NORTHWESTERN UNIVERSITY

Dynamic Decision Making Under Uncertainty

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

For the degree

DOCTOR OF PHILOSOPHY

Field of Psychology

By

Ben Reuveni

EVANSTON, ILLINOIS

June 2020

ABSTRACT

Dynamic decision-making is a complex process that relies on our ability to generate, evaluate, and implement a variety of strategies. Understanding how people navigate this process is a difficult problem that requires a wide range of methodologies. This study details a combination of behavioral experiments, computational modeling, and neuroimaging that complement each other in describing how people engage in dynamic decision-making under uncertainty. To create opportunities to observe such decision-making participants were taught two categories by sorting sine-wave gratings selected by an adaptive, real-time computational model, PINNACLE 2.0. During the protocol, participants were presented with a challenging, dynamic decision-making task that experimentally prolongs strategy exploration while fMRI data were collected. During this task, participants displayed a broad range of behavior indicative of a variety of explicit strategy use and evaluation. Accounting for the results of this experiment proved challenging for existing computational models of category learning. In developing better accounts of the data, and how people navigated this task, we rule out several previously successful models. The models considered include exemplar, and rule-based models that include parallel and sequential strategy representations, incremental rule modification and rule replacement mechanisms, and single-step and hierarchical rule structures. Through competitive model fitting, we further the development of the PINNACLE architecture culminating in version 2.1a. This version represents the best account of participant behavior to-date. Finally, by contrasting successful and unsuccessful learning on this task, we describe preliminary evidence of the neural correlates of decision-making under uncertainty. How well people do on the task is a function of their performance expectation. Those with high expectations engage in more strategy generation, evaluation, and replacement, and tend to succeed by finding better rules whereas those with lower expectations tend to settle for less successful ones.

ACKNOWLEDGEMENTS

Writing this document was only the last step in a near decade long journey. Of course, a full account of how I got here would see me thanking errant electrons, primordial protozoa and altruistic apes. I think that would add yet another hundred pages to this already wordy work. As a more succinct alternative, I would like to take some space to thank the numerous people who directly influenced me as I navigated my own journey of decision-making under uncertainty.

The first person I would like to acknowledge is Professor Matthew J. Crossley. When I first arrived at Berkeley, I knew that diving into research would be central to my development as a scientist. I still remember agonizing over whether to apply to a position posted by Dr. Crossley because it required that I learn to code. That was 2012 and here we are, at the beginning of a paper in no small part devoted to computational models. Dr. Crossley was not only my first academic mentor, but also was and continues to be a great friend. He introduced me to category learning, memory systems interactions, and scientific research. Who knows where I would be if I had not replied to that listing?

In trying to gain competency in programming, I was lucky enough to have a close friend I met years before while running an online play-by-email/wiki roleplaying game. Earl Brown took me under his wing and spent countless hours tutoring me, helping me work through the frustrations and mental knots that arise when trying to piece together something as complex as a psychological experiment as your first programming project. Without Earl, there is no question my productivity would not have been a tenth of what it was. I am happy to say that his fingerprints can still be found in my code today.

The first lab I ever joined was Professor Richard B. Ivry's Cognac lab. Little did I know I had just been accepted to work under one of the founding fathers of cognitive neuroscience. Prof. Ivry consistently dazzled me with his ability to find simplicity in complex tangles of data and ideas. His kindness and interpersonal warmth never wavered, and I have held him as a standard ever since.

Deciding which lab to join for graduate school was one decision I would not like to make again. It was one of those hard decisions where the difficulty lay in the fact that each option was so attractive. Ultimately, my choice was clear, and I joined Professor Paul J. Reber's lab. Perhaps the single most important deciding factor in my choice was and still is the fact that Professor Reber's office is quite literally a paper-airplane's toss away from his graduate students. Over the last six years Professor Reber and I spent countless hours talking about everything from why minors should be allowed to vote, to the intricacies of DOTA2, to the relative merits of chess openings and how to quantify the happiness curve. All these were just brain-candy as he calls them. They were supplementary to the profound knowledge and insight he shared with me about human psychology, the scientific method, category learning, implicit and explicit memory systems, computational modeling, decision-making, and topics I have learned and forgotten since. Professor Reber was always been willing to chat "for a moment" which inevitably turn into several hour conversations about things I did not even realize I was interested in. He is not only a meticulous and knowledgeable scientist but also a warm and kind mentor. There is no question that without his guidance and help I would be a far worse scientist than I am today.

Anyone who has been to graduate school knows that few survive this process without social support. Alongside my professional development were several friends who helped me through the good times, and the bad. Sadie Witkowski (soon to be Dr. Witkowski!) who I

literally met on my first day and who has been my steadfast friend and gym-partner throughout graduate school. Your kindness, insight, wit, and determination have helped me immensely. My friends: Amanda Kramer, Alexander (Sandy) LaTourrette, Chris Petsko, Erin Anderson, Francisco Maravilla, you have all surrounded me with kindness and support across the years. Cat Han and Max Smith, you both have given me so much to think about, I can't picture my graduate experience without your companionship.

Science is the incremental process of standing on the shoulders of giants and reaching ever higher. My reach would not extend as far without the support and advice of faculty members outside my lab. My committee: Professors Kenneth A. Paller, Satoru Suzuki, Caterina Gratton have been a wellspring of insight and advice. Their doors have always been open to talk about difficulties in data or process, working out computational complexity, or brainstorming ideas on how to move forward. In addition, Professor Marcia Grabowecky has been a constant source of advice and knowledge, always willing to talk and lend her considerable knowledge.

Finally, I would like to acknowledge those people in my life that formed the foundation upon which my education took place. My mom and dad, my two brothers, and my partner Amy M. Ellifritz. Without them I would not have been able to accomplish all that I have. Their support both professionally and interpersonally have impacted me far more than a few words can convey.

To all of you, and those I did not explicitly mention, thank you.

TABLE OF CONTENTS

LIST OF TABLES AND EQUATIONS

LIST OF FIGURES

Chapter 1: Introduction

Buy or sell? Accept or reject? Malignant or benign? And the all-important vanilla or chocolate? We are often faced with complex decisions in changing environments with access to limited options. To make matters worse the solution to these problems frequently requires exploring many possible strategies such as using rules, going with your gut, or similarity-based judgements in a short amount of time. The way people solve these problems can involve a complex process of strategy generation and evaluation. When this process succeeds, appropriate solutions are quickly found. Yet when it fails, valuable time can be spent exploring irrelevant strategies while overlooking partially successful ones. To study this complex process, we modified a well-established category learning task. This task required participants to adapt to changing task demands by searching for appropriate strategies when old ones failed. The task begins by presenting people with a simple image and asking a simple question: Is this an ‘A’ or a ‘B’?

Much of the research on how people navigate such a task has focused on characterizing the learning process associated with acquiring categories. However, less attention has been spent on the decision-making component that leads to successful or unsuccessful learning. Unfortunately, we are not able to analyze people’s thought process directly. Instead, we need a way to plausibly account for how participants navigate this complex and dynamic decision-making task. To achieve this, we leverage a substantial body of work on memory systems, decision-making, computational modeling, and neuroimaging.

Multiple Memory Systems

Memory systems research (Squire, 1992; Squire & Dede, 2015*)* has provided substantial evidence for separate neural systems supporting learning and memory in the brain. Explicit memory can be acquired from a single experience, is consolidated over an extended period of time, and depends critically on neurobiological processes within the medial temporal lobe (MTL) memory system, particularly the hippocampus (Reber 2008; Ashby, 2013). When these structures are damaged, the effect is a profound deficit in the acquisition of new declarative memories (anterograde amnesia).

However, patients with MTL damage exhibit intact learning on a variety of tasks, indicating selective preservation of some specific forms of memory (Squire, 1992; Knowlton, Mangels & Squire, 1996). This type of memory can be described as the product of learning without awareness, known as implicit memory. Implicit memory is mostly acquired gradually, is largely unavailable to conscious awareness, does not depend on the MTL memory system and is generally inflexible to generalization. It is knowledge we express effortlessly, does not require working memory, and can arise as an intuition or "gut feeling" (Reber, Beeman, Paller, 2013).

In combining these two types of memory, we adopt a parallel, competitive, and interactive view of these two memory systems in which each system learns independently, and continuously vies for control of behavior. However, given our data this study is primarily focused on explicit rule discovery, hypothesis testing, and decision-making rather than the interplay between systems.

Decision-Making

Whether one should continue using a specific strategy or abandon it in search of a better one depends on the context and scope of the problem being solved. What is good enough in one context may not be in another. For example, consider a scenario in which you discover a strategy that reliably leads to 60% success where chance is 50%. Consideration of whether or not to use this strategy is in part a function of whether one believes there is a better strategy to be found, and what the potential gain or loss is while attempting to find it. In circumstances where the stakes are high, such as in survival situations or where substantial amounts of money are involved, the rational course of action is to stick with the safe (60%) option. Even if one believes there is a better solution, the possibility of failure may not be worth the risk. On the other hand, if the penalty for low performance is not too costly, and one believes a better solution can be found, it is rational to sacrifice near-term performance for the possibility of discovering a better solution.

The question then becomes: what level of performance is “good enough”? In our example, we used 60%, but why should someone settle for 60% and not 55%, or 75%? One solution to this question comes from Herbert Simon who, in 1956, introduced the term “satisficing”. He argued that finding truly optimal solutions to most problems is unrealistic due to the sheer complexity of the world, our limited mental capacity, and a lack of information. Instead, people seem to settle on solutions that are both *satisfy* their aspirations, and *suffice* in accomplishing their goal, and thus *satisfice*. This simple yet intuitive framework for thinking about how people evaluate strategies has been a powerful tool in understanding decision-making, especially under situations of uncertainty. The overall strategy of satisficing seems to apply across a wide range of human behavior. However, the question of what the individual satisficing standard is appears to be an individual difference that varies from person to person, and from task to task.

Computational Modeling

The human ability to reason is a cornerstone of science and a primary way in which we learn. Unfortunately, human reasoning is not without its limitations. Beginning with Kahneman and Tversky (1975) and more recently Hintzman (1991) a long list of ways in which human reasoning fails seems to only get longer. Of note are limitations in working memory and our ability to reason about complicated interacting structures (e.g. White 2008). In addition, it is often the case that our understanding of a process is only as useful as our ability to communicate it clearly to others. Conveying mathematically accurate descriptions of a theory or hypothesis is invaluable in advancing the state of understanding.

A practical solution to these limitations in the form of computational modeling was well described by Farrell and Lewandowsky (2010). In their argument they laid out various limitations that confront the pursuit of science as well as how computational modeling provides solutions to them. Addressing the concerns above, they point out that computational models are far less limited in terms of working memory and do not suffer from failures of memory retrieval under normal circumstances. In addition, the nature of computational operations avoids the limitations of complexities due to higher order interactions, massively parallel computations or handling distributed structures. Computational models must be programmed meticulously in order to function which allows for clear descriptions of a theory to be shared across people, thus avoiding issues of misunderstandings or underspecified hypotheses.

Moreover, when using computational models to describe data, the evaluation of how well a model accounts for the data happens in the context of alternative models. Through this competitive process of comparing and developing better models, we can systematically rule out alternative accounts for behavior. These formal systems behave identically regardless of who runs them and what their personal beliefs about the solution might be, thus avoiding the ever-present confirmation bias. Finally, though not exhaustively, the process of developing rigorous accounts of behavior often leads to unanticipated insights into processes and mechanisms that were not apparent earlier thus furthering our understanding. To quote Herbert Simon, *"If we can simulate it we have learned something about an important human activity".*

The current work examines a series of hypotheses about how people engage in a complex decision-making task under uncertainty. Several plausible accounts are ruled out based on available data and a new framework for explaining this process as well as its neural correlates are described. Understanding how people navigate the decision-making process is critical to many real-world applications from better education practices, to improved healthcare, to economic success and interpersonal relationships. By modeling this process we not only gain deeper insights into what is happening, but also gain a framework for testing hypotheses for how to improve outcomes across these various domains.

Chapter 2: Prior Work

The laboratory category learning task provides a deceptively simple model of a rich and complex set of cognitive processes. In a typical design, the participant sits down to the task naïve to the underlying category structure. A stimulus is presented on the screen that can be a sine-wave grating (or Gabor patch: Nomura et al., 2007), a simple single line (Ashby & Waldron, 1999), an angle bracket (Ashby & Gott, 1988), a circle with a line through it (Maddox & Ashby, 1993), an artificial animal (Smith & Grossman, 2008), etc. As a participant, you are directed to determine which of a small number of categorical options the stimulus belongs. The most common designs use two categories (2AFC) or a *member* or *non-member* discrimination (*A*, *not A* designs). The categories themselves can be given either arbitrary labels (*A*, *B*) or semantic ones (*builder*, *digger*). Generally little to no information is provided for the answer, so as a participant you are then posed with the problem of identifying the correct label, perhaps guessing initially, and then learning to choose more accurately over the course of many trials, usually guided by feedback.

Considering the position of the participant who is early in the learning process with only a few trials of experience, it is clear that there are a number of different potential approaches that they could take. On trial 10, a participant might use an explicit memory comparison to prior examples they have seen, i.e., “this one looks like one I saw before that was an *A*, so maybe this one is an *A* too.” In addition, participants (often undergraduates) who have been instructed that there are two categories or types of stimuli are very likely searching for rules that make sense of the stimuli they have seen, e.g., “the ones with more stripes seem to be *B*’s.” The rule discovery process is a cognitively complex one, involving hypothesis generation and testing against the experimental stimuli over trials, remembering prior outcomes, etc. The neuropsychological evidence on memory systems reviewed in Chapter 1 indicates that a third process is also likely to be involved: implicit learning of the underlying category stimulus dimension statistics that presents as a hunch or high-accuracy guessing about category membership without an ability to verbalize the information used to guide the response. While this study was initially concerned with both explicit and implicit processes and their interactions, the data suggest that participants primarily engaged in explicit strategy generation and use during the task. We thus focus our attention on the processes of explicit rule generation and evaluation.

In the following sections we provide background information on decision-making under uncertainty as well as the systems neuroscience of this process. We then present several prominent cognitive computational models that have previously been used to study the category learning process. We discuss several approaches including single, hybrid, and multiple system models. These models all appear to converge on the need for multiple forms of knowledge representation to adequately describe human behavior across a broad set of data. This body of work provides a solid foundation for the current study aimed at furthering our understanding of the dynamic decision-making process.

Decision-Making Under Uncertainty

Much of the research on visual category learning has focused on the various learning processes by which participants come to know the category structure (i.e. explicit / implicit). However, less work has focused on the decision-making aspect of this task. Specifically, how do participants generate and test hypotheses that inform learning, and how do they respond to feedback? Ideally, participants would aggregate information and methodically test a variety of potential solutions to find the best possible strategy. Unfortunately, several limitations stand in the way of such a strategy. Explicit strategy use is often necessary for early task performance or following situations where the task changes since any implicit knowledge previously gained may no longer apply. To that end, a good mental model of the task and stimulus space helps direct the search for appropriate strategies. Unfortunately, participants navigating this task must rely on working memory to create such models and attempt to map out the category structures. This process is intentionally hindered both by the rapid pace of the task and the complexity of the category structure which make keeping track of all the information they previously saw difficult or impossible. Since there is not enough time to explore every possible solution, and given the information constraints, participants must come up with alternative strategies for success.

One approach for how people generate and evaluate different strategies comes from work done by Herbert Simon on rational decision making. In a 1956 paper Simon argued that people generally lack the capacity, information, and time to truly optimize their strategies, as is the case here. Instead, he introduced the concept of *satisficing.* As noted earlier the idea of satisficing is that instead of exploring all possible solutions in an attempt to optimize the outcome, people select a solution that is both sufficient for the task and that satisfies their aspiration, hence the term satisficing. This simple principle provides a powerful and intuitive account for how people actually behave in many real-world circumstances. For example, suppose we are in the market for a new shirt. From a rational decision-making framework, an optimal strategy would be to visit every clothing store and gather information on every shirt, then compute the optimal purchase choice while maximizing utility and minimizing cost. This is obviously not what people do. Instead, we may visit a few stores, look at several options and finally settle on a shirt that is good enough for our needs. There is certainly a better shirt to be found, but we select one that is both sufficient (it is a shirt) and satisfies our criteria (aesthetics, price, quality, etc.). We satisfice.

This principle can similarly be applied to situations in which participants search for a strategy that makes sense of the evidence they have seen thus far, such as in category learning experiments. If the goal were to discover the best possible solution to the task, several highly successful, though not perfect, solutions would need to be discarded along the way in the hope that a better one would be found. This is not what people seem to do during these tasks. Instead they seem to adopt a strategy that works most of the time and tolerate mistakes even though it demonstrates that their rule is not as good as it could be. They satisfice.

Systems Neuroscience of Decision-Making

Decades of research have identified key areas of the brain associated with various aspects of decision-making. No single area of the brain is solely responsible for such a complex process. Rather, various regions contribute to this process based on the particular nature of the task. Even so, there seem to be common denominators that support the types of computations necessary to evaluate between alternatives. A key area identified in the decision-making process across numerous studies is the dorsolateral prefrontal cortex (DLPFC). This area is well situated to send and receive inputs from a variety of brain areas, including the basal ganglia, the frontal cortex and primary and secondary association areas of neocortex, including parietal, and occipital areas (Tekin & Cummings, 2002; Dosenbach et al., 2007). It has been shown to play crucial roles in a variety of attention-demanding cognitive tasks such as decision-making (van’t Wout et al., 2005; Heekeren, Marrett, Bandettini & Ungerleider, 2006; Nomura & Reber, 2012), working-memory (Curtis & D’Esposito, 2003), reasoning (Goel & Dolan, 2004), and problem solving (Barbey &

Barsalou, 2009). Another key area found in several studies of decision-making, especially those involving a perceptual component is the parietal cortex. Several studies have shown its involvement in working-memory tasks (Bunge, Hazeltine, Scanlon, Rosen & Gabrieli, 2002; Olesen, Westerberg & Klingberg, 2004; Curtis, 2006), and perceptual decision-making (Heekeren et al., 2008; Andersen & Cui, 2009) through fronto-parietal connectivity.

Several theories describing their interactions in support of decision-making suggest that they collaborate in the maintenance of working-memory and goal-directed behavior (notably WM: Curtis & D'Esposito, 2003; goal-directed: Rangel, Camerer & Montague, 2008). Across these views, DLPFC and Parietal areas coordinate in the maintenance of task relevant explicit memory and perceptual information in service of accomplishing the task at hand. Regarding category learning, this would likely manifest as tracking both stimulus information, coordinating long term memory access and retrieval, as well as maintaining and monitoring relevant explicit strategies. These strategies would be updated based on new evidence or may be replaced with new representations of possible solutions.

Models of Category Learning

The majority of category learning studies result in a behavioral trace of button presses that represent both the learning and decision-making process. We are thus confronted with the difficult task of inferring the mental processes that drove these behaviors from a pattern of 'A' and 'B' responses. Despite the potential complexity of the approaches used to learn the category structure, the fact that these categories have precise definitions within a defined set of perceptual dimensions makes the mathematical analysis of choice behavior tractable. This has led to a variety of analytical and simulation-based modeling approaches that instantiate specific hypotheses about category learning theory. An early, influential example of this is the 'context' model described in Medin & Schaffer (1978) and generalized as 'exemplar theory' in Nosofsky (1986). This general approach aligns well with the explicit strategy described earlier in which new stimuli are compared to the previously seen exemplars. At the same time, Ashby & Gott (1988) described a mathematical characterization of how rules can be used to partition the stimulus space that was eventually formalized as 'decision bound theory.' While a great deal of debate ensued comparing and contrasting these approaches (e.g., Maddox & Ashby 1993; McKinley & Nosofsky, 1996), the fact that the use of both rules and exemplars aligns well with subjective experience (of both processes) also led to the development of hybrid models of category learning that included both of these strategies. In this section we describe several prominent models that have been used to account for a variety of category learning data to provide context for the work described in the following chapters. Importantly, these models constitute hypotheses as to the process by which humans learn categories. They are mathematical formulations of the proposed mechanisms people use when confronted with such tasks. As such they make specific predictions as to the type of behavior we should expect if people indeed use similar mechanisms. When behavior generated by the model disagrees with participant behavior, it serves to disprove the current hypothesis and indicates that a more robust model is necessary. A model whose data agrees with participant behavior allows us to test future predictions through additional behavioral and neuroimaging data.

