRs.

fMRI Data Preprocessing

fMRI data were preprocessed and analyzed using the Analysis of Functional Neuroimages software package (AFNI) (Cox 1996). Functional scans were corrected for slice acquisition using sinc-interpolation. Volume registration using 6-parameter rigid-body transformation, to account for head motion, and normalization into Talairach space using 12-parmeter affine transformation were performed in a single interpolation step. Data were resampled to 3mm isotropic voxels. Data were iteratively smoothed to achieve a final full-width half-max Gaussian kernel of 6mm. Signal intensity was normalized by individual voxel to percent signal change.

Data Analysis

Behavioral

Clinical characteristics were compared using Student's t-tests for independent samples, with Welch correction for unequal variances, and Mann-Whitney-Wilcox tests for ordered measures. For each individual, a hyperbolic discount rate (k) was estimated from their choice data, per Session. Log-transformed discount rates, $\log(k)$, were analyzed using a linear mixed-effects model, with main predictors of interest Diagnosis, Session, and their

interaction and with random intercepts and Session slopes for each participant, also repeated with age and IQ included in the model. This method models individual variability and is robust to missing data, and was selected due to the different sample sizes across different stages of analysis. Log-transformed discount rates were compared using the lme4, and afex packages in the R-statistics language (R-Core-Team 2012; Bates et al. 2014; Singmann 2013). *P*-values were determined using conditional *F*-tests with Kenward-Roger correction of degrees-of-freedom, as implemented in the Anova function (with Type III F-tests) from the package car (Fox and Weisberg 2010); this function calls the KRmodcomp function from package pbkrtest; (Højsgaard 2014)). The interaction was examined using post-hoc tests of Diagnosis for Session 1 and 2, and tests of Session for AN and HC. Lower values of log(*k)* indicate less steep discount rates, or a greater preference for the LL reward. We examined the relationship of discounting with various clinical characteristics by testing the correlation of the discount rate with age, IQ, body mass index, duration of illness, eating disorder examination score, and the time to discharge after weight was restored.

Fitting Discount Models

We estimated the probability of making a larger-later choice given the choice subjective values on a given trial.

$$p = \frac{1}{1 + e\text{^}(-\beta(SV_{LL} - SV_{SS}))}$$

Where p is probability, β is the slope parameter, and SV is the reward (SS or LL) subjective value. SV was determined using three different discounting

models (Mazur, Nevin, and Rachlin 1987; McKerchar et al. 2009; Benhabib, Bisin, and Schotter 2010):

$$SV = \frac{A}{1 + kT} \qquad hyperbolic$$

$$SV = Ae^{-kT} \qquad exponential$$

$$SV = \begin{cases} A & if\ T = 0 \\ Abe^{-dT} & if\ T > 0 \end{cases} \qquad quasihyperbolic$$

Where *A* is the offered amount, *k* is a 1-parameter discount rate, *b* and *d* are the 2-parameter discount rates, and *T* is the time to the reward in years.

These parameters were estimated to minimize the negative log-likelihood of individual choice probability using MATLAB's fmincon minimizing function (MATLAB 2012). A pseudo-R^2 was generated by comparing the fit against that of a random chooser, $p = 0.5$ ((Doll et al. 2009), Figure 4.2). Though the quasi-hyperbolic model had a significantly better fit than the other two models using the signed Wilcox rank sum test, there was no difference between models using an unsigned test (hyperbolic $p = 0.25$, exponential $p = 0.33$). As such, the hyperbolic model was chosen to describe the results as a single discount rate is simpler to interpret, and is the most commonly described discount model in the literature. Participants for whom the fitted model was not better than chance were excluded (Green, Fry, and Myerson 1994; Reynolds and Schiffbauer 2004), set at 0.15 to exclude as few imaging participants as possible.

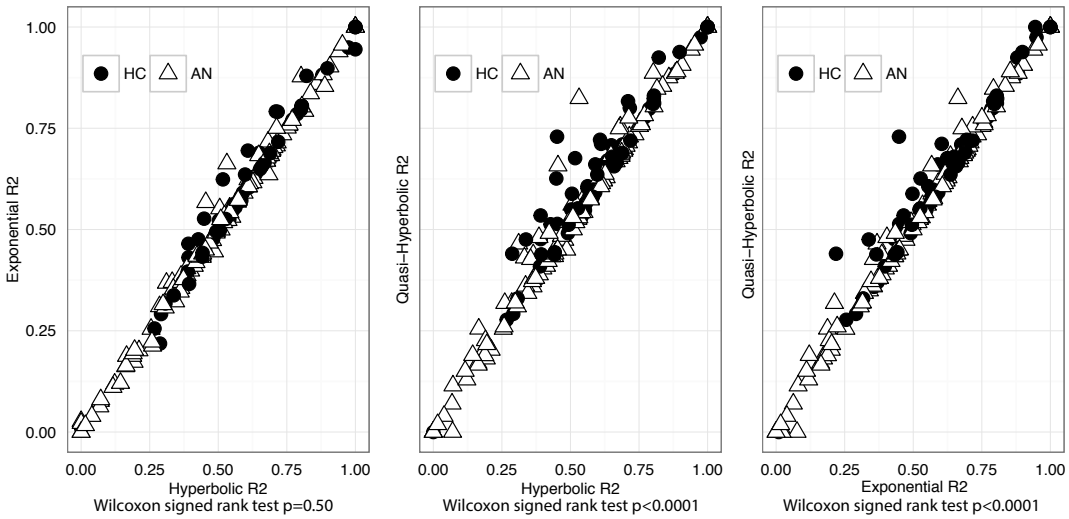

Figure 4.2 Pseudo-R^2 Comparisons of Each Discounting Model. All participants who completed the task are included. A pseudo-R^2 below 0.15 for the hyperbolic model was selected as being insufficiently different from random to use in the analysis.

Inter-temporal Choice Analysis

Another approach to analyzing the delay discounting behavior is to use a generalized linear mixed-effects model for choice behavior from each trial. The model was similar to the one for discounting behavior, but additional predictors were added that model Immediacy (NowNotnow), smaller-sooner amount (SS_Amount), time difference between options (TimeDiff), and relative difference between reward (RelDiff100) for the given trials. The continuous predictors were scaled such that they had a mean of zero and a standard deviation of 1. We followed the advice of (Barr et al. 2013), and used a maximal random-effects structure: the repeated-measures nature of the data was accordingly modeled by including a per-participant random adjustment to the fixed intercept ("random intercept"), as well as per-participant random adjustments to the Session, NowNotnow, SS_amount, TimeDiff, RelDiff100,

and four interaction (Session:NowNotnow, Session:SS_amount, Session:TimeDiff, and Session:RelDiff100) slopes ("random slopes"); in addition, we included all possible random correlation terms among the random effects. *P*-values were determined using the Likelihood Ratio Tests as implemented in the mixed function of the afex package.

Response-time analysis

Response times may provide some insight into the type of strategy being used by individuals. We used a linear mixed-effects model from trial data using Diagnosis, Session, Choice, their interactions, and including a full random-effects structure, similar to what was described for the inter-temporal choice analysis. This analysis was also repeated including the absolute difference in subjective value between the two options presented (LL-SS). The subjective value was determined with the hyperbolic discount function of each value, using participant specific discount rates and trial specific delays. Absolute value was utilized, as it is the magnitude of the difference, and not the direction, that captures the difficulty of a given trial.

fMRI

Single-subject analysis on preprocessed data was done using a general linear model (GLM). Each subject had a design matrix with 22 regressors: baseline, trend, and quadratic signal to capture shifts in signal change for each of the 4 runs(12), motion parameters(6), and 4 trial-specific regressors that reflected participants' choices (SS-NOW, SS-NOT-NOW, LL-NOW, LL-NOT-NOW). These trial regressors were convolved with a duration-modulated (by trial response time) block hemodynamic-response function. Trials with greater

than 2mm of motion (as well as the preceding and following trial) were censored, and scans with greater than 10% of trials censored were excluded from further analysis. In an attempt to account for the change in subjective values seen across sessions, we added an additional amplitude-modulated regressor for the difference in subjective value between LL and SS options of a given trial according to the subject-specific, behaviorally determined discount rate (z-normalized).

Our main analysis of interest of the group data was performed using linear mixed-effects modeling using AFNI's 3dLME function (Chen et al. 2013), which is robust to small amounts of missing data. Regression coefficients from the individual analysis estimated from fewer than 12 trials were excluded from the group analysis. The model included fixed-effect terms: Choice (SS or LL), Immediacy (NOW or NOT-NOW), Diagnosis (AN or HC), Session (1 or 2), and all their possible interactions, covariates (age and IQ), and a random intercept for each participant. For regions in which the Choice by Session by Diagnosis interaction term was identified as significant, we extracted the average SS and LL first- level regression coefficients of each individual for each session. These were used to further examine how these regions differed between groups through post-hoc t-tests. This group analysis was repeated using the individual regression coefficients from when absolute difference in subjective value was included.

RESULTS

Participants are described in Table 1. Of the initially enrolled 106 participants, we excluded 1 HC and 1 individual with AN when it was discovered during further screening that they did not meet inclusion criteria.

Table 4.1 Demographic and Clinical Characteristics of Participants

	HC (n=39)		AN (n=59)*				
	Mean ± SD		Mean ± SD		*t*	*df*	*p*
Time 1	n=39		n=54**				
Age (years)	24.7±7.6		25.0±7.5		-0.17	91	0.87
BMI (kg/m^2)	21.7±1.9		16.6±1.5		-3.13	62.5	<0.005
Education (years)	15.1±3.0		14.1±2.1		1.72	63.9	0.09
Eating Disorder Examination	0.08±0.10		3.1±1.4†		-15.3	53.7	<0.001
WTAR estimated IQ	108.5±11.8		107.9±8.0		0.26	56.6	0.79
Time 2	n=31		n=43				
Days between sessions	58.5±35.4		52.6±15.6		0.87	38.8	0.39
BMI (kg/m^2)	21.9±2.0		20.4±0.7		-2.76	50.3	<0.01
EDE	0.1±0.1		2.1±1.3		-13.9	42.7	<0.001
	N	(%)	N	(%)	X^2	df	p
Caucasian	27	69.2	53	89.8	2.5	1	0.11
Female	37	94.9	57	96.6	0.20	1	0.65

WTAR=Weschler Test of Adult Reading, EDE=Eating Disorder Examination, BMI=Body Mass Index
*At Time 1, AN-R n=25 and AN-BP n=29. At Time 2, AN-R n=22 and AN-BP n=21.
†EDE scores were significantly different between AN-R and AN-BP (3.7±1.1 vs 2.6±1.5, respectively, p=0.004).
**Three individuals' data were excluded at Time 1, based on the behavioral algorithm. These individuals were included at Time 2. Two subjects participated only at Time 2.

Subset of the above participants who also provided fMRI data

	HC (n=21)	AN (n=25)*			
	Mean ± SD	Mean ± SD	*t*	*df*	*p*
Time 1	n=21	n=23			
Age (years)	20.7±2.8	19.3±2.5	1.6	42	0.10
BMI (kg/m^2)	21.4±1.8	16.8±1.4	3.1	39.9	<0.005
Education (years)	14.1±2.2	13.2±2.0	1.4	40.6	0.17
Eating Disorder Examination	0.08±0.11	3.09±1.58†	-9.1	22.5	<0.001
WTAR estimated IQ	109.3±10.2	104.7±8.1	1.6	34	0.13
Time 2	n=16	n=18			
Days between sessions	53.3±30.2	45.4±11.8	1.0	19.1	0.34
BMI (kg/m^2)	21.7±1.8	20.2±0.6	3.3	31.8	<0.005
EDE	0.1±0.1	2.3±1.4	-9.6	17.1	<0.001

* At Time 1, AN-R n=13 and AN-BP n=10. At Time 2, AN-R n=9, and AN-BP n=9.
†EDE scores did not differ between subtypes.
**Two individuals data were excluded at Time 1, based on the behavioral algorithm. These individuals were included at Time 2. One HC and one subject with AN had fMRI data that was excluded due to motion, their behavioral data were included in analyses.

Data from 6 participants (all AN) were excluded because their behavior was not distinguishable from random choice. The 6 excluded participants did not differ significantly from the included AN group in clinical characteristics, and

are not included in any of the analyses below. The final sample included 98 participants (39 HC and 59 AN); of these, 30 HC and 37 AN provided task data at both Sessions. Mean duration of illness among AN was 8.6±6.9 years, with a history of 0-15 (mean=2.6) prior hospitalizations. There were 28 individuals with AN-R and 31 individuals with AN-BP in the full sample. There were no significant differences in clinical characteristics between the groups who completed one versus two sessions. As shown in Table 1, there was a small difference in BMI between AN and HC at Time 2, likely related to the narrow BMI range among weight-restored AN.

The age range for individuals who participated in the fMRI portion of the study was designed to be narrow in order to obtain a more homogeneous sample (22 HC, 26 AN). As such, they were younger than non-scanned participants (19.8±2.7 years versus 29.8±7.2 years, t_{96}=8.97, p<0.0001), with no difference between AN and HC. All fMRI participants were female; two with AN were left-handed. There were no other significant differences between the fMRI and behavioral participants.

Behavioral Results

Discount rate results

Results of the delay discounting task showed a significant effect of Session ($F_{1,71.4}$=12.0, p=0.0009) and a Diagnosis by Session interaction effect ($F_{1,71.4}$=19.45, p<0.0001) on the discount rates, log(*k*) (Table 4.2). Individuals with AN had a significantly lower mean discount rate than HC at Session 1(t_{91}=2.25, p=0.027), a significant increase between Sessions (t_{36}=-4.6, p<0.0001), and did not differ from HC at Session 2 (t_{72}=-0.84, p=0.40)(Figure 4.3). There was no difference in HC between sessions (t_{29}=1.22, p=0.23). The

same pattern was seen when age and IQ were included in the linear mixed-effects model (Table 4.3). Delay discounting results followed the same pattern, with no significant differences among the subsets in and out of the scanner. Discount rate and change in discount rate were not significantly associated with measures of illness severity or other variables that could potentially affect the value of reward (Table 4.4).

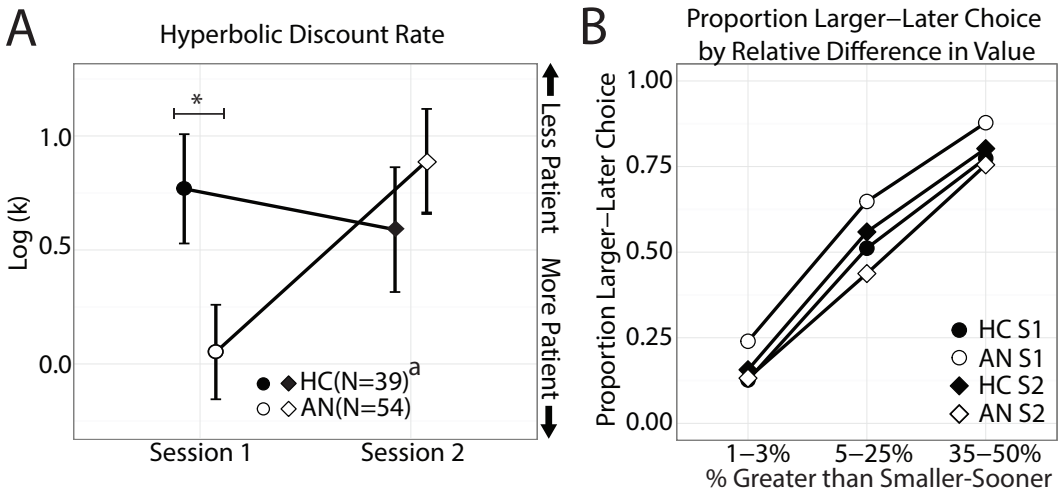

Figure 4.3 Individuals with AN have lower discount rates than HC only when underweight. (A) The log-transformed discount rates (per unit years) are shown for individuals with AN and HC at Sessions 1 and 2. Lower log-transformed discount rates indicate less steep discounting, i.e., a preference for larger-later over smaller-sooner options. (B) The proportion of trials that the larger-later option was chosen is shown for the AN and HC groups at each Session, separated into three bins indicating how much greater the larger-later choice was than then smaller-sooner choice in percentage terms. The AN group shows an overall decrease in the proportion of trials that they chose the delayed option, rather than for a specific subset of trials.
[a] Session 2 sample size (N=31 HC, 43 AN)
[*] $p<0.05$ (error bars are SEM).

Table 4.2 Discount rate as a function of diagnosis (Dx) and session

Effect	Estimate	*F*-Stat	ndf	ddf	*p*-value
Intercept	0.68	482.23	1	95.85	<0.0001
Dx	0.06	.45	1	95.85	0.50
Session	-0.14	12.01	1	71.43	0.0009
Dx:Session	0.17	19.45	1	71.43	<0.0001

Table 4.3 Discount rate as a function of diagnosis (Dx), session, age (z-normalized zAge) and IQ (z-normalized zIQ)

Effect	Estimate	*F*-Stat	ndf	ddf	*p*-value
Intercept	0.63	15.52	1	84.99	<0.0001
Dx	0.14	.72	1	84.99	0.40
Session	-0.21	14.52	1	61.46	0.0003
zAge	0.31	2.70	1	88.38	0.10
zIQ	-0.38	4.89	1	83.66	0.030
Dx:Session	0.28	25.76	1	61.46	<0.0001
Dx:zAge	0.14	0.53	1	88.38	0.47
Session:zAge	-0.12	3.16	1	62.28	0.080
Dx:zIQ	-0.14	0.67	1	83.66	0.42
Session:zIQ	0.09	1.89	1	60.38	0.17
zAge:zIQ	-0.00	0.00	1	83.00	0.98
Dx:Session:zAge	0.18	6.45	1	62.28	0.014
Dx:Session:zIQ	-0.02	0.17	1	60.38	0.68
Dx:zAge:zIQ	-0.01	0	1	83.00	0.97
Session:zAge:zIQ	0.11	4.27	1	59.84	0.043
Dx:Session:zAge:zIQ	-0.05	0.90	1	59.84	0.35

As a secondary analysis, discount rates were compared between AN-R, AN-BP and HC in the mixed-effects model and showed a significant Session effect ($F_{1,71.4}$=24.23, p<0.0001) and a significant Diagnosis by Session interaction ($F_{1,71.3}$=10.82, p<0.0001). Compared with HC at Session 1, the AN-R group had lower discount rates (i.e. more patience) (t_{62}=2.48, p=0.016), and no difference from HC at Session 2 (t_{51}=0.096, p=0.92). Individuals with AN-BP did not differ significantly from HC or AN-R at Session 1 (HC: t_{66}=1.37, p=0.175, AN-R: t_{52}=-1.09, p=0.28) or Session 2 (HC: t_{50}=-1.62, p= 0.11 and AN-R: t_{41}=-1.56, p=0.13). Both AN-R (t_{17}=-3.38p=0.004) and AN-BP (t_{18}=-3.40,p=0.003) groups changed significantly across Sessions.

Table 4.4 Additional delay discounting statistical analyses

Log(k)	All Participants				Scanned Subset			
	Session 1		Session 2		Session 1		Session 2	
Diagnosis	***n***	**Mean±SD**	***n***	**Mean±SD**	***n***	**Mean±SD**	***n***	**Mean±SD**
HC	39	0.77±1.50	31	0.59±1.52	22	0.60±1.45	17	0.46±1.47
AN	54	0.05±1.52	43	0.89±1.50	24	0.23±1.54	19	0.97±1.13

Scanned Subset Significance Testing[b]	**Statistic**	***p*-value**
Diagnosis by Session Interaction	$F_{1,34.2} = 7.39$	$p = 0.010$
Diagnosis at Session 1	$t_{44} = 0.82$	$p = 0.42$
Diagnosis at Session 2	$t_{34} = -1.17$	$p = 0.25$
Session:AN paired	$t_{16} = -2.58$	$p = 0.020$
Session:HC paired	$t_{16} = 0.6$	$p = 0.56$

Behavioral Subset Significance Testing	**Statistic**	***p*-value**
Diagnosis by Session Interaction	$F_{1,34.0} = 11.10$	$p = 0.0021$
Diagnosis at Session 1	$t_{45} = 2.32$	$p = 0.025$
Diagnosis at Session 2	$t_{36} = -0.15$	$p = 0.88$
Session:AN paired	$t_{19} = -3.83$	$p = 0.0011$
Session:HC paired	$t_{12} = 1.06$	$p = 0.31$

Log(k) correlation		**Session 1**			**Session 2**			**Change in Log(k)**		
		***t*-test**	***r***	***p***	***t*-test**	***r***	***p***	***t*-test**	***r***	***p***
Age	HC	2.19	0.34	0.04	1.57	0.28	0.13	0.33	0.06	0.75
	AN	-0.38	-0.05	0.71	1.53	0.23	0.13	2.79	0.42	0.008
IQ	HC	-2.33	-0.37	0.03	-2.59	-0.45	0.02	-1.82	-0.33	0.08
	AN	-0.31	-0.04	0.75	-0.26	-0.04	0.80	0.28	0.05	0.78
BMI	HC	-0.38	-0.06	0.71	0.13	0.02	0.90	-1.73	-0.31	0.094
	AN	0.27	0.04	0.79	0.55	0.09	0.58	0.88	0.15	0.38
Duration of Illness	AN	-0.95	-0.13	0.35	0.55	0.09	0.59	1.75	0.28	0.09
EDE Score	AN	0.72	0.10	0.48	0.56	0.09	0.58	0.24	0.04	0.81
Time to Discharge	AN				-0.14	-0.02	0.89	0.38	0.06	0.71

		HC		**AN**			
Household Income	**Level**	***n***	**(%)**	***n***	**(%)**	**W**	***p***
< $10,000	1	5	13.2	9	16.7	1222	0.12
$10,000-19,999	2	4	10.5	7	13		
$20,000-34,999	3	7	18.4	13	24.1		
$35,000-49,999	4	4	10.5	6	11.1		
$50-000-99,999	5	4	10.5	12	22.2		
$100,000-199,999	6	10	26.3	4	7.4		
> $200,000	7	4	10.5	3	5.6		
Employment	**Level**	***n***	**(%)**	***n***	**(%)**	**W**	***p***
	None	14	35.9	36	61	1354	0.10
	Part-time	17	43.6	10	16.9		
	Full-Time	8	20.5	13	22		

Inter-temporal choice results

There were no group differences in immediacy *(p* = 0.15), time difference (*p* = 0.23), or relative difference (*p* = 0.34), but a marginal effect of SS amount (*p* = 0.055). This suggests that there were no systematic session-independent differences between AN and HC in how variation in the timing and in the amounts of the rewards affected their inter-temporal choices, or to put it differently, there was no evidence that individuals with AN might have neglected the difference in the delays or amounts between the two options.

Table 4.5 Choice as a function of diagnosis (Dx), session, immediacy (Now-Notnow), smaller-sooner amount (zSS_Amount), relative difference between smaller-sooner and larger-later amount (zRelDiff100), and time difference between delays (zTimeDiff)

Effect	Estimate	Chi-sq	*p*-value
Intercept	1.17	8.17	0.0043
Dx	-0.25	0.32	0.57
Session	-0.72	14.07	0.0002
NowNotnow	0.05	0.52	0.47
zSS_Amount	1.19	119.38	<0.0001
zRelDiff100	3.58	111.88	<0.0001
zTimeDiff	-0.36	41.31	<0.0001
Dx:Session	0.83	20.34	<0.0001
Dx:zNowNotnow	-0.09	2.12	0.15
Dx:zSS_Amount	0.14	3.67	0.055
Dx:zRelDiff100	0.23	0.93	0.34
Dx1:zTimeDiff	-0.06	1.45	0.23
Session:NowNotnow	0.05	1.44	0.23
Session:zSS_Amount	-0.01	0.06	0.81
Session:zRelDiff100	-0.44	9.66	0.0019
Session:zTimeDiff	0.02	0.42	0.52
Dx:Session:NowNotnow	0.09	3.52	0.061
Dx:Session:zSS_Amount	0.04	0.77	0.38
Dx:Session:zRelDiff100	0.33	6.68	0.0098
Dx:Session:zTimeDiff	0.05	1.78	0.18

The marginally significant effect of SS amount suggests that AN had an attenuated magnitude effect (which is the effect that, everything else being

equal, larger amounts of money lead to increased patience). This latter result might suggest that AN tend to be less sensitive to increasing outcome magnitudes; but given that this effect is only marginally significant, we are hesitant to make strong conclusions. However, there is a significant Diagnosis-by-Session-by-Relative Difference interaction (p = 0.0098). This term similarly suggests that that the change in preference for the larger-later choice that occurred in the AN group across sessions might depend on being more responsive to changes in relative difference once weight restored. This would suggest that the AN group were somewhat less aware or responsive to the relative differences in outcome magnitudes when underweight.

Response time results

Overall response times quickened across sessions for both groups (mean difference=-311ms, SD=815ms, t_{68}=3.17, p=0.002), with no diagnosis-by-session interaction (p=0.67; Table 4.6). Response times for HC did not differ between SS and LL choices across sessions (difference=-50ms, SD=819ms, t_{27}=-0.32, p=0.75). Individuals with AN showed a significant shift in response time between sessions: At Session 1, AN were slower for SS choices than for LL choices, and after treatment, responses were faster for SS choices than for LL choices (difference=-336ms, SD=797ms, t_{34}=2.5, p=0.018)(Figure 4.4). When absolute difference in subjective value of the two options was included in the analysis, to account for choice difficulty, the pattern was the same, but only at trend level (p=0.063)(i.e., slower response for SS choices in the underweight phase, Table 4.7).

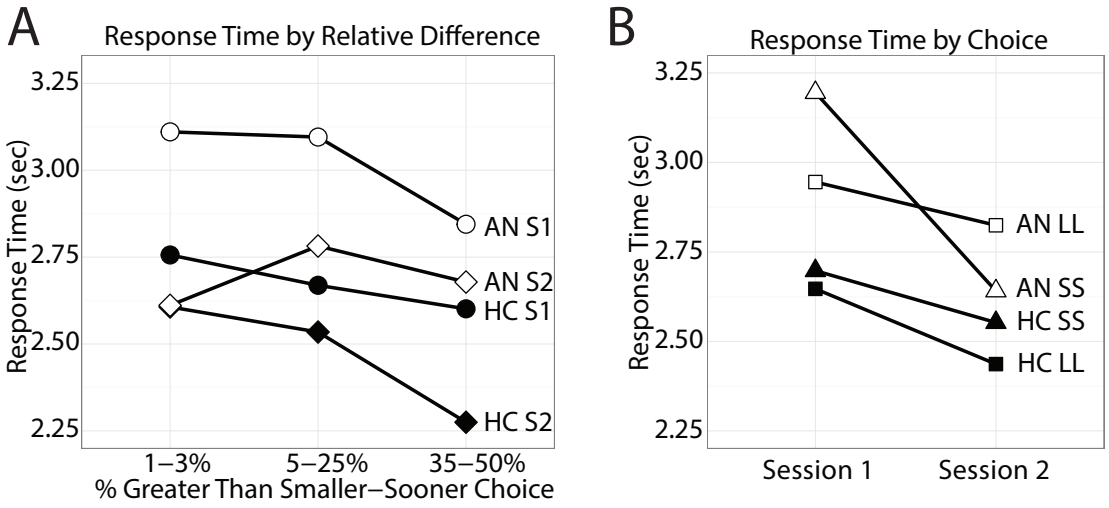

Figure 4.4 Individuals with AN chose smaller-sooner options more slowly than larger-later options when underweight, and switched when weight restored. (A) Response time by percent that the larger-later (LL) option is greater than the smaller-sooner (SS) option. This shows that both groups quickened their responses across session, and for some trial types more than the other. (B) Response time by session, split by SS and LL choice. The AN group shows a significant switch from being slower during SS than during LL choices when underweight, to being faster during SS than during LL choices once weight restored.

Table 4.6 Response time as a function of diagnosis (Dx), session, and choice (SS or LL)

Effect	Estimate	Chi-square	*p*-value
Intercept	3055.3	221.28	0
Dx	-63.2	0.37	0.54
Session	177.3	12.67	0.0004
Choice	-8.1	0.02	0.88
Dx:Session	-19.9	0.18	0.67
Dx:Choice	-41.4	0.66	0.42
Session:Choice	2.9	0.02	0.90
Dx:Session:Choice	-59.8	5.86	0.016

Table 4.7 Response time as a function of diagnosis (Dx), session, choice (SS or LL), and the absolute difference in subject value (z-normalized, zAbsPresDiff)

Effect	Estimate	Chi-square	*p*-value
Intercept	2988.5	221.64	<0.0001
Dx	-83.5	0.72	0.40
Session	158.0	10.3	0.0013
Choice	-73.0	2.72	0.099
zAbsPresDiff	-233.0	81.64	<0.0001
Dx:Session	-7.8	0.03	0.86
Dx:Choice	-35.3	0.65	0.42
Session:Choice	-11.6	0.2	0.65
Dx:Session:Choice	-44.8	3.46	0.063

Imaging Results

Imaging analyses probed the behavioral finding of differences in preference for delayed rewards between AN and HC. Analyses examined differences in neural activity between LL and SS choices. There was a significant Choice (SS/LL) by Diagnosis (AN/HC) by Session (S1/S2) interaction in multiple brain regions, including the striatum bilaterally, the dorsal anterior cingulate cortex (dACC), the right dlPFC (rdlPFC), and the right parietal lobule (rPar)(Figure 4.5, Table 4.8). We compared the differences in Choice (LL minus SS) activity between diagnostic groups at each Session in these regions. At Session 1, HC showed no difference between LL and SS activity in any of these regions, whereas individuals with AN showed lower LL relative to SS activity in the striatum and dACC. At Session 2, HC showed lower LL relative to SS activity in the dACC, rdlPFC, and rPar, whereas individuals with AN showed greater LL relative to SS activity in the striatum, dACC, and rdlPFC. HCs showed a significant change across sessions: LL minus SS activity was smaller at Session 2 than Session 1 in the striatum, dACC, rdlPFC, and rPar, whereas AN showed the opposite change in these regions, with LL minus SS activity being greater at Session 2 than Session 1. These differences appear to be

driven only by a decrease in LL activity in HC and by a concurrent increase in LL and decrease in SS activity in individuals with AN (Table 4.9).

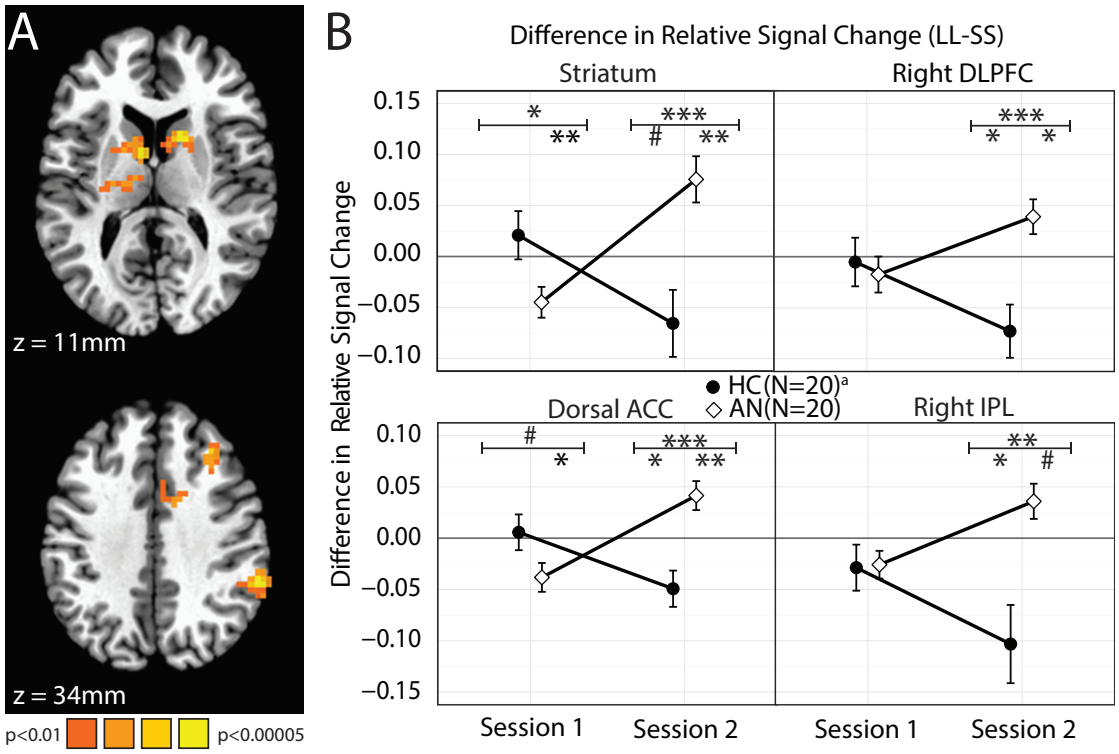

Figure 4.5 Individuals with AN have altered neural activity as compared with HC for larger-later (LL) versus smaller-sooner (SS) choices in cingulo-striatal and fronto-parietal circuitry. (A) Areas with a significant interaction effect of Choice (LL or SS), Diagnosis (AN or HC), and Session (1 or 2) (whole-brain corrected p<0.01, individual voxel threshold p<0.01, spatial extent >= 41 voxels; ST2). (B) Mean contrasts of LL minus SS choice neural activity between Diagnosis (AN or HC) and Session (1 or 2) in regions identified in the interaction effect. A positive value indicates greater neural activity when making LL choices than when making SS choices.
[a] Session 2 sample size (17AN, 14HC).
[#] p<0.10, [*] p<0.05, [**] p<0.01, [***] p<0.001 (error bars are SEM) Symbols below the horizontal bar indicate the test of the LL-SS contrast, those above indicate t-tests between diagnostic group of this contrast (Table 4.9).

Table 4.8 Activation maps of individuals with anorexia nervosa (AN) and healthy controls (HC) during the delay discounting task

Region	x	y	z	voxels	volume	*F*-stat
Interaction Effect of Choice (SS vs LL), Diagnosis (AN vs HC), and Session (1 vs 2) Cluster FWE-corrected $p < 0.01$, individual voxel threshold 0.01, size ≥ 41 voxels						
Left Putamen	+25.5	+19.5	+8.5	89	2.4 cm^3	20.3
Left Striatum	+7.5	-1.5	+11.5	75	2.0 cm^3	20.6
Right Striatum	-16.5	-7.5	+11.5	69	1.9 cm^3	21.7
Dorsal Anterior Cingulate	+7.5	-19.5	+29.5	66	1.8 cm^3	13.2
Right Parietal Cortex	-55.5	+40.5	+35.5	60	1.6 cm^3	19.6
Right Dorsolateral Prefrontal Cortex	-31.5	+25.5	+32.5	44	1.2 cm^3	14.2
Absolute Difference in Subjective Value						
Interaction Effect of Choice (SS vs LL), Diagnosis (AN vs HC), and Session (1 vs 2) Cluster FWE-correct $p < 0.01$, individual voxel threshold 0.005, size ≥ 31 voxels						
Left Striatum	+16.5	-4.5	+8.5	64	1.7 cm^3	17.2
Right Striatum	-16.5	-13.5	+5.5	52	1.4 cm^3	15.6
Left Putamen	+31.5	+16.5	-3.5	37	1.0 cm^3	14.6

Talairach-Tournoux coordinates.
FWE, family-wise error.

Table 4.9 Comparisons of the Difference in Neural Activity between Larger-Later and Smaller-Sooner Choices

Table 4.9 (Continued) Region	Test	Statistic	*p*-value
Clusters Identifying Differences by Diagnosis and Session (main test)			
Bilateral Anterior Caudate	Diagnosis:Session	$F_{1,30.9} = 48.0$	<0.0001
	Diagnosis:Session 1	$t_{38} = 2.45$	0.019
	Diagnosis:Session 2	$t_{29} = -3.66$	0.0010
	Session:AN paired	$t_{13} = -5.51$	<0.0001
	Session:HC paired	$t_{13} = 4.13$	0.0012
	AN Session 1	$t_{19} = -2.97$	0.0079
	AN Session 2	$t_{16} = 3.41$	0.0036
	HC Session 1	$t_{19} = 1.00$	ns
	HC Session 2	$t_{13} = -2.00$	0.066
Dorsal Anterior Cingulate (dACC)	Diagnosis:Session	$F_{1,31.3} = 29.7$	<0.0001
	Diagnosis:Session 1	$t_{38} = 1.92$	0.062
	Diagnosis:Session 2	$t_{29} = -4.01$	0.0004
	Session:AN paired	$t_{13} = -4.5$	0.0006
	Session:HC paired	$t_{13} = 2.51$	0.026
	AN Session 1	$t_{19} = -2.76$	0.012
	AN Session 2	$t_{16} = 2.92$	0.010
	HC Session 1	$t_{19} = 0.28$	Ns
	HC Session 2	$t_{13} = -2.73$	0.017
Right Dorsolateral Prefrontal Cortex (rdlPFC)	Diagnosis:Session	$F_{1,31.4} = 15.0$	0.0005
	Diagnosis:Session 1	$t_{38} = 0.41$	Ns
	Diagnosis:Session 2	$t_{29} = -3.71$	0.0009
	Session:AN paired	$t_{13} = -3.22$	0.0067
	Session:HC paired	$t_{13} = 2.08$	0.058
	AN Session 1	$t_{19} = -1.12$	Ns
	AN Session 2	$t_{16} = 2.31$	0.035
	HC Session 1	$t_{19} = -0.33$	Ns
	HC Session 2	$t_{13} = -2.78$	0.016
Right Inferior Parietal Cortex (rPar)	Diagnosis:Session	$F_{1,30.3} = 21.2$	<0.0001
	Diagnosis:Session 1	$t_{38} = -0.15$	Ns
	Diagnosis:Session 2	$t_{29} = -3.53$	0.0014
	Session:AN paired	$t_{13} = -2.88$	0.013
	Session:HC paired	$t_{13} = 3.27$	0.0061

Table 4.9 (Continued) Region	Test	Statistic	*p*-value
	AN Session 1	t_{19} = -1.97	0.063
	AN Session 2	t_{16} = 2.09	0.053
	HC Session 1	t_{19} = -1.32	Ns
	HC Session 2	t_{13} = -2.71	0.018
Clusters Identifying Differences by Diagnosis and Session with Absolute Difference in Subjective Value			
Right Striatum	Diagnosis:Session	$F_{1,35.8}$ = 46.8	<0.0001
	Diagnosis:Session 1	t_{38} = 3.00	0.0047
	Diagnosis:Session 2	t_{29} = -6.95	<0.0001
	Session:AN paired	t_{13} = -2.69	0.019
	Session:HC paired	t_{13} = 8.64	<0.0001
	AN Session 1	t_{19} = -1.48	0.16
	AN Session 2	t_{16} = 3.79	0.0016
	HC Session 1	t_{19} = 3.21	0.0046
	HC Session 2	t_{13} = -6.04	<0.0001
Left Striatum	Diagnosis:Session	$F_{1,36.2}$ = 47.1	<0.0001
	Diagnosis:Session 1	t_{38} = 3.40	0.0016
	Diagnosis:Session 2	t_{29} = -5.7	<0.0001
	Session:AN paired	t_{13} = -2.94	0.012
	Session:HC paired	t_{13} = 5.78	<0.0001
	AN Session 1	t_{19} = -2.01	0.058
	AN Session 2	t_{16} = 3.24	0.0051
	HC Session 1	t_{19} = 2.76	0.013
	HC Session 2	t_{13} = -4.52	0.0006
Left Putamen	Diagnosis:Session	$F_{1,37.1}$ = 38.9	<0.0001
	Diagnosis:Session 1	t_{38} = 2.82	0.0076
	Diagnosis:Session 2	t_{29} = -6.31	<0.0001
	Session:AN paired	t_{13} = -2.39	0.032
	Session:HC paired	t_{13} = 6.2	<0.0001
	AN Session 1	t_{19} = -1.38	0.18
	AN Session 2	t_{16} = 3.88	0.0013
	HC Session 1	t_{19} = 2.71	0.014
	HC Session 2	t_{13} = -5.29	0.0002

When absolute difference in subjective value between SS and LL (to account for choice difficulty) was added as a regressor in the first level

analysis, the same pattern of activity remained in the striatum, but there were no longer any group and session differences in fronto-parietal activity (Figure 4.6). This suggests that the anterior cingulate, dorsolateral prefrontal, and parietal cortex activity differences seen in the main analysis were due in part to the different changes in subjective value for each group at each time point, and perhaps therefore the differences in difficulty that were experienced across sessions. It is important to note that the main analysis likely already accounted for some of these differences in group difficulty to some degree, due to the inclusion of the duration modulated regressor of response time. Response times were negatively correlated with trial absolute difference in subjective value (r = -0.231, p < 0.0001).

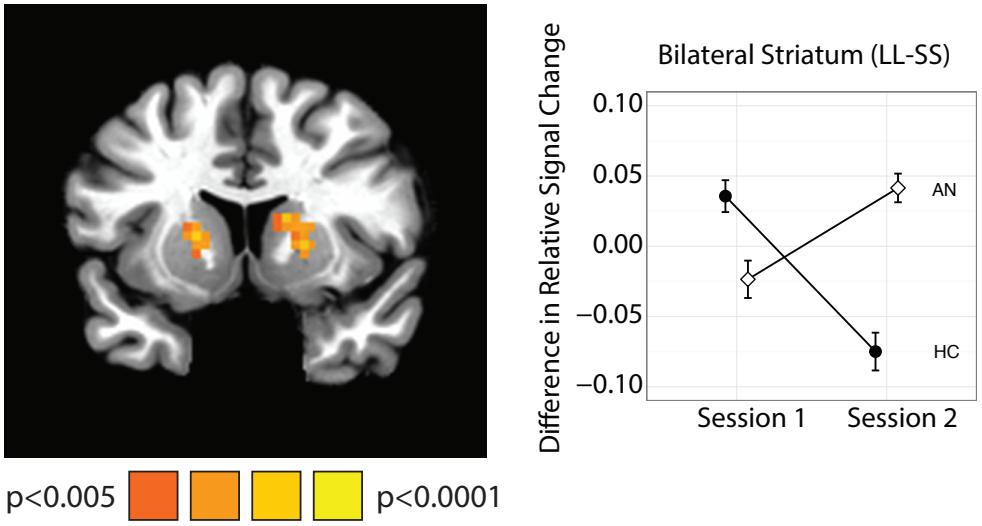

Figure 4.6 Group Analysis when the absolute difference in subjective value was included in the single subject general linear model. The diagnosis-by-session-by-choice interaction seen in the bilateral striatum reported in the main text is also present when including the regressor of absolute difference in subjective value. The dorsolateral prefrontal cortex, parietal cortex, and anterior cingulate no longer pass threshold correction.

DISCUSSION

This study provides behavioral and neural data on monetary delay discounting from a large sample of acutely ill individuals with AN, tested before and after weight restoration, as well as a comparison with healthy peers. We replicated our previous result that in the underweight state, individuals with AN discount the value of a reward over time significantly *less steeply* than healthy peers (Steinglass et al. 2012). Specifically, individuals with AN selected a larger reward delivered after a delay more often than HC, a behavior commonly interpreted as indicating self-control. Additionally, underweight AN responded more slowly when choosing the smaller, earlier options than the larger, delayed options. Once weight-restored, individuals with AN showed normalized discount rates (i.e., less tendency to delay reward as their health improved), and quickened response times when choosing earlier options, suggesting a change in how the choices are perceived.

Neural activation patterns also differed from HC, though not in the expected ways. We predicted that individuals with AN might show increased neural activity compared to HC in regions associated with executive control (e.g. dlPFC), which has been shown to subserve the tendency to choose delayed rewards among HC (Figner et al. 2010). Instead, underweight AN showed relatively less activity than HC during delayed compared to earlier choices in the dACC and striatum, regions associated with multiple aspects of cognition and behavior. After behavior normalized with weight restoration, neural activity then differed between groups, specifically with differences in the cingulo-striatal and fronto-parietal systems to delayed versus earlier choices. However, when subjective value was included, capturing an aspect of choice

difficulty for each individual, there were no observable group or session differences in this fronto-parietal circuit. In other words, results in these fronto-parietal regions vary by analytical approach (and may reflect changes in subjective difficulty that come with changes in discounting behavior); thus the most conservative interpretation would be that the main difference between groups is in striatal activity. These results suggest that phenotypic "excessive self-control" in AN might not result from executive-control circuit hyperactivity in the prefrontal cortex, but rather appears mainly associated with differences in striatal activity.