We separate these models into three classes to illustrate the wide range of options that have been explored to-date. These classes are: 1) single process models that rely on a single knowledge structure to account for behavior, 2) hybrid models that incorporate both similarity-and rule-based judgements, and 3) multi-system models that simulate distinct and parallel explicit and implicit processes.

Single System Models

If guided by parsimony, a legitimate question arises about whether a purely single-system computational model can account for the wide range of data on visual category learning. Proponents of such single system accounts of category learning have advanced models that account for observed behavior on a wide variety of tasks. However, in order to achieve this, these models must rely on differing internal knowledge structures that can arise from a single mechanism given different parameter values that cause the model to behave in qualitatively different ways. While they include a single mechanism to account for behavior, this mechanism results in more than one representation of the problem space thus demonstrating a necessity for more than one process. Two such models are the Generalized Context Model (GCM; Nosofsky, 1984, 1986, 1991; Nosofsky & Johansen, 2000), and SUSTAIN (Love & Medin, 1998a, 1998b; Love, Markman & Yamauchi, 2000; Love, Medin & Gureckis, 2004).

GCM

The Generalized Context Model (GCM; recently Nosofsky & Johansen, 2000) is a similarity-based exemplar model that uses a multidimensional shaping mechanism to convert stimulus dimensions into a perceptual space. It predicts that each stimulus a person encounters is stored in long-term memory. When a new stimulus is encountered, a similarity score is computed between it, and all previously experienced stimuli. An attentional weighting mechanism is included to bias the system towards relevant dimensions based on feedback either exaggerating or minimizing differences between stimuli. For example, a stimulus with two dimensions (line length and line orientation) might be associated with a category based only on line-length. People would thus learn to weight similarity across the length dimension more heavily than the orientation dimension, resulting in rule-like behavior.

To illustrate a family of challenges that arise for a single system exemplar-similarity model, we consider one that was raised by Nosofsky & Johansen (2000) in which the model initially failed to match participant behavior on a category vs. similarity task. Participants were asked to imagine 3-inch diameter circles and decide if they were more similar to a pizza or a quarter, and then categorize them. Most participants said the stimulus was more similar to a quarter but categorized it as a pizza. These results make sense because while a 3-inch disk is small like a quarter, there are no 3-inch quarters. On the other hand, there are 3-inch pizzas (e.g. pizza bites). Attempting to model this behavior proved problematic due to the nature of knowledge representation within the model. Because quarters are all very similar in size, the variability of that category is quite small compared to the high variability of pizzas. As a result, any stimulus that is equidistant from boundary examples of each category (i.e. the largest quarter vs. the smallest pizza) will always be mathematically more similar to the less variable category (in this case, quarters). This is because the stimulus being considered is closer to the center of mass of the less variable category. This led the model to respond *quarter* for both judgments (similarity, and category membership). Mathematically, the failure of the model to account for participant behavior stemmed from the parameter that governs how quickly similarity decreases as a function of distance. To overcome this, the authors made an allowance for a similarity sensitivity parameter value specific to each category, thus allowing differential consideration of a single stimulus for each category. While it may seem like a minor alteration, the resulting knowledge representations (i.e. sharp- vs. variable, gist-like categories) are reflective of fundamentally different processes in the brain corresponding to cortical vs. hippocampal representations. This issue is illustrative of a broader set of problems that systems with a single knowledge representation encounter in which different a single strategy must suffice for a variety of tasks.

SUSTAIN

SUSTAIN is a clustering model of human category learning that predicts that people follow five basic principles in how they learn: 1) simple is better, 2) cluster based on similarity, 3) learning involves both unsupervised and supervised processes 4) inferred category structures are shaped by feedback, and 5) clusters compete to control behavior. Following these principles, SUSTAIN begins with a relevant stimulus dimension set (either featural or spatial) and a single cluster representing a simple hypothesis that governs category membership. In addition, an attentional weighting mechanism helps direct the model towards relevant features and down-weight irrelevant ones. As the model engages with the task, it gradually adds complexity in response to surprising events such as negative feedback to an incorrect decision (supervised learning) or upon encountering a stimulus significantly different from any category (unsupervised learning). If a stimulus is encountered that is not similar enough to be considered part of an existing cluster, the model creates a new one to accommodate the stimulus. If the new cluster is found to be a more reliable predictor of future category membership, its attentional weighting is modified such that it comes to dominate the response competition versus the previous cluster. In this way, SUSTAIN suggests that people discover the minimally sufficient complex rule for solving the given task. SUSTAIN has been very successful in accounting for a variety of human behavioral data.

SUSTAIN, though considered a single-system model, employs two distinct processes. This can be seen in the process of creating new clusters. As described, when a stimulus is encountered, a similarity score is computed for all existing clusters. Mathematically, this is indistinguishable from a prototype (i.e. gist-like, cortical-based) calculation. If a stimulus is encountered that is dissimilar enough from all previous clusters, the model is capable of instantiating a new cluster based on a single trial. This is in essence single-trial episodic encoding, a process that is thought to occur in the hippocampus. Thus, SUSTAIN is capable of one process that is qualitatively similar to gist-like similarity judgments (i.e. cortical) and another process of rapid single-trial encoding (i.e. hippocampal).

Interim Summary

In the two models just discussed, we see a need for more than one form of knowledge representation in order to provide an accurate account of human performance. The necessity to incorporate two processes, as well as the neuroanatomical evidence for distinct memory systems, leads to an important theoretical question. Why would the brain need two distinct and potentially redundant systems that seem to violate principles of parsimony? First described by McClelland, McNaughton & O’Reilly (1995) the Complementary Learning Systems (CLS) model provides a compelling answer.

Complementary Learning Systems

A long-standing fundamental limitation of neural network models of learning was that they exhibited *catastrophic interference* where newly acquired knowledge tended to completely overwrite earlier learning (McCloskey & Cohen, 1989). O'Reilly et al. showed that this phenomenon was an inevitable outcome of systems that employ highly overlapping knowledge representations that allow, for example, for generalization and inference. In the human brain, the cortex is thought to store such generalized (gist-ike) representations of knowledge and is known to learn gradually across days, months, and years as evidenced by anterograde amnesia observed in certain memory disorders involving the medial temporal lobe (MTL) structures (Scoville & Milner, 1957; Penfield & Milner; 1958; Zola-Morgan, Squire & Amara, 1986; Victor & Agamanolis, 1990; Rempel-Clower et al., 1996; Gold & Squire, 2006; Pascual et al., 2013). Their insight and solution to this problem of catastrophic interference was the existence of an additional, structurally distinct, system with complementary learning properties. This system uses sparse, non-overlapping representations that are robust to interference from later learning and is well represented in the medial-temporal lobe, specifically the hippocampus. Thus, in order to incorporate new knowledge in the cortex while avoiding the problem of overwriting existing knowledge, newly acquired information is rapidly encoded in the hippocampus. The hippocampus then slowly trains the cortex, allowing it to extract a more generalized representation of the information in a process of consolidation.

Hybrid Systems

Embracing the need for multiple knowledge representations has led to a variety of hybrid and multi-system models that incorporate several strategies in service of learning and behavior. In this section we describe two hybrid systems that utilize several explicit strategies to represent human behavior: RULEX and ALCOVE & ATRIUM.

RULEX

The RULEX model (Nosofsky, Palmeri & McKinley, 1994) of category learning is a stochastic process in which the model attempts to discover increasingly complex rules governing category membership based on the stimulus dimensions and their combinations. It predicts that people begin with a simple one-dimensional rule and use it until it fails. They then either attempt other simple rules or adopt a more complex one. If no single-dimension or combination rule can account for all the data, exceptions begin to be stored such as to keep the rule as simple as possible while maintaining high accuracy. Exceptions are, therefore, only stored as a last resort after attempts to abstract an appropriate rule has failed. Generally, RULEX has been applied to category learning with two mutually exclusive categories.

ALCOVE & ATRIUM

ALCOVE is a three-layered connectionist model of category learning. In it, the input layer corresponds to stimulus features, the middle, *hidden*, layer corresponds to exemplars, and the output layer corresponds to category responses. This model is explicitly based on the exemplar-based “generalized context category learning model” (GCM: Nosofsky, 1986; 1987). It predicts that when a new stimulus is encountered, people compute an attentional-weighted similarity score between it and every previously stored exemplar. People then assign the stimulus to a particular category based on similarity.

ATRIUM is a hybrid connectionist model that predicts that people use rule- and exemplar-based representations during learning and was first described by Kruschke & Erickson, 1994. The rule module contains both category and rule nodes. Each category node is connected to a number of rule nodes equal to the number of primary stimulus dimensions by learned, weighted connections. Thus, each category node learns to associate optimal rules based on feedback. The exemplar module is a full implementation of the ALCOVE model (Kruschke, 1992). A competitive mixture of experts gating mechanism combines the activity of both modules to compute the probability that the exemplar module is used for the current trial. Essentially, people select between a similarity- and rule-based strategy that maximizes performance. Finally, feedback is backpropagated through the model to modify weights in order to reduce future errors (note that overconfident responses are not considered errors; i.e. humble teachers) based on gradient descent.

Interim Summary

While the two models above employ a hybrid of rules and exemplars, they do so with the implication that both strategies are explicit. That is, participants are fully aware of the strategy they are using and can report it when asked. An alternative approach to accounting for human behavior is to account for both explicit and implicit processes. The idea of modeling category learning with separate processes for different memory types was first proposed (presciently) by Ashby et al. in the Competition between Verbal and Implicit Systems theory (COVIS; 1998).

Multiple Memory Systems

COVIS

The COVIS model predicts that people use both explicit and implicit processes by allowing for separate memory representations that emerge from explicit (verbalizable) processes such as rules, and implicit (non-declarative stimulus-response) learning. Further, it predicts that these systems operate competitively across the learning process and that people are biased towards explicit rules via a gating mechanism. Formally, COVIS focuses on modeling neurobiological circuits in the basal ganglia and prefrontal cortex in the form of a spiking model that attempts to simulate individual and group level neuronal activity in order to produce behavior based on known neuroanatomical connections. COVIS has been shown to account for a wide range of category learning phenomena (Ashby & Valentin, 2017), including data that are challenging for single-system models (Ashby & Ell, 2002).

PINNACLE

A second framework that takes a similar approach is the Parallel Interactive Neural Networks Active in Competitive Learning architecture (PINNALCE; Nomura, Maddox & Reber, 2007; Nomura & Reber, 2012; PINNACLE 2.0; Reuveni & Reber, *in prep*). PINNACLE predicts that people rely on separate information processing streams for explicit hypothesis testing and implicit learning with a decision-making mechanism that arbitrates between them. On each trial, explicit and implicit processes produce an independent prediction as to stimulus category membership, and the most confident of the two is selected to drive behavior. Feedback is then used to update each process' internal category representation. This model has been successfully used to guide analysis of functional neuroimaging data. In its original form, PINNACLE 1.0 incorporated an explicit module that tested a one-dimensional rule and an implicit module that tested a more complex two-dimensional rule. The model predicted that people spontaneously transition from relying on a simple explicit strategy, to a more complex implicit one over the course of the experiment.

Interim Summary

Setting theoretical assumptions of underlying neurophysiology aside for a moment, we note that all the models described so far share a commonality of employing more than one processes in order to account for behavior. In RULEX, there is a process for evaluating exemplar similarity (i.e. *RULE*), and a process for accounting for violations of those similarity-based judgments (i.e. *X*). In the ATRIUM model, there are similarly two processes (rule and exemplar modules). Both COVIS and PINNACLE employ two distinct, independent processes that map onto explicit and implicit strategies for solving category learning tasks. The success of these models in accounting for observed behavior across a wide range of paradigms suggests that a hybrid approach is well suited to fully capture categorization behavior.

Studying Dynamic Decision Making

Evidence both from converging two-process computational approaches as well as decades of behavioral and neuroanatomical work strongly imply that the brain uses two qualitatively different processes in categorization. Thus, we now turn our attention to the difficult problem of how these systems generate and evaluate different strategies: the decision-making process. Several questions arise in a framework where multiple strategies are available. 1) How do people navigate the decision-making process given a potentially large set of possible solutions? 2) How is feedback used to update and select different strategies? 3) What is the nature of the strategy representation (i.e. parallel or sequential, modifying existing rules or generating new ones, similarity judgements, one-step or hierarchical strategies)?

Before investigating these questions, a method for prolonging the period of strategy uncertainty during the category learning task is needed to provide more opportunities to observe the decision-making process. Traditional category learning experiments require participants to learn categories using random or pseudo-random stimulus selection in order to expose participants to the entire category throughout the experiment. To our knowledge, no paradigm currently exists that is aimed at intentionally prolonging the period of strategy exploration during category learning. The next chapter details the development of this new paradigm along with a detailed description of PINNACLE 2.0 that was developed to account for how people navigated this process and how it was used to provide trial-by-trial predictions of the observed strategy exploration. Chapter 4 describes two experiments that used PINNACLE 2.0's predictions to drive adaptive stimulus selection in the new paradigm. Results of those experiments necessitated further development of the PINNACLE model. Chapter 5 describes neuroimaging data gathered while participants performed the task and describes preliminary efforts to identify neural correlates of the decision-making process. In addition, further refinement to the PINNACLE model is described model based on these findings. Finally, Chapter 6 provides a discussion of the results.

Chapter 3: Prior Behavioral Work & PINNACLE 2.0

How do people converge on successful strategies in a new environment when a range of potential solutions exist? How do they adapt when environmental demands change, and those strategies no longer work? To understand these complex processes, we presented participants with increasingly difficult category learning tasks in which a both one- and two-dimensional explicit strategies as well as an implicit strategy are appropriate. We began with traditional diagonal category learning paradigms in which participants spontaneously explore a variety of explicit as well as implicit strategies and culminate with a new paradigm and dynamic protocol that experimentally induces strategy switching.

If it were as simple a matter as introspection or asking a person how they accomplished these tasks psychology would be a much easier field. A major difficulty in studying the decision-making process is the fact that on any given trial it is impossible to recover which strategy was used or what new rule was created by participants. In addition, developing a theory of how people navigate the dynamic decision-making process requires formal descriptions of the hypothesized underlying processes. To address both points we developed a computational cognitive model, PINNACLE 2.0. This model is capable of accounting for a wide range of behaviors that produces group- and individual-level data and provides trial-by-trial predictions as to the mental state of participants. The updated version represents a significant improvement over its predecessor, PINNACLE 1.0, which was unable to account for the range of these data. PINNACLE 2.0 describes a theory for how participants transition from one-dimensional to two-dimensional rules, both explicit and implicit. It does so by incorporating multiple strategies, a simple yet effective decision-making mechanism, and neurally plausible feedback processing.

The more successful the model is at accounting for participant behavior, the more confidence we have that the model approximates the processes used by people. Importantly, a successful model generates specific testable hypotheses about the participant's behavior and mental state.

In experiment 1, we present an examination of task parameter variations on a classic visual category learning task as part of instantiating reliable learning of both rule-based (RB, explicit) and information-integration (II, sometimes implicit) types of categories. The goal was to re-instantiate the task and manipulate the initial difficulty and learning rate to observe both explicit-explicit and explicit-implicit strategy switching over a larger number of trials than prior paradigms had allowed. In addition, we sought to characterize the way in which participants approached solving the task using detailed verbal self-report of strategies by participants. During this process we discovered a higher than anticipated rate of successful explicit rule use during an II task, which has been traditionally thought of as an implicit learning task. Participants learned the diagonal categorization boundary used as a signature of implicit learning but verbally described a two-step explicit rule. This phenomenon has not been previously reported, but few prior studies have carefully examined the type of learning with a structured post-session interview. This result implies that some prior reports of II category learning that depend solely on analysis of choice behavior may actually reflect an explicit-II approach rather than implicit knowledge expression.

In experiment 2, we introduced a new variation on the standard behavioral paradigm, *Falling Cat*, that used movement of the stimuli to create a greater sense of task engagement and urgency. This served to further obscure more successful strategies thus prolonging the period of strategy exploration. This paradigm then served as the basis for experiment 3, which introduces the *Dynamic Cat* protocol that experimentally induces strategy switching. In this approach, participants are guided towards using an incomplete one-dimensional rule-based strategy at the beginning of training by selectively showing stimuli consistent with that rule. Over the next several hundred trials, the true structure of the task is gradually revealed and requires participants to explore a variety of strategies as they search for a better alternative.

In the second part of the chapter we describe a computational cognitive model, Parallel Interactive Neural Networks Active in Competitive Learning (PINNACLE) 2.0, capable of accounting for a wider range of data than its previous 1.0 version including the Dynamic Cat protocol. We describe a general architecture for the model, the changes made to the previous version and how they aid in accounting for this broader set of data. PINNACLE 2.0 provides theory-driven hypotheses about the cognitive mental state of participants derived from their behavioral choice data. PINNACLE is a predictive information-processing model that can simulate group performance data that matches human performance from the same task conditions (stimuli and trial order). In addition, PINNACLE 2.0 can be yoked to an individual participant's record of specific choice behavior to make predictions about their internal state on each trial. The ability to estimate strategy use in real time during learning was used to guide stimulus selection towards increased strategy switching in experiments 5 and 6, described in the following chapter.

Experiment 1 – Static Cat

The goal of experiment 1 was to modify the classic RB/II visual category learning task to create more opportunities for participants to switch between one-dimensional strategies as well as between one- and two-dimensional ones over approximately one hour. Previous work by Nomura & Reber (2012) used a diagonal category structure and identified a relatively low rate of spontaneous strategy switching between explicit and implicit strategies. In order to encourage more instances of strategy switching we manipulated both trial length and category difficulty such that participants would not achieve ceiling performance, while allowing more time to abandon simple explicit rules and adopt more complex ones or shift to implicit decision-making. For this experiment, our criteria for success were: 1) an exclusion rate less than 30%, 2) between 70% and 80% accuracy on the last block (50 trials) after 300 trials of learning. In addition, we sought to quantify the number of participants able to verbally report an appropriate explicit strategy indicating implicit learning. This metric would guide development of our model by informing likely strategies participants used while solving the task.

Here we report the experiment parameter variations of a diagonal category structure that were tested to identify a protocol that reliably produced learning and complex strategy use within a one-hour protocol. The combination of relatively low rates of spontaneous strategy switching in participants within the first hour of learning as evidenced by low accuracy performance, and our strict criteria lead us to test eleven conditions. Across conditions, we varied several factors: 1) experimental timings including the duration of stimulus presentation, response time, and feedback presentation 2) category offset from the category boundary (i.e. how distinct the two categories are from each other), 3) within-category variability (i.e. how large the standard deviation of each category distribution is), 4) number of trials (600 vs. 300), 5) feedback-delay & dual-task.

Participants

One-hundred and eight participants were recruited from the Northwestern University research participation pool. All participants had normal or corrected to normal vision, were required to be over 18, and were provided informed consent as well as a post-experiment debriefing session as compliant with Northwestern University's Institutional Review Board. In category learning paradigms there is a concern that participants who have previously taken part in similar category learning experiments may not reflect the same sort of learning that a naïve participant might due to their previous experience with the stimuli and/or stimulus space. To mitigate this concern, participants were only allowed to participate once regardless of which experiment they were assigned to.

Procedure

All data were gathered on laboratory windows-based computers with 23" flat screens at 60Hz. Responses were made on a keyboard using the 'd' and 'k' keys. The experiment was written in python 2.7 using PsychoPy libraries (Peirce, 2007). Whenever possible, testing rooms were restricted to a single participant per session with no more than 2 performing the experiment at a time. Instructions were provided both verbally, and in writing presented on the screen, were self-paced, and read as follows: *"In this experiment, you will be shown a series of images. These images belong to either category A or category B. Categorize each image by pressing "d" for A, or "k" for B. Please note that you have* X *second to make your decision. Press any key to begin."* Where X was either 1, or 5 seconds depending on the condition.

Learning Rates & Accuracy

Across the three experiments reported in this chapter, we assessed learning by dividing trials into blocks of 50 and computing the average accuracy for each block. Participants who failed to achieve 60% accuracy on the final block of learning were excluded from further analysis. Exclusion rates are reported per experiment and condition.

Verbal report & Explicit II

On all but two of the experimental conditions in experiment 1, we employed a structured post-experimental interview to identify participants that used an explicit, verbalizable strategy during the task. This additional metric was added after two rounds of initial pilot testing, when we noticed that a proportion of participants would spontaneously report an accurate strategy they used during the post-experiment debriefing session. While the conventional wisdom of II category learning was that participants were using an implicit strategy, a proportion of participants demonstrated successfully using an explicit strategy (*explicit II*) that led to robust learning (e.g. “if it is thick it’s an *A*, unless it is also very tilty, then it’s a *B*” and vice versa.”)

The post-experimental interview was divided into two parts: 1) a verbal interview, and 2) a drawing component. In the verbal interview, participants were prompted to report any strategies they might have used during the experiment without explicitly asking them for strategy use (i.e. we do not use the words “strategy”, “rules”, “how did you solve”, etc.) so as not to create an expectation of a strategy if they did not have one in mind already. We began by asking the participant to describe their experience in the experiment while making notes as they talk. This usually resulted in participants talking about what they thought about the stimuli, and their experience in attempting to learn the categories. If participants mentioned any strategies they tried such as “the thickness of the line seemed to matter” we prompted them to talk more about that, and whether it seemed to work for them. This sort of prompting continued until the participant had nothing more to say. Afterwards, we asked the participant to imagine that a new participant was about to perform the experiment and that they had the opportunity to give them a head-start by giving them tips on how to successfully sort the stimuli. What advice would they give other than to stay attentive and alert? Finally, we asked participants whether there is anything else they would like to tell us about their experience.

In the drawing portion, we provided a blank X by Y chart space with notations on the X-axis for bar thickness (thick on left, thin to the right), and orientation on the Y-axis (vertical at 0, horizontal at the top). We explained how stimuli map to this space by selecting points on the chart and drawing the corresponding stimulus. For example, at (0,0) we drew a circle with thick vertical bars. At to top right of the chart we drew a circle with thin horizontal bars. We continued to provide examples until the participant understood the mapping. Once both the participant and the interviewer were satisfied, we provided a new, blank chart adjacent to the one just used and asked the participant to fill the chart by imagining what a circle would look like at multiple points, but instead of drawing the circle, to write down "A" if they thought that stimulus would belong to category "A" or write a "B" if it belonged to category "B". In this way they had the opportunity to express non-verbal knowledge of the space. Successful knowledge expression on this task produced groups of hand-written "A"s and "B"s that were roughly divided by a diagonal line.

Explicit or Implicit strategy use?

An important discovery that arose from the interview process was that not all participants who demonstrated successful learning did so by using an implicit strategy. Previous visual category learning studies have traditionally associated success on diagonal category learning tasks as indicative of implicit strategy use. Results from the interview process show that a percentage of participants were able to articulate an appropriate verbal rule that matched their successful choice behavior. We classify these participants has having used an *explicit II* strategy.

Alternatively, in cases where participants 1) learn the task to at least 60% accuracy 2) fail to articulate an explicit strategy that matches their choice behavior during the interview are classified as using an implicit strategy since they seem to lack awareness of how they achieved their performance. For example, participants we take to have used an implicit strategy may report that they guessed, or that their fingers knew what to do, or that there is in fact no way to solve the task, or they may report using a simple explicit rule (i.e. thick | thin) but their accuracy levels and choice behavior indicate that they were taking both dimensions into account.

Materials

Stimuli were circular sine wave gratings (Figure 3.1; Top) that varied in spatial frequency (thickness of lines) and orientation (tilt of lines). Stimulus space ranged from 0-1 in arbitrary units. This stimulus space can be thought of as a Cartesian X by Y coordinate space in which the X-axis corresponds to Spatial Frequency (low S.F = thick bars, high S.F. = thin bars) and the Y-axis corresponds to Orientation (low orientation = more vertical, high orientation = more horizontal). All stimuli were non-linearly transformed as described in (Treutwein, Rentschler & Caelli, 1989), which roughly equate the salience of each dimension (see Appendix A for details).

All conditions used a category structure in which stimuli were divided into two categories based on a diagonal category boundary (X = Y) in which the formal rule is X > Y = ‘B’; Y > X = ‘A’ (Figure 3.1; Bottom). This type of category structure requires that both dimensions be taken into account in order to achieve optimal performance. Each category (‘A’ and ‘B’) was constructed by generating stimuli according to a normal distribution centered along a diagonal spaced at a fixed distance from the category boundary. Thus, each category is defined by two numbers: 1) offset from the category boundary (i.e. half the distance between the categories), and

2) the standard deviation of the category (i.e. how spread out it is). Each participant saw one of 10 pregenerated stimulus orders that were generated pseudo-randomly such that the entire space was sampled every 10 stimuli, and no more than 3 stimuli from the same category could appear in succession to avoid creating perceived order effects.

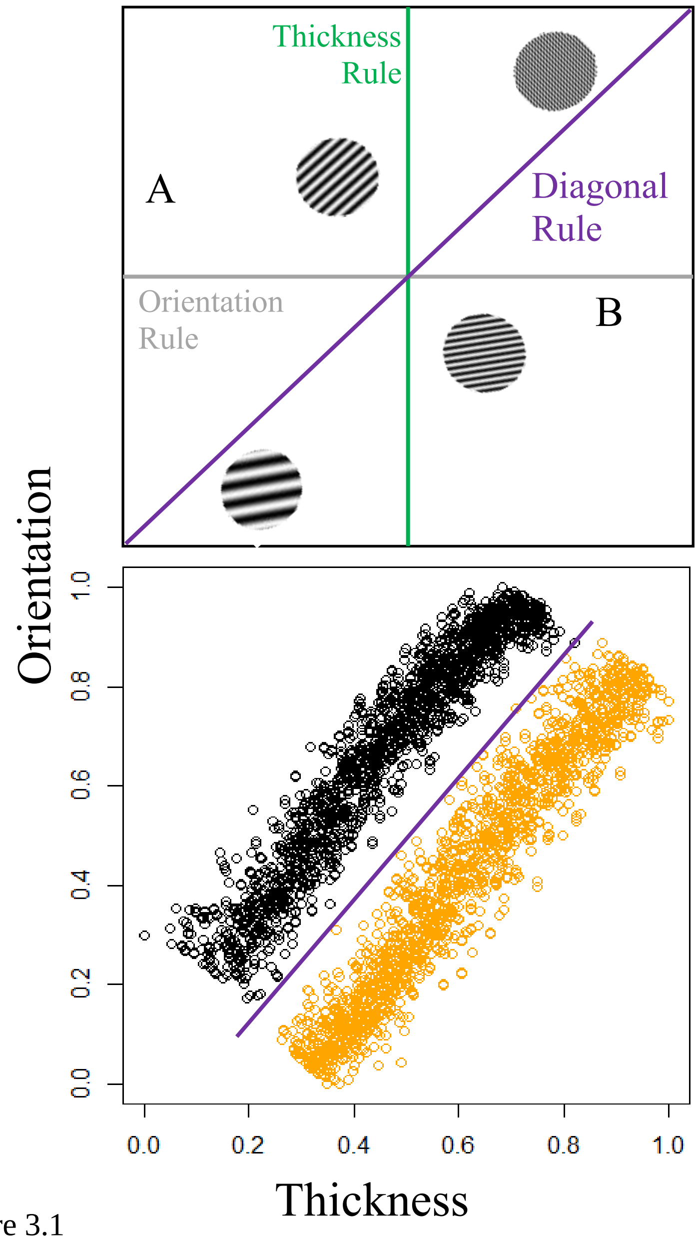


Figure 3.1

Top: examples of how identical stimuli can be divided into different categories. Grey: orientation rule based on Y-axis. Green: thickness rule based on X-axis. Purple: diagonal rule based on both dimensions.

Bottom: Example of a diagonal category distribution in perceptual space. Each point represents a set of coordinates that corresponds to a unique stimulus.

Procedure – Static Cat

Each trial began with a 250ms white fixation cross that subtended 0.382° visual angle in the center of the screen followed by a circular sinewave grating that subtended 6.1° visual angle for a variable length of time presented in the center of the screen. This was then followed by a random noise square mask subtending 6.1° visual angle that totally obscured the stimulus for 250ms. The participant was then prompted with “A or B?” for either 5,000ms, 2,000ms or 1,000ms based on the condition. If a response was not made in time, they were prompted with “Too slow, please make your selection faster next time.” Feedback was presented either immediately, or after a 2,750ms delay, for either 500ms or 1,000ms depending on the condition in the form of a circular cartoon thumbs-up on a green background for correct trials, or a similar thumbs-down on a red background for incorrect trials that subtended 7.63° visual angle presented in the center of the screen. In the final block, no feedback was provided to measure category knowledge without any additional learning. Finally, a blank screen was shown for 250ms as an inter-trial-interval (ITI) before the next trial began. Self-terminated breaks were offered every 100 trials in which participants were asked to keep them short, but no timer was enforced.

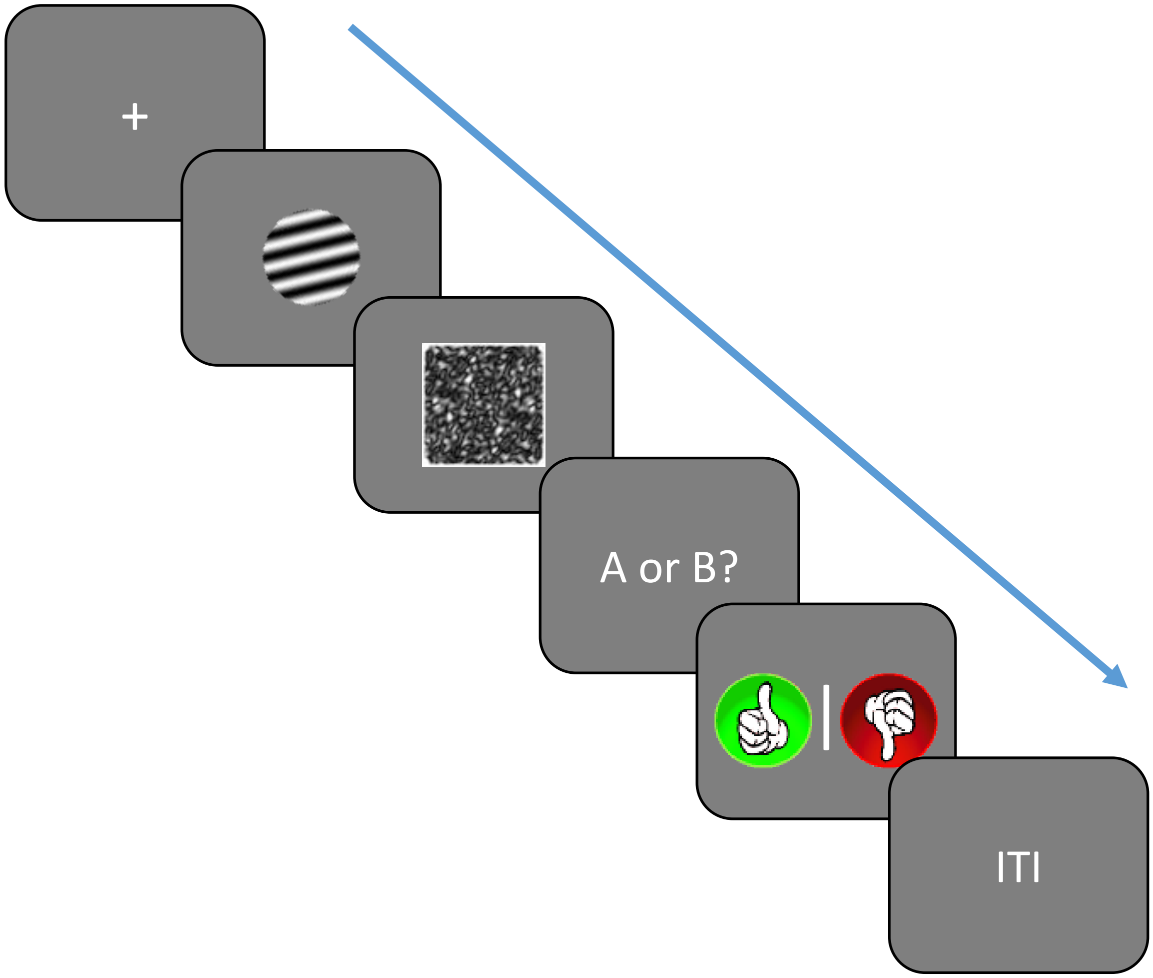


Figure 3.2

Standard visual category learning paradigm “Static Cat”. A stimulus is shown followed by a mask. A response is then gathered, and feedback is shown. ITIs can be jittered or not to accommodate neuroimaging techniques.

Conditions

The conditions tested in experiment 1 are summarized in table 3.1. In addition to varying the category distributions, we varied the number of trials (600, 400, or 300) because the ultimate goal was to perform the task in a functional magnetic resonance imaging (fMRI) machine which requires a jittered inter-trial-interval (ITI) in order to deconvolve the blood-oxygen dependent (BOLD) signal for analysis thus lengthening each trial. Given considerations for participant task engagement, as well as the cost of scanning, we aimed for an hour-long protocol which restricted the number of trials possible.

Finally, in several conditions, to further support the possibility that participants were performing the task implicitly we added a feedback-delay condition which has been shown to selectively impair implicit learning and not explicit learning (Maddox, Ashby, Bohil, 2003). Thus, if learning persisted under these conditions, the likelihood of explicit learning is higher which would make the condition less viable. In one condition, we tested a tone-counting dual-task which should impair explicit learning, but not implicit learning based on working-memory interference.

Results

Eleven conditions are reported in Table 3.1, participants who achieved less than 60% accuracy in the final block were excluded to ensure that our evaluation was reflective of the learning process. The offset: 10, sd: 4 (10 / 4) 300 trial condition produced an appropriate rate of learning (74%) with a low rate of exclusion (22%). It also produced a comparatively low proportion of participants who demonstrated explicit rule verbalizability (43%). While other conditions provided higher estimates of learning, lower exclusion, or lower rates of explicit II rule report, the 10/4 – 300 condition provided the best combination of the three.

	Participants		Category Structure			Timings		Results	
	n	Excld	Offset	SD	Trials	Stim dur.	FB dur.	Final Acc	% explict II rule
1	7	2	10	4	300	500ms	1,000ms	74%	43%
2	7	1	10	7	300	500ms	1,000ms	85%	57%
3	6	1	10	7	300ב	500ms	1,000ms	89%	83%
4	8	1	10	7	600	5,000ms	500ms	89%	N/A
5	6	6	8	7	600	500ms	500ms	83%	N/A
6	10	0	8	7	300	500ms	1,000ms	76%	50%
7	6	2	10	4	300ב	250ms	500ms	75%	50%
8	7	10	4	5	600	500ms	500ms	74%	57%
9	6	2	6	7	300ב	500ms	1,000ms	72%	17%
10	10	2	6	7	300	500ms	1,000ms	68%	10%
11	7	1	13	3	500ן	500ms	500ms	63%	0%
N =	108								

Table 3.1

Results of eleven conditions tested in Experiment 1. Columns are: n – # of subjects after exclusion. Excld – # of excluded participants. Offset – distance from the category boundary to the center of the category distribution. SD – the standard deviation of each category from the offset midline Trials - # of trials. Stim dur. – stimulus presentation duration. FB dur. – feedback presentation duration. Final Acc – performance accuracy on the last 50 trials. % explicit II rule - % of participants that were able to articulate an appropriate two-dimensional rule based on 4 independent raters with an inter-rater reliability of 75% or higher. ב – Delay condition. ן – dual-task condition.

Summary

In experiment 1 we identified experimental timings as well as a category structure that produced robust learning (~70-80%) within a 1-hour, 300 trial protocol. The higher than anticipated rate of explicit II strategy use indicating that participants transitioned from relying on an explicit- to an implicit-strategy as well as discovering an appropriate two-dimensional explicit rule. A shortcoming of the classic paradigm is its static nature, which is both boring and affords participants more time to think about their decisions which may have contributed to the higher than anticipated rates of explicit II strategy use. To improve upon this design, we developed a new paradigm aimed at both engaging participants in the task as well as providing a greater sense of urgency associated with the decision. These changes also led to a relatively lower rate of explicit-II strategy use.

Experiment 2 – Falling Cat

Experiment 2 tested a new visual category learning paradigm, *Falling Categories*. This protocol incorporates motion, and a more gamified feel to both encourage greater participant engagement as well as adding a sense of urgency surrounding the decision-making process.

Participants

Sixty-two participants were recruited from the Northwestern University research participation pool. All participants had normal, or corrected to normal vision, were provided full informed consent in accordance with Northwestern University's IRB protocols, and received course credit for their participation.

Methods & Materials

Experiment 2 tested two stimulus speeds that corresponded to 1,000ms and 1,500ms. In addition, the target stimulus and feedback stimulus sizes were reduced to 3.82° visual angle. ITI times were removed in order to speed up the overall pacing of the experiment and to allow for more trials.