How, then, should we best understand increased "patience" in AN? The current study yields three sets of results. One, among AN in the underweight state, discount rate was abnormal, responses were slowed to the earlier choices, and cingulo-striatal activity was lower than HC during delayed choices relative to earlier choices. Two, with weight restoration, discount rate normalized and response times shifted to being faster for the earlier choices. Three, with weight restoration, neural activity in the cingulo-striatal and fronto-parietal circuits increased during delayed relative to earlier choices in the AN group, whereas activity decreased for the HC group. Taking these results together suggests a new hypothesis: the tendency to prefer larger, delayed rewards in the acutely ill state of AN may reflect a state-specific shift in decision-making. We can further speculate that acutely ill individuals with AN may be relying on choice strategies with reduced cognitive demands. While we cannot address this with the data in this study, perhaps choices amongst the underweight AN group are more habit driven (Walsh 2013), choosing to delay as a default response. Alternatively, the evaluation of delay and outcome-magnitude information might be changed (Haber and Knutson 2009;

J. Peters and Büchel 2011; Kurth-Nelson, Bickel, and Redish 2012) compared to healthy controls.

Faster response times can be an indication of a more automatic response (Schneider and Shiffrin 1977; Keramati, Dezfouli, and Piray 2011). For the AN group, response times were slower during earlier versus delayed choices when underweight, which reversed with treatment. This suggests that the delayed choice may be the default option and choosing earlier rewards required more deliberation. Considering the clinical phenomena, where delay of eating is likely rewarded initially, it may be that delay of gratification is incrementally reinforced and becomes a habitual choice (Walsh 2013), which may be amplified in the setting of starvation. This has yet to be tested in AN.

Although this interpretation is speculative, prior research suggests it is worthy of testing. Malnourishment is known to lead to many cognitive changes (Keys et al. 1950), and cognitive deficits have frequently been observed in AN (Steinglass and Glasofer 2011). Furthermore, chronic starvation in animals has been shown to alter reward processing (Carr 2002). The current data suggest that starvation may interact with the pathology of the illness to alter decision-making in ways that contribute to its entrenchment and create challenges in treatment. These data seem to differ from the reward-enhancing effects of acute hunger in food-related (Goldstone et al. 2009) and monetary paradigms (Wang and Dvorak 2010), yet may relate to the literature that shows hunger does not lead to increased risk tolerance (Levy, Thavikulwat, and Glimcher 2013). Cingulo-striatal circuits have been suggested to play a role in modulation of basic reward signals (Haber and Knutson 2009). The hypoactivity in the dACC and striatum during *delayed* choices among

underweight AN suggests a possible deficit in complex decision-making during delay discounting.

The absence of longitudinal imaging studies of delay discounting in HC makes interpreting the pattern of neural signal in HC across sessions difficult. One possibility is that neural activity decreases with task familiarity. Prior studies show variable neuroimaging results among HC, some with similarities to ours at either Session 1 or Session 2 (Kable and Glimcher 2007; Luo et al. 2009), while others show differences (Wittmann, Leland, and Paulus 2007; Christakou, Brammer, and Rubia 2011)—likely related to differences in task design. Whereas HC showed less neural activity, individuals with AN showed increased activity during larger-later choices upon repeat administration of the task. Inpatient treatment may improve health such that, after treatment, individuals with AN are able to engage in more deliberate decision-making. These cognitive processes may be necessary in AN for making consistent healthy choices. Individuals with long-term remission of AN showed no behavioral difference from HC in a recent delay discounting study (Wierenga et al. 2014), suggesting that normalized discount rates persist once weight is restored.

The majority of research on delay discounting in psychiatry has suggested that discount rates are steeper than normal in behavioral disorders (e.g., substance abuse disorders) (Koffarnus et al. 2013). AN thus appears unusual in being characterized by the opposite behavior. Additionally, one study has shown that less steep discount rates were associated with more lethal suicidal behavior (Dombrovski et al. 2011), and another reported lower rates among individuals with obsessive compulsive personality disorder (Pinto et al. 2013), a personality disorder often comorbid with AN. Our behavioral and

neuroimaging results suggest that abnormally low discount rates in AN warrant further study. For example, does discount rate relate to maladaptive food choice, and does it predict response to treatment, as seen in disorders associated with high impulsivity and steeper discounting rates (Yoon et al. 2007; MacKillop and Kahler 2009)?

The patients in this study were all receiving inpatient treatment, raising the question as to whether AN and HC differ because of context, such that monetary rewards are less valuable during inpatient treatment. However, patients showed similar behavior between cash and gift card trials, and their discount rates showed no correlation with time to discharge, which mitigates this concern. Nevertheless, we cannot rule out the possibility that discounting behavior was influenced by the prospect of leaving the inpatient setting. Our main interest was to compare AN and HC and accordingly, our study was not powered to make strong empirical conclusions regarding the AN subtypes, particularly in the fMRI sample. It may be that self-control differs between these groups, a possibility that deserves attention in future research. Discounting preferences have been shown to differ across development; as such we age-matched our groups, and there was no change to the results when age and IQ were included as covariates in the analysis.

In conclusion, these novel behavioral and brain imaging results illustrate how delay discounting differs among individuals with AN, pre- and post- treatment, compared with healthy peers. Our results suggest that self-control, as measured by a delay discounting task, is selectively altered in the acutely ill, underweight state rather than a trait-like abnormality of AN, and that this alteration is not due to heightened dorsolateral prefrontal cortex activity, as one might have expected based on previous work in healthy individuals

(Figner et al. 2010). Thus, the “iron determination” (Bruch 1978) manifested by individuals with AN is perhaps not the result of persistent executive control, a cognitively demanding approach that may be too challenging for an undernourished brain. Rather, these findings may indicate a maladaptive rule-based or automatic tendency to select the larger, delayed option when undernourished. Treatment and weight restoration may facilitate the switch to cognitively more demanding strategies. This aberrant decision-making warrants exploration specifically as related to choices about eating and suggests new directions for understanding the basic mechanisms of AN.

LIMITATIONS

Unlike the previous two chapters, this experiment consists only of a decision phase and no learning phase, and as such the computational model used is somewhat different. This is similar in nature to the analysis in Chapter 3 in which the behavior was fit only to the decisions of the test phase. The algorithm that was used for action-selection in this chapter was equivalent to those of the previous chapters. However, rather than being used to make inferences about different learning strategies from these decisions, this algorithm is used to predict an individual discount rate that leads to such decision-making. The action-selection model, just as in the previous chapters, is a fair representation of the underlying neural process of decision-making, that of contrasting the value of two options. However, the determination of the subjective value of each of the options is less representative of an underlying neural process. In this section, I discuss limitations in the discounting model used, participant subgroups, and imaging of delay discounting in general.

Limitations of the hyperbolic discounting model

The discount rate is the primary behavioral measure reported in this study, and yet there are some crucial limitations to its use. In the methods section, three alternative discounting models were presented—hyperbolic, exponential, and double exponential—all of which resulted in fits and discount rates that were highly correlated to one another. The hyperbolic model captures the phenomena of preference reversals in which the estimated value of two variably delayed rewards changes as one of them becomes immediately available (Laibson 1997; Story et al. 2014). The exponential is an internally consistent discounting function that is invariant to amount of delay. The double exponential, or quasi-hyperbolic, is also able to explain preference reversals and is often discussed using the dual-systems framework – with a habitual, steep discounting system and a goal-directed, flatter discounting system (McClure et al. 2004; Kurth-Nelson, Bickel, and Redish 2012).

For most participants, these models explain a large portion of their variance, as captured by the pseudo-R^2 measure (discussed previously in Chapter 2 limitations section). However, it must be remarked again that of the participants that had the lowest fits and were excluded due to the cutoff, all six were individuals with anorexia nervosa. We examined whether any clinical measurements could explain why this subset might be different from the rest, but did not discover anything in the attributes we measured. This might suggest that there may be some variants in AN that either could not perform the task, or chose not to. Generally, discounting has been shown to be quite stable over time (Kirby 2009), even to the point that it has been suggested that it be used as an individual trait marker (Odum 2011a). However, although this model fits behavioral data fairly well and these discount functions are able to

reflect this paradoxical preference reversal behavior, this does not mean that they provide a good cognitive model of decision-making (Kirby 1997). It is unlikely that humans have a mathematical constant from which the subjective value of certain options can be derived given reward magnitude and delay to receipt. Rather, research suggests that different areas of the brain are sensitive to reward and delay information, and it has been proposed that the relative weight given each of these signals, as well as their integration, may result in behavior that is well fit by a hyperbolic discounting model (Ballard and Knutson 2009). Such an account could also explain how various contextual and framing manipulations can bias an individual's decision-making in ways that are not appropriately accounted for in a standard hyperbolic discounting model (J. Peters and Büchel 2011; Koffarnus et al. 2013). For example, increasing the reward magnitudes, making the immediate choice the default one, performing a secondary task, and using real money all lead to increased discounting of the delayed options, and imagining or drawing attention to the future can decrease discounting (J. Peters and Büchel 2010; N. Cooper et al. 2013). Such manipulations likely change the cognitive representation of the magnitude and delay themselves, rather than some mathematical discounting function. Given that delay discounting is malleable, it is important that our analysis is able to reflect this behavior.

Searching for cognitively relevant approaches to modeling discounting

Given the limitations of using a hyperbolic discounting model as a description for the underlying evaluative process of delay discounting, the search for alternative analytic approaches is warranted. In our study, we used a generalized linear mixed effects regression analysis, often dubbed

intertemporal choice analysis (Figner et al. 2010), as an alternative method of analyzing the choice data from the delay discounting task. This analysis estimates the influence of a variety of within subject factors (e.g. magnitude and delay, relative difference in magnitude and delay, immediacy, framing, and session, and their interactions) and between subjects factors (e.g. diagnosis, age, and IQ in this case) on the probability of choosing the delayed option. Unlike the hyperbolic model, this approach provides a slightly better representation for what the brain is thought to be doing—considering the various aspects of each stimulus to reach a decision. In our study, both the hyperbolic and intertemporal analysis provided the same overall interpretation that underweight individuals with anorexia nervosa had lower discounting than healthy controls and that they discounted at a normal level once weight restored. However, the intertermporal choice analysis offered a more nuanced explanation of what leads to delay preference. As might have been expected, all participants were more likely to delay with increasing value difference and decreasing delay difference, but also with increasing magnitude of the earlier available option. This last effect has often been noted in the literature, that the discounting changes when the range of rewards is shifted, which is not predicted by the hyperbolic model (Baker, Johnson, and Bickel 2003; M. W. Johnson and Bickel 2002). Additionally, there was an indication that the AN group became more sensitive to the relative difference between the reward magnitudes of the presented options with weight restoration. Thus, certain decision attributes influence the subjective value of a delayed choice that are not accounted for in a hyperbolic discounting model. While intertemporal choice analysis does highlight important decision features, and to some extent the degree to which they are important, this analytic approach does not

provide a formalized computational model for how these decisions may be formed. Simply stating that relative difference is important for discounting is an unsatisfactory description of decision-making, and we should consider alternative models that attempt to model the underlying neural processes thought to support such behavior.

There have only been a few efforts to define new computational models that could potentially reflect a neural process that still allows for the known idiosyncrasies of delay discounting decision-making. One such effort uses a computational model akin to the model-based algorithm discussed in Chapter 2. Here, a cognitive map representing many different delay times and reward magnitudes is searched through in a step-wise fashion (Kurth-Nelson, Bickel, and Redish 2012). The probability that such an algorithm would find the given reward-delay pair in a constrained number of steps reflects behavior that resembles true discounting curves. Furthermore, by limiting the number of steps, biasing the slope of the random search, or by changing the starting point of this search algorithm, this model is able to reproduce many of the state dependent discounting effects that hyperbolic discounting fails to capture. Furthermore, this model predicts that imagining or drawing attention to the future would decrease the search distance and therefore lower the amount of discounting, and such an effect has been observed in healthy adults (J. Peters and Büchel 2010; N. Cooper et al. 2013). The random search process presented in this model is somewhat difficult to imagine being pursued by the brain; however, it is important to test and refine new models when the old ones have been shown to be incomplete, even if many will eventually be found to be inappropriate.

Concerns regarding the differences between participant samples

Although, on the whole, the study reproduced that underweight individuals with AN discounted less steeply than healthy controls (Steinglass et al. 2012), there were some differences between the subsamples. Namely, when separating the group that only performed the task behaviorally and those that performed the task in the scanner, there was no diagnosis difference in the scanner group at session one when the AN group was underweight (see Table 4.4). However, the group interaction effect and change across session effect within the AN group were still significant. Furthermore, the AN group was consistent between the scanned and behavioral groups, it was actually the healthy control group that differed between the two contexts. This is somewhat concerning, and suggests that the effect may not be as robust as we thought. In fact, a recent article failed to reproduce the previous and current result (Ritschel et al. 2015), but there were many differences between the task designs. Their task used only hypothetical rewards, they had a fixed immediate option (not presented), presented trials in a non-random fashion (enabling internal consistency), and used far fewer trials. For the presentation structure they used a highly regularized adaptive algorithm, within each of five different delay blocks, 10 trials were presented according to the following algorithm: if the immediate amount was chosen, the delayed amount increased by half the difference between the immediate and delayed rewards in the next trial and if the delayed amount was chosen, it decreased by half the difference between both rewards in the next trial. Given this pattern, participants could have learned to move to a high value range and then simply switch back and forth. However, it may also be that the result we have reported is not very robust.

In our AN sample we included both restricting and binge-purging subtypes. There is some indication in the data that these have slightly different discounting preferences, with the AN-R group showing less discounting than the AN-BP group. Additionally, once weight restored, the AN-BP group show higher discounting relative to the adults, suggesting they may become more impulsive. While there is some concern with combining subtypes, this can be difficult to avoid, especially when dealing with young patients who may not yet have crossed-over from AN-R to AN-BP (Eddy et al. 2008). This also another concern with the study discussed above, as they make no mention of what subtype distribution was present in their sample, if any. This is something that should be pursued in a larger sample to see if the differences hold, and then it should be decided whether the distinction in meaningful, both for this task and clinically.

Limitations to neuroimaging in delay discounting

Due in part to the inadequacies of modeling delay discounting behavior, there are some concerns in imaging discounting decisions that limit the interpretation of the results. Depending on the specific task structure, behavioral analysis, or imaging analysis used, neuroimaging studies of delay discounting have had varied and sometimes conflicting results and interpretations (McClure et al. 2004; Kable and Glimcher 2007). Our study, which closely matched the design of McClure et al., was unable to replicate the result that separate neural networks supported decisions for immediate options versus delayed options. This led to a major worry that an error occurred during some stage of the collection or analysis the data. However, when reanalyzing to look for subjective value signals as done by Kable and

Glimcher, we found activation in the expected regions (the medial prefrontal cortex and ventral striatum, Figure 4.7), in both groups at both sessions, with no differences between them. This suggested that participants were indeed doing the task as neural signals in value regions matched individual estimates of subjective value.

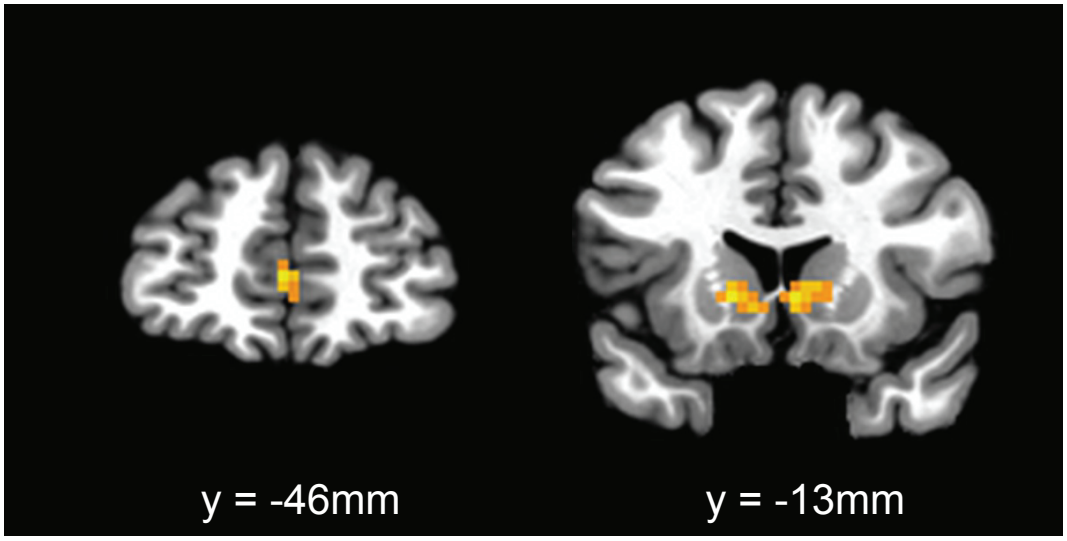

Figure 4.7 Subjective value signal of AN and HC participants from both sessions. Signal in the medial prefrontal cortex and ventral striatum correlated with an amplitude modulated regressor that tracked trial by trial subjective value of both options summed together.

In contrast, the dorsolateral prefrontal cortex has been consistently shown to be active during delayed choices (McClure et al. 2007; Kable and Glimcher 2007; J. Peters and Büchel 2011), specifically in a region overlapping with an area active during working memory tasks (Wesley and Bickel 2013). This was the basis of our hypothesis that the AN group might have heightened activity in this region. However, in our study, neither the HC nor AN group showed heightened activity in this region for delayed choice at session 1. While the lack of heightened activity in this region for individuals with anorexia nervosa, might suggest that dorsolateral regions are not

supporting heightened control, it is difficult to draw this conclusion given the lack of signal in this region for the HC group.

There are multiple difficulties in interpreting the imaging from session 2. First, when the subjective value regressor is included, there are no differences observed in the frontoparietal regions, suggesting this may simply reflect the change in perceived subjective value. Second, we see neural changes in our healthy controls across sessions with no change in behavior. As we are unaware of a longitudinal imaging study of delay discounting in healthy controls, there is limited ability to understand our observed session differences. These differences might suggest that the test retest reliability for discounting imaging findings is low, or that the network used to perform the same task changes over time, perhaps due to familiarity with the task.

Another issue with our study is that we showed and varied both options whereas most current imaging studies of delay discounting keep the immediate option at a fixed value, off screen, only showing the value of the variable delay option. The advantage to having only one changed value is that a single regressor tracks many different value signals. More specifically, a regressor for the subjective value of the delayed options (which changed trial-by-trial) also tracks sum, difference, and ratio between the both options (Kable and Glimcher 2007). This is not possible in our study, and we would need multiple regressors, that are highly collinear to get the same effect. Thus, our study was not optimized to pull out subjective value signals out of the analysis, and we had to be selective in which signal we would include as an amplitude modulated regressor. We settled on absolute-value of the difference between regressors as that might best track the relative difficulty of a decision.

These are important issues in the data analysis, and should have been worked out in a healthy population before performing the experiment on a special population. It is good practice to determine the analysis and effects of various manipulations in a healthy population first, taking extra precaution to make sure that the results are consistent with what is reported in the field, before extending a study into a special population (Stephan and Mathys 2014).

CHAPTER 5: Individual and developmental differences in decision-making: Conclusions

Advances in the study of decision-making, from psychological, neural, and computational sciences, have provided a foundation for understanding the strategies individuals employ in evaluating and choosing actions. Characterizations of these decision strategies have distinguished between a simple, habitual process and a more complex, goal-directed process, and it is proposed that behavior results from a mixture of these two strategies (Dolan and Dayan 2013). The balance between these decision-making systems may change over development, or may be disrupted in psychiatric populations, such that one system, namely habit, is generally favored. It has been appreciated that developmental and psychiatric populations differ in their ability to make goal-directed decisions, but the mechanisms that lead to these differences have not been fully elucidated (Posner and Rothbart 2000; D. Lee 2013). There is hope that insights into these underlying evaluative processes, and how they may be perturbed, can be applied to improvements in diagnosing and researching numerous overlapping psychopathologies (Insel et al. 2010), as well as their developmental basis (Casey, Oliveri, and Insel 2014). However, until there is further evidence that these analytic approaches yield constructive insights for understanding behavior across development, it is too early to draw strong conclusion from these models.

Discussion of goal-directed decision-making across development

The central question that was probed in this thesis was how the recruitment of decision-making strategies in developmental and psychiatric

populations may differ from healthy adults. In Chapter 2, “Model-based learning emerges across development,” we studied how information about a task structure was incorporated into decision-making strategies across development. The language used in this chapter was framed under the computational terminology of model-based and model-free reinforcement learning algorithms. The model-based algorithm reflects a process in which an individual makes goal-directed decisions by using a cognitive representation of the task structure and outcome information. The model-free algorithm reflects a process in which an individual makes habitual decisions, relying solely on feedback about previous actions to update and track the value of possible actions. Adaptive decision-making requires the appropriate balance of these two systems in various contexts. The central task in the study was designed with a probabilistic transition and reward structure such that habitual and goal-directed strategies, as well as a mixture of the two, would be valid approaches to the task. Whereas a model-free, habitual pattern of behavior was observed across all age groups, the recruitment of task information into a model-based, goal-directed strategy emerged in adolescence and continued to increase into adulthood. In other words, whereas adolescents and adults showed a mixture of strategies, children favored using the model-free, habitual strategy.

However, this lack of a model-based strategy in children was not due to an inability to understand the task structure. Indeed, children, like the adolescents and adults, were explicitly aware of the task structure when asked upon completion of the task and showed implicit awareness by slowing their responses after a rare transition. The reason that this knowledge went unused cannot be determined from our data; it could have been a deliberate strategy selection, it could reflect indifference to the task goals, it could reflect a

childhood tendency to rely on habitual responses, or it could be an inability to incorporate this knowledge into goal-directed actions. As recent adult studies of this task have shown that working memory load and stress lead to a reduction in the use of a model-based, goal-directed strategy (Otto, Gershman, et al. 2013; Otto, Raio, et al. 2013), working memory issues are a likely candidate for more model-free decisions in the younger participants. Indeed, it has long been observed that working-memory improves across development, possibly due to gains in capacity or efficiency (Case, Kurland, and Goldberg 1982; Casey et al. 1995; Ullman, Almeida, and Klingberg 2014). However, it is difficult to conclude that children "failed" to use a goal-directed strategy, as model-free and model-based approaches were both valid. Regardless of the underlying cause, the use of a goal-directed strategy in this task only emerged in adolescence.