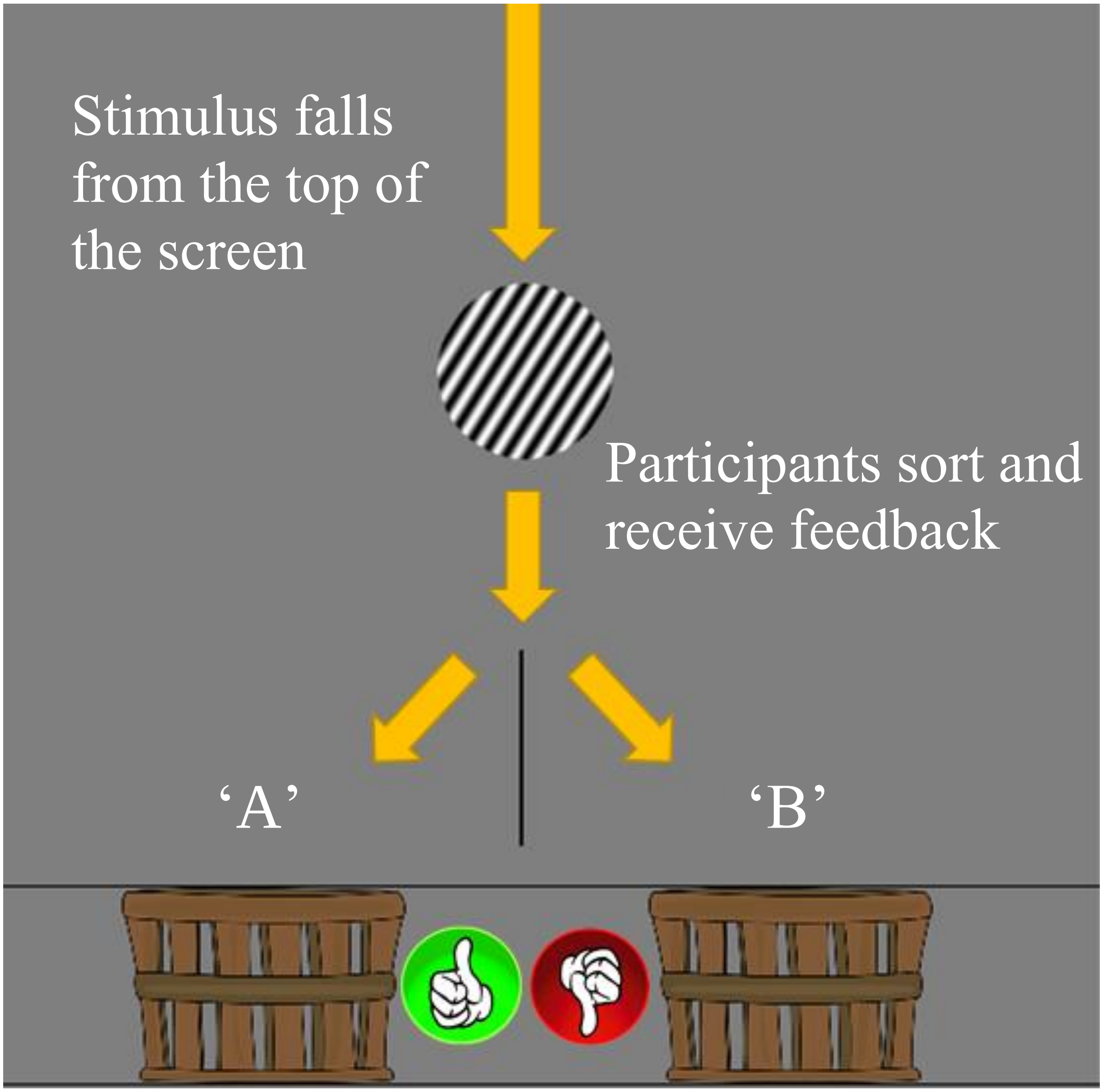


Figure 3.3

Depiction of a trial design in "Falling Cat". Stimuli drop from the top of the screen giving participants 1,500ms to respond by pressing one of two buttons. Feedback appears for 500ms following the decision before the next trial begins.

Procedure

In this new paradigm *Falling Cat*, on each trial a stimulus appears at the top of the screen and immediately begins to fall at a fixed rate. Participants have to respond before the stimulus reaches the lever and are instructed to sort each stimulus into one of two categories (‘A’ or ‘B’) by pressing one of two buttons on a keyboard (‘d’ or ‘k’) and thus manipulating a lever (black line, Figure 3.3) on the screen. Once the stimulus is sorted, participants are provided immediate visual feedback in the form of either a green thumbs up or red thumbs down at the bottom of the screen for 500ms. If a response is not made in time, participants are prompted with “Too slow – Please make your selection faster next time”. The time required to complete an individual trial is much quicker, allowing for category learning protocols with up to 1,000 trials within an hour-long session.

Conditions

Experiment 2 tested five conditions across two category structures: RB (1 condition) and II (4 conditions). The RB condition was included as a control condition to ensure the protocol could also produce learning of a very simple category. The II conditions included two stimulus motion speed conditions: 1,000ms (2 conditions of a 13/3 category structure from experiment 1) and 1,500ms (3 conditions, one of a 13/3 category structure, and 1 10/4 category structures). We included the 13/3 distribution in this experiment for the possibility of changing the profile of strategy use from mostly explicit in experiment 1, to more implicit. The 10/4 condition was shown to be the best overall condition based on the same criteria as in experiment 1.

Results

RB condition: n=10 (1 excluded), 600 trials, 1,500ms stimulus presentation time, group accuracy on last block = 93%, 100% verbal strategy reported.

II conditions: **1)** n=9 (3 excluded), 13/3 distribution, 1,000 trials, 1,000ms stimulus presentation time, group accuracy on last block = 75%, 0% explicit II strategy reported. **2)** n=5 (1 excluded), 13/3 distribution, 600 trials, 1,000ms stimulus presentation time, group accuracy on last block = 68%, no verbal report was collected. **3)** n=12 (0 excluded), 13/3 distribution, 1,000 trials, 1,500ms stimulus presentation, group accuracy on last block = 88%, 66% explicit II strategy reported. **4)** n=20 (4 excluded), 10/4 distribution, 1,000 trials, 1,500ms, stimulus presentation, group accuracy on last block = 75%, 18% explicit II strategy reported.

Summary

In experiment 2, we tested a new experimental paradigm that incorporated motion to create an increased sense of urgency, coupled with a more engaging cover-story and gamified trial design. This paradigm led to equivalent levels of participant accuracy while reducing the rate of explicit II report, thus allowing for a better model of both explicit-explicit and explicit-implicit strategy switching. Even with high accuracy performance, and low rates of explicit verbal report, there was still a concern that the number of instances of strategy switching per participant would be too low to allow for reliable statistical analyses. We therefore set out to develop a protocol that would experimentally manipulate strategy use to increase the number of such occurrences.

Experiment 3 – Dynamic Cat

Experiment 3 built upon experiments 1 & 2 by modifying the stimulus selection and presentation order to experimentally induce a strategy switch from a simple to more complex strategies. This novel experimental protocol *Dynamic Categories* can produce a shift from an initially successful unidimensional explicit strategy to a more complex two-dimensional strategy or in some cases an implicit one. Being able to experimentally manipulate strategy switching in this way is critical to studying the dynamic decision-making as it provides a robust testbed for neuroimaging, as well as future behavioral tests and manipulations. Here we report a pilot condition with 1,000 trials, and a follow-up with 545 trials that replicate these results.

Participants

Forty-one participants were recruited from the Northwestern University research participation pool. All participants had normal, or corrected to normal vision, were provided full informed consent in accordance with Northwestern University's IRB protocols, and received course credit for their participation.

Materials & methods

All methods and materials in experiment 3 were identical to those in experiment 2 except for the following: 1) number of trials were either 1,000 (pilot) or 545 (replication) in order to accommodate a jittered ITI in anticipation of neuroimaging. 2) A jittered ITI of between 2-8 seconds was added. 3) Most critically, the stimulus sampling method was altered to experimentally induce a shift from an explicit to an implicit strategy. In classic experimental designs (including experiments 1 & 2 reported here), stimuli are sampled pseudo-randomly to ensure the entire category space is sampled evenly across the entire experiment. In Dynamic Cat we control the sampling of stimuli in three phases to encourage strategy switching:

Phase 1: Bait

For the first 100 trials of the *Dynamic Cat* protocol, participants were exclusively shown stimuli sampled from areas of each category that conform to a simple "thick = *A*, thin = *B*" rule (black and orange areas of Figure 3.4). This quickly establishes a simple explicit strategy in all participants as well as high accuracy.

Phase 2: Switch

During the next 100 trials stimuli were sampled such that they disconfirmed the simple rule learned in phase 1 (green and purple areas in Figure 3.4). We used a ratio of 3:1 disconfirming to confirming stimuli such that participants saw far more trials in which their initially successful rule did not work. This sudden change in the stimuli participants experience leads to a reliable drop in accuracy as well as a period of strategy exploration in which a variety of strategies are explored in an attempt to recover from the change.

Phase 3: Learning

From trial 200 till the end of the experiment, the ratio of disconfirming to confirming stimuli was reduced to 2:1 in order to more evenly sample the space while still encouraging participants to abandon their initial explicit rule as participants tend to perseverate on an explicit rule even when the preponderance of evidence suggests it no longer works.

As in experiments 1 & 2, stimuli were sampled such that a maximum of 3 consecutive stimuli from the same category could be shown in a row. In addition, stimuli were sampled at a minimum distance from the previously seen stimulus to ensure uniform sampling of both categories.

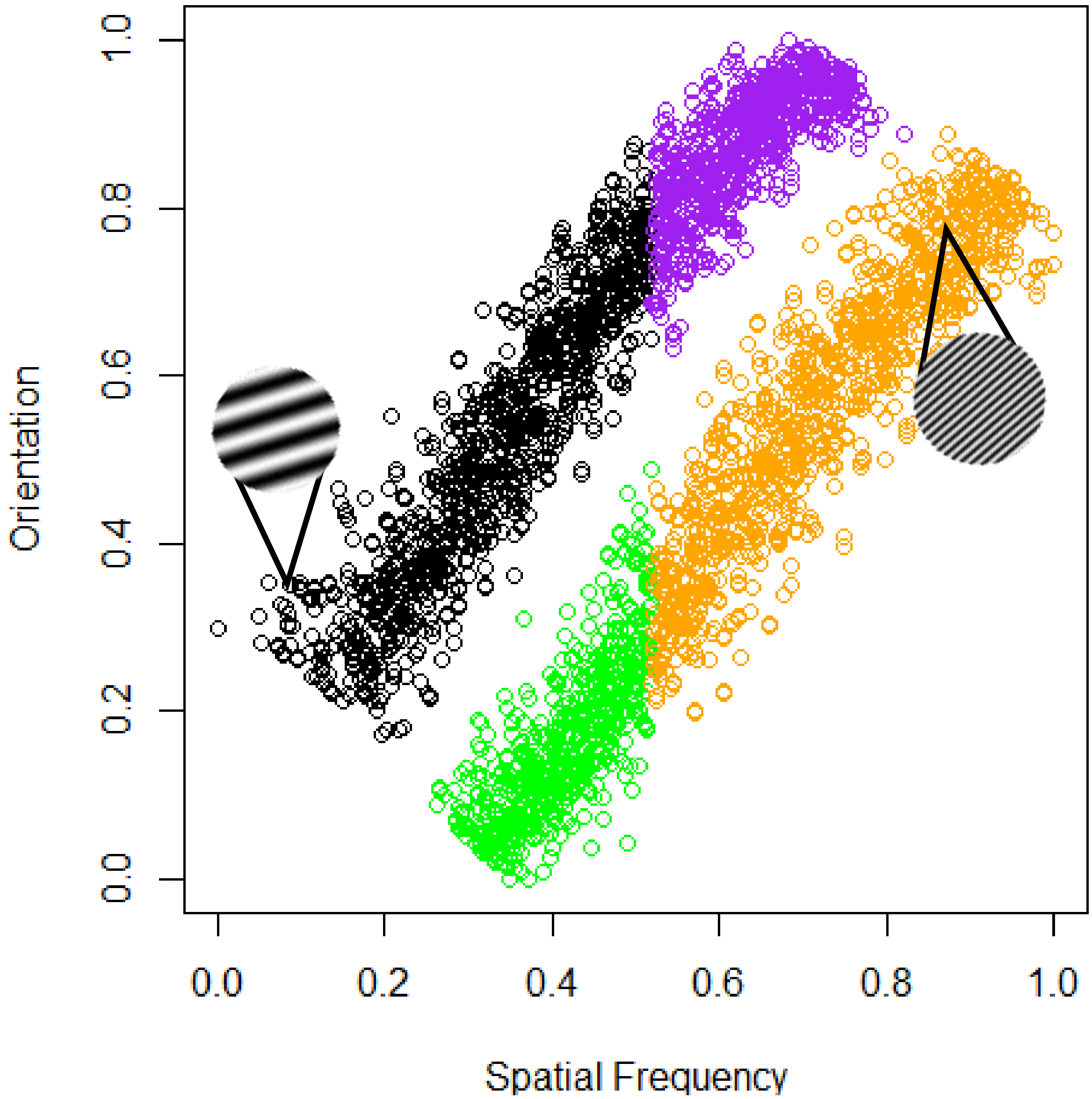


Figure 3.4

An example diagonal distribution identical to figure 3.1 (bottom) with areas of the category space colored based on whether they confirm (black and orange) or disconfirm (purple and green) a simple explicit rule based on thickness. Sine-wave gratings represent example stimuli from the indicated areas.

Procedure

All procedures were identical to those in experiment 2 except for the stimulus presentation time (1,500ms or 1,000ms) and a either no ITI or a jittered ITI ranging from 2-8s in order to accommodate the proposed fMRI neuroimaging experiment.

Results

Pilot: n=11 (2 excluded), 1,000 trials, 10/4 distribution, 1,500ms stimulus presentation time, group accuracy on last block = 80%, 22% explicit II strategy reported.

Replication: Accuracy results are shown in Figure 3.5. n=30 (9 excluded), 545 trials, 10/4 distribution, 1,000ms stimulus presentation time, group accuracy on last block = 80%, 38% explicit II strategy reported.

In Figure 3.5 we see very accurate performance as participants identify the simple explicit rule used to bait them over the first 100 trials. Entering the Switch phase, we see a sharp drop in accuracy as the rule is heavily disconfirmed. In the Learning phase, when the ratio of disconfirming-to-confirming stimuli is eased to 2:1 to allow a more even sampling of the category space, we see a gradual increase in accuracy as a proportion of participants settle into a more complex, and sometimes implicit, strategy.

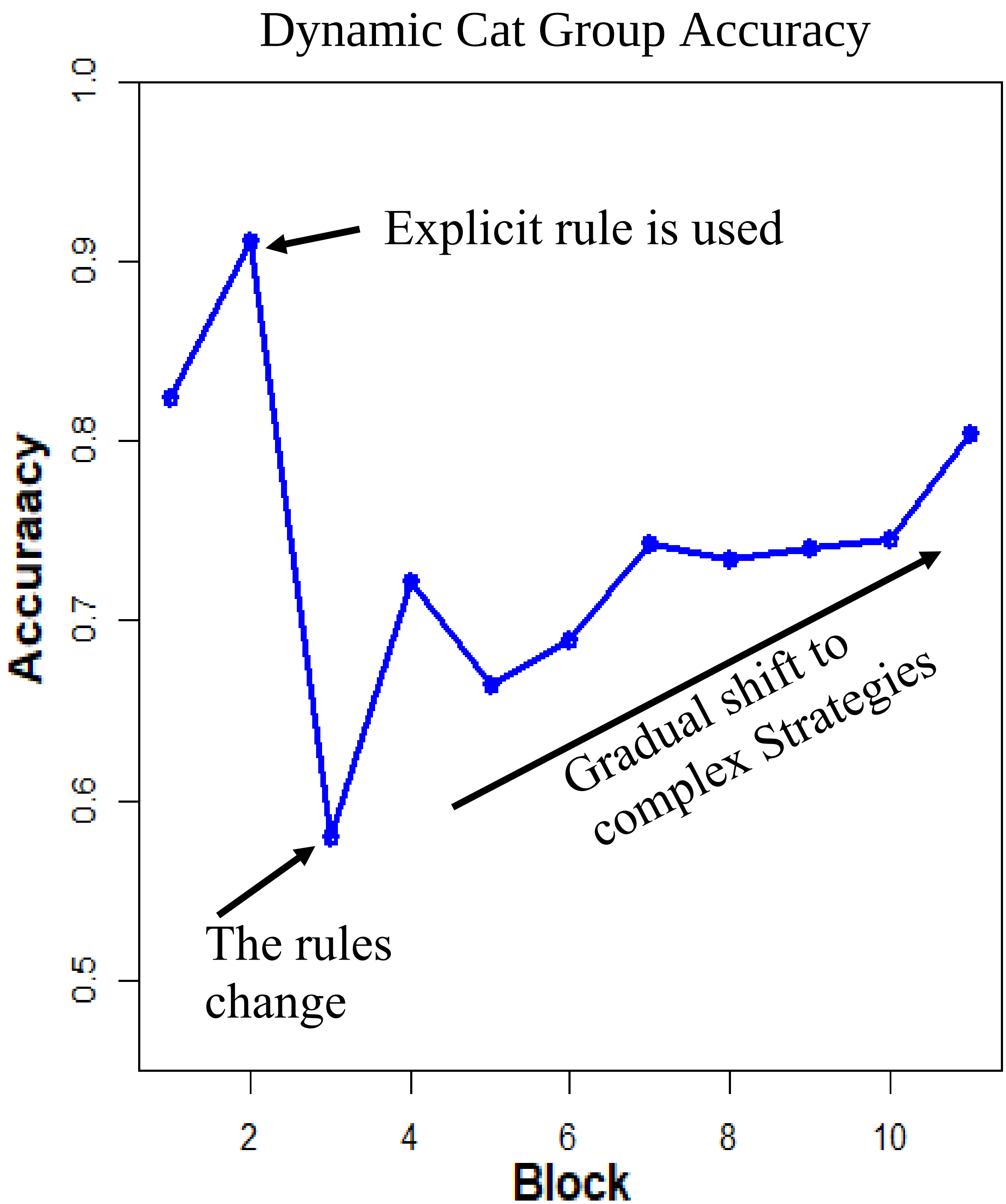


Figure 3.5

Group accuracy for experiment 3. Includes data from the 1,000 and 545 trial condition: n = 30. In the Bait phase (blocks 1-2) participants all perform at a high level of accuracy by using a simple thick | thin rule. In the Switch phase (blocks 3-4) participants all experience a reduction in accuracy. The Learning phase (blocks 5-end) show a gradual increase in performance as more complex strategies are used.

Summary

In experiment 3, we introduced a new behavioral paradigm, Dynamic Cat, in which participants are initially invited to discover a simple one-dimensional explicit rule based on line thickness. This was done by selectively sampling stimuli from the appropriate regions of the diagonal category. In the second phase, this rule was then shown to be inaccurate by sampling stimuli from different regions of the category space that both disconfirm the rule and expose the true underlying category that requires the use of both line thickness, and orientation. Participants then gradually switched strategies from simple, to more complex rules over time with few participants able to explicitly report the strategy they used. Across both experiments participants learned the categories while exhibiting a learning profile indicative of strategy switching.

PINNACLE 2.0

The data obtained using the new Dynamic Cat protocol exposed several limitations within the original PINNACLE 1.0 framework that proved insufficient to capture human learning. The greater need to change from one- to two-dimensional strategies highlighted weaknesses in the mechanisms used for system-specific learning. In addition, limitations in the available one-dimensional hypotheses to consider within the explicit system, and feedback processing made the model unviable. PINNACLE 2.0, described here, builds on the original PINNACLE 1.0 description and improves our ability to describe human learning behavior. Importantly, it does this by allowing for separate learning rates for explicit and implicit systems, an explicit system that can rapidly shift its criteria based on evidence, and reward-prediction-error (RPE) weighted feedback which allows for faster strategy shifting.

PINNACLE (Parallel Interactive Neural Networks Active in Competitive Learning; Figure 4.1) is a systems neuroscience inspired architecture for understanding interactions among memory systems and competing strategies during complex learning and decision making. The original PINNACLE 1.0 model was described in Nomura & Reber (2008, 2012) and reflected an initial effort towards an integrated model based on findings of separate neural systems for explicit and implicit category learning. While prior work had shown a clear dissociation between category learning systems that mirrored the dissociation between implicit and explicit memory systems, little attention had been paid to how these two different systems would interact in a cognitively healthy participant who would have simultaneous access to both kinds of processing. PINNACLE 1.0 added a hypothesized *Decision Module* to resolve competition between systems and showed a reasonable match to the available human behavioral data.

Architecture

The PINNACLE architecture incorporates separate and competitive information processing streams that reflect the separate neural bases for explicit rule-based learning and implicit II learning. Here explicit strategies are modeled as one-dimensional rules based on the fact that participants readily report simple strategy use. Prior visual category learning research has shown that as the complexity of the category increases, fewer participants are able to articulate rules that allow for high accuracy performance. Based on this, we model the implicit system as a two-dimensional (diagonal) rule. It is important to note that our data shows that participants are also able to approximate this rule explicitly and as such use of this more complex rule may be reflective of either explicit or implicit strategies. A competition resolution mechanism is necessary for selecting a single response even though each system is independently attempting to predict the category of the stimulus. We termed this conflict resolution area, the Decision Module, reflecting the hypothesis that this process is carried out in a separate region of the brain to adjudicate between the category membership estimates from each of the explicit and implicit systems. In PINNACLE 1.0, a simple mechanism for this process was conjectured and this both provided a fit to data and appeared to match up with neuroimaging data for trials associated with high competition.

Some process like the Decision Module is necessary for any multi-system or hybrid strategy model of categorization. In addition, these approaches pose a question about the process of learning from feedback, specifically assigning credit across systems after a correct or incorrect response. The simplest approach of only allowing the competition-winning model to both drive the response and learn from feedback was not used in PINNACLE for two reasons. First, as a model of the development of expertise, PINNACLE needs to be capable of a strategy shift from an explicit RB approach to an automatic, habitual II approach, which cannot happen if the II system does not learn from feedback even when RB is driving behavior. Second, comparative model fitting of different feedback approaches found that allowing both systems to learn in parallel simply fit human choice behavior better (Nomura & Reber, 2012). The apparently straightforward solution of having both systems learn independently raises a difficulty in terms of the neuroanatomical underpinnings of reward processing in the non-controlling (*off*) system. For example, if the implicit system was not selected but had the correct prediction and the controlling system was wrong, negative feedback is provided and no dopamine should be released. How then, does a basal-ganglia based, dopamine dependent implicit system learn in these cases? While the current study is not well situated to answering this question, the PINNACLE framework provides a good foundation for future research in this direction.

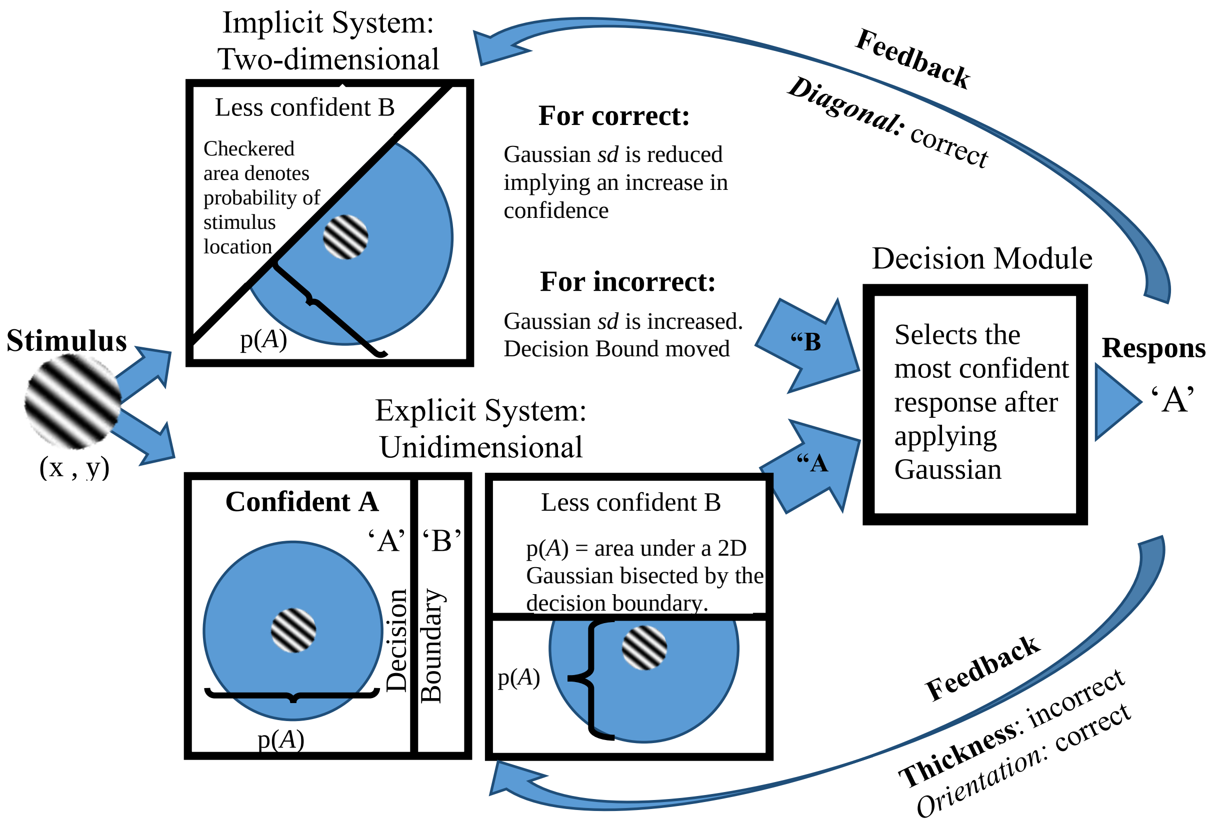


Figure 3.6

Cartoon of PINNACLE 2.0. Stimulus information flows to two memory systems. The lower stream represents the explicit system which evaluates simple one-dimensional category boundary hypotheses. The upper stream represents the implicit system which evaluates a two-dimensional category boundary hypothesis. Each system provides a prediction as to category membership. Information from both systems then flows to a decision module where a decision is made about stimulus category membership based on relative confidence. Feedback is provided, and each system updates its internal category representation based on reward prediction error weighted learning.

Several aspects of the architecture were updated in the development of PINNACLE 2.0 to provide a better fit to data from the new Dynamic Cat protocol.

1) The explicit RB system was enhanced to allow for one-dimensional learning on either category dimension (previously, the RB subsystem could only learn categories based on line thickness). This raised some additional complexity by making explicit learning itself function as a multi-system model. However, it is clear that humans can hypothesize, learn and use a wide array of explicit strategies, even for a simple visual categorization task, so this enhancement increases the connection of the model to human cognition.
2) The simple reinforcement learning mechanism in PINNACLE 1.0 did not fit the learning profile exhibited by participants over the Dynamic Cat protocol that forced greater strategy switching. Therefore, the mechanism for feedback learning that was enhanced to incorporate reward prediction error (RPE; recently Schultz, Stauffer & Lak, 2017; for computational models: Maia, 2009; Maia & Frank, 2011), using weighted feedback that better matched participant behavior.

Three critical mechanisms of PINNACLE 2.0 allow us to reproduce participant behavior on the Dynamic Cat task: 1) a decision module that arbitrates two distinct systems (explicit \ implicit) capable of guiding behavior based on a measure of confidence. 2) A credit assignment mechanism that dictates how each system learns in light of feedback, and 3) Reward prediction error (RPE) weighted learning with separate learning rates for explicit and implicit systems. Together, these mechanisms allow PINNACLE 2.0 to display rapid explicit learning when confronting simple category structures, while simultaneously developing a more complex representation of the category space using a slower, implicit system. When task demands change, PINNACLE 2.0 is able to dynamically switch strategies to maintain optimal performance.

Since PINNACLE 2.0 is an information processing model that aims to perform experiments like participants it is divided into several sub-systems through which information flows. Broadly speaking the information flows in the following order on a given trial: Perception → Cognition/Decision-Making → Response → Feedback & Learning. In the next several sections we detail each sub system's operation and how it contributes to PINNACLE 2.0's success.

Perception

PINNACLE 2.0 perceives a stimulus as a (X, Y) coordinate in an idealized perceptual space in which each dimension of the stimulus (i.e. thickness, orientation) corresponds to a different axis (see Chapter 3, Figure 3.1 for an example). Each stimulus dimension has random Gaussian noise applied to it to simulate a degree of uncertainty as to where in the space the stimulus is exactly located. This mechanism allows PINNACLE 2.0 to achieve a stable, non-ceiling plateau in cases of category overlap, or when stimuli from each category are very close to the category boundary. The *sd* of the Gaussian noise applied to each stimulus dimension is taken as a free parameter.

Memory Systems

Stimulus information is then forwarded simultaneously to an explicit, and implicit memory module. Both modules use a decision bound theory (DBT; Ashby and Townsend, 1986) framework to represent categories in this perceptual space. DBT formalizes category separation with a decision boundary that bisects the space into two categories (e.g. 'A' and 'B') that can be thought of as a specific hypothesis as to what defines each category. For example, a category might be defined by the thickness of the stimulus bars formalized as a linear vertical decision boundary with a specific value along the X-axis. In order to provide a probability-estimate of category membership, a perceptual shaping (PS) function is applied to each stimulus in the form of a bivariate Gaussian distribution centered on the stimulus coordinates. A single initial value for the standard deviation of this symmetrical 2D Gaussian is used for both systems and is taken as a free parameter and is fit to the data. Thus, category membership is calculated as the area under the Gaussian that falls on one side of the decision boundary referred to as p*(A)* (probability of 'A') according to formula 4.1.

$$p(A) = \iint_{-\infty\,-\infty}^{\infty\;DB} e^{-\frac{1}{2}\left(\frac{DB-(loc_x, loc_y)}{PS}\right)^2} dx \tag{3.1}$$

Where PS is the perceptual shaping parameter, and DB is the position of the decision boundary A *p(A)* > 0.5 constitutes evidence for category 'A' and vice versa for category 'B'.

Based on the available stimulus dimensions, as well as the fact that a majority of participants readily report these two explicit hypotheses, the explicit system tests two simple one-dimensional rule-based (RB) hypotheses: thickness, and orientation, formalized as a linear vertical (X-axis; *RB X*), and linear horizontal (Y-axis; *RB Y*) decision boundary respectively. For each stimulus, each hypothesis provides a *p(A)* and the most confident is selected as the output of the explicit system. It is worth noting that we acknowledge that a far more sophisticated explicit system, capable of generating and testing complex hypotheses could be implemented. Nevertheless, we were able to account for the observed data from experiments 1-3 with these simple models.

The implicit system tests a single linear diagonal hypothesis with a slope fixed at 1 and a variable intercept that integrates information across both dimensions (*II*). The decision to implement a simple implicit system was driven by the desire to maintain interpretability, as well as the fact that such a simple mechanism is sufficient to capture human behavior. The initial values of the three decision boundaries are set equal to the respective values of the first stimulus PINNACLE 2.0 encounters. As our data shows, some participants were able to articulate two-dimensional rules commensurate with a diagonal strategy. As such, the diagonal strategy is taken to be reflective of an overall more complex strategy which in some cases can be implicit and in others explicit.

Decision and Response

A competitive two-system model requires placing these memory modules within an architecture that allows a single response to be made. On each trial, both the explicit and implicit systems each provide an estimation of p*(A)* but only one response can be made by the simulation. To account for this PINNACLE 2.0 models two flows of information into a *Decision Module* where the selection of a response is made. A variety of options for how the sources of information could be evaluated were considered and a simple model of selection between systems weighted by confidence and the addition of random normally distributed noise was found to provide a good fit to human category learning. For this calculation, confidence is defined as distance from chance (0.5) such that a system predicting *p(A)* at 0.9 would be preferred over predictions such as 0.6 and 0.4 (also note that a prediction of 0.1 is preferred over 0.6 since it is further from chance). The *sd* of the Gaussian used to generate random noise (Decision Module Noise) is taken as a free parameter that is fit to behavior (and remains fixed across performance).

Of theoretical note, our implementation of a confidence based probabilistic selection mechanism embeds a strong hypothesis of fully independent and competitive systems. Alternative hypotheses that accomplish the same function are so-called *mixture of experts* mechanisms that sum evidence across the two systems, or a *hard-gating* mechanism that shunts control of one system while another is active (i.e. explicit overrides implicit while active). Nomura and Reber (2012) compared the current competitive implementation with a mixture-of-experts mechanism and found the latter provided a poorer fit during competitive model fitting suggesting that the explicit and implicit systems work at least largely independently.

Feedback Credit Assignment

The resulting architecture highlighted a new consideration in attempting to simulate human behavior related to incorporating feedback. Both the explicit and implicit systems independently attempt to learn the category structure on each trial (typically with the one that reflects the experimental structure being more successful) but the actual feedback provided is tied to the response selection of the decision module. This poses a puzzle that will arise when strategy switching is necessary. For example, in Dynamic Cat, initially PINNACLE will learn to rely on the explicit system's vertical line strategy. However, in the *switch* phase this strategy will cease to work, and in fact the implicit diagonal strategy may make more accurate predictions. This leads to a situation in which the non-controlling system is correct, but the feedback provided to the system is based on an incorrect prediction made by the controlling system. How should feedback be processed by the non-controlling system? Curiously, variations on mechanisms for handling feedback that allow the non-controlling system to learn based on its own predictions, independent of the controlling system, fit human behavior better (Nomura &

Reber, 2012). These results are challenging with respect to a basal ganglia based, dopamine dependent implicit system in terms of a plausible neural mechanism for this kind of feedback process. Specifically, strengthening confidence is hypothesized to depend on dopamine, but no dopamine would be released when the controlling system leads to an incorrect response. Questions about how feedback is handled in an integrated system are only revealed by instantiating and rigorously evaluating models such as PINNACLE and constitute an important motivation for the simulation modeling approach used here to build our theory of human category learning.

$$PS = PS * \begin{cases} (1 - (RPE \ * \ PS_{lr})), & if\ correct \\ \dfrac{PS}{1 - (RPE * PS_{lr})}, & if\ incorrect \end{cases} \tag{3.2}$$

where

$$RPE = \begin{cases} |1 - p(A)|, & for\ stimuli\ that\ are\ As \\ |p(A)|, & for\ stimuli\ that\ are\ Bs \end{cases}$$

A separate perceptual shaping learning rate (PS lr) is taken as a free parameter for the explicit and implicit systems respectively with the explicit parameter typically being much larger than the implicit PS lr parameter.

Reward Processing & Learning

In the context of the model, we define learning as the process of gaining confidence in a correct hypothesis as to what defines each category. As such, PINNACLE 2.0 modifies its internal representation of the category structure (i.e. modify the decision boundary), as well as its confidence in each hypothesis being tested (i.e. the *sd* of the PS parameter) based on its behavior, and the feedback provided. To do this, PINNACLE 2.0 uses a reward prediction error (RPE) weighted function to modify its internal representation of the categories being learned according to equation 4.2.

Incorporating RPE allows a strategy to gain or lose confidence proportional to its current confidence as task demands change. This is critical for modeling Dynamic Cat data as previously successful strategies must be abandoned in light of new evidence and increases competition between systems.

Correct feedback

When processing correct feedback for a given hypothesis, PINNACLE 2.0 does not modify the decision boundary since it was supported. The value of the Perceptual Shaping (PS) parameter for each correct hypothesis is reduced respectively to reflect an increase in confidence since more of the distribution will now fall to one side of the decision boundary.

Incorrect feedback

When processing incorrect feedback, the value of the PS parameter for each hypothesis is increased to reflect a loss in confidence since less of the distribution will fall on one side of the decision boundary.

The implicit system modifies its decision boundary incrementally by a fixed amount in the direction that would have produced the correct response. The amount that the decision boundary is modified represents the learning rate for that system. It is taken as a free parameter that is fit to the data and remains constant throughout the experiment. This value is typically very low thus facilitating slow and gradual learning in the implicit system.

The decision boundaries for explicit hypotheses, however, are modified to be the relevant value of the current stimulus (e.g. the vertical decision boundary intercept takes on the value of the X coordinate of the current stimulus). This approach of “anchoring” the rule to the current stimulus represents updating the hypothesis to the most extreme instance of the category seen to-date. The idea of effectively instantiating a new instance of the rule rather than incrementally modifying the previous one is a fundamentally different hypothesis as to how people explore solutions. The anchoring approach allows us to simulate the rapid explicit learning of these categories observed in participant data in which very few trials are required to achieve a near optimal decision boundary. Of note, versions of PINNACLE in which the explicit system modified its decision boundary incrementally by a fixed amount produced worse fits to participant data.

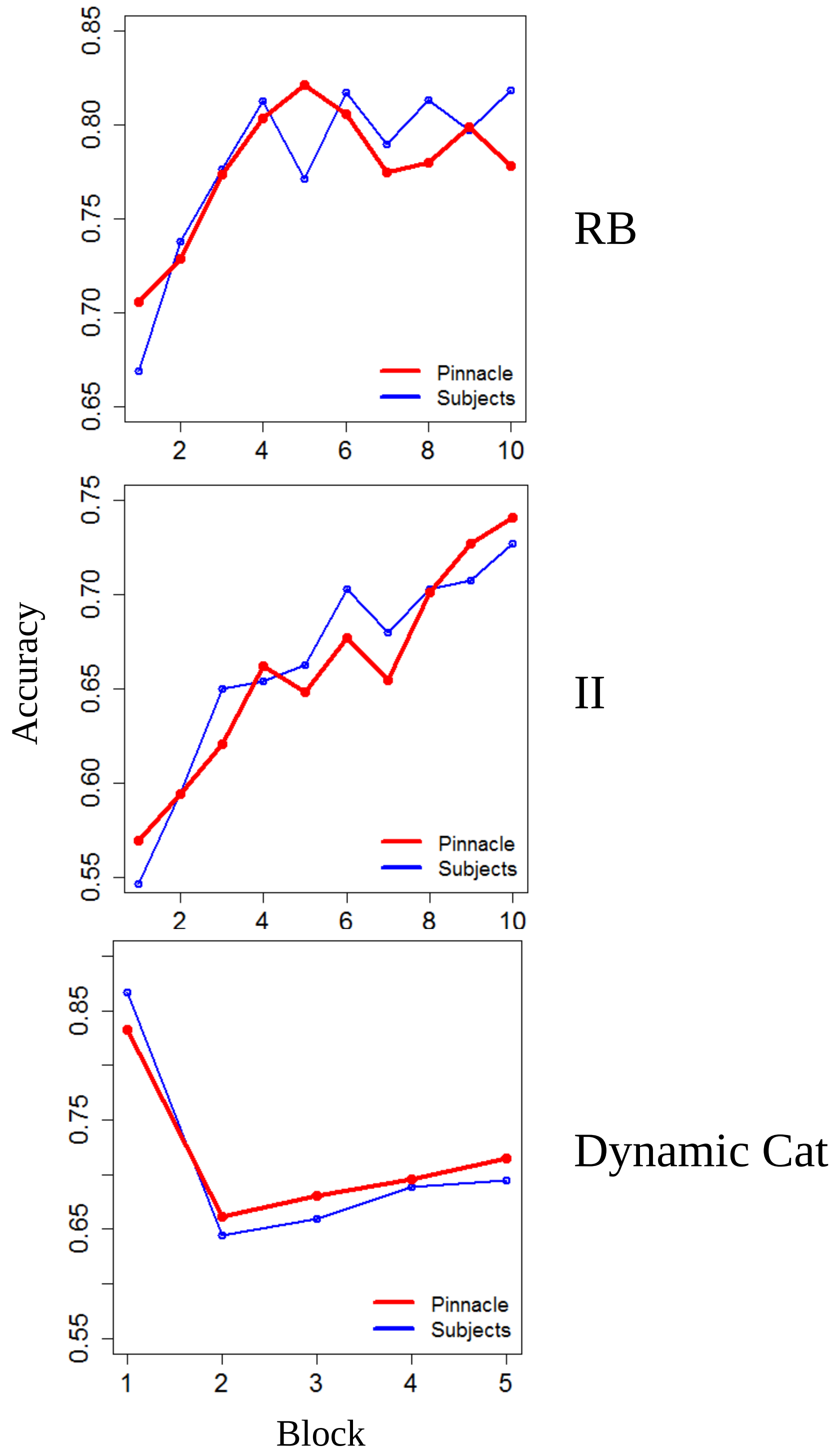


Figure 3.2

PINNACLE 2.0 fits to participant group data across three conditions reported in Chapter 3. Left: RB condition (n=10) from experiment 2. Middle: 10/4 Falling Cat II condition (n=20) from experiment 2. Right: Dynamic Cat condition (n=30) reported in experiment 3. All blocks are of 100 trials.