In Chapter 3, "Experiential learning outweighs instruction prior to adulthood," (published 2015, Decker, Lourenco, Doll, & Hartley. 2015. *Cognitive and Affective Behavioral Neuroscience*), we explored whether the influence of instruction on experiential learning changed across development. Our task modified a common reward learning experimental design. Normally, the task consists of a learning phase in which participants choose between stimuli presented in fixed pairs and learn their relative value through positive and negative feedback, followed by a test phase in which participants choose between all possible stimulus pairs and receive no feedback, which serves as an assessment of the learned values. The key manipulation is that participants were given inaccurate instruction that one stimulus had a high value when its true value was low. In such a design, the model-based ability to represent and recruit this information may be observed in two acts of biased behavior,

following the instruction initially, and continuing to show a preference for the instructed stimulus despite experiencing contradictory feedback.

Children, adolescents, and adults were initially biased toward choosing the inaccurately instructed stimulus. However, all age groups were able to recruit experiential feedback and eventually learned to preferentially select the higher-valued stimulus for all instructed and uninstructed pairs, showing no residual instruction bias in the learning phase. During the test phase, however, behavior suggested qualitative differences across development in how instruction influenced choices. Consistent with previous findings (Doll et al., 2009; Staudinger & Buchel, 2013), we found that adults exhibited an instruction-consistent bias, suggesting that a continued representation of this inaccurate instruction distorted their feedback-based value learning. In contrast, both children and adolescents showed minimal influence of instruction on test phase performance, and two potential accounts might explain this difference. It is possible that the representation of the one-time instruction was more susceptible to the interfering effects of experiential information (Casey, Jones, and Somerville 2011), or that they integrated positive and negative feedback in a more objective, model-free manner during the learning phase. These data suggest that when explicit instruction or advice conflicts with experiential feedback about the value of an action, children and adolescents weight their own experience more heavily than the instruction. This might have important implications for how adults attempt to guide adolescent behavior regarding risky decisions (Reyna and Farley 2006), especially as public information campaigns have been shown to be largely ineffective (Ennett et al. 1994; Trenholm et al. 2007).

When these two studies are considered together, it becomes clear that the developmental emergence of model-based, goal-directed strategies differs across various types of decision-making. As the ability to represent the complexity of a decision increases, goal-directed evaluative strategies may come to predominate over habits. This pattern has been observed at various ages in many studies of cognitive development (Piaget 1954; Zelazo, Frye, and Rapus 1996; Gerstadt, Hong, and Diamond 1994; Kirkham, Cruess, and Diamond 2003; Klossek, Russell, and Dickinson 2008). However, the absence of a goal-directed strategy does not imply an inability to represent goal-relevant rules. Indeed, both in our task and others', children have been shown to be aware of the rules but fail to incorporate that knowledge into their actions (Strommen 1973; Zelazo, Frye, and Rapus 1996; Kirkham, Cruess, and Diamond 2003). Finding manipulations of the task that either promote or prevent the recruitment of task knowledge, thereby driving either goal-directed or habitual decisions, may be helpful in understanding what factors can influence the decision-making behavior of adolescents in more relevant contexts. The objective is difficult, however, as an ideal task would be simple and engaging enough that younger participants might engage in goal-directed decisions, but not too simple that adults find and use the optimal strategy. We believe that the relative recruitment of transition information is a promising target, and plan to continue exploring it from different perspectives.

Computational modeling limitations in developmental decision-making

We used computational modeling in an attempt to get closer to determining what the underlying evaluative processes of decision-making might be and how they might differ across development. No model is capable of describing

all human behavior, but the hope is that by finding a model that captures behavior sufficiently well, we can scrutinize its components as a reflection of an underlying neural process (Daw 2011). So, what was gained by using computational models in these developmental studies over more traditional behavioral analysis methods? To answer this question we first needed to determine whether the chosen modeling approach was valid by examining how well the model explained the observed data relative to a random model (Camerer and Hua Ho 1999; Daw 2011). The models from both Chapter 2 and 3 explained more variance than a random model for most participants; however, these fits improved with age, highlighting an inability of the model to capture younger participant behavior as accurately as adult behavior. Why might this be? Perhaps the wrong model is being applied to describe the younger participants' behavior, or they may be choosing more randomly (either purposefully or unintentionally) than adults, or both. Without an alternative hypothesis, however, it is difficult to generate a meaningful alternative model (Nassar and Gold 2013). Randomness may, additionally, reflect increased exploratory behavior in children and adolescents (Daw et al. 2006; Tymula et al. 2012), which can be viewed as an important aspect of learning to become independent (Spear 2000). Thus, it may be incorrect to assert that these models are false because they explain little variance, as the randomness may be an appropriate feature of their behavior. In our studies, many of the younger participants' behavior was fit equivalently well by simple reinforcement-learning models without model-based modification parameters, compared to those that included those parameters. While this could reflect the children's use of the simpler model, it is difficult to use as confirmatory

evidence that children are not using a different goal-directed, model-based strategy, especially given the low total variance explained.

Given the limited amount of variance that the models explained, it is unclear how much the resulting individual parameter estimates can be trusted. If the low level of explained variance is due to noisy or random choosing, then the parameter estimates may be fairly trustworthy, and would become increasingly so with more data, because the model still reflects the underlying structure of the data. However, if the behavior has an underlying structure that goes unexplained by a given computational model, that can lead to the systematic biasing of the parameter estimates (Nassar and Gold 2013). To an extent, running simulations and confirming that the same pattern of results emerges can help determine the quality of parameter estimates. This was done for the study in Chapter 2, and revealed that, on average, the same pattern of results was obtained when using the parameter estimates from the data. However, some participants had a low estimated value of the deterministic (inverse-temperature) parameter, which measures the degree to which decisions were deterministic or random. A simulation that used the lowest value that was estimated for that parameter would occasionally produce a pattern of behavior that did not match the behavioral profile generally observed in the tasks (a model-free, model-based, or mixed strategy). Furthermore, when higher levels of randomness (low levels of the deterministic parameter) were used, the resulting simulations were less likely to return the original input parameters. The issues of overall fit and parameter estimation limit the ability to interpret the data and draw conclusions about their meaning. We have accordingly avoided a close comparison of the parameter estimate differences between age groups, and rather focused on

the overall pattern of behavior. Nevertheless, defining and formalizing a process model that describes both learning and action-selection is an attempt to constrain the possible interpretations of more traditional behavioral analysis, and computational modeling approaches should be continued despite the limited insight observed in these studies. Computational analysis of behavior, especially for developmental behavior, is still in its early stages, and as these analyses improve, our interpretation of them will hopefully become less superficial.

Other measures that capture goal-directed behavior in development

Going forward, if the computational modeling approach is to be relevant for developmental populations, it will be important to determine how other cognitive processes, and individual traits, in addition to age, contribute to differences in decision-making. Cognitive control is proposed to support goal-directed behavior by coordinating the internal representation of those goals and associated rules, and likely coordinates with the hippocampus to support planning and prospection (Miller and Cohen 2001), and is thought to emerge in various stages across development (Munakata, Snyder, and Chatham 2012). Various measures have been shown to reflect these abilities, including but not limited to working memory capacity, future-orientation, and IQ. We failed to collect these measures in the studies presented here, but it is clear that attention and working memory play a role in goal-directed, model-based behavior (Rueda, Posner, and Rothbart 2005; Otto, Raio, et al. 2013). Therefore, we are collecting these measures in current and future studies. Alternatively, it would be useful to consider modifications to the task designs presented here, or to use altogether new tasks. For example, would changing

the reward structure such that the task favors following a model-based strategy promote more goal-directed behavior in the younger participants? By adapting the design appropriately, making it simpler, more engaging, or more rewarding to use a goal-directed strategy, it may be possible to elicit a goal-directed strategy in children, and from there more modifications could be done to determine where the breakdown from using goal-directed to habitual strategies is occurring. Testing how various measures of cognitive control and individual levels of intelligence relate to goal-directed behavior will help provide traction for understanding the evaluative processes of decision-making.

Discussion of goal-directed decision-making in anorexia nervosa

The recruitment of decision-making strategies was also examined in individuals with anorexia nervosa, a disorder that emerges in adolescence (Kaye, Fudge, and Paulus 2009). In Chapter 4, “On weight and waiting: delay discounting in anorexia nervosa pretreatment and posttreatment,” (published 2015, Decker, Figner, & Steinglass. (2015). *Biological Psychiatry)*, we addressed the perturbed decision-making in anorexia that may underlie the ability to forgo basic caloric needs in favor of the future goal of thinness. We used a delay discounting task in which an individual makes a series of choices between smaller amounts of money that are available sooner or larger amounts that are available later. This task has been used as an assessment of self-control (Steinberg et al. 2009). Showing a preference for the smaller, sooner rewards (steeper discounting) has been associated with the heightened impulsivity in various disorders (H. de Wit 2009), and a preference for the delayed rewards (less discounting) has been associated with healthier

outcomes later in life (Mischel, Shoda, and Peake 1988; Story et al. 2014). However, an extreme preference for delayed outcomes has been observed in anorexia nervosa, and when juxtaposed to severe food restriction, suggests that the underlying perturbation in decision-making may be part of the pathology of the disease (Steinglass et al. 2012). We examined individuals with anorexia nervosa as they performed this task in an MRI scanner, before and after treatment, and compared them to healthy controls to investigate the underlying circuitry of this decision process in anorexia nervosa.

Individuals with anorexia nervosa (AN) showed many changes with treatment. The underweight AN group showed significantly less discounting than the healthy controls (HC), but their discounting increased to normal levels once weight restored. This suggests that the unhealthy decision to seek distant goals may be exacerbated with worsening disease. Additionally, underweight individuals with AN responded more slowly when choosing the smaller earlier options than the larger delayed options, but showed a switch once weight restored. This may suggest that choosing the earlier options required more deliberation when underweight, but that this was facilitated with treatment. Rather than the AN group showing the expected heightened neural signal in the dorsolateral prefrontal cortex, an executive control region previously shown to support preferentially choosing delayed options (Figner et al. 2010), the AN and HC groups differed primarily in their striatal signals. The initial diminished striatal signal for delayed choices relative to sooner choices in the underweight AN group changed to heightened striatal signal for delayed relative to sooner choices once weight restored, and the HC group showed the opposite pattern, reflecting that there may be changes in reward processing before and after treatment. When the behavioral and imaging results are

considered together, they argue that the phenotypic self-control of AN may not result from canonical cognitive self-control. An alternative hypothesis is that their behavior reflects habitual decision-making, in which the initial goal-directed choice to forgo eating gets reinforced—by observed weight loss or social approval—and over time, forms a habit (Walsh 2013). Furthermore, as starvation has been associated with an increased reliance on habitual over goal-directed actions (Keys et al. 1950; Godier and Park 2014), the formation of a habit to delay rewards could be exacerbated by progressive weight loss, leading to a feedback loop in which waiting for future weight loss becomes less of a mental burden. The existence of such a feedback loop may explain the therapeutic effect of weight restoration treatment.

Interpreting phenotypic self-control behavior as a habit, when the opposite behavior is normally considered impulsive, is counterintuitive, and so this habit hypothesis must be further examined. The response time data also provide an indication that delay goes from being the default option to one that requires more deliberation. However, the idea that anorexia nervosa would lead to preferentially habitual choices is likely. As stress and working memory taxation led to more model-free behavior in the two-stage task (Otto, Raio, et al. 2013), stress has been shown to decrease the ability to shift attention, possibly due to disruptions to a prefrontal-parietal network (Liston, McEwen, and Casey 2009). Given that anorexia nervosa is a disease of high stress and anxiety (Kaye et al. 2004), it would not be surprising to find a general shift to more model-free, habit based strategies, as has been seen in other disorders or compulsivity (Voon et al. 2015). Perhaps the best approach to examining whether the delay preference is a habitual decision would be to design a task made to disentangle habitual or goal-directed decision-making from delay

preference. Many tasks, like those presented in Chapters 2 and 3, might be used to probe whether individuals with anorexia nervosa tend to use a model-free or model-based approach. However, even if such studies were to show that individuals with anorexia nervosa generally respond habitually it would not prove that their tendency to delay is also habitual. Perhaps a task could be used in which delaying is clearly a suboptimal decision. Ideally, a task could vary both the difficulty and optimality of delaying, which would allow the dissociation between habits and delay preference. That self-control may result from a form of habitual behavior is an intriguing possibility, and developing ways to probe this more closely is warranted.

Limitations in delay discounting analysis

The behavioral result that lower discounting in underweight individuals with anorexia nervosa normalized with weight restoration was reported primarily using a hyperbolic discounting model, and there are issues regarding this model's behavioral relevance. A major reason for its use is that it captures behavior well in pigeons (Ainslie 1974), rats (Reynolds, de Wit, and Richards 2002), healthy adolescent and adult humans (Green, Fry, and Myerson 1994; Steinberg et al. 2009; Odum 2011b) and those with various disorders (Bickel et al. 2007; Dombrovski et al. 2011), specifically in regards to the phenomena of preference reversal (Story et al. 2014). However, this hyperbolic model is unlikely to reflect an underlying neural process despite fitting behavior well. Although some cognitive process likely exists that considers the value and delay of each reward and leads to decisions that are correlated with hyperbolic discounting rate (Ballard and Knutson 2009), various contextual and framing manipulations have been shown to influence discounting that are not predicted

by a hyperbolic discounting model (J. Peters and Büchel 2011; Koffarnus et al. 2013). For example, increasing magnitude, making the immediate choice the default one, performing a secondary task, and using real money all make individuals discount delayed values more. Again, such manipulations may make it difficult to recruit and use information in a model-based, goal-directed fashion. To look at the importance of such factors, we performed an additional regression analysis. While the results mostly coincided with the discounting model, there was also a suggestion that individuals with anorexia nervosa were more sensitive to the relative difference in value between options after they were weight restored relative to when they were underweight. This shows that certain attributes of the decision can influence the subjective value of a delayed choice that are not accounted for in the hyperbolic discounting model.

Given the issues with the hyperbolic and similar discounting models, there are current efforts to find new computational models that reflect neural processes that could support discounting decisions. One such effort uses a computational model akin to the model-based algorithm discussed in Chapter 2. By searching through a cognitive representation of potential future events, this model measures the probability that a delayed option gets sampled as a function of time to delay, which resembles real discounting data (Kurth-Nelson, Bickel, and Redish 2012). Furthermore, this model predicts that imagining or drawing attention to the future would decrease the search distance and therefore lower the amount of discounting, and such an effect has been observed in healthy adults (J. Peters and Büchel 2010; N. Cooper et al. 2013). Perhaps the emphasis that individuals with anorexia nervosa place on the future draws their attention more to the delayed option. As the use of the hyperbolic discounting model is limited in its ability to explain the underlying

neural evaluative processes, it is important to consider alternative computational models if progress is to be made. The testing and refining of new models is required when the old ones are shown to be incomplete, even if many will eventually be found to be inappropriate.

Due in part to the inadequacies of modeling delay discounting behavior, there are some concerns in imaging discounting decisions that limit the interpretation of these results as well. Depending on the specific task structure, behavioral analysis, or imaging analysis used, neuroimaging studies of delay discounting have had varied and sometimes conflicting results and interpretations (McClure et al. 2004; Kable and Glimcher 2007). Our study, which closely matched the design of McClure et al., was unable to replicate that study's result that separate neural networks supported decisions for immediate options versus delayed options. In contrast, the dorsolateral prefrontal cortex has been consistently shown to be active during delayed choices (McClure et al. 2007; Kable and Glimcher 2007; J. Peters and Büchel 2011), specifically in a region overlapping with an area active during working memory tasks (Wesley and Bickel 2013). The lack of heightened activity in this region for individuals with anorexia nervosa, given that it is so commonly seen to be involved, would suggest that some other mechanism is likely involved in anorexia nervosa. However, it is difficult to be confident in this conclusion given that the healthy controls in our study also failed to show activity in this region. A large issue with our study is that we showed and varied both options, whereas most current imaging studies of delay discounting keep the immediate option at a fixed value and off screen, only showing the delay and value of the other option. This is problematic because there is collinearity between factors that track the changing value and subjective value of each

option, which is ameliorated by a simpler task design with only one changing value. Another issue is that we see neural changes in our healthy controls across sessions with no change in behavior. As we are unaware of a longitudinal imaging study of delay discounting in healthy controls, there is limited ability to understand our observed session differences. These differences might suggest that the test retest reliability for discounting imaging findings is low, or that the network used to perform the same task changes over time, perhaps due to familiarity with the task. These are important issues, and before an imaging study is to be performed in a special population, it would be good practice to determine the analysis and effects of various manipulations in a healthy population first, taking extra precaution to make sure that the results are consistent with what is reported in the field (Stephan and Mathys 2014).

Future directions for using computational modeling in developmental and psychiatric populations

There are important limitations when using computational modeling to understand decision-making in both developmental and psychiatric populations, but there may be certain ways in which modeling may still prove useful. The biggest limitation comes from trying to extend a theoretical model about the underlying processes of decision-making that was developed in healthy adults, and is still being refined, to populations that are known to differ in that decision domain. Often, as is the case for the developmental studies presented in this thesis, the finding is that younger participants are not using the same model-based evaluative process as adults, and it is then assumed that they are using a simpler default model. It is also possible, however, that

they are using a different cognitive representation of the task to make goal-directed decisions that the model fails to capture (Nassar and Gold 2013). Even if it is the case that children use a simpler model than adults, this observation is unsurprising and does not in itself offer suggestions as to how we might guide children and adolescents to improve their decision-making, nor inform how the transition to more goal-directed decision-making occurs. However, if the framework of the evaluative process is sufficiently correct, testing whether task manipulations exist that could push decision-making closer or farther from that of adults may generalize to more relevant contexts. Similarly, tasks that were developed for children and adolescents could receive similar computational attention. Because computational modeling provides a formalized operation by which an individual learns, evaluates, and makes a decision, it can offer predictions as to how a specific manipulation might change behavior. It is worthwhile to pursue many lines of inquiry in the hope that we might gain traction in the understanding of the underlying evaluative processes of decision-making.

Another hope of computational neuroscience is that these methods can be used to extend the knowledge of the function that various neural networks may play in decision-making. The ventral striatum had long been known to be involved in reward processing (Taylor and Robbins 1986), but neurophysiological and neuroimaging studies extending computational reinforcement learning models provided evidence that it does so by tracking the difference between the reward that was received and what was predicted (Schultz, Dayan, and Montague 1997; Elliott et al. 2000; Pagnoni et al. 2002). This extension, from knowledge of involvement to evidence of its functional role, fueled research which led to a more nuanced understanding of reward

based learning (Schultz 2013), and how this learning signal may differ in developmental (Galvan et al. 2006; J. R. Cohen et al. 2010) and psychiatric (Huys et al. 2014; G. K. W. Frank et al. 2012) populations. This is an example of the promise that the computational study of decision-making offers. As the networks that are involved in various types of decision-making and their functions are resolved, perhaps it will become feasible to describe how the process gets disrupted in a psychiatric disorder, and how best to treat it (van der Meer, Kurth-Nelson, and Redish 2012).

The current initiative by the National Institutes of Mental Health is to find new ways of diagnosing and researching psychiatric disorders that depend less on classification by symptom clustering and more on genetic, psychological, and neuroimaging criteria that are thought to be involved in the underlying pathologies of psychiatric diseases (Insel et al. 2010; Cuthbert and Insel 2013). Furthermore, as adolescence is the stage when most psychiatric disorders first emerge (F. S. Lee et al. 2014), it is critical to understand the developmental trajectory of the behaviors and corresponding neural circuitry that are being proposed to be researched in this initiative (Casey, Oliveri, and Insel 2014). The current diagnosis system is imprecise and this imprecision is likely a contributor to the lack of progress over the past few decades in the development of new treatments for psychiatric disorders, and the hope is that more precise diagnoses could lead to “individualized” treatment. As such, there has been much interest in using computational studies of decision-making to investigate pathophysiological processes and their relation to behavior (Maia and Frank 2011; Montague et al. 2012; D. Lee 2013; van der Meer, Kurth-Nelson, and Redish 2012; Stephan and Mathys 2014). Reinforcement-learning paradigms have received particular focus due to the

well-studied role that dopamine plays both in reward learning and various psychiatric disorders (Maia and Frank 2011), but the same approach should prove useful for other neurocognitive processes and disorders. This new initiative calls for research that spans across multiple psychiatric disorders by focusing on how they may be linked by perturbations in common neurocognitive domains. However, such an approach should be equally applicable to the study of the development of decision-making, which may prove especially useful as adolescence is the time when many psychiatric disorders develop (Kessler et al. 2005; Somerville, Jones, and Casey 2010; F. S. Lee et al. 2014). An important component in this research initiative is that the cognitive construct being examined have high construct validity to a specific behavioral function, and that there is strong evidence that this function is reflected in specific neural substrate (Cuthbert and Insel 2013). Thus, if the goal is to study decision-making when it is not fully developed or has gone awry, we should be using tasks that have been well characterized in developmental or psychiatric populations, or carefully modifying tasks that are highly trusted in healthy adults, so as to be confident that a known process is being studied. In this thesis, we used various decision-making tasks and computational modeling in an effort to determine whether they were useful in describing differences in decision-making across development and in a psychiatric population. These approaches have provided a preliminary indication of developmental and psychiatric differences in decision-making processes, and by further refining the computational models as well as drawing connections from research in other cognitive domains, it may be possible to better characterize the computational nature of these differences.

REFERENCES

Ainslie, G W. 1974. “Impulse Control in Pigeons.” *Journal of the Experimental Analysis of Behavior* 21 (3) (May): 485–9.

Akaike, H. 1974. “A New Look at the Statistical Model Identification.” *IEEE Transactions on Automatic Control* 19. doi:10.1109/TAC.1974.1100705.

American Psychiatric Association. 2013. “Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.” *American Psychiatric Association*.

American-Psychiatric-Association. 2013. “Diagnostic and Statistical Manual of Mental Disorders (5th Ed.).” *American Psychiatric Association*.

Amso, Dima, and Juliet Davidow. 2012. “The Development of Implicit Learning from Infancy to Adulthood: Item Frequencies, Relations, and Cognitive Flexibility.” *Developmental Psychobiology* 54: 664–673. doi:10.1002/dev.20587.

Anderson, John R. 1990. *Cognitive Psychology and Its Implications.* WH Freeman/Times Books/Henry Holt & Co.

Arcelus, Jon, and AJ Mitchell. 2011. “Mortality Rates in Patients with Anorexia Nervosa and Other Eating Disorders: A Meta-Analysis of 36 Studies.” *Archives of General Psychiatry* 68 (7): 724–731.

Arnsten, Amy F T. 2009. “Stress Signalling Pathways That Impair Prefrontal Cortex Structure and Function.” *Nature Reviews. Neuroscience* 10 (juNe): 410–422. doi:10.1038/nrn2648.

Aron, Adam R, and Russell a Poldrack. 2006. “Cortical and Subcortical Contributions to Stop Signal Response Inhibition: Role of the Subthalamic Nucleus.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 26 (9): 2424–2433. doi:10.1523/JNEUROSCI.4682-05.2006.

Ashby, F G, L a Alfonso-Reese, a U Turken, and E M Waldron. 1998. “A Neuropsychological Theory of Multiple Systems in Category Learning.” *Psychological Review* 105 (3): 442–481.

Attia, E, C Haiman, B T Walsh, and S R Flater. 1998. “Does Fluoxetine Augment the Inpatient Treatment of Anorexia Nervosa?” *The American Journal of Psychiatry* 155 (4) (April): 548–51.

Ayduk, Ozlem, Rodolfo Mendoza-Denton, Walter Mischel, Geraldine Downey, Philip K. Peake, and Monica Rodriguez. 2000. “Regulating the Interpersonal Self: Strategic Self-Regulation for Coping with Rejection Sensitivity.” *Journal of Personality and Social Psychology* 79 (5): 776–792. doi:10.1037//0022-3514.79.5.776.

Baddeley, Alan. 2010. “Working Memory.” *Current Biology* 20: R136–R140. doi:10.1016/j.cub.2009.12.014.

Baddeley, Alan, and Graham Hitch. 1974. “Working Memory.” *Psychology of Learning and Motivation*. doi:10.1016/j.cub.2009.12.014.

Baker, Forest, Matthew W Johnson, and Warren K Bickel. 2003. “Delay Discounting in Current and Never-before Cigarette Smokers: Similarities and Differences across Commodity, Sign, and Magnitude.” *Journal of Abnormal Psychology* 112 (3): 382–392. doi:10.1037/0021-843X.112.3.382.

Ballard, Kacey, and Brian Knutson. 2009. “Dissociable Neural Representations of Future Reward Magnitude and Delay during Temporal Discounting.” *NeuroImage* 45 (1): 143–150. doi:http://dx.doi.org/10.1016/j.neuroimage.2008.11.004.

Balleine, Bernard W, Mauricio R Delgado, and Okihide Hikosaka. 2007. “The Role of the Dorsal Striatum in Reward and Decision-Making.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 27 (31): 8161–8165. doi:10.1523/JNEUROSCI.1554-07.2007.

Balleine, Bernard W, and John P O’Doherty. 2009. “Human and Rodent Homologies in Action Control: Corticostriatal Determinants of Goal-Directed and Habitual Action.” *Neuropsychopharmacology* 35 (1): 48–69.

Barkley, R a, G Edwards, M Laneri, K Fletcher, and L Metevia. 2001. “Executive Functioning, Temporal Discounting, and Sense of Time in Adolescents with Attention Deficit Hyperactivity Disorder (ADHD) and Oppositional Defiant Disorder (ODD).” *Journal of Abnormal Child Psychology* 29 (6) (December): 541–56.

Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. “Random Effects Structure for Confirmatory Hypothesis Testing: Keep It Maximal.” *Journal of Memory and Language* 68 (3) (April): 255–278. doi:10.1016/j.jml.2012.11.001.

Bates, Douglas, Martin Maechler, Ben Bolker, and Steven Walker. 2014. "lme4: Linear Mixed-Effects Models Using Eigen and S4." *R Package Version 1.1-6*.

Bayer, Hannah M, and Paul W Glimcher. 2005. "Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal." *Neuron* 47 (1): 129–141.

Bechara, Antoine. 2005. "Decision Making, Impulse Control and Loss of Willpower to Resist Drugs: A Neurocognitive Perspective." *Nature Neuroscience* 8 (11): 1458–1463. doi:10.1038/nn1584.

Bechara, Antoine, H Damasio, and a R Damasio. 2000. "Emotion, Decision Making and the Orbitofrontal Cortex." *Cerebral Cortex (New York, N.Y. : 1991)* 10: 295–307. doi:10.1093/cercor/10.3.295.

Benhabib, Jess, Alberto Bisin, and Andrew Schotter. 2010. "Present-Bias, Quasi-Hyperbolic Discounting, and Fixed Costs." *Games and Economic Behavior* 69 (2) (July): 205–223. doi:10.1016/j.geb.2009.11.003.

Benoit, Roland G, Sam J Gilbert, and Paul W Burgess. 2011. "A Neural Mechanism Mediating the Impact of Episodic Prospection on Farsighted Decisions." *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 31 (18) (May 4): 6771–9. doi:10.1523/JNEUROSCI.6559-10.2011.

Bickel, Warren K, Michelle L Miller, Richard Yi, Benjamin P Kowal, Diana M Lindquist, and Jeffery a Pitcock. 2007. "Behavioral and Neuroeconomics of Drug Addiction: Competing Neural Systems and Temporal Discounting Processes." *Drug and Alcohol Dependence* 90 Suppl 1 (September): S85–91. doi:10.1016/j.drugalcdep.2006.09.016.

Bickel, Warren K, Amy L Odum, and G J Madden. 1999. "Impulsivity and Cigarette Smoking: Delay Discounting in Current, Never, and Ex-Smokers." *Psychopharmacology* 146 (4) (October): 447–54.

Biele, Guido, Jörg Rieskamp, and Richard Gonzalez. 2009. "Computational Models for the Combination of Advice and Individual Learning." *Cognitive Science* 33 (2) (March): 206–42. doi:10.1111/j.1551-6709.2009.01010.x.

Biele, Guido, Jörg Rieskamp, Lea K Krugel, and Hauke R Heekeren. 2011. "The Neural Basis of Following Advice." *PLoS Biology* 9 (6) (June): e1001089. doi:10.1371/journal.pbio.1001089.

Bischoff-Grethe, Amanda, Kelly M Goedert, Daniel T Willingham, and Scott T Grafton. 2004. “Neural Substrates of Response-Based Sequence Learning Using fMRI.” *Journal of Cognitive Neuroscience* 16: 127–38. doi:10.1162/089892904322755610.

Bischoff-Grethe, Amanda, Danyale McCurdy, Emily Grenesko-Stevens, Laura E Zoe Irvine, Angela Wagner, Wai-Ying Wendy Yau, Christine Fennema-Notestine, et al. 2013. “Altered Brain Response to Reward and Punishment in Adolescents with Anorexia Nervosa.” *Psychiatry Research* 214 (3) (December 30): 331–40. doi:10.1016/j.pscychresns.2013.07.004.

Bocanegra, Bruno R, and Bernhard Hommel. 2014. “When Cognitive Control Is Not Adaptive.” *Psychological Science* 25 (6) (April 23): 1249–1255. doi:10.1177/0956797614528522.

Bornstein, Aaron M, and Nathaniel D Daw. 2011. “Multiplicity of Control in the Basal Ganglia: Computational Roles of Striatal Subregions.” *Current Opinion in Neurobiology* 21 (3) (June): 374–80. doi:10.1016/j.conb.2011.02.009.

Bornstein, Aaron M., and Nathaniel D. Daw. 2012. “Dissociating Hippocampal and Striatal Contributions to Sequential Prediction Learning.” *European Journal of Neuroscience* 35 (May 2011): 1011–1023. doi:10.1111/j.1460-9568.2011.07920.x.

Botvinick, M M, T S Braver, D M Barch, C S Carter, and J D Cohen. 2001. “Conflict Monitoring and Cognitive Control.” *Psychological Review* 108 (3): 624–652. doi:10.1037/0033-295X.108.3.624.

Brand, Matthias, Elke Kalbe, Kirsten Labudda, Esther Fujiwara, Josef Kessler, and Hans J. Markowitsch. 2005. “Decision-Making Impairments in Patients with Pathological Gambling.” *Psychiatry Research* 133: 91–99. doi:10.1016/j.psychres.2004.10.003.

Brovelli, Andrea, Bruno Nazarian, Martine Meunier, and Driss Boussaoud. 2011. “Differential Roles of Caudate Nucleus and Putamen during Instrumental Learning.” *NeuroImage* 57 (4) (August 15): 1580–90. doi:10.1016/j.neuroimage.2011.05.059.

Bruch, Hilde. 1978. *The Golden Cage: The Enigma of Anorexia Nervosa*. New York, Vintage Books.

Bunge, Silvia A, and Philip David Zelazo. 2006. “A Brain-Based Account of the Development of Rule Use in Childhood.” *Current Directions in*

Psychological Science 15 (3): 118–121. doi:10.1111/j.0963-7214.2006.00419.x.

Camerer, Colin, and Teck Hua Ho. 1999. “Experience-Weighted Attraction Learning in Normal Form Games.” *Econometrica* 67 (March 1997): 827–874. doi:10.1111/1468-0262.00054.

Cardinal, R N, D R Pennicott, C L Sugathapala, T W Robbins, B J Everitt, and C L Lakmali. 2001. “Impulsive Choice Induced in Rats by Lesions of the Nucleus Accumbens Core.” *Science* 292: 2499–2501. doi:10.1126/science.1060818.

Carr, KD. 2002. “Augmentation of Drug Reward by Chronic Food Restriction: Behavioral Evidence and Underlying Mechanisms.” *Physiology & Behavior* 76: 353–364.

Carter, R. McKell, Justin R. Meyer, and Scott a. Huettel. 2010. “Functional Neuroimaging of Intertemporal Choice Models: A Review.” *Journal of Neuroscience, Psychology, and Economics* 3 (1): 27–45. doi:10.1037/a0018046.

Case, Robbie, D.Midian Kurland, and Jill Goldberg. 1982. “Operational Efficiency and the Growth of Short-Term Memory Span.” *Journal of Experimental Child Psychology* 33: 386–404. doi:10.1016/0022-0965(82)90054-6.

Casey, BJ. 2005. “Frontostriatal and Frontocerebellar Circuitry Underlying Cognitive Control.” In *Developing Individuality in the Human Brain: A Tribute to Michael I. Posner. Decade of Behavior*, edited by Ulrich Mayr, Edward Awh, and Steven W Keele. Washington DC: American Psychological Association.

Casey, BJ. 2014. “Beyond Simple Models of Self-Control to Circuit-Based Accounts of Adolescent Behavior.” *Annual Review of Psychology*: 1–25. doi:10.1146/annurev-psych-010814-015156.

Casey, BJ, J D Cohen, P Jezzard, R Turner, D C Noll, R J Trainor, J Giedd, D Kaysen, L Hertz-Pannier, and J L Rapoport. 1995. “Activation of Prefrontal Cortex in Children during a Nonspatial Working Memory Task with Functional MRI.” *NeuroImage*. doi:10.1006/nimg.1995.1029.

Casey, BJ, Rebecca M Jones, Liat Levita, Victoria Libby, Siobhan Pattwell, Erika Ruberry, Fatima Soliman, and Leah H Somerville. 2010. “The Storm and Stress of Adolescence: Insights from Human Imaging and Mouse

Genetics." *Developmental Psychobiology* 52 (3): 225–235. doi:10.1002/dev.20447.The.

Casey, BJ, Rebecca M. Jones, and Leah H Somerville. 2011. "Braking and Accelerating of the Adolescent Brain." *Journal of Research on Adolescence* 21 (1): 21–33. doi:10.1111/j.1532-7795.2010.00712.x.

Casey, BJ, Mary Ellen Oliveri, and Thomas Insel. 2014. "A Neurodevelopmental Perspective on the Research Domain Criteria (RDoC) Framework." *Biological Psychiatry* 76 (5): 350–353. doi:10.1016/j.biopsych.2014.01.006.

Casey, BJ, Leah H Somerville, Ian H Gotlib, Ozlem Ayduk, Nicholas T Franklin, Mary K Askren, John Jonides, et al. 2011. "Behavioral and Neural Correlates of Delay of Gratification 40 Years Later." *Proceedings of the National Academy of Sciences of the United States of America* 108 (36) (September 6): 14998–5003. doi:10.1073/pnas.1108561108.

Chabris, Christopher F, David Laibson, Carrie L Morris, Jonathon P Schuldt, and Dmitry Taubinsky. 2008. "Individual Laboratory-Measured Discount Rates Predict Field Behavior." *Journal of Risk and Uncertainty* 37 (2-3) (December 1): 237–269. doi:10.1007/s11166-008-9053-x.

Chechik, G, I Meilijson, and E Ruppin. 1998. "Synaptic Pruning in Development: A Computational Account." *Neural Computation* 10: 1759–1777. doi:10.1162/089976698300017124.

Chein, Jason, Dustin Albert, Lia O Brien, Kaitlyn Uckert, and Laurence Steinberg. 2012. "Peers Increase Adolescent Risk Taking by Enhancing Activity in the Brain's Reward Circuitry." *Developmental Science* 14 (2): 1–16. doi:10.1111/j.1467-7687.2010.01035.x.Peers.

Chen, Gang, Ziad S Saad, Jennifer C Britton, Daniel S Pine, and Robert W Cox. 2013. "Linear Mixed-Effects Modeling Approach to FMRI Group Analysis." *NeuroImage* 73: 176–190.

Chikazoe, Junichi, Seiki Konishi, Tomoki Asari, Koji Jimura, and Yasushi Miyashita. 2007. "Activation of Right Inferior Frontal Gyrus during Response Inhibition across Response Modalities." *Journal of Cognitive Neuroscience* 19: 69–80. doi:10.1162/jocn.2007.19.1.69.

Christakou, Anastasia, Mick Brammer, and Katya Rubia. 2011. "Maturation of Limbic Corticostriatal Activation and Connectivity Associated with Developmental Changes in Temporal Discounting." *NeuroImage* 54 (2) (January 15): 1344–54. doi:10.1016/j.neuroimage.2010.08.067.

Clarke, Hannah F, Trevor W Robbins, and Angela C Roberts. 2008. “Lesions of the Medial Striatum in Monkeys Produce Perseverative Impairments during Reversal Learning Similar to Those Produced by Lesions of the Orbitofrontal Cortex.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 28 (43): 10972–10982. doi:10.1523/JNEUROSCI.1521-08.2008.

Cohen, J D, W M Perlstein, T S Braver, L E Nystrom, D C Noll, J Jonides, and E E Smith. 1997. “Temporal Dynamics of Brain Activation during a Working Memory Task.” *Nature* 386: 604–608. doi:10.1038/386604a0.

Cohen, Jessica R, Robert F Asarnow, Fred W Sabb, Robert M Bilder, Susan Y Bookheimer, Barbara J Knowlton, and Russell A Poldrack. 2010. “A Unique Adolescent Response to Reward Prediction Errors.” *Nature Neuroscience* 13 (6): 669–671.

Cohen, P, J Cohen, and J Brook. 1993. “An Epidemiological Study of Disorders in Late Childhood and Adolescence--II. Persistence of Disorders.” *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 34 (6): 869–877. doi:10.1111/j.1469-7610.1993.tb01095.x.

Cools, Roshan. 2011. “Dopaminergic Control of the Striatum for High-Level Cognition.” *Current Opinion in Neurobiology* 21 (3) (June): 402–7. doi:10.1016/j.conb.2011.04.002.

Cooper, N, J W Kable, B K Kim, and G Zauberman. 2013. “Brain Activity in Valuation Regions While Thinking about the Future Predicts Individual Discount Rates.” *Journal of Neuroscience* 33 (32): 13150–13156.

Cooper, Zafra, and Christopher Fairburn. 1987. “The Eating Disorder Examination: A Semi-structured Interview for the Assessment of the Specific Psychopathology of Eating Disorders.” *International Journal of Eating Disorders* 8: 1–8.

Corbetta, Maurizio, and Gordon L Shulman. 2002. “Control of Goal-Directed and Stimulus-Driven Attention in the Brain.” *Nature Reviews. Neuroscience* 3 (March): 201–215. doi:10.1038/nrn755.

Corbit, Laura H., and Bernard W. Balleine. 2003. “The Role of Prelimbic Cortex in Instrumental Conditioning.” *Behavioural Brain Research* 146: 145–157. doi:10.1016/j.bbr.2003.09.023.

Cox, R W. 1996. “AFNI: Software for Analysis and Visualization of Functional Magnetic Resonance Neuroimages.” *Computers and Biomedical Research, an International Journal* 29 (3) (June): 162–73.

Crone, Eveline a, L Bullens, E a a van der Plas, E J Kijkuit, and P D Zelazo. 2008. “Developmental Changes and Individual Differences in Risk and Perspective Taking in Adolescence.” *Development and Psychopathology* 20 (4) (January): 1213–29. doi:10.1017/S0954579408000588.

Crone, Eveline a, Carter Wendelken, Sarah Donohue, Linda van Leijenhorst, and Silvia a Bunge. 2006. “Neurocognitive Development of the Ability to Manipulate Information in Working Memory.” *Proceedings of the National Academy of Sciences of the United States of America* 103 (24): 9315–9320. doi:10.1073/pnas.0510088103.

Crone, Eveline a., and Ronald E. Dahl. 2012. “Understanding Adolescence as a Period of Social–affective Engagement and Goal Flexibility.” *Nature Reviews Neuroscience* 13 (September): 636–650. doi:10.1038/nrn3313.

Curtis, C. E. 2006. “Prefrontal and Parietal Contributions to Spatial Working Memory.” *Neuroscience* 139: 173–180. doi:10.1016/j.neuroscience.2005.04.070.

Cuthbert, Bruce N, and Thomas R Insel. 2013. “Toward the Future of Psychiatric Diagnosis: The Seven Pillars of RDoC.” *BMC Medicine* 11: 126. doi:10.1186/1741-7015-11-126.

Dalley, Jeffrey W., Barry J. Everitt, and Trevor W. Robbins. 2011. “Impulsivity, Compulsivity, and Top-Down Cognitive Control.” *Neuron* 69 (4): 680–694. doi:10.1016/j.neuron.2011.01.020.

Daniel, Reka, and Stefan Pollmann. 2010. “Comparing the Neural Basis of Monetary Reward and Cognitive Feedback during Information-Integration Category Learning.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 30 (1): 47–55. doi:10.1523/JNEUROSCI.2205-09.2010.

Dauer, William, and Serge Przedborski. 2003. “Parkinson’s Disease.” *Neuron* 39: 889–909. doi:10.1016/S0896-6273(03)00568-3.

Davidson, M. C., J. C. Horvitz, N. Tottenham, J. a. Fossella, R. Watts, a. M. Uluğ, and BJ Casey. 2004. “Differential Cingulate and Caudate Activation Following Unexpected Nonrewarding Stimuli.” *NeuroImage* 23: 1039–1045. doi:10.1016/j.neuroimage.2004.07.049.

Daw, Nathaniel D. 2011. “Trial-by-Trial Data Analysis Using Computational Models.” *Decision Making, Affect, and Learning: Attention and Performance XXIII*: 555. doi:10.1093/acprof:oso/9780199600434.003.0001.

Daw, Nathaniel D, Samuel J Gershman, Ben Seymour, Peter Dayan, and Raymond J Dolan. 2011. “Model-Based Influences on Humans’ Choices and Striatal Prediction Errors.” *Neuron* 69 (6): 1204–1215.

Daw, Nathaniel D, Yael Niv, and Peter Dayan. 2005. “Uncertainty-Based Competition between Prefrontal and Dorsolateral Striatal Systems for Behavioral Control.” *Nature Neuroscience* 8 (12): 1704–1711.

Daw, Nathaniel D, John P O’Doherty, Peter Dayan, Ben Seymour, and Raymond J Dolan. 2006. “Cortical Substrates for Exploratory Decisions in Humans.” *Nature* 441 (June): 876–879. doi:10.1038/nature04766.

Daw, Nathaniel D, and Daphna Shohamy. 2008. “The Cognitive Neuroscience of Motivation and Learning.” *Social Cognition* 26 (5): 593–620.

De Wit, Harriet. 2009. “Impulsivity as a Determinant and Consequence of Drug Use: A Review of Underlying Processes.” *Addiction Biology* 14 (1) (January): 22–31. doi:10.1111/j.1369-1600.2008.00129.x.

De Wit, Sanne, Holly R. Standing, Elise E. Devito, Oliver J. Robinson, K. Richard Ridderinkhof, Trevor W. Robbins, and Barbara J. Sahakian. 2012. “Reliance on Habits at the Expense of Goal-Directed Control Following Dopamine Precursor Depletion.” *Psychopharmacology* 219: 621–631. doi:10.1007/s00213-011-2563-2.

Decker, Johannes H, Frederico S Lourenco, Bradley B Doll, and Catherine A Hartley. 2015. “Experiential Reward Learning Outweighs Instruction prior to Adulthood.” *Cognitive, Affective & Behavioral Neuroscience* (January 13). doi:10.3758/s13415-014-0332-5.

Delazer, M., H. Sinz, L. Zamarian, and T. Benke. 2007. “Decision-Making with Explicit and Stable Rules in Mild Alzheimer’s Disease.” *Neuropsychologia* 45 (8): 1632–1641. doi:10.1016/j.neuropsychologia.2007.01.006.

Den Ouden, Hanneke E M, Nathaniel D Daw, Guillén Fernandez, Joris a Elshout, Mark Rijpkema, Martine Hoogman, Barbara Franke, and Roshan Cools. 2013. “Dissociable Effects of Dopamine and Serotonin on Reversal Learning.” *Neuron* 80: 1090–100. doi:10.1016/j.neuron.2013.08.030.

Deserno, Lorenz, Quentin J M Huys, Rebecca Boehme, Ralph Buchert, and Hans-jochen Heinze. 2015. “Ventral Striatal Dopamine Reflects Behavioral and Neural Signatures of Model-Based Control during Sequential Decision Making.” *Proceedings of the National Academy of Sciences* 112 (5): 1595–1600. doi:10.1073/pnas.1417219112.

Diamond, Adele. 2006. “The Early Development of Executive Functions.” *Lifespan Cognition: Mechanisms of Change*: 70–95.

Diamond, Adele, Natasha Kirkham, and Dima Amso. 2002. “Conditions under Which Young Children Can Hold Two Rules in Mind and Inhibit a Prepotent Response.” *Developmental Psychology* 38 (3): 352–362. doi:10.1037//0012-1649.38.3.352.

Dias-ferreira, Eduardo, João C Sousa, Irene Melo, Pedro Morgado, Ana R Mesquita, João J Cerqueira, Rui M Costa, and Nuno Sousa. 2009. “Chronic Stress Causes Frontostriatal Reorganization and Affects Decision-Making.” *Science* 325: 621–625.