Model fits and Strategy Diagnosis

PINNACLE 2.0 is capable of reproducing group, and individual level data on simple one-dimensional (vertical) categories, two-dimensional (diagonal) categories, as well as on the Dynamic Cat task in experiment 3. Figure 3.2 - A shows group level model fits. All three fits were produced using the same set of free parameters. For each condition, PINNACLE 2.0 produced individual performance traces equal to the number of participants in the condition thus representing group level data. Behavior was then averaged across 100 trials and plotted.

In addition to group-level modeling, PINNACLE 2.0 is capable of producing choice behavior similar to that of participants, the model can be fit to individual participants by maximizing the likelihood that the model will produce the same choice behavior as that participant. This provides a trial-by-trial prediction as to which strategy was used by each participant – thickness, orientation, or diagonal.

Adaptive Tutoring

PINNCLE 2.0 can be fit to individual participant data on the order of several seconds which allows us to employ the model during the Dynamic Cat experiment and optimize its fit during the ITI periods. In this way, we obtain a real-time prediction of which strategy a participant is currently using (RB X, RB Y, or II). Using this information, we can adaptively select stimuli for the participant based on a number of criteria designed to facilitate a transition from unidimensional rules to two-dimensional explicit, and sometimes implicit strategies.

Beginning after the first block (80 trials; to have enough data to provide an initial fit), on each trial a prediction of which strategy was the most confident, as well as the predicted values of the decision boundary locations are provided by PINNACLE 2.0 based on the pattern of participant's choice behavior produced thus far. If the most confident strategy was a one-dimensional rule (RB X or RB Y), then a stimulus was selected that would be incorrectly classified by that strategy but classified correctly by the diagonal strategy. For example, if the most confident strategy is RB X (which is true in virtually all cases after the first block of Dynamic Cat), then a stimulus is selected that violates this rule, but conforms to that participant's predicted diagonal decision boundary. Additional consideration is given to verify that 1) stimuli are sampled far enough away from the last stimulus in that area of space to ensure uniform sampling of the space. 2) A mixture of *easy* and *hard* stimuli are shown based on their distance from the decision boundary with the assumption that stimuli further from the decision boundary are easier than those closer to it.

To validate this method, we replicated experiment 3 (Dynamic Cat) and substituted the previously used pre-determined stimulus presentation order for the adaptive stimulus selection method driven by PINNACLE's predictions in experiments 4 and 5.

Summary

Across three category learning paradigms we discovered an unanticipated propensity in participants to use more complex rules than previous assumed. This suggests a greater role for complex explicit decision-making than was accounted for in previous models. The Dynamic Cat protocol introduced the first attempt, to our knowledge, of experimentally inducing uncertainty into the category learning process thus driving strategy exploration. This feature allows us to develop increasingly accurate accounts of the decision-making process.

PINNACLE 2.0 was developed to better account for these data and represents a more complete theory of human cognition than its previous version. It predicts that people use both one- and two-dimensional rules to judge category membership across both explicit and implicit memory systems. These strategies independently compete for control of behavior as people process imperfect stimulus information due to intrinsic neural noise and incorporate feedback proportional to their confidence. This allows people to transition between simple strategies, and from simple to complex strategies based on task demands. Its success in accounting for a range of group- and individual-level behavior suggests it likely reflects a plausible account of how people navigate this process.

Chapter 4: Dynamic Cat & PINNACLE 2.1

Following the success of experiment 3 and PINNACLE 2.0, two additional datasets were gathered using the Dynamic Cat protocol. Rather than the orderly transition from simple to complex rules seen in experiment 3 and modeled by PINNACLE 2.0, participants displayed a wide range of behaviors in response to the changing environmental demands. The first dataset, experiment 4, utilized PINNACLE 2.0's predictions to drive adaptive stimulus selection during the task. Results indicated a mixture of transitions between one- and two-dimensional rules as well as explicit to implicit strategy switching, similar to results from experiment 3. For the second dataset, experiment 5, the Dynamic Cat task was performed while neuroimaging data were collected. Behavioral results of experiment 5 show a significant departure from previous results in several ways: 1) an unexpectedly high rate of participants (18/32) failed to achieve greater than 60% performance on the final block of learning; 2) participants (10/14) who did achieve higher than 60% accuracy were much more likely to verbally describe an appropriate two-dimensional rule; and 3) participants who did not achieve greater than 60% accuracy displayed prolonged periods of unsuccessful strategy exploration without evidence of a loss of motivation or disengagement from the task.

These results expose several important limitations in PINNACLE 2.0's approach. First, the data indicate that participants considered a broader set of explicit strategies in attempting to solve the task than the ones modeled by PINNACLE 2.0. To address this, we explored a wider variety of explicit strategies including one-dimensional, quadrant, and exemplar-based strategies. In addition, the high rates of explicit II strategy use indicated that relatively few participants adopted implicit strategies or even a rule approximating a diagonal line. This suggests a smaller role for such a strategy in the model. Second, accounting for these data given a broader set of explicit strategies requires an update to both the decision-making mechanism as well as the way in which knowledge representations are modified.

To address these limitations, we further developed PINNACLE to version 2.1, such that it could accommodate the richer set of behaviors seen across these experiments, thus providing a more accurate model of how participants navigate this dynamic and challenging task. In developing PINNACLE 2.1, specific attention was paid to capturing both the variability of performances as well as the wide range of behaviors exhibited by participants.

Experiment 4 – Adaptive Tutor Pilot

Participants

Nineteen participants were recruited from the Northwestern University research participation pool. All participants had normal, or corrected to normal vision, were provided full informed consent in accordance with Northwestern University's IRB protocols, and received course credit for their participation.

Methods, Materials, and Procedure

All methods, materials, and procedures were identical to experiment 3 except for the method of stimulus selection. In this experiment, PINNACLE 2.0 was fit to participant choice behavior on all trials for which a response was made (data from non-responses were omitted from model fitting) during the jittered ITI periods. Its predictions were used to select a stimulus that discouraged an explicit strategy, while encouraging an implicit one as described earlier.

Results

Accuracy results were not significantly different from those in experiment 3. n=19 (6 excluded), 480 trials, 10/4 distribution, 1,000ms stimulus presentation time, group accuracy on last block = 78%. Critically, using this method, we found only 2 of the 19 participants were able to verbally report an explicit II strategy (10%).

Experiment 5 – Dynamic Cat with Neuroimaging

Methods, Materials, and Procedure

Thirty-five participants were recruited from the community in order to obtain a more diverse sample and performed the same task as in experiment 4 while in a fMRI scanner. Seven participants were excluded from data analysis, and their data was not used: participants and exclusion criterion are described in the Exclusions section.

Our neuroimaging protocol used a Siemens 3T fMRI scanner. During our Dynamic Cat task, we ran a T2*-weighted echo planar (EPI) scan covering 35 interleaved axial slices (TR = 2200 ms, TE = 21 ms, multiband = 2, voxel size = 2×2×2 mm) for 218 volumes in each of six scans. For anatomical localization, high-resolution, 3D MP-RAGE T1-weighted scans (voxel size = 1×1×1 mm; 128 axial slices) were collected for each participant following the functional runs. Neuroimaging work is discussed in Chapter 5.

Behavioral Data

All methods and procedures were identical to experiment 4 with the exceptions that the experiment was done in a fMRI scanner, recruitment was done more broadly from the community, and that the compensation was higher given the added requirement of neuroimaging. All participants were provided informed consent and a preliminary medial screen before agreeing to participant in the study. No health complications due to the scanner were observed or reported, however, one participant stopped the experiment at the end of the first block due to feelings of uneasiness.

Results

Results of experiment 5 differed significantly from those of experiments 3 and 4 in three distinct ways. First, the unexpectedly high rate of participants who could verbally report an explicit-II strategy. In experiments 3 and 4, participants who achieved >60% accuracy on the final block of learning were largely unable to report an explicit-II rule. In experiment 5, 10 of the 14 participants who achieved >60% on block 5 were able to articulate the rule. Second, there was an unexpectedly high rate of participants (18/32) who failed to achieve greater than 60% accuracy on block 5. This is a significant deviation in terms of the number of excluded participants when compared to experiment 3 (9/30) and experiment 4 (6/19). Third, the range of accuracy performance seen by the end of the experiment. A summary of these results is shown in Figure 4.1.

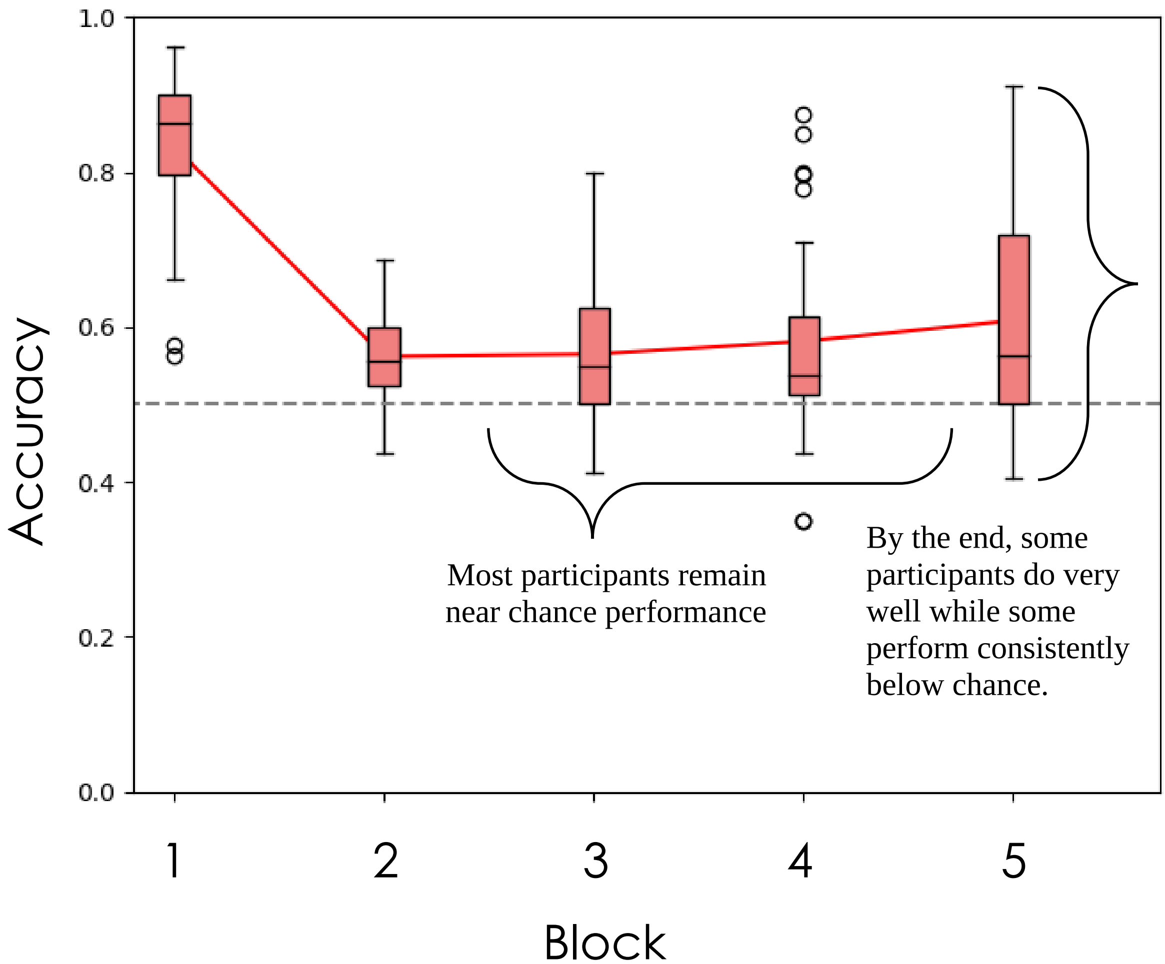


Figure 4.1

Accuracy by Block results for experiment 5. Participants show the expected behavior during the Bait (block 1) and Switch (block 2) phases. However, by the end of learning most participants remain at approximately 60% accuracy. These results suggest a prolonged period of unsuccessful strategy exploration compared to results of experiments 3 & 4.

Analysis of Choice Behavior

In visual category learning participants who fail to achieve greater than chance performance (typically 55%, 60% or 65% accuracy) by the end of the experiment are traditionally excluded from analysis. The motivation to exclude these participants is often that the researchers are interested in studying learning behavior and presumably these participants did not learn. In some cases, the decision to exclude these participants is further supported through competitive model fitting. This process attempts to account for their choice behavior through several models corresponding to different hypothesized rules such as DBT models based on individual dimensions as well as combinations of them (i.e. RB-X, RB-Y, diagonal, conjunctive, etc.). In addition, a so-called Random Responder model is fit which simulates either biased or unbiased random guessing throughout the experiment. When participants are best fit by this Random Responder model it is taken as support of the claim that these participants can be reasonably excluded from further analysis since they either failed to engage with the task or failed to learn it.

In experiment 5, 18/32 participants failed to achieve greater than 60% accuracy on Block 5. While it is tempting to simply exclude these participants, several factors lead us to believe the behavior observed here is not the product of so-called “random responding”. In the following section we look for evidence that these non-learners disengaged from the task. In cases where participants seem to respond in stimulus-independent ways we exclude those trials. Of the seven excluded participants, three were found to have disengaged form the task significantly throughout a majority of the experiment and were thus excluded from further data analyses. Of the remaining 28 participants a total of 16% of trials were omitted based on the criteria described in the following sections.

Engagement

While in many instances chance performance may be attributed to a lack of motivation or engagement, participants in this study all report both being frustrated by the task, as well as trying to solve it throughout its duration. In addition, participants recruited for fMRI studies tend to be highly motivated and are compensated well for their time. Our participants did not show signs of falling asleep or prolonged periods of non-responses (aside from one participant who was excluded for an abundance of non-responses) that would indicate they were disengaged from the task. Total non-responses accounted for 1.14% of the data.

Random Responding

An established body of evidence demonstrates that humans are bad at producing random numbers (Ginsburg & Karpiuk, 1994; Rabinowitz et al., 1989), and when they attempt to it leads to marked increases in their reaction times on dual-tasks indicating that attempting to produce random responses acts as a working-memory task (Brown & Marsden, 1991). This is likely because people attempt to track the answers they have generated in an attempt to ‘balance’ the set such that it appears more random. In many cases this leads to an under production of long strings of the same response (i.e. 8 A’s in a row), and an over production of alternating responses (A, B, A, B).

Reaction Times

To test whether learners and non-learners differed significantly on reaction time (RT), we performed an ANOVA using reaction time as the dependent variable, and block number and learners / non-learners as the independent variable. There was a reliable difference between learners' RT (M=0.789, SD=0.120) and non-learners (M=0.735, SD=0.115) such that non-learners responded approximately 0.054ms faster: $F(1,190)=10.53, p=0.0014$. There was no reliable difference between the two groups across blocks: $F(5,190)=1.19, p=0.316$.

Stimulus-Independent Strategies

Still, a reliable difference in reaction time may indicate differential engagement in the task, and in order to further rule out the possibility that participants were disengaged from the task we performed behavioral analyses on their response patterns in an attempt to identify plausible maladaptive behaviors. Critically, these strategies should lead to responses that are independent on the stimulus. Three obvious ways of disengaging from the experiment we tested in which responses are independent of the stimulus were 1) repeatedly pressing the same button (e.g. "A", "A", "A", "A", "B", "B", "B", "B", etc.), 2) consistently alternating responses from one to the other (e.g. "A", "B", "A", "B", "A", "B", etc.), 3) responding to the current stimulus based on the previous stimulus label (e.g. "The last stimulus was an "A", so this one is a "B" or vice versa.).

Method:

There are currently no established criteria for identifying trials of interest for exclusion regarding these strategies. Importantly, any sequence of responses to be excluded must be longer than one would reasonably expect by chance thus indicating a deliberate strategy. For example, a sequence of five 'A's in a row happens somewhat frequently in a randomly generated set of 480 responses (~8% of trials), whereas a sequence of sixteen 'A's is considerably less common (<0.05% of trials). In order to support the claim that this was an intentional strategy rather than an incidental occurrence, the pattern should exceed expected rates. To characterize the frequency of expected occurrences of these strategies in our experiment, assuming true random responding, we ran a Monte Carlo simulation by generating a dataset of 10,000 instances of 480 (the number of trials in our experiment) random responses. These data were then used to create an expected distribution of two strategy frequencies. We then compared both learner and non-learner data to this expected frequency distribution. Figure 4.1 shows the distributions for both single response and alternating response strategies.

Repeated Response

The first maladaptive strategy we tested for was one in which the participant generates multiple trials of the same response. This strategy does not depend on stimulus information and indicates a lack of engagement with the task and would thus be grounds for exclusion. As seen in Figure 4.2 (top) learners and non-learners both under-produce sequences compared to the expected frequency up to length 8 but do not differ significantly from each other until length

15+. We thus excluded any block containing a sequence of 15 or more resulting in 4/102 blocks in two participants being excluded in the non-learners and no blocks in the learners.

Alternating Responses

The second maladaptive strategy we tested for was one in which the participant generates an alternating pattern of A's and B's. This strategy also does not depend on stimulus information and indicates a lack of engagement with the task and would thus be grounds for exclusion. As seen in Figure 4.2 (bottom) learners and non-learners match simulated data until length 7 sequences after which they deviate from the expected distribution. However, they do not deviate from each other until lengths of 15+. Following the criteria for repeated responses, we excluded any block containing a sequence of 15 or more resulting in 13/102 blocks in ten participants being excluded in the non-learners and 3/84 blocks in three participants in the learners. In addition, participant 27 was entirely excluded from the non-learners due to having sequences of 15 or more in 5/6 blocks and participant 10 was entirely excluded for having 39% of trials excluded.

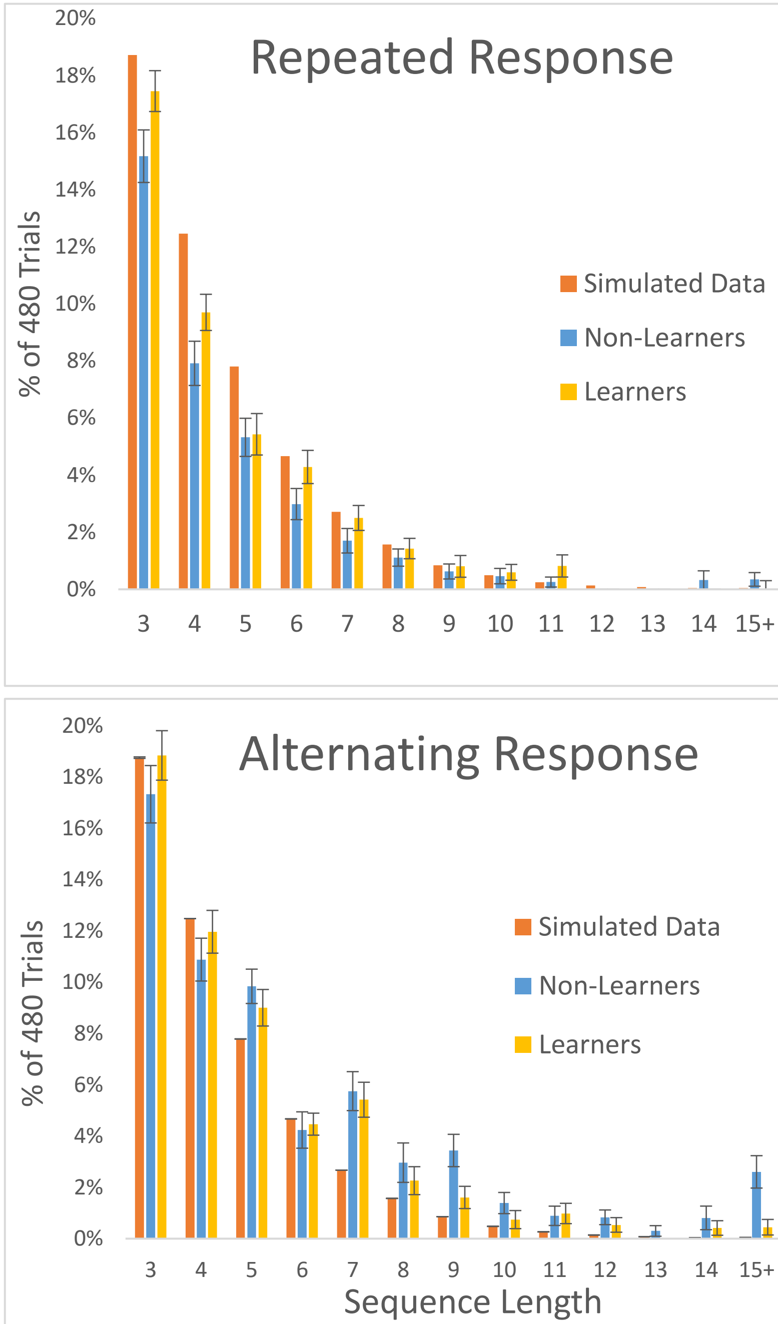


Figure 4.2: Frequency distributions of 10,000 Monte Carlo simulations (orange), the bottom half of participants (blue) and top half of learners (yellow).

Top: Distributions of sequences of the same response ('AAA…' or 'BBB…'). Participants under-produce compared to expected frequencies until lengths of 8 but learners do not differ from non-learners. Non-learners and over-produce compared to learners at lengths of 15+.

Bottom: Distributions of sequences of alternating patterns ('ABAB…'). Both participant groups over-produce at lengths greater than 8 compared to expected data and non-learners over-produce compared to learners at lengths of 15+.

Transitional Probabilities and Response Bias

Having excluded the participants and trials above we examined the transitional probability structure of the data. This analysis tested for evidence that participants exhibited a bias for same-same (i.e. responding 'A' after 'A' or 'B' after 'B') or same-different (i.e. responding 'A' after 'B' or 'B' after 'A') responses or whether they showed evidence of statistical learning of the true underlying stimulus label distributions. To characterize both the stimulus label orders and the participant's response patterns, we divided the number of same-different labels and responses by same-same labels and responses to get an odds-ratio that represents the rate of responding same-different compared to same-same. Overall, trial orders had approximately 6% greater tendency towards same-different trials as a consequence of limiting long consecutive same-same streaks during stimulus presentation. Participants showed a slightly stronger tendency to prefer same-different responses 11-20%. In addition, the increased tendency of participants to prefer same-different responses compared to the trial orders suggests this was not a consequence of statistical learning of the trial order frequencies since they do not match the rate but rather a response bias.

Exclusions

In sum, seven participants were excluded from analyses in experiment 5. As described earlier, two participants (10 & 27) were excluded for an abundance of trials accounted for by an apparent alternating response strategy. Participant 29 was excluded for failing to respond to 19% of trials. Participant 13 was excluded due to technical difficulties in the scanner during block 1 such that they could only respond to category 'A'. Participant 16 was excluded after it was discovered that they required glasses for corrected to normal vision after having completed the experiment without corrective eyewear. Participant 19 withdrew from the experiment after 1 block. Participant 8 was excluded due to concerns with their structural MRI scan. The concern was forwarded to the medical contact through the scanning center. Finally, we excluded any trail for which no response was made across all remaining participants.

Decision-Making under Uncertainty

The Dynamic Cat task affords us an excellent opportunity to study the decision-making process in a dynamic environment, especially during periods of uncertainty. The task requires that participants search a potentially broad space of possible strategies in a relatively short amount of time. As such, participants must generate, use, evaluate and discard strategies in short order to succeed in the task. Rather than a convergent profile of behavior on the task in which most participants perform roughly the same way, our participants displayed a wide variety of behaviors. Some participants were hardly affected by the change in task demands during the switch and maintain a consistently high level of accuracy throughout the experiment. These participants seem to have easily adapted to the change and exhibited an efficient process of strategy exploration. Other participants showed intermediate success in which they seem to gradually discover increasingly better strategies as evidence accumulated. Yet others experienced significant difficulties in finding appropriate strategies, often failing to do so for the rest of the experiment following the switch. Whether an appropriate solution never occurred to them, or they were profoundly unlucky, these participants showed a persistent pattern of trial-and-error. As seen with PINNCALE 1.0 and 2.0, understanding this range of processes is greatly aided by computational models that attempt to capture this behavior and account for as much of the data as possible using specific and detailed mechanisms.

PINNACLE 2.1

PINNACLE 2.0 is a model of rational learning that was designed to account for a wide range of data on category learning and decision-making. It is capable of producing human-like performance, even on a complex task like Dynamic Cat. However, matching the experiment 5 dataset requires that not only learning behavior be accounted for, but also prolonged periods of near chance performance following high initial accuracy. PINNACLE 2.0 cannot achieve this since it incorporates an objectively accurate strategy (diagonal boundary) that incrementally gains accuracy and confidence even when not controlling behavior. As such, over time the model is guaranteed to both find an optimal strategy and control behavior. The only way to avoid this outcome is to effectively set the learning rate for that strategy, which in some cases represents implicit learning to zero. This modification simulates something like a dopamine or learning disorder. Since we do not believe our participants have neurological deficits - as evidenced at the very least by their ability to learn in block 1 - this is not an appropriate solution to the problem. In addition, doing so would limit the model's ability to match the most successful learners. Thus, given PINNACLE 2.0's implementation there is a trade-off between matching the best or the worst learners. The variability and wide range of performance exhibited by participants cannot be accounted for by PINNACLE 2.0.

The high percentage of non-learners in experiment 5 demonstrate that participants can fall into periods of prolonged unsuccessful strategy exploration despite trying to succeed. This poses a general problem for models of learning. These models can account for non-learning behavior by effectively shutting off learning, or through adding significant noise to either perception or decision-making. While these methods can allow such models to account for some of the performance seen in experiment 5, they are not appropriate solutions for these data. Because participants initially learn a simple strategy in the Bait phase, they demonstrate they are capable of rapid learning. This precludes modeling solutions that rely on near-zero learning rates, or high levels of system noise. We thus need a model capable of rapid early learning during the Bait phase, a reduction in accuracy during the Switch, and either successful or unsuccessful learning in the Learning phase.

Importantly, this model should be able to account for the wide range of participant performance without the need to modify its parameters for each participant's trial order. The idea is to create a model of the general process participants engage in rather than custom-make a solution for each participant. While the latter is certainly an interesting endeavor, creating a general framework that captures the rich variability that participants display is an important first step towards understanding the general process of dynamic decision-making. In this section, we consider and evaluate several candidate models in order to address these needs. The result is PINNACLE 2.1. Since the evidence for implicit or diagonal strategy use in experiment 5 is quite low, PINNACLE 2.1 focuses primarily on explicit strategy use. Thus, the implicit module is effectively ignored for the purpose of fitting these data. To further explore the role of a diagonal strategy we incorporate it in some of our models. Based on competitive model fitting, we show that it predicted to be a minor contributor to the set of strategies considered by participants. Future work should aim to reincorporate an implicit module with mechanisms for preferring explicit over implicit strategies during periods of targeted strategy search.

Datasets

To build the most robust model of what participants do during this task, we aggregated data from experiments 3, 4, and 5 after filtering participants based on the previously described exclusion criteria for a total of 78 participants. As a brief reminder, experiment 3 included a set of 30 participants each of whom completed 500 trials of the Dynamic Cat task with 100 trials in the Bait phase, 100 trials in the Switch phase, and 300 in the Learning phase. As such, we divided their data into blocks of 100 trials for the purpose of averaging with a total of 5 blocks. In this experiment during the Switch phase, stimuli that disconfirmed the simple rule from the Bait phase were sampled at a ratio of 3:1 disconfirming to confirming stimuli and then 2:1 for the Learning phase.

Experiments 4 & 5 had 19 and 28 participants respectively, each of which completed 480 trials of the Dynamic Cat task with 80 trials in the Bait phase, 80 trials in the Switch phase and 320 trials in the Learning phase. As such, we divided their data into blocks of 80 trials for the purpose of averaging with a total of 6 blocks. On the final block of the experiment feedback was withheld from the participants. This was intended as a pure test of their category knowledge without the ability for supervised learning. However, the adaptive tutor was mistakenly left enabled during these trials and thus continued to attempt to dissuade them from any simple strategies they might be using. Since this was not an intended aspect of the experiment, and in order to normalize block counts across the three datasets, we report results from the first five blocks of the experiment rather than all six. Accounting for the sixth block of the fMRI dataset is an interesting challenge that will be left to future work.

Candidate Models

Here we describe in detail the eight models we considered in an attempt to fit data from the current dataset. Each model is detailed in terms of specific mechanisms as well as the mental model it is intended to capture. We then evaluate how each model performs across our task describing limitations or successes during the Bait phase, Switch phase and Learning phase. As with PINNACLE 2.0 we organize the architecture into a standard flow of information: Perception → Cognition/Decision-Making → Response → Feedback & Learning. Across all these models both the perceptual and response systems were identical. While PINNACLE 2.0 incorporated perceptual noise in the form of a random Gaussian number added to the X and Y coordinates of the incoming stimulus, PINNCLE 2.1 simplifies this approach and directly maps the stimulus to perceptual space. The various models we considered differ in two primary ways: 1) the specific knowledge representations used as explicit strategies, and 2) how they update those knowledge representations based on feedback. In addition, in cases where more than one strategy is available, an alternative decision-making mechanism to the one used in PINNACLE 2.0 is introduced. Table 4.1 summarizes the different models we considered while developing PINNACLE 2.1. The following sections detail each model in terms of the mechanisms and intended cognitive state.

Parallel versus Sequential Systems

In PINNACLE 2.0, three strategies (RBx, RBy, Diagonal) operate in parallel and compete for control of behavior on each trial. When confronted with a new stimulus, PINNACLE 2.0 considers each strategy before selecting a response. While powerful, such an architecture comes with necessary complexity. For example, it needs to keep track of every strategy available to it, and the question of how to select the winner must be solved. An alternative approach is a sequential system in which a single strategy is considered at any given time. For example, instead of selecting among RBx, RBy, or a Diagonal, the system might only consider RBx for a time, then switch to considering RBy. Sequential systems have the advantage of avoiding the problem of keeping potentially large numbers of alternatives in mind during the decision-making process. Explicitly evaluating even just three strategies seems unlikely given the limited response window available to participants. The models we considered for PINNACLE 2.1 included both parallel and sequential systems.

Single Step versus Hierarchical Rules

In PINNACLE 2.0, each strategy is comprised of a single step rule. These rules only check whether the stimulus is on one side of the boundary or the other in order to assign a category. A natural extension of this is a hierarchical process in which a series of rules are checked in the decision process to reach an answer. Such a process allows for more elaborate strategies to be constructed at the cost of complexity. Given our preference for simplicity and maintaining interpretability of our models, we only implemented hierarchical structures in sequential models, and restricted them to two-steps.

Rule Updating versus Rule Replacement

In PINNACLE 2.0, each strategy is updated based on feedback. The diagonal rule is incremented towards an optimal position, the RBx and RBy boundaries are moved to the position of the current stimulus, and in all three cases the perceptual shaping (PS) parameter is increased or decreased acting as a proxy for confidence in the rule. Thus, regardless of the position of the boundary, the RBx rule is considered a single strategy that the participant gains or loses confidence in while modifying it. Such a model predicts that people effectively edit, or overwrite existing strategies – "Oh, the boundary should be a bit more towards thick stimuli than I thought." An alternative to this approach is to replace rules rather than update old ones. Under such a model instead of incrementing the decision-boundary by some amount, a new boundary is instantiated – "Oh, that old rule didn't work, I think this new one is the right one." A main difference between these approaches is whether rules are edited, or new rules are created. The models we considered for PINNACLE 2.1 included both rule updating and rule replacement mechanisms.

	Parallel or Sequential	Knowledge Representation	Representation Updating
PINNACLE 2.0	Parallel	Thickness, Orientation & Diagonal Rules	Incremental Learning PS grows when incorrect & shrinks when correct.
Exemplar	Sequential	Similarity-based	Stores previous examples
Basic RBx	Sequential	Thickness Rule	No Learning
Incremental RBx	Sequential	Thickness Rule	Incremental Learning
Incremental RBx | RBy	Parallel	Thickness & Orientation Rules	Incremental Learning PS grows when incorrect & shrinks when correct.
Satisficing Anchored	Sequential	Thickness & Orientation Rules	Rules replaced when no longer good enough
Satisficing Anchored w/ Bias	Sequential	Thickness & Orientation Rules	Rules replaced when no longer good enough
Incremental RBx w/ Exceptions	Hierarchical	Thickness & Orientation Rules w/ quadrant exceptions	Incremental Learning. Exceptions replaced when not good enough
Satisficing Anchored w/ Exceptions	Hierarchical	Thickness, Orientation & Diagonal Rules w/ quadrant exceptions	Rules & Exceptions are replaced when not good enough

Table 4.1: Descriptions of nine models (rows) in terms of parallel or sequential systems (column 1), mechanisms of knowledge representations (column 2) and how those knowledge representations are updated (column 3).

Exemplar Model

The first model we considered is the Nosofsky exemplar model. As discussed in chapter 2, an exemplar model is one in which new stimuli are compared to previously seen examples on the basis of nearness or similarity. Thus, if the current stimulus is nearer in perceptual space to previously seen examples of category A than category B, the stimulus is assigned to category A. “This stimulus looks more like the ‘A’s I’ve seen than the ‘B’s. Exemplar models are necessarily one-step, sequential, and continuously update a single strategy in the form of storing previously seen stimuli.

Exemplar models are successful at characterizing various learning profiles and are applicable in situations where stimuli can be mapped to a perceptual space such that distance can be calculated. Modeling the current dataset provides an interesting and difficult challenge since any solution must be capable of rapid successful early learning followed by a drop in accuracy and then either prolonged, stable sub-60% performance, or gradual learning to approximately 80% accuracy. Exemplar models are capable of rapid, gradual, or no learning over prolonged periods of time. However, creating a model that is capable of producing a mix of all three behaviors is not possible without the addition of significant complexity or by customizing the model on a per-participant basis. In short, this is because learning models are good at learning. A model capable of incorporating feedback such that it learns rapidly in the Bait phase (block 1) will similarly be capable of learning rapidly during the learning phase (blocks 3-5). On the other hand, a model that does not learn rapidly will fail to match performance in the Bait phase. As such, a model that accounts for participant success cannot also account for participant failure.

Architecture

Figure 4.3 depicts a cartoon of how a decision is made in the exemplar model. A stimulus is encountered and mapped to perceptual space (thickness on X, orientation on Y). The distance to each category of previously seen exemplars is then calculated. The distance scores from both categories are compared and the current stimulus is assigned to the nearest category. In our models, the distance to each category was calculated using the linalg.norm() function in the Numpy library which calculates the L^2 norm (Euclidean norm) though other distance calculations did not modify our results significantly. Feedback is obtained, and the current stimulus is added to the list of previously seen exemplars for the category it was assigned to. In this model, learning is the act of storing the stimulus as a previously seen exemplar and happens on every trial. The feedback signal has no influence on the mechanisms of the model. The distance and similarity functions remain static throughout the experiment.

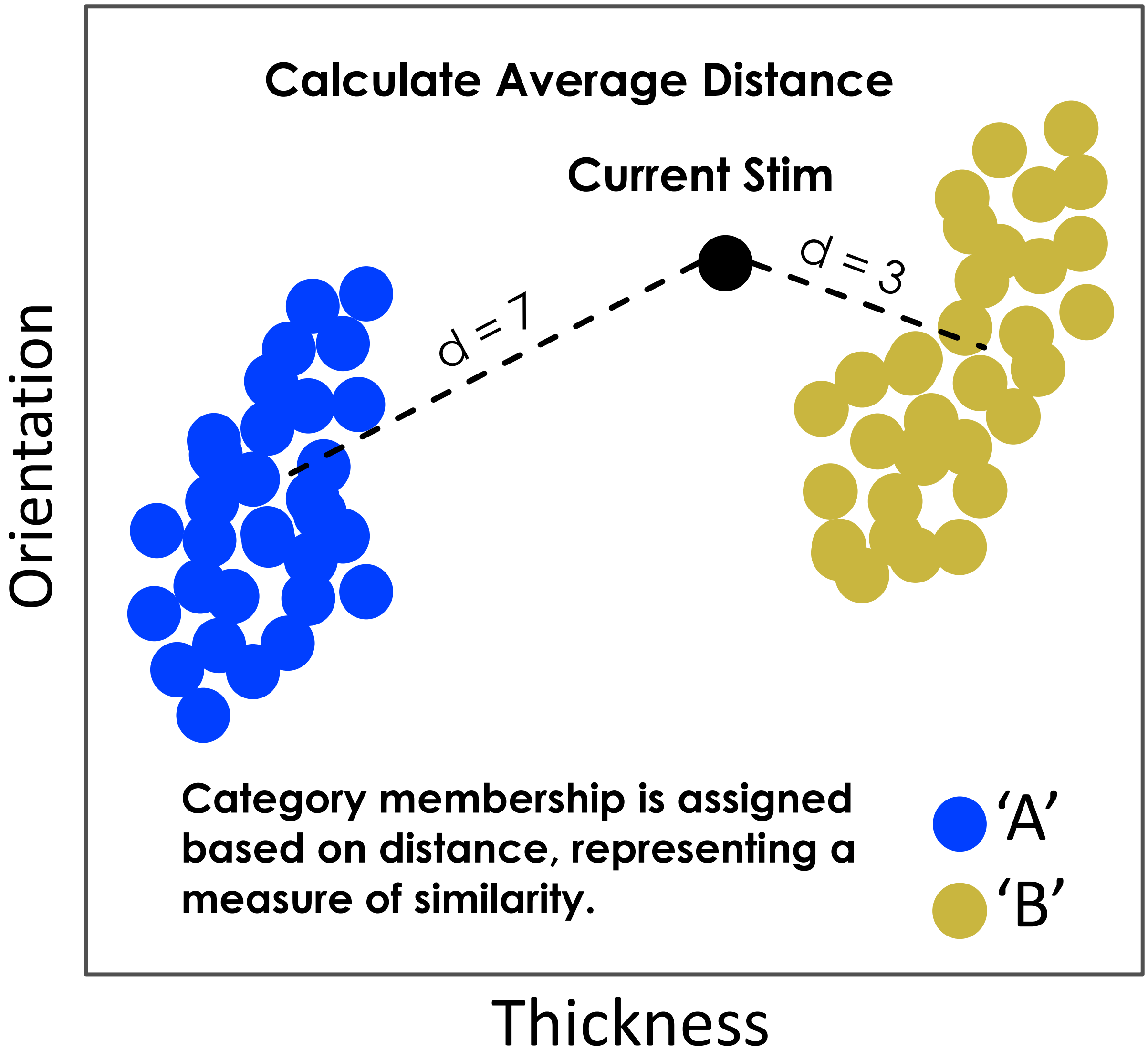


Figure 4.3:

A cartoon representation of the exemplar model's decision-making process. A stimulus is mapped to perceptual space. Average distance to each category is calculated using Euclidean distance and a response is made based on the closest category. Feedback is obtained and the stimulus is stored in the correct category regardless of the actual decision.