Dickinson, Anthony. 1985. “Actions and Habits: The Development of Behavioural Autonomy.” *Phiolosophical Transactions of the Royal Society of London* 308: 67–78.

Dickinson, Anthony. 1994. “Instrumental Conditioning.” In *Animal Learning and Cognition*, 45–79.

Dijksterhuis, Ap, Maarten W Bos, Loran F Nordgren, and Rick B van Baaren. 2006. “On Making the Right Choice: The Deliberation-without-Attention Effect.” *Science (New York, N.Y.)* 311: 1005–1007. doi:10.1126/science.1121629.

Dolan, Ray J., and Peter Dayan. 2013. “Goals and Habits in the Brain.” *Neuron* 80 (2): 312–325. doi:10.1016/j.neuron.2013.09.007.

Doll, Bradley B, Kent E Hutchison, and Michael J Frank. 2011. “Dopaminergic Genes Predict Individual Differences in Susceptibility to Confirmation Bias.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 31 (16) (April 20): 6188–98. doi:10.1523/JNEUROSCI.6486-10.2011.

Doll, Bradley B, W Jake Jacobs, Alan G Sanfey, and Michael J Frank. 2009. “Instructional Control of Reinforcement Learning: A Behavioral and Neurocomputational Investigation.” *Brain Research* 1299 (November 24): 74–94. doi:10.1016/j.brainres.2009.07.007.

Doll, Bradley B, Dylan a Simon, and Nathaniel D Daw. 2012. “The Ubiquity of Model-Based Reinforcement Learning.” *Current Opinion in Neurobiology* 22 (6) (December): 1075–81. doi:10.1016/j.conb.2012.08.003.

Doll, Bradley B, James A Waltz, Jeffrey Cockburn, Jaime K Brown, Michael J Frank, and James M Gold. 2014. “Reduced Susceptibility to Confirmation

Bias in Schizophrenia." *Cognitive, Affective & Behavioral Neuroscience* (January 31). doi:10.3758/s13415-014-0250-6.

Dombrovski, Alexandre Y, Katalin Szanto, Greg J Siegle, Meredith L Wallace, Steven D Forman, Barbara Sahakian, Charles F Reynolds, and Luke Clark. 2011. "Lethal Forethought: Delayed Reward Discounting Differentiates High- and Low-Lethality Suicide Attempts in Old Age." *Biological Psychiatry* 70 (2) (July 15): 138–44. doi:10.1016/j.biopsych.2010.12.025.

DuBrow, S., and L. Davachi. 2014. "Temporal Memory Is Shaped by Encoding Stability and Intervening Item Reactivation." *Journal of Neuroscience* 34 (42): 13998–14005. doi:10.1523/JNEUROSCI.2535-14.2014.

Durstewitz, Daniel, and Jeremy K. Seamans. 2008. "The Dual-State Theory of Prefrontal Cortex Dopamine Function with Relevance to Catechol-O-Methyltransferase Genotypes and Schizophrenia." *Biological Psychiatry* 64: 739–749. doi:10.1016/j.biopsych.2008.05.015.

Eaton, Danice K., Laura Kann, Steve Kinchen, James Ross, Joseph Hawkins, William a. Harris, Richard Lowry, et al. 2006. "Youth Risk Behavior Surveillance — United States, 2005." *Morbidity and Mortality Weekly Report* 55: 1 – 108. doi:10.1016/S0002-9343(97)89440-5.

Eddy, Kamryn T., David J. Dorer, Debra L. Franko, Kavita Tahilani, Heather Thompson-Brenner, and David B. Herzog. 2008. "Diagnostic Crossover in Anorexia Nervosa and Bulimia Nervosa: Implications for DSM-V." *American Journal of Psychiatry* 165: 245–250. doi:10.1176/appi.ajp.2007.07060951.

Edwards, Ward. 1954. "The Theory of Decision Making." *Psychological Bulletin* 51 (4): 380–417. doi:10.1037/h0053870.

Elliott, R, R Elliott, K J Friston, K J Friston, R J Dolan, and R J Dolan. 2000. "Dissociable Neural Responses in Human Reward Systems." *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 20 (16): 6159–65. doi:20/16/6159 [pii].

Engelmann, Jan B, Sara Moore, C Monica Capra, and Gregory S Berns. 2012. "Differential Neurobiological Effects of Expert Advice on Risky Choice in Adolescents and Adults." *Social Cognitive and Affective Neuroscience* 7 (5) (June): 557–67. doi:10.1093/scan/nss050.

Ennett, S. T., N. S. Tobler, C. L. Ringwalt, and R. L. Flewelling. 1994. "How Effective Is Drug Abuse Resistance Education? A Meta-Analysis of

Project DARE Outcome Evaluations." *American Journal of Public Health* 84 (9): 1394–1401. doi:10.2105/AJPH.84.9.1394.

Eppinger, Ben, Maik Walter, Hauke R Heekeren, and Shu-Chen Li. 2013. "Of Goals and Habits: Age-Related and Individual Differences in Goal-Directed Decision-Making." *Frontiers in Neuroscience* 7 (December) (January): 253. doi:10.3389/fnins.2013.00253.

Ernst, Monique, and Martin P. Paulus. 2005. "Neurobiology of Decision Making: A Selective Review from a Neurocognitive and Clinical Perspective." *Biological Psychiatry* 58: 597–604. doi:10.1016/j.biopsych.2005.06.004.

Evans, Jonathan. 1984. "Heuristic and Analytic Processes in Reasoning." *British Journal of Psychology*. doi:10.1111/j.2044-8295.1984.tb01915.x.

Figner, Bernd, Daria Knoch, Eric J Johnson, Amy R Krosch, Sarah H Lisanby, Ernst Fehr, and Elke U Weber. 2010. "Lateral Prefrontal Cortex and Self-Control in Intertemporal Choice." *Nature Neuroscience* 13 (5) (May): 538–9. doi:10.1038/nn.2516.

Finn, Amy S, Margaret a Sheridan, Carla L Hudson Kam, Stephen Hinshaw, and Mark D'Esposito. 2010. "Longitudinal Evidence for Functional Specialization of the Neural Circuit Supporting Working Memory in the Human Brain." *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 30 (33): 11062–11067. doi:10.1523/JNEUROSCI.6266-09.2010.

Flegal, Kristin E, and Michael C Anderson. 2008. "Overthinking Skilled Motor Performance: Or Why Those Who Teach Can't Do." *Psychonomic Bulletin & Review* 15 (5): 927–932. doi:10.3758/PBR.15.5.927.

Fox, John, and Sanford Weisberg. 2010. *An R Companion to Applied Regression*. Sage.

Frank, Guido K, Ursula F Bailer, Shannan E Henry, Wayne Drevets, Carolyn C Meltzer, Julie C Price, Chester a Mathis, et al. 2005. "Increased Dopamine D2/D3 Receptor Binding after Recovery from Anorexia Nervosa Measured by Positron Emission Tomography and [11c]raclopride." *Biological Psychiatry* 58 (11) (December 1): 908–12. doi:10.1016/j.biopsych.2005.05.003.

Frank, Guido K W, Jeremy R Reynolds, Megan E Shott, Leah Jappe, Tony T Yang, Jason R Tregellas, and Randall C O'Reilly. 2012. "Anorexia Nervosa and Obesity Are Associated with Opposite Brain Reward

Response." *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology* 37 (9) (August): 2031–46. doi:10.1038/npp.2012.51.

Frank, Michael J, Ahmed a Moustafa, Heather M Haughey, Tim Curran, and Kent E Hutchison. 2007. "Genetic Triple Dissociation Reveals Multiple Roles for Dopamine in Reinforcement Learning." *Proceedings of the National Academy of Sciences of the United States of America* 104 (41): 16311–16316. doi:10.1073/pnas.0706111104.

Frank, Michael J, and Randall C O'Reilly. 2006. "A Mechanistic Account of Striatal Dopamine Function in Human Cognition: Psychopharmacological Studies with Cabergoline and Haloperidol." *Behavioral Neuroscience* 120 (3): 497–517. doi:10.1037/0735-7044.120.3.497.

Frank, Michael J, Lauren C Seeberger, and Randall C O'reilly. 2004. "By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism." *Science (New York, N.Y.)* 306 (December): 1940–1943. doi:10.1126/science.1102941.

Gallistel, C. 1989. "Animal Cognition: The Representation Of Space, Time And Number." *Annual Review of Psychology* 40: 155–189. doi:10.1146/annurev.psych.40.1.155.

Galván, Adriana. 2014. "Insights about Adolescent Behavior, Plasticity, and Policy from Neuroscience Research." *Neuron* 83 (2) (July): 262–265. doi:10.1016/j.neuron.2014.06.027.

Galvan, Adriana, Todd A Hare, Cindy E Parra, Jackie Penn, Henning Voss, Gary Glover, and BJ Casey. 2006. "Earlier Development of the Accumbens Relative to Orbitofrontal Cortex Might Underlie Risk-Taking Behavior in Adolescents." *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 26 (25): 6885–6892. doi:10.1523/JNEUROSCI.1062-06.2006.

Gardner, Margo, and Laurence Steinberg. 2005. "Peer Influence on Risk Taking, Risk Preference, and Risky Decision Making in Adolescence and Adulthood: An Experimental Study." *Developmental Psychology* 41 (4): 625–635. doi:10.1037/0012-1649.41.4.625.

Gazzaniga, Michael S. 2004. *The Cognitive Neurosciences*. MIT press.

Gershman, Samuel J, Arthur B Markman, and a Ross Otto. 2014. "Retrospective Revaluation in Sequential Decision Making: A Tale of Two

Systems." *Journal of Experimental Psychology* 143 (1) (February): 182–94. doi:10.1037/a0030844.

Gerstadt, C L, Y J Hong, and a Diamond. 1994. "The Relationship between Cognition and Action: Performance of Children 3 1/2-7 Years Old on a Stroop-like Day-Night Test." *Cognition* 53: 129–153. doi:10.1016/0010-0277(94)90068-X.

Giedd, Jay N, Jonathan Blumenthal, Neal O Jeffries, FX Castellanos, Hong Liu, Alex Zijdenbos, Tomás Paus, Alan C Evans, and Judith L Rapopor. 1999. "Brain Development during Childhood and Adolescence: A Longitudinal MRI Study." *Nature Neuroscience* 2 (10): 861–863.

Gilbert, Daniel T, and Timothy D Wilson. 2007. "Prospection: Experiencing the Future." *Science (New York, N.Y.)* 317 (September): 1351–1354. doi:10.1126/science.1144161.

Gladwin, Thomas E, and Bernd Figner. 2014. "' Hot' Cognition and Dual Systems: Introduction, Criticisms, and Ways Forward." In *Frontiers of Cognitive Psychology Series: Neuroeconomics, Judgment and Decision Making*, edited by E Wilhelms and VF Reyna, 1–39. Frontiers of Cognitive Psychology Series.

Glimcher, Paul W, and Ernst Fehr. 2013. *Neuroeconomics: Decision Making and the Brain*. Academic Press.

Godier, Lauren R., and Rebecca J. Park. 2014. "Compulsivity in Anorexia Nervosa: A Transdiagnostic Concept." *Frontiers in Psychology* 5 (July): 1–18. doi:10.3389/fpsyg.2014.00778.

Gogtay, Nitin, Jay N Giedd, Leslie Lusk, Kiralee M Hayashi, Deanna Greenstein, a Catherine Vaituzis, Tom F Nugent, et al. 2004. "Dynamic Mapping of Human Cortical Development during Childhood through Early Adulthood." *Proceedings of the National Academy of Sciences of the United States of America* 101 (21) (May 25): 8174–9. doi:10.1073/pnas.0402680101.

Goldstone, Anthony P, Christina G Prechtl de Hernandez, John D Beaver, Kinan Muhammed, Charlotte Croese, Gabriel Bell, Giuliana Durighel, et al. 2009. "Fasting Biases Brain Reward Systems towards High-Calorie Foods." *The European Journal of Neuroscience* 30 (8) (October): 1625–35. doi:10.1111/j.1460-9568.2009.06949.x.

Goto, Yukiori, and Anthony a Grace. 2005. “Dopaminergic Modulation of Limbic and Cortical Drive of Nucleus Accumbens in Goal-Directed Behavior.” *Nature Neuroscience* 8 (6): 805–812. doi:10.1038/nn1471.

Grafton, S T, E Hazeltine, and R Ivry. 1995. “Functional Mapping of Sequence Learning in Normal Humans.” *Journal of Cognitive Neuroscience*. doi:10.1162/jocn.1995.7.4.497.

Graybiel, a M. 1995. “Building Action Repertories: Memory and Learning Functions of the Basal Ganglia.” *Current Opinion in Neurobiology* 5: 733–741.

Green, Leonard, Astnd F Fry, and Joel Myerson. 1994. “Discounting of Delayed Rewards: A Life-Span Comparison.” *Psychological Science* 5 (1) (January): 33–36. doi:10.1111/j.1467-9280.1994.tb00610.x.

Grillner, Sten, and Brita Robertson. 2015. “The Basal Ganglia Downstream Control of Brainstem Motor Centres—an Evolutionarily Conserved Strategy.” *Current Opinion in Neurobiology* 33: 47–52. doi:10.1016/j.conb.2015.01.019.

Grillner, Sten, Peter Wallén, Kazuya Saitoh, Alexander Kozlov, and Brita Robertson. 2008. “Neural Bases of Goal-Directed Locomotion in Vertebrates-An Overview.” *Brain Research Reviews* 57 (L): 2–12. doi:10.1016/j.brainresrev.2007.06.027.

Haber, Suzanne N, and Brian Knutson. 2009. “The Reward Circuit: Linking Primate Anatomy and Human Imaging.” *Neuropsychopharmacology* 35 (1): 4–26.

Hagler, Donald J., and Martin I. Sereno. 2006. “Spatial Maps in Frontal and Prefrontal Cortex.” *NeuroImage* 29: 567–577. doi:10.1016/j.neuroimage.2005.08.058.

Hampshire, Adam, Roger R Highfield, Beth L Parkin, and Adrian M Owen. 2012. “Fractionating Human Intelligence.” *Neuron* 76 (6): 1225–1237.

Hare, Todd A, Colin F Camerer, and Antonio Rangel. 2009. “Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System.” *Science (New York, N.Y.)* 324 (5927) (May 1): 646–8. doi:10.1126/science.1168450.

Hassabis, Demis, Dharshan Kumaran, Seralynne D Vann, and Eleanor a Maguire. 2007. “Patients with Hippocampal Amnesia Cannot Imagine New Experiences.” *Proceedings of the National Academy of Sciences of*

the United States of America 104: 1726–1731. doi:10.1073/pnas.0610561104.

Hayes, Richard L. 1993. “A Facilitative Role for Counselors in Restructuring Schools.” *The Journal of Humanistic Education and Development* 31 (4): 156–162.

Hayes, Steven C, A J Brownstein, R D Zettle, I Rosenfarb, and Z Korn. 1986. “Rule-Governed Behavior and Sensitivity to Changing Consequences of Responding.” *Journal of the Experimental Analysis of Behavior* 45 (3): 237–256. doi:10.1901/jeab.1986.45-237.

Heatherton, Todd F., and Dylan D. Wagner. 2011. “Cognitive Neuroscience of Self-Regulation Failure.” *Trends in Cognitive Sciences* 15 (3): 132–139. doi:10.1016/j.tics.2010.12.005.

Henriksson, M, Hillevi M Aro, Mauri J Marttunen, Martti E Heikkinen, E Isometsä, Kimmo I Kuoppasalmi, and J Lönnqvist. 1993. “Mental Disorders and Comorbidity in Suicide.” *American Journal of Psychiatry* 150 (6): 935–940. doi:10.1111/j.1600-0447.1996.tb09855.x.

Højsgaard, U H S. 2014. “Pbkrtest: Parametric Bootstrap and Kenward Roger Based Methods for Mixed Model Comparison.”

Hudson, James I, Eva Hiripi, Harrison G Pope, and Ronald C Kessler. 2007. “The Prevalence and Correlates of Eating Disorders in the National Comorbidity Survey Replication.” *Biological Psychiatry* 61 (3) (February 1): 348–58. doi:10.1016/j.biopsych.2006.03.040.

Huys, Quentin J M, Philippe N. Tobler, Gregor Hasler, and Shelly B. Flagel. 2014. *The Role of Learning-Related Dopamine Signals in Addiction Vulnerability*. *Progress in Brain Research*. 1st ed. Vol. 211. Elsevier B.V. doi:10.1016/B978-0-444-63425-2.00003-9.

Imperati, Davide, Stan Colcombe, Clare Kelly, Adriana Di Martino, Juan Zhou, F Xavier Castellanos, and Michael P Milham. 2011. “Differential Development of Human Brain White Matter Tracts.” *PloS One* 6 (8) (January): e23437. doi:10.1371/journal.pone.0023437.

Insel, Thomas, Bruce Cuthbert, Majorie Garvie, Robert Heinssen, Daniel S. Pine, Kevin Quinn, Charles Sanislow, and Phillip Wang. 2010. “Research Domain Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders.” *The American Journal of Psychiatry* 167 (July): 748–751.

Jacobs, Je, and Pa Klaczynski. 2002. “The Development of Judgment and Decision Making during Childhood and Adolescence.” *Current Directions in Psychological Science* 11: 145–149.

Jitsumori, Masako. 1993. “Category Discrimination of Artificial Polymorphous Stimuli Based on Feature Learning.” *Journal of Experimental Psychology: Animal Behavior Processes* 19 (3): 244–254. doi:10.1037/0097-7403.19.3.244.

Joel, Daphna, Yael Niv, and Eytan Ruppin. 2002. “Actor–critic Models of the Basal Ganglia: New Anatomical and Computational Perspectives.” *Neural Networks* 15 (4): 535–547.

Johnson, M H. 2001. “Functional Brain Development in Humans.” *Nature Reviews. Neuroscience* 2 (July): 475–483. doi:10.1038/35081509.

Johnson, Matthe W, and Warren K Bickel. 2002. “Within-Subject Comparison of Real and Hypothetical Money Rewards in Delay Discounting.” *Journal of the Experimental Analysis of Behavior* 77 (2) (March): 129–46. doi:10.1901/jeab.2002.77-129.

Jones, Rebecca M, Leah H Somerville, Jian Li, Erika J Ruberry, Alisa Powers, Natasha Mehta, Jonathan Dyke, and BJ Casey. 2014. “Adolescent-Specific Patterns of Behavior and Neural Activity during Social Reinforcement Learning.” *Cognitive, Affective & Behavioral Neuroscience* 14: 683–97. doi:10.3758/s13415-014-0257-z.

Kable, Joseph W, and Paul W Glimcher. 2007. “The Neural Correlates of Subjective Value during Intertemporal Choice.” *Nature Neuroscience* 10 (12): 1625–1633.

Kahneman, Daniel. 2003. “A Perspective on Judgment and Choice: Mapping Bounded Rationality.” *The American Psychologist* 58 (9): 697–720. doi:10.1037/0003-066X.58.9.697.

Kaye, Walter H, Cynthia M Bulik, Laura Thornton, Nicole Barbarich, and Kim Masters. 2004. “Comorbidity of Anxiety Disorders with Anorexia and Bulimia Nervosa.” *The American Journal of Psychiatry* 161 (12) (December): 2215–21. doi:10.1176/appi.ajp.161.12.2215.

Kaye, Walter H., Julie L. Fudge, and Martin Paulus. 2009. “New Insights into Symptoms and Neurocircuit Function of Anorexia Nervosa.” *Nature Reviews Neuroscience* 10 (8) (July 15): 573–584. doi:10.1038/nrn2682.

Keating, Charlotte, Alan J Tilbrook, Susan L Rossell, Peter G Enticott, and Paul B Fitzgerald. 2012. “Reward Processing in Anorexia Nervosa.” *Neuropsychologia* 50 (5) (April): 567–75. doi:10.1016/j.neuropsychologia.2012.01.036.

Kelly, P H, P W Seviour, and S D Iversen. 1975. “Amphetamine and Apomorphine Responses in the Rat Following 6-OHDA Lesions of the Nucleus Accumbens Septi and Corpus Striatum.” *Brain Research* 94: 507–522. doi:10.1016/0006-8993(75)90233-4.

Keramati, Mehdi, Amir Dezfouli, and Payam Piray. 2011. “Speed/accuracy Trade-off between the Habitual and the Goal-Directed Processes.” *PLoS Computational Biology* 7 (5) (May): e1002055. doi:10.1371/journal.pcbi.1002055.

Kessler, Ronald C, Patricia Berglund, Olga Demler, Robert Jin, Kathleen R Merikangas, and Ellen E Walters. 2005. “Lifetime Prevalence and Age-of-Onset Distributions of.” *Arch Gen Psychiatry* 62 (June 2005): 593–602. doi:10.1001/archpsyc.62.6.593.

Keys, Ancel, Josef Brožek, Austin Henschel, Olaf Mickelsen, and Henry Longstreet Taylor. 1950. *The Biology of Human starvation.(2 Vols).* Univ. of Minnesota Press.

Kirby, Kris N. 2009. “One-Year Temporal Stability of Delay-Discount Rates.” *Psychonomic Bulletin & Review* 16 (3) (June): 457–62. doi:10.3758/PBR.16.3.457.

Kirby, Kris N. 1997. “Bidding on the Future: Evidence against Normative Discounting of Delayed Rewards.” *Journal of Experimental Psychology: General* 126: 54–70. doi:10.1037/0096-3445.126.1.54.

Kirkham, Natasha Z., Loren Cruess, and Adele Diamond. 2003. “Helping Children Apply Their Knowledge to Their Behavior on a Dimension-Switching Task.” *Developmental Science* 6: 449–467. doi:10.1111/1467-7687.00300.

Kirkham, Natasha Z., Jonathan a. Slemmer, and Scott P. Johnson. 2002. “Visual Statistical Learning in Infancy: Evidence for a Domain General Learning Mechanism.” *Cognition* 83: 4–5. doi:10.1016/S0010-0277(02)00004-5.

Klossek, U M H, J Russell, and Anthony Dickinson. 2008. “The Control of Instrumental Action Following Outcome Devaluation in Young Children

Aged between 1 and 4 Years." *Journal of Experimental Psychology. General* 137 (1): 39–51. doi:10.1037/0096-3445.137.1.39.

Knutson, Kristine M., Jacqueline N. Wood, and Jordan Grafman. 2004. "Brain Activation in Processing Temporal Sequence: An fMRI Study." *NeuroImage* 23: 1299–1307. doi:10.1016/j.neuroimage.2004.08.012.

Koffarnus, Mikhail N, David P Jarmolowicz, E Terry Mueller, and Warren K Bickel. 2013. "Changing Delay Discounting in the Light of the Competing Neurobehavioral Decision Systems Theory: A Review." *Journal of the Experimental Analysis of Behavior* 99 (1) (January): 32–57. doi:10.1002/jeab.2.

Kolb, B. 1984. "Functions of the Frontal Cortex of the Rat: A Comparative Review." *Brain Research* 320: 65–98. doi:10.1016/0165-0173(84)90018-3.

Kolb, D.A. 1984. *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, NJ. doi:10.1016/B978-0-7506-7223-8.50017-4.

Kotsoni, Eleni, Dana Byrd, and BJ Casey. 2006. "Special Considerations for Functional Magnetic Resonance Imaging of Pediatric Populations." *Journal of Magnetic Resonance Imaging* 23 (6): 877–886.

Kurth-Nelson, Zeb, Warren K Bickel, and A David Redish. 2012. "A Theoretical Account of Cognitive Effects in Delay Discounting." *The European Journal of Neuroscience* 35 (7) (April): 1052–64. doi:10.1111/j.1460-9568.2012.08058.x.

Laibson, D. 1997. "Golden Eggs and Hyperbolic Discounting." *The Quarterly Journal of Economics* (May).

Lau, Brian, and Paul W Glimcher. 2005. "Dynamic Response-by-Response Models of Matching Behavior in Rhesus Monkeys." *Journal of the Experimental Analysis of Behavior* 84 (3): 555–579. doi:10.1901/jeab.2005.110-04.

Lee, Daeyeol. 2013. "Decision Making: From Neuroscience to Psychiatry." *Neuron* 78 (2): 233–248. doi:10.1016/j.neuron.2013.04.008.

Lee, Daeyeol, Hyojung Seo, and Min Whan Jung. 2012. "Neural Basis of Reinforcement Learning and Decision Making." *Annual Review of Neuroscience* 35 (1): 287–308.

Lee, Francis S, Hakon Heimer, N Giedd, Edward S Lein, Nenad Šestan, Daniel R Weinberger, and B J Casey. 2014. “Adolescent Mental Health—Opportunity and Obligation Emerging Neuroscience Offers Hope for Treatments.” *Science* 346 (6209): 547–549.

Levy, Dino J, Amalie C Thavikulwat, and Paul W Glimcher. 2013. “State Dependent Valuation: The Effect of Deprivation on Risk Preferences.” *PloS One* 8 (1) (January): e53978. doi:10.1371/journal.pone.0053978.

Li, Jian, Mauricio R Delgado, and Elizabeth a Phelps. 2011. “How Instructed Knowledge Modulates the Neural Systems of Reward Learning.” *Proceedings of the National Academy of Sciences* 108 (1) (January 4): 55–60. doi:10.1073/pnas.1014938108.

Liston, Conor, B S McEwen, and BJ Casey. 2009. “Psychosocial Stress Reversibly Disrupts Prefrontal Processing and Attentional Control.” *Proceedings of the National Academy of Sciences of the United States of America* 106 (3): 912–917. doi:10.1073/pnas.0807041106.

Liston, Conor, Richard Watts, Nim Tottenham, Matthew C Davidson, Sumit Niogi, Aziz M Ulug, and BJ Casey. 2006. “Frontostriatal Microstructure Modulates Efficient Recruitment of Cognitive Control.” *Cerebral Cortex (New York, N.Y. : 1991)* 16 (4) (April): 553–60. doi:10.1093/cercor/bhj003.

Lowe, C F, a Beasty, and R P Bentall. 1983. “The Role of Verbal Behavior in Human Learning: Infant Performance on Fixed-Interval Schedules.” *Journal of the Experimental Analysis of Behavior* 39 (I): 157–164. doi:10.1901/jeab.1983.39-157.

Lucantonio, Federica, Daniele Caprioli, and Geoffrey Schoenbaum. 2014. “Transition from ‘Model-Based’ to ‘Model-Free’ Behavioral Control in Addiction: Involvement of the Orbitofrontal Cortex and Dorsolateral Striatum.” *Neuropharmacology* 76 Pt B (January): 407–15. doi:10.1016/j.neuropharm.2013.05.033.

Lukacs, Paul M., Kenneth P. Burnham, and David R. Anderson. 2010. “Model Selection Bias and Freedman’s Paradox.” *Annals of the Institute of Statistical Mathematics* 62: 117–125. doi:10.1007/s10463-009-0234-4.

Luksys, Gediminas, Wulfram Gerstner, and Carmen Sandi. 2009. “Stress, Genotype and Norepinephrine in the Prediction of Mouse Behavior Using Reinforcement Learning.” *Nature Neuroscience* 12 (9): 1180–1186. doi:10.1038/nn.2374.

Luna, Beatriz, and John a. Sweeney. 2004. “The Emergence of Collaborative Brain Function: fMRI Studies of the Development of Response Inhibition.” *Annals of the New York Academy of Sciences* 1021: 296–309. doi:10.1196/annals.1308.035.

Luo, S, G Ainslie, L Giragosian, and J R Monterosso. 2009. “Behavioral and Neural Evidence of Incentive Bias for Immediate Rewards Relative to Preference-Matched Delayed Rewards.” *Journal of Neuroscience* 29 (47): 14820–14827.

MacKillop, J, and CW Kahler. 2009. “Delayed Reward Discounting Predicts Treatment Response for Heavy Drinkers Receiving Smoking Cessation Treatment.” *Drug and Alcohol Dependence* 104 (3): 197–203. doi:10.1016/j.drugalcdep.2009.04.020.Delayed.

Maia, Tiago V, and Michael J Frank. 2011. “From Reinforcement Learning Models to Psychiatric and Neurological Disorders.” *Nature Neuroscience* 14 (2): 154–162. doi:10.1038/nn.2723.

Manes, Facundo, Barbara Sahakian, Luke Clark, Robert Rogers, Nagui Antoun, Mike Aitken, and Trevor Robbins. 2002. “Decision-Making Processes Following Damage to the Prefrontal Cortex.” *Brain : A Journal of Neurology* 125: 624–639. doi:10.1093/brain/awf049.

Marchette, Steven a, Arnold Bakker, and Amy L Shelton. 2011. “Cognitive Mappers to Creatures of Habit: Differential Engagement of Place and Response Learning Mechanisms Predicts Human Navigational Behavior.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 31 (43) (October 26): 15264–8. doi:10.1523/JNEUROSCI.3634-11.2011.

Marr, D. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. *San Francisco, Freeman & Co*. WH Freeman and Co. San Francisco.

MATLAB. 2012. *Version 8.0.0.783 (R2012b)*. Natick, Massachusetts: The MathWorks Inc.

Mazur, James E, JA Nevin, and H Rachlin. 1987. “An Adjusting Procedure for Studying Delayed Reinforcement.” In *Quantitative Analyses of Behavior: The Effect of Delay and Intervening Events on Reinforcement Value.*, Vol 5, 55–73. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

McClure, Samuel M, Gregory S Berns, and P Read Montague. 2003. "Temporal Prediction Errors in a Passive Learning Task Activate Human Striatum." *Neuron* 38 (2) (April 24): 339–46.

McClure, Samuel M, Keith M Ericson, David I Laibson, George Loewenstein, and Jonathan D Cohen. 2007. "Time Discounting for Primary Rewards." *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 27 (21): 5796–5804. doi:10.1523/JNEUROSCI.4246-06.2007.

McClure, Samuel M, David I Laibson, George Loewenstein, and Jonathan D Cohen. 2004. "Separate Neural Systems Value Immediate and Delayed Monetary Rewards." *Science (New York, N.Y.)* 306 (5695) (October 15): 503–7. doi:10.1126/science.1100907.

McKerchar, Todd L, Leonard Green, Joel Myerson, T Stephen Pickford, Jade C Hill, and Steven C Stout. 2009. "A Comparison of Four Models of Delay Discounting in Humans." *Behavioural Processes* 81 (2) (June): 256–9. doi:10.1016/j.beproc.2008.12.017.

McNamee, D., M. Liljeholm, O. Zika, and John P O'Doherty. 2015. "Characterizing the Associative Content of Brain Structures Involved in Habitual and Goal-Directed Actions in Humans: A Multivariate fMRI Study." *Journal of Neuroscience* 35 (9): 3764–3771. doi:10.1523/JNEUROSCI.4677-14.2015.

Merikangas, Kathleen Ries, Jian-Ping He, Marcy Burstein, Sonja a Swanson, Shelli Avenevoli, Lihong Cui, Corina Benjet, Katholiki Georgiades, and Joel Swendsen. 2010. "Lifetime Prevalence of Mental Disorders in U.S. Adolescents: Results from the National Comorbidity Survey Replication--Adolescent Supplement (NCS-A)." *Journal of the American Academy of Child and Adolescent Psychiatry* 49 (10): 980–989. doi:10.1016/j.jaac.2010.05.017.

Meshi, Dar, Guido Biele, Christoph W. Korn, and Hauke R. Heekeren. 2012. "How Expert Advice Influences Decision Making." *PLoS ONE* 7 (11). doi:10.1371/journal.pone.0049748.

Mesulam, M M. 1990. "Large-Scale Neurocognitive Networks and Distributed Processing for Attention, Language, and Memory." *Annals of Neurology*. doi:10.1002/ana.410280502.

Metcalfe, J, and W Mischel. 1999. "A Hot/cool-System Analysis of Delay of Gratification: Dynamics of Willpower." *Psychological Review* 106: 3–19. doi:10.1037/0033-295X.106.1.3.

Michalczuk, R, H Bowden-Jones, A Verdejo-Garcia, and L Clark. 2011. "Impulsivity and Cognitive Distortions in Pathological Gamblers Attending the UK National Problem Gambling Clinic: A Preliminary Report." *Psychological Medicine* 41: 2625–2635. doi:10.1017/S003329171100095X.

Miller, E K, and J D Cohen. 2001. "An Integrative Theory of Prefrontal Cortex Function." *Annual Review of Neuroscience* 24 (January): 167–202. doi:10.1146/annurev.neuro.24.1.167.

Mischel, W, Y Shoda, and P K Peake. 1988. "The Nature of Adolescent Competencies Predicted by Preschool Delay of Gratification." *Journal of Personality and Social Psychology* 54 (4) (April): 687–96.

Mischel, W, Y Shoda, and M I Rodriguez. 1989. "Delay of Gratification in Children." *Science (New York, N.Y.)* 244 (4907) (May 26): 933–8.

Montague, P. Read, Raymond J. Dolan, Karl J. Friston, and Peter Dayan. 2012. "Computational Psychiatry." *Trends in Cognitive Sciences* 16 (1): 72–80. doi:10.1016/j.tics.2011.11.018.

Munakata, Y., H. R. Snyder, and C. H. Chatham. 2012. "Developing Cognitive Control: Three Key Transitions." *Current Directions in Psychological Science* 21: 71–77. doi:10.1177/0963721412436807.

Naneix, Fabien, Alain R Marchand, Georges Di Scala, Jean-Rémi Pape, and Etienne Coutureau. 2012. "Parallel Maturation of Goal-Directed Behavior and Dopaminergic Systems during Adolescence." *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 32 (46) (November 14): 16223–32. doi:10.1523/JNEUROSCI.3080-12.2012.

Nassar, Matthew R., and Joshua I. Gold. 2013. "A Healthy Fear of the Unknown: Perspectives on the Interpretation of Parameter Fits from Computational Models in Neuroscience." *PLoS Computational Biology* 9 (4): 1–6. doi:10.1371/journal.pcbi.1003015.

Nickerson, Raymond S. 1998. "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises." *Review of General Psychology* 2 (2): 175–220. doi:10.1037/1089-2680.2.2.175.

Nigg, Joel, and BJ Casey. 2005. "An Integrative Theory of Attention-Deficit/ Hyperactivity Disorder Based on the Cognitive and Affective Neurosciences." *Development and Psychopathology* 17: 785–806. doi:10.1017/S0954579405050376.

O'Doherty, John P, Peter Dayan, Karl Friston, Hugo Critchley, and Raymond J. Dolan. 2003. "Temporal Difference Models and Reward-Related Learning in the Human Brain." *Neuron* 38: 329–337. doi:10.1016/S0896-6273(03)00169-7.

O'Doherty, John P, Peter Dayan, Johannes Schultz, Ralf Deichmann, Karl Friston, and Raymond J Dolan. 2004. "Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning." *Science (New York, N.Y.)* 304 (April): 452–454. doi:10.1126/science.1094285.

Odum, Amy L. 2011a. "Delay Discounting: Trait Variable?" *Behavioural Processes* 87 (1): 1–9. doi:10.1016/j.beproc.2011.02.007.Delay.

Odum, Amy L. 2011b. "Delay Discounting: I'm a K, You're a K." *Journal of the Experimental Analysis of Behavior* 96 (3): 427–439.

Olesen, Pernille J, Helena Westerberg, and Torkel Klingberg. 2004. "Increased Prefrontal and Parietal Activity after Training of Working Memory." *Nature Neuroscience* 7 (1): 75–79. doi:10.1038/nn1165.

Ostlund, Sean B, and Bernard W Balleine. 2007. "Orbitofrontal Cortex Mediates Outcome Encoding in Pavlovian but Not Instrumental Conditioning." *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 27 (18): 4819–4825. doi:10.1523/JNEUROSCI.5443-06.2007.

Otto, a Ross, Samuel J Gershman, Arthur B Markman, and Nathaniel D Daw. 2013. "The Curse of Planning: Dissecting Multiple Reinforcement-Learning Systems by Taxing the Central Executive." *Psychological Science* 24 (5) (May): 751–61. doi:10.1177/0956797612463080.

Otto, a Ross, Candace M Raio, Alice Chiang, Elizabeth a Phelps, and Nathaniel D Daw. 2013. "Working-Memory Capacity Protects Model-Based Learning from Stress." *Proceedings of the National Academy of Sciences of the United States of America* 110 (52) (December 24): 20941–6. doi:10.1073/pnas.1312011110.

Otto, a Ross, Anya Skatova, Seth Madlon-Kay, and Nathaniel D Daw. 2015. "Cognitive Control Predicts Use of Model-Based Reinforcement Learning." *Journal of Cognitive Neuroscience* 27 (2): 319–333. doi:10.1162/jocn.

Pagnoni, Giuseppe, Caroline F Zink, P Read Montague, and Gregory S Berns. 2002. "Activity in Human Ventral Striatum Locked to Errors of Reward Prediction." *Nature Neuroscience* 5: 97–98. doi:10.1038/nn802.

Paulesu, E, C D Frith, and R S Frackowiak. 1993. “The Neural Correlates of the Verbal Component of Working Memory.” *Nature* 362: 342–345. doi:10.1038/362342a0.

Paus, Tomás, Matcheri Keshavan, and Jay N Giedd. 2008. “Why Do Many Psychiatric Disorders Emerge during Adolescence?” *Nature Reviews. Neuroscience* 9 (DeceMBeR): 947–957. doi:10.1038/nrn2513.

Pennartz, C M a, R Ito, P F M J Verschure, F P Battaglia, and T W Robbins. 2011. “The Hippocampal-Striatal Axis in Learning, Prediction and Goal-Directed Behavior.” *Trends in Neurosciences* 34 (10) (October): 548–59. doi:10.1016/j.tins.2011.08.001.

Pereboom, Derk. 2001. “Living without Free Will.” *Cambridge Studies in Philosophy*: xxiii, 231 p. doi:10.1017/CBO9780511498824.

Pessiglione, Mathias, Ben Seymour, Guillaume Flandin, Raymond J Dolan, and Chris D Frith. 2006. “Dopamine-Dependent Prediction Errors Underpin Reward-Seeking Behaviour in Humans.” *Nature* 442 (August): 1042–1045. doi:10.1038/nature05051.

Peters, Jan, and Christian Büchel. 2010. “Episodic Future Thinking Reduces Reward Delay Discounting through an Enhancement of Prefrontal-Mediotemporal Interactions.” *Neuron* 66 (1): 138–148.

———. 2011. “The Neural Mechanisms of Inter-Temporal Decision-Making: Understanding Variability.” *Trends in Cognitive Sciences* 15 (5) (May): 227–39. doi:10.1016/j.tics.2011.03.002.

Peters, Sabine, Barbara R Braams, Maartje E J Raijmakers, P Cédric M P Koolschijn, and Eveline A Crone. 2014. “The Neural Coding of Feedback Learning across Child and Adolescent Development.” *Journal of Cognitive Neuroscience*: 1705–1720. doi:10.1162/jocn.

Pfeifer, Jennifer H, and Nicholas B Allen. 2012. “Arrested Development? Reconsidering Dual-Systems Models of Brain Function in Adolescence and Disorders.” *Trends in Cognitive Sciences* 16 (6): 322–329. doi:10.1016/j.biotechadv.2011.08.021.Secreted.

Pfeifer, Jennifer H, Carrie L Masten, William E Moore, Tasha M Oswald, John C Mazziotta, Marco Iacoboni, and Mirella Dapretto. 2011. “Entering Adolescence: Resistance to Peer Influence, Risky Behavior, and Neural Changes in Emotion Reactivity.” *Neuron* 69 (5): 1029–1036. doi:10.1038/nature11130.Reduced.

Pfeiffer, Brad E, and David J Foster. 2013. “Hippocampal Place-Cell Sequences Depict Future Paths to Remembered Goals.” *Nature* 497 (7447): 74–9. doi:10.1038/nature12112.

Piaget, J. 1954. *The Construction of Reality in the Child*. New York: Basic Books.

Pinto, Anthony, Joanna E Steinglass, Ashley L Greene, Elke U Weber, and H Blair Simpson. 2013. “Capacity to Delay Reward Differentiates Obsessive-Compulsive Disorder and Obsessive Compulsive Personality Disorder.” *Biological Psychiatry* 75 (8) (November 4): 1–7. doi:10.1016/j.biopsych.2013.09.007.

Posner, Michael I, Mary Jo Nissen, and William C Ogden. 1978. “Attended and Unattended Processing Modes: The Role of Set for Spatial Location.” *Modes of Perceiving and Processing Information* 137: 158.

Posner, Michael I, and M K Rothbart. 2000. “Developing Mechanisms of Self-Regulation.” *Development and Psychopathology* 12: 427–441. doi:10.1017/S0954579400003096.

Posner, Michael I., and Steven E. Petersen. 1990. “The Attention System of the Human Brain.” *Annual Review of Neuroscience* 13: 25–42. doi:10.1146/annurev.ne.13.030190.000325.

R-Core-Team. 2012. *R: A Language and Environment for Statistical Computing*. Edited by R Foundation for Statistical Computing. Vol. 172. Vienna, Austria. doi:10.1002/jae.

Rauch, Scott L., Paul J. Whalen, Cary R. Savage, Tim Curran, Adair Kendrick, Halle D. Brown, George Bush, Hans C. Breiter, and Bruce R. Rosen. 1997. “Striatal Recruitment during an Implicit Sequence Learning Task as Measured by Functional Magnetic Resonance Imaging.” *Human Brain Mapping* 5: 124–132. doi:10.1002/(SICI)1097-0193(1997)5:2<124::AID-HBM6>3.0.CO;2-5.

Raznahan, Armin, Phillip W Shaw, Jason P Lerch, Liv S Clasen, Deanna Greenstein, Rebecca Berman, Jon Pipitone, Mallar M Chakravarty, and Jay N Giedd. 2014. “Longitudinal Four-Dimensional Mapping of Subcortical Anatomy in Human Development.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (4) (January 28): 1592–7. doi:10.1073/pnas.1316911111.

Reber, Arthur S. 1989. “Implicit Learning and Tacit Knowledge.” *Journal of Experimental Psychology: General* 118 (3): 219–235. doi:10.1037//0096-3445.118.3.219.

Redgrave, P., T. J. Prescott, and K. Gurney. 1999. “The Basal Ganglia: A Vertebrate Solution to the Selection Problem?” *Neuroscience* 89 (4): 1009–1023. doi:10.1016/S0306-4522(98)00319-4.

Reyna, Valerie F, and Frank Farley. 2006. “Risk and Rationality in Adolescent Decision Making: Implications for Theory, Practice, and Public Policy.” *Psychological Science in the Public Interest (Wiley-Blackwell)* 7 (1): 1–44. doi:10.1111/j.1529-1006.2006.00026.x.

Reynolds, Brady, Harriet de Wit, and Jerry B Richards. 2002. “Delay of Gratification and Delay Discounting in Rats.” *Behavioural Processes* 59 (3): 157–168. doi:http://dx.doi.org/10.1016/S0376-6357(02)00088-8.

Reynolds, Brady, and Ryan Schiffbauer. 2004. “Measuring State Changes in Human Delay Discounting: An Experiential Discounting Task.” *Behavioural Processes* 67 (3) (November 30): 343–56. doi:10.1016/j.beproc.2004.06.003.

Richfield, E K, R Twyman, and S Berent. 1987. “Neurological Syndrome Following Bilateral Damage to the Head of the Caudate Nuclei.” *Annals of Neurology* 22: 768–771. doi:10.1002/ana.410220615.

Ritschel, F., J. a. King, D. Geisler, L. Flohr, F. Neidel, I. Boehm, M. Seidel, et al. 2015. “Temporal Delay Discounting in Acutely Ill and Weight-Recovered Patients with Anorexia Nervosa.” *Psychological Medicine*: 1–11. doi:10.1017/S0033291714002311.

Rodriguez, P. F., a. R. Aron, and R. a. Poldrack. 2006. “Ventral-Striatal/nucleus-Accumbens Sensitivity to Prediction Errors during Classification Learning.” *Human Brain Mapping* 27: 306–313. doi:10.1002/hbm.20186.

Rothemund, Y, C Buchwald, P Georgiewa, G Bohner, H-C Bauknecht, M Ballmaier, B F Klapp, and R Klingebiel. 2011. “Compulsivity Predicts Fronto Striatal Activation in Severely Anorectic Individuals.” *Neuroscience* 197 (December 1): 242–50. doi:10.1016/j.neuroscience.2011.09.016.

Rubia, Katya, Anna B Smith, Eric Taylor, and Michael Brammer. 2007. “Linear Age-Correlated Functional Development of Right Inferior Fronto-Striato-Cerebellar Networks during Response Inhibition and Anterior Cingulate

during Error-Related Processes." *Human Brain Mapping* 28 (11) (November): 1163–77. doi:10.1002/hbm.20347.