Rule-Based models

The next six models we consider are based on the Ashby decision bound theory (DBT) discussed in chapter 2. These rule-based models are successful at describing category learning as evidenced by PINNACLE 1.0 & 2.0 as well as an abundance of prior studies. They capture the concept of explicit rules based on some criterion for deciding between alternatives as in: “The deciding factor is thickness. If thicker than some value, A, else B.” These models include parallel and sequential systems. Of the sequential models, we considered both single step and hierarchical versions. We also considered models that update existing strategies or replace them.

Basic RBx

The simplest model we considered was a sequential, single step model that does not learn. Figure 4.4 depicts a cartoon of how this model makes a decision. A stimulus is encountered and mapped to perceptual space (thickness on X, orientation on Y). A decision boundary based on some feature or combination of features that bisects the space is used to evaluate the stimulus and assign it a category label. In this model the boundary is based on the X axis: Rule Based X or RBx. The stimulus can either be on one side of the boundary or the other (left or right for RBx). Formally, this is calculated with the “less than”, “less than or equal to”, “greater than or equal to”, and “greater than” ($<$, $\leq$, $\geq$, $>$) operators. For example, consider a RBx rule with a boundary at 0.5 in which all stimuli less than or equal to the boundary are labeled ‘A’ and those greater than the boundary are labeled ‘B’. A stimulus with the coordinates (0.2, 0.6) is encoutered. Since we are evaluating a RBx rule, the Y coordinate is ignored. Thus the stimulus X value: 0.2 is compared to the boundary: 0.5 as:

IF $0.2 \leq 0.5$: ‘A’, ELSE ‘B’.

We therefore describe the rule using two pieces of information: 1) The comparator (i.e. $<$, $\leq$, $\geq$, $>$) and 2) the category label associated with that comparator (i.e. whether a stimulus that meets the comparator criteria is labeled as ‘A’ or ‘B’).

In all our Rule Based (RB) models the initial position of the decision boundary is set between the first ‘A’ and ‘B’ stimuli it encounters. For all trials before that, the model randomly assigns a category label. In the case of the basic RB model, the RBx rule always has the comparator $\leq$ and the label is ‘A’. Thus for all instances of the model, stimuli *less than or equal to* the boundary are assigned to category ‘A’ and those *greater than* the boundary are ‘B’. Feedback is then obtained and recorded. In its most basic form, this model does not learn and thus feedback is only used to track its performance.

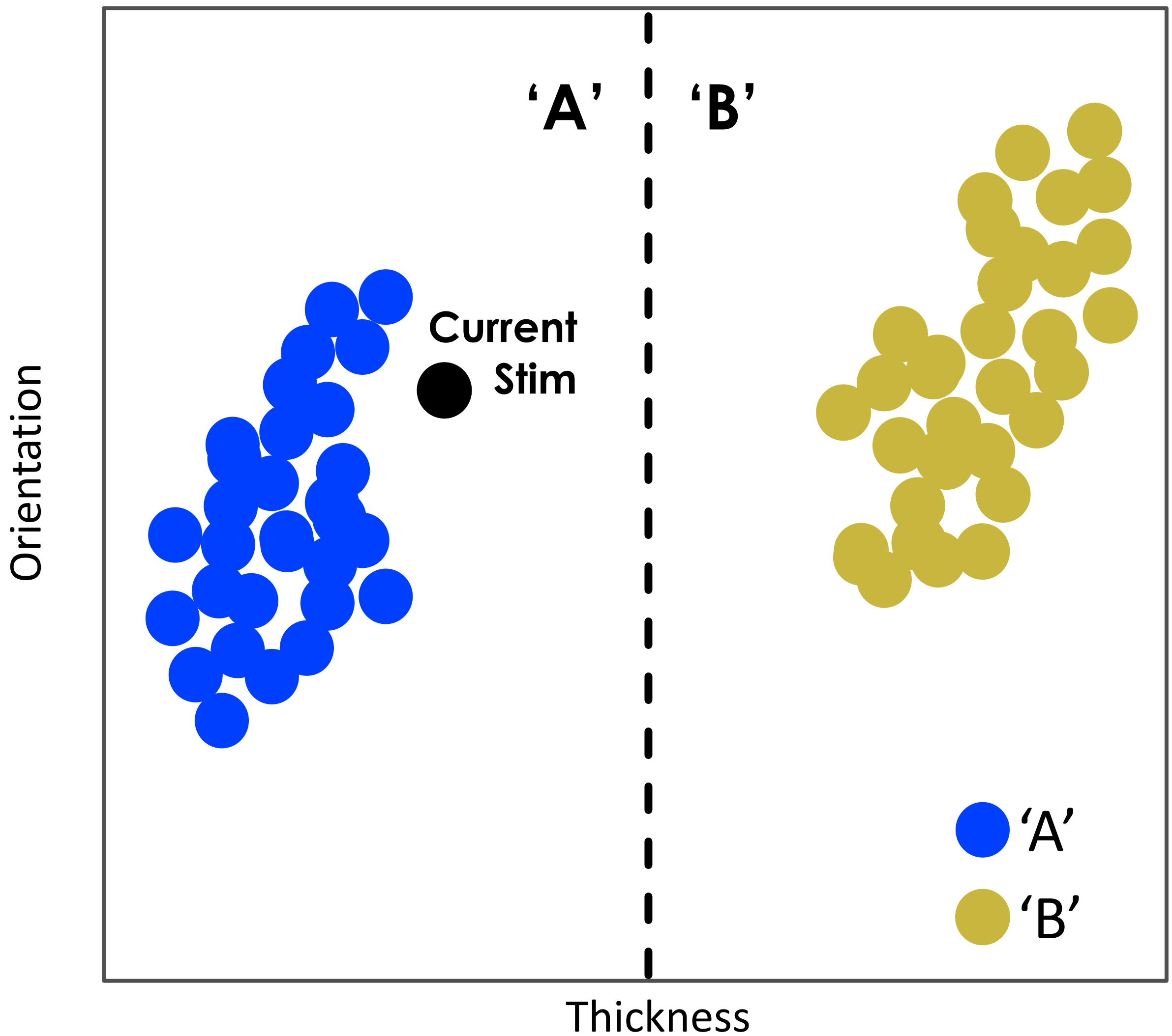


Figure 4.4:

A cartoon representation of a rule-based model's decision-making process. A stimulus is mapped to perceptual space. A linear decision-boundary bisects the space into two categories. A category is assigned based on which side of boundary the stimulus is on. Feedback is strictly used to track performance.

Incremental Learning - RBx

The next model we considered was also a single step, sequential model that updates an existing rule by modifying its decision boundary incrementally in response to feedback as shown in Figure 4.5. In this model feedback is used to improve subsequent predictions. When incorrect, the boundary is incremented by a fixed learning rate in the direction that would have produced the correct answer. The learning rate is taken as a free parameter and a single value is found that provides the best fit across all trial orders. Thus, the incremental learning model simulates learning as continuously modifying a single rule. While it is unlikely that participants truly perseverate on a single rule endlessly trying to optimize it, this model has nonetheless shown success in the past at accounting for category learning behavior.

Incremental Learning – RBx & RBy

The next model we considered was a single step parallel model that updates existing rules by modifying their decision boundaries incrementally in response to feedback. This model tested both RBx and RBy rules for each stimulus. The RBx rule labeled all stimuli *less than or equal to* ($\leq$) the boundary as 'A's and those *greater than* ($>$) the boundary 'B's. The RBy rule labeled all stimuli *greater than or equal to* ($\geq$) the boundary 'A's and all those *less than* ($<$) the boundary 'B's. Since this model involves more than one possible solution, we implemented a simple decision-making as in PINNACLE 2.0.

On each trial each rule provided a *p(A)* value by calculating the area under the curve of bivariate Gaussian centered on the stimulus and bisected by the decision-boundary. The resulting predictions were compared and the most confident one was selected to drive behavior. If correct, the standard deviation (*sd)* on that rule's Gaussian was reduced to simulate an increase in confidence since more of the distribution would fall on one side of the boundary. When incorrect the *sd* was increased and the boundary was incremented in the direction that would have produced the correct answer. Thus, these two strategies compete with each other for behavior with the most confident among them driving behavior.

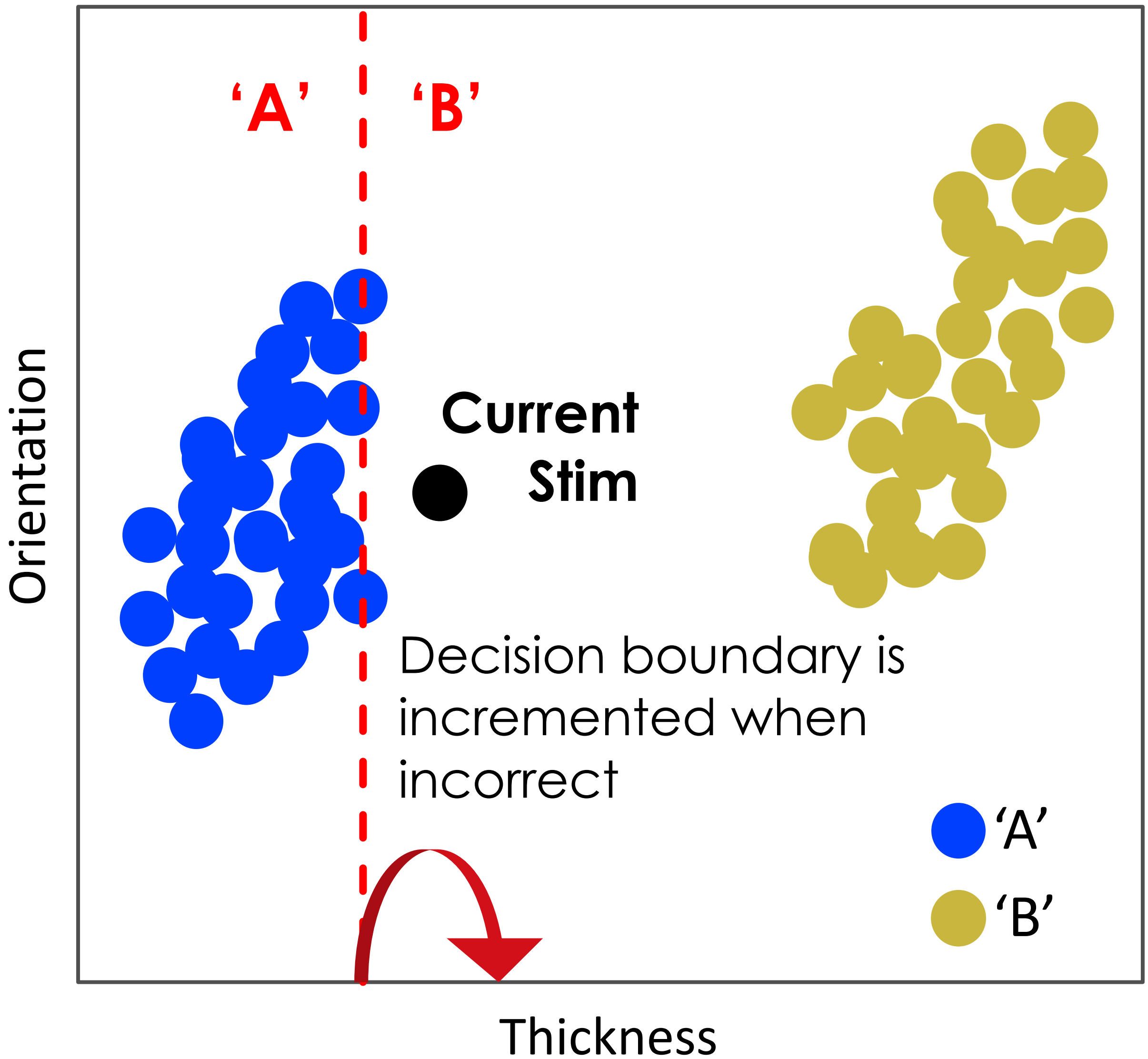


Figure 4.5:
A cartoon representation of an incremental learning rule-based model. Each time the model is wrong, the decision-boundary is incremented in the directed that would have produced the correct answer.

In the case above, the black dot denotes the current stimulus. Since its X-value is greater than that of the decision-boundary it is assigned to category ‘B’ on the right (orange). However, this is incorrect and so the boundary is incremented by a fixed amount in the direction that would have produced the correct answer.

Satisficing Anchored Model

Our current dataset confronts us with the difficult requirement of producing both learning and non-learning within a single framework. In this case we want a model that can produce both successful and unsuccessful rules and thus either exhibit learning or non-learning under the same set of parameters. A simple solution is to increase the possible set of rules available to the model, and to allow for random chance to influence the results. As discussed earlier, an alternative to incremental learning in which a single rule is continually modified is one in which the rule itself is replaced with a new one based on how well it is doing. This intuitive meta-strategy of using a rule until it is deemed no longer good enough and then replacing it with a new rule is called *satisficing*, as introduced in chapter 2. Instead of searching for the perfect rule, we settle for a rule that meets our criterion for "good enough".

The next model we considered was a single-step sequential model that replaces its rule rather than updating an old one. Rather than consider multiple competing strategies, only a single rule is evaluated on any given trial. To computationally formalize the satisficing concept, this criterion is defined as the *base-rule satisficing threshold*. For the following single step satisficing models, a base-rule satisficing threshold of 80% was found to be the optimal value. This means that according to our model, participants' standard for a good enough rule was one that was 80% or better; a rule that performs worse is replaced.

In the incremental learning model, the emphasis is placed on learning with respect to rule modification, in this model the emphasis is shifted to decision-making (i.e. generating and evaluating rules) while learning takes the form of tracking an individual rule's success (i.e. how well it has performed in the past). On each trial a response is made based on which side of the boundary the stimulus is on. Note that since only one rule is considered at any given time there is no need for a *p(A)* estimate. After feedback is obtained the overall success of the current rule is updated and is then evaluated with respect to the *base-rule satisficing threshold* to decide whether it is good enough to keep using, or if it should be replaced. Figure 4.5 depicts a cartoon example of this process.

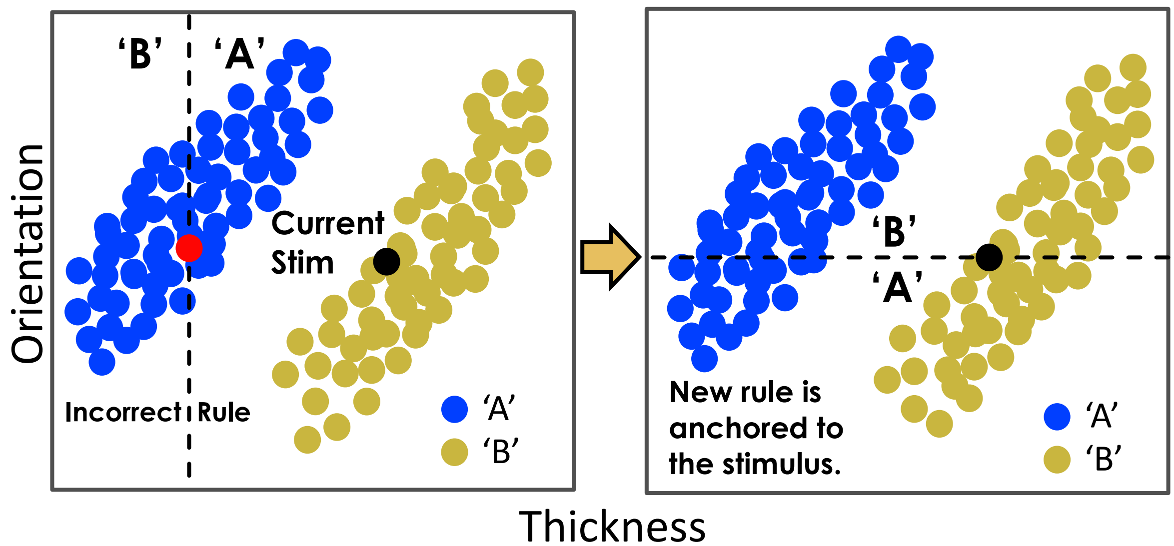


Figure 4.6: A cartoon representation of a satisficing anchored rule-based model. Each time the model is wrong, the current rule (left panel: in black) is evaluated based on its cumulative accuracy. If the accuracy is below a threshold value, the rule is replaced.

In the case above, the black dashed line represents the current decision boundary anchored to the stimulus (red dot) used to generate it. In this case the rule is of the form:

IF stimulus_X $\leq$ Boundary: 'B' ELSE 'A'.

The black dot denotes the current stimulus being evaluated. According to the rule the stimulus is assigned to category 'A' which is incorrect. This causes the cumulative accuracy of the current rule to drop below threshold and thus the rule is replaced with a new one (in the right panel) centered on the current stimulus. The new rule in this case was randomly selected as an orientation rule and has a randomly determined comparator, in this case $\geq$. Since the stimulus label is 'B' the new rule is of the form:

IF stimulus_Y $\geq$ Boundary: 'B' ELSE 'A'.

Rule generation

In the non-learning and incremental learning models the thickness rule was always of the same form: IF stimulus_X ≤ boundary: ‘A’ ELSE: ‘B’. The comparator was always ≤ and the label was always ‘A’. For orientation category ‘A’ was always above the line and ‘B’ below: IF stimulus_Y ≥ boundary: ‘A’ ELSE: ‘B’. In the current model we allow for a wider range of possible rules to be considered. This is both to allow the model the possibility of success and failure and based on participant verbal report of trying a variety of different rules. Figure 4.6 depicts an example of evaluating and generating a new rule. If the current rule’s cumulative accuracy falls below the *base-rule satisficing threshold* it is replaced. When generating rules we simulate the idea that the participant is in a state of uncertainty with respect to the state of the task. Perhaps the experimenter has surreptitiously switched all ‘A’s to be called ‘B’s and vice versa. Perhaps the rule is now based on orientation, or perhaps it is no longer an equal split across categories. All the participant can be sure about is the current stimulus. Rules are therefore *anchored* to the current stimulus such that whatever rule is created will accurately categorize the current stimulus. This is done by using information about the stimulus during rule generation. The model then generates a new rule in the following steps:

1) Randomly decide between a thickness or orientation rule. If a thickness rule is selected, a vertical boundary is created centered on the stimulus with the stimulus’s X coordinate. For orientation rules a horizontal boundary is centered on the stimulus with the stimulus’s Y coordinate. Thus the decision boundary in each case is anchored to the current stimulu’s position.
2) The model randomly decides whether the comparator will be ≤ or ≥.

3) The model assigns the current stimulus label to that side of the boundary.

These simple steps guarantee that the current stimulus is assigned its actual label since it meets the criteria of the rule (≤ or ≥ to the boundary and thus assigned its own label). For example if a stimulus with corrdinates (0.2, 0.7) and a label of 'B' is encountered and a new rule must be generated the model may, at random, select an orientation rule with a ≤