Rubia, Katya, Anna B Smith, James Woolley, Chiara Nosarti, Isobel Heyman, Eric Taylor, and Mick Brammer. 2006. "Progressive Increase of Frontostriatal Brain Activation from Childhood to Adulthood during Event-Related Tasks of Cognitive Control." *Human Brain Mapping* 27 (12) (December): 973–93. doi:10.1002/hbm.20237.

Rueda, M Rosario, Michael I Posner, and Mary K Rothbart. 2005. "The Development of Executive Attention: Contributions to the Emergence of Self-Regulation." *Developmental Neuropsychology* 28 (2) (January): 573–94. doi:10.1207/s15326942dn2802_2.

Russel, W Ritchie. 1948. "Functions of the Frontal Lobes." *The Lancet*: 356–360. doi:10.1093/brain/26.2.215.

Saffran, Jenny R, Richard N Aslin, and Elissa L Newport. 1996. "Statistical Learning by 8-Month-Old Infants." *Science* 274: 1926–1928.

Sareen, J, B J Cox, T O Afifi, R de Graaf, G J Asmundson, M ten Have, and M B Stein. 2006. "Anxiety Disorders and Risk for Suicidal Ideation and Suicide Attempts." *Archives of General Psychiatry* 62: 1249–1257. doi:10.1001/archpsyc.62.11.1249.

Scheres, Anouk, Chandra Tontsch, Allison Lee Thoeny, and Antonia Kaczkurkin. 2010. "Temporal Reward Discounting in Attention-Deficit/hyperactivity Disorder: The Contribution of Symptom Domains, Reward Magnitude, and Session Length." *Biological Psychiatry* 67 (7) (April 1): 641–8. doi:10.1016/j.biopsych.2009.10.033.

Schneider, W, and RM Shiffrin. 1977. "Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention." *Psychological Review* 84 (1): 1–66.

Schoenbaum, Geoffrey, Michael P. Saddoris, and Thomas a. Stalnaker. 2007. "Reconciling the Roles of Orbitofrontal Cortex in Reversal Learning and the Encoding of Outcome Expectancies." *Annals of the New York Academy of Sciences* 1121: 320–335. doi:10.1196/annals.1401.001.

Schönberg, Tom, Nathaniel D Daw, Daphna Joel, and John P O'Doherty. 2007. "Reinforcement Learning Signals in the Human Striatum Distinguish Learners from Nonlearners during Reward-Based Decision Making." *The Journal of Neuroscience : The Official Journal of the Society for*

Neuroscience 27 (47): 12860–12867. doi:10.1523/JNEUROSCI.2496-07.2007.

Schultz, Wolfram. 1986. “Responses of Midbrain Dopamine Neurons to Behavioral Trigger Stimuli in the Monkey.” *Journal of Neurophysiology* 56 (5) (November): 1439–61.

———. 2013. “Updating Dopamine Reward Signals.” *Current Opinion in Neurobiology* 23 (2) (April): 229–38. doi:10.1016/j.conb.2012.11.012.

Schultz, Wolfram, P Dayan, and P R Montague. 1997. “A Neural Substrate of Prediction and Reward.” *Science* 275: 1593–1599. doi:10.1126/science.275.5306.1593.

Shanks, David R., Richard J. Tunney, and John D. McCarthy. 2002. “A Re-Examination of Probability Matching and Rational Choice.” *Journal of Behavioral Decision Making* 15 (March): 233–250. doi:10.1002/bdm.413.

Sharp, Carla, John Monterosso, and P Read Montague. 2012. “Neuroeconomics: A Bridge for Translational Research.” *Biological Psychiatry* 72 (2) (July 15): 87–92. doi:10.1016/j.biopsych.2012.02.029.

Shohamy, Daphna, and R. Alison Adcock. 2010. “Dopamine and Adaptive Memory.” *Trends in Cognitive Sciences* 14 (10): 464–472. doi:10.1016/j.tics.2010.08.002.

Shohamy, Daphna, and Nicholas B Turk-Browne. 2013. “Mechanisms for Widespread Hippocampal Involvement in Cognition.” *Journal of Experimental Psychology. General* 142 (4): 1159–70. doi:10.1037/a0034461.

Shteingart, Hanan, and Yonatan Loewenstein. 2014. “Reinforcement Learning and Human Behavior.” *Current Opinion in Neurobiology* 25: 93–98. doi:10.1016/j.conb.2013.12.004.

Siegel, L S, and E B Ryan. 1989. “The Development of Working Memory in Normally Achieving and Subtypes of Learning Disabled Children.” *Child Development* 60 (4): 973–980. doi:10.2307/1131037.

Simons-Morton, Bruce, Neil Lerner, and Jeremiah Singer. 2005. “The Observed Effects of Teenage Passengers on the Risky Driving Behavior of Teenage Drivers.” *Accident Analysis and Prevention* 37: 973–982. doi:10.1016/j.aap.2005.04.014.

Singmann, Henrik. 2013. “Afex: Analysis of Factorial Experiments.”

Skatova, Anya, Patricia a Chan, and Nathaniel D Daw. 2013. “Extraversion Differentiates between Model-Based and Model-Free Strategies in a Reinforcement Learning Task.” *Frontiers in Human Neuroscience* 7 (September) (January): 525. doi:10.3389/fnhum.2013.00525.

Skinner, Bf. 1938. “The Behavior of Organisms: An Experimental Analysis.” *The Psychological Record*: 486.

Sloman, Steven a. 1996. “The Empirical Case for Two Systems of Reasoning.” *Psychological Bulletin* 119 (1): 3–22. doi:10.1037/0033-2909.119.1.3.

Smittenaar, Peter, Thomas H B FitzGerald, Vincenzo Romei, Nicholas D Wright, and Raymond J Dolan. 2013. “Disruption of Dorsolateral Prefrontal Cortex Decreases Model-Based in Favor of Model-Free Control in Humans.” *Neuron* 80 (4) (November 20): 914–9. doi:10.1016/j.neuron.2013.08.009.

Somerville, Leah H. 2013. “Special Issue on the Teenage Brain: Sensitivity to Social Evaluation.” *Current Directions in Psychological Science* 22: 121–127. doi:10.1177/0963721413476512.

Somerville, Leah H, and BJ Casey. 2010. “Developmental Neurobiology of Cognitive Control and Motivational Systems.” *Current Opinion in Neurobiology* 20 (2): 271–277. doi:10.1016/j.conb.2010.01.006.

Somerville, Leah H, T Hare, and BJ Casey. 2011. “Frontostriatal Maturation Predicts Cognitive Control Failure to Appetitive Cues in Adolescents.” *Journal of Cognitive Neuroscience* 23 (9): 2123–2134. doi:10.1162/jocn.2010.21572.Frontostriatal.

Somerville, Leah H, Rebecca M. Jones, and B. J. Casey. 2010. “A Time of Change: Behavioral and Neural Correlates of Adolescent Sensitivity to Appetitive and Aversive Environmental Cues.” *Brain and Cognition* 72 (1): 124–133. doi:10.1016/j.bandc.2009.07.003.

Spear, L P. 2000. *The Adolescent Brain and Age-Related Behavioral Manifestations. Neuroscience and Biobehavioral Reviews*. Vol. 24.

Stanovich, Keith E. 2003. “Is Probability Matching Smart? Associations between Probabilistic Choices and Cognitive Ability.” *Memory & Cognition* 31 (2): 243–251. doi:10.3758/BF03194383.

Staudinger, Markus R., and Christian Büchel. 2013. “How Initial Confirmatory Experience Potentiates the Detrimental Influence of Bad Advice.” *NeuroImage* 76: 125–133. doi:10.1016/j.neuroimage.2013.02.074.

Steinberg, Laurence, Sandra Graham, Lia O'Brien, Jennifer Woolard, Elizabeth Cauffman, and Marie Banich. 2009. "Age Differences in Future Orientation and Delay Discounting." *Child Development* 80 (1): 28–44.

Steinberg, Laurence, and Kathryn C Monahan. 2007. "Age Differences in Resistance to Peer Influence." *Developmental Psychology* 43 (6): 1531–1543. doi:10.1037/0012-1649.43.6.1531.Age.

Steinglass, Joanna E, Bernd Figner, Staci Berkowitz, H Blair Simpson, Elke U Weber, and B Timothy Walsh. 2012. "Increased Capacity to Delay Reward in Anorexia Nervosa." *Journal of the International Neuropsychological Society : JINS* 18 (4) (July): 773–80. doi:10.1017/S1355617712000446.

Steinglass, Joanna E, and Deborah R Glasofer. 2011. "Eating Disorders and the Brain." In *Neuropsychology.*, 106–121. Wiley Online Library.

Stephan, Klaas Enno, and Christoph Mathys. 2014. "Computational Approaches to Psychiatry." *Current Opinion in Neurobiology* 25 (April): 85–92. doi:10.1016/j.conb.2013.12.007.

Story, Giles W, Ivo Vlaev, Ben Seymour, Ara Darzi, and Raymond J Dolan. 2014. "Does Temporal Discounting Explain Unhealthy Behavior? A Systematic Review and Reinforcement Learning Perspective." *Frontiers in Behavioral Neuroscience* 8 (March) (January): 76. doi:10.3389/fnbeh.2014.00076.

Strack, Fritz, and Roland Deutsch. 2004. "Reflective and Impulsive Determinants of Social Behavior." *Personality and Social Psychology Review : An Official Journal of the Society for Personality and Social Psychology, Inc* 8 (3): 220–247. doi:10.1207/s15327957pspr0803_1.

Strang, Nicole M, Jason M Chein, and Laurence Steinberg. 2013. "The Value of the Dual Systems Model of Adolescent Risk-Taking." *Frontiers in Human Neuroscience* 7 (May): 223. doi:10.3389/fnhum.2013.00223.

Strommen, Ellen a. 1973. "Verbal Self-Regulation in a Children's Game : Impulsive Errors on 'Simon Says.'" *Child Development* 44 (4): 849–853. doi:10.2307/1127737.

Sullivan, Patrick F. 1995. "Mortality in Anorexia." *American Journal of Psychiatry*: 1073–1074.

Suominen, K H, E T Isometsä, M M Henriksson, a I Ostamo, and J K Lönnqvist. 1998. "Inadequate Treatment for Major Depression Both before

and after Attempted Suicide." *The American Journal of Psychiatry* 155 (December): 1778–1780.

Sutton, Richard S, and Andrew G Barto. 1998. *Reinforcement Learning: An Introduction*. MIT Press. doi:10.1109/TNN.1998.712192.

Tanji, Jun, and Eiji Hoshi. 2001. "Behavioral Planning in the Prefrontal Cortex." *Current Opinion in Neurobiology* 11: 164–170. doi:10.1016/S0959-4388(00)00192-6.

Taylor, Jane R., and Trevor W. Robbins. 1986. "6-Hydroxydopamine Lesions of the Nucleus Accumbens, but Not of the Caudate Nucleus, Attenuate Enhanced Responding with Reward-Related Stimuli Produced by Intra-Accumbens D-Amphetamine." *Psychopharmacology* 90 (t 986): 390–397. doi:10.1007/BF00179197.

Thomas, Kathleen M, Ruskin H Hunt, Nathalie Vizueta, Tobias Sommer, Sarah Durston, Yihong Yang, and Michael S Worden. 2004. "Evidence of Developmental Differences in Implicit Sequence Learning: An fMRI Study of Children and Adults." *Journal of Cognitive Neuroscience* 16: 1339–1351. doi:10.1162/0898929042304688.

Thompson-Schill, Sharon L, Michael Ramscar, and Evangelia G Chrysikou. 2009. "When a Little Frontal Lobe Goes a Long Way." *CURRENT DIRECTIONS IN PSYCHOLOGICAL SCIENCE* 18 (5): 259–263.

Thorndike, Edward L. 1898. "Animal Intelligence: An Experimental Study of the Associative Processes in Animals." *Psychological Monographs: General and Applied* 2 (4): i–109. doi:10.1097/00005053-190001000-00013.

Tolman, Edward Chace. 1948. "Cognitive Maps in Rats and Men." *Psychological Review* 55 (4): 189–208. doi:10.1037/h0061626.

Tolman, Edward Chace, and Charles H Honzik. 1930. "Introduction and Removal of Reward, and Maze Performance in Rats." *University of California Publications in Psychology*.

Trenholm, Christopher, Barbara Devaney, Ken Fortson, Lisa Quay, Justin Wheeler, and Melissa Clark. 2007. *Impacts of Four Title V , Section 510 Abstinence Education Programs: Final Report*. Mathematica Policy Research, Inc. Princeton, NJ.

Tricomi, Elizabeth, Bernard W. Balleine, and John P O'Doherty. 2009. "A Specific Role for Posterior Dorsolateral Striatum in Human Habit

Learning." *European Journal of Neuroscience* 29 (April): 2225–2232. doi:10.1111/j.1460-9568.2009.06796.x.

Turk-Browne, Nicholas B, Brian J Scholl, Marcia K Johnson, and Marvin M Chun. 2010. "Implicit Perceptual Anticipation Triggered by Statistical Learning." *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 30 (33): 11177–11187. doi:10.1523/JNEUROSCI.0858-10.2010.

Tversky, Amos, and Daniel Kahneman. 1971. "Belief in the Law of Small Numbers." *Psychological Bulletin* 76 (2): 105–110. doi:10.1037/h0031322.

Tymula, Agnieszka, Lior a Rosenberg Belmaker, Amy K Roy, Lital Ruderman, Kirk Manson, Paul W Glimcher, and Ifat Levy. 2012. "Adolescents' Risk-Taking Behavior Is Driven by Tolerance to Ambiguity." *Proceedings of the National Academy of Sciences* 109 (16) (October 16): 17135–17140. doi:10.1073/pnas.1207144109.

Ullman, H., R. Almeida, and T. Klingberg. 2014. "Structural Maturation and Brain Activity Predict Future Working Memory Capacity during Childhood Development." *Journal of Neuroscience* 34 (5): 1592–1598. doi:10.1523/JNEUROSCI.0842-13.2014.

Valentin, Vivian V, Anthony Dickinson, and John P O'Doherty. 2007. "Determining the Neural Substrates of Goal-Directed Learning in the Human Brain." *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 27 (15): 4019–4026. doi:10.1523/JNEUROSCI.0564-07.2007.

Van den Bos, Wouter, Michael X Cohen, Thorsten Kahnt, and Eveline a Crone. 2012. "Striatum-Medial Prefrontal Cortex Connectivity Predicts Developmental Changes in Reinforcement Learning." *Cerebral Cortex* 22 (6) (June): 1247–55. doi:10.1093/cercor/bhr198.

Van den Bos, Wouter, Berna Güroğlu, Bianca G van den Bulk, Serge A R B Rombouts, and Eveline A Crone. 2009. "Better than Expected or as Bad as You Thought? The Neurocognitive Development of Probabilistic Feedback Processing." *Frontiers in Human Neuroscience* 3 (December) (January): 52. doi:10.3389/neuro.09.052.2009.

Van der Meer, Matthijs, Zeb Kurth-Nelson, and a David Redish. 2012. "Information Processing in Decision-Making Systems." *The Neuroscientist* 18 (4) (August): 342–59. doi:10.1177/1073858411435128.

Van Kuyck, Kris, Nathalie Gérard, Koen Van Laere, Cindy Casteels, Guido Pieters, Loes Gabriëls, and Bart Nuttin. 2009. “Towards a Neurocircuitry in Anorexia Nervosa: Evidence from Functional Neuroimaging Studies.” *Journal of Psychiatric Research* 43 (14) (September): 1133–45. doi:10.1016/j.jpsychires.2009.04.005.

Vanderschuren, Louk J M J, and Barry J Everitt. 2004. “Drug Seeking Becomes Compulsive after Prolonged Cocaine Self-Administration.” *Science (New York, N.Y.)* 305 (August): 1017–1019. doi:10.1126/science.1098975.

Voon, V, K Derbyshire, C Rück, M a Irvine, Y Worbe, J Enander, L R N Schreiber, et al. 2015. “Disorders of Compulsivity: A Common Bias towards Learning Habits.” *Molecular Psychiatry* 20: 345–352. doi:10.1038/mp.2014.44.

Voorn, Pieter, Louk J M J Vanderschuren, Henk J. Groenewegen, Trevor W. Robbins, and Cyriel M a Pennartz. 2004. “Putting a Spin on the Dorsal-Ventral Divide of the Striatum.” *Trends in Neurosciences* 27 (8): 468–474. doi:10.1016/j.tins.2004.06.006.

Wagner, Allan R., and Robert a. Rescorla. 1972. “Inhibition in Pavlovian Conditioning: Application of a Theory.” *Inhibition and Learning*.

Wagner, Angela, Howard Aizenstein, Vijay Venkatraman, Julie Fudge, J Christopher May, Laura Mazurkewicz, Guido K W Frank, et al. 2007. “Altered Reward Processing in Women Recovered From From Anorexia Nervosa.” *American Journal of Psychiatry* 164 (December): 1842–1849.

Wahlstrom, Dustin, Paul Collins, Tonya White, and Monica Luciana. 2010. “Developmental Changes in Dopamine Neurotransmission in Adolescence: Behavioral Implications and Issues in Assessment.” *Brain and Cognition* 72 (1) (February): 146–59. doi:10.1016/j.bandc.2009.10.013.

Walsh, BT. 2013. “The Enigmatic Persistence of Anorexia Nervosa.” *American Journal of Psychiatry* 170 (May): 477–484.

Wang, X T, and Robert D Dvorak. 2010. “Sweet Future: Fluctuating Blood Glucose Levels Affect Future Discounting.” *Psychological Science* 21 (2) (February): 183–8. doi:10.1177/0956797609358096.

Watkins, Christopher J C H, and Peter Dayan. 1992. “Technical Note Q , -Learning.” *Machine Learning* 292: 279–292. doi:10.1007/978-1-4615-3618-5_4.

Wechsler, D. 2001. *Weschler Test of Adult Reading*. *San Antonio, TX: The Psychological Corporation*.

Wesley, Michael J, and Warren K Bickel. 2013. “Remember the Future II: Meta-Analyses and Functional Overlap of Working Memory and Delay Discounting.” *Biological Psychiatry* 75 (6) (September 13): 435–448. doi:10.1016/j.biopsych.2013.08.008.

Wierenga, CE, Amanda Bischoff-Grethe, BS Melrose, Z Irvine, BS Torres, UF Bailer, A Simmons, et al. 2014. “Hunger Does Not Motivate Reward in Women Remitted from Anorexia Nervosa.” *Biological Psychiatry* doi:10.101: In Press.

Wilson, Robert C, and Yael Niv. 2013. “Is Model Fitting Necessary for Model-Based fMRI?” *RLDM 2013*.

Wilson, Robert C., Yuji K. Takahashi, Geoffrey Schoenbaum, and Yael Niv. 2014. “Orbitofrontal Cortex as a Cognitive Map of Task Space.” *Neuron* 81 (2): 267–278. doi:10.1016/j.neuron.2013.11.005.

Wittchen, H.-U., C. B. Nelson, and G. Lachner. 1998. “Prevalence of Mental Disorders and Psychosocial Impairments in Adolescents and Young Adults.” *Psychological Medicine* 28: 109–126. doi:10.1017/S0033291797005928.

Wittmann, Marc, David S Leland, and Martin P Paulus. 2007. “Time and Decision Making: Differential Contribution of the Posterior Insular Cortex and the Striatum during a Delay Discounting Task.” *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale* 179 (4) (June): 643–53. doi:10.1007/s00221-006-0822-y.

Worthy, Darrell a, Jessica a Cooper, Kaileigh a Byrne, Marissa a Gorlick, and W Todd Maddox. 2014. “State-Based versus Reward-Based Motivation in Younger and Older Adults.” *Cognitive, Affective & Behavioral Neuroscience*: 1208–1220. doi:10.3758/s13415-014-0293-8.

Wunderlich, Klaus, Peter Smittenaar, and Raymond J Dolan. 2012. “Dopamine Enhances Model-Based over Model-Free Choice Behavior.” *Neuron* 75 (3) (August 9): 418–24. doi:10.1016/j.neuron.2012.03.042.

Yin, Henry H, and Barbara J Knowlton. 2006. “The Role of the Basal Ganglia in Habit Formation.” *Nature Reviews Neuroscience* 7 (6): 464–476.

Yin, Henry H, Barbara J Knowlton, and Bernard W Balleine. 2004. “Lesions of Dorsolateral Striatum Preserve Outcome Expectancy but Disrupt Habit

Formation in Instrumental Learning." *European Journal of Neuroscience* 19 (August 2003): 181–190. doi:10.1046/j.1460-9568.2003.03095.x.

Yin, Henry H, Sean B Ostlund, and Bernard W. Balleine. 2008. "Reward-Guided Learning beyond Dopamine in the Nucleus Accumbens: The Integrative Functions of Cortico-Basal Ganglia Networks." *European Journal of Neuroscience* 28 (8): 1437–1448. doi:10.1111/j.1460-9568.2008.06422.x.Reward-guided.

Yin, Henry H, Sean B Ostlund, Barbara J Knowlton, and Bernard W Balleine. 2005. "The Role of the Dorsomedial Striatum in Instrumental Conditioning." *European Journal of Neuroscience* 22: 513–523. doi:10.1111/j.1460-9568.2005.04218.x.

Yin, Henry H., and Barbara J. Knowlton. 2004. "Contributions of Striatal Subregions to Place and Response Learning." *Learning & Memory* 11: 459–463. doi:10.1101/lm.81004.

Yoon, Jin H, Stephen T Higgins, Sarah H Heil, Rena J Sugarbaker, Colleen S Thomas, and Gary J Badger. 2007. "Delay Discounting Predicts Postpartum Relapse to Cigarette Smoking among Pregnant Women." *Experimental and Clinical Psychopharmacology* 15 (2) (April): 176–86. doi:10.1037/1064-1297.15.2.186.

Zastrow, Arne, Stefan Kaiser, Christoph Stippich, Stephan Walther, Wolfgang Herzog, Kate Tchanturia, Aysenil Belger, Matthias Weisbrod, Janet Treasure, and Hans-Christoph Friederich. 2009. "Neural Correlates of Impaired Cognitive-Behavioral Flexibility in Anorexia Nervosa." *The American Journal of Psychiatry* 166 (5) (May): 608–16. doi:10.1176/appi.ajp.2008.08050775.

Zelazo, Philip David, Douglas Frye, and Tanja Rapus. 1996. "An Age-Related Dissociation between Knowing Rules and Using Them." *Cognitive Development* 11: 37–63. doi:10.1016/S0885-2014(96)90027-1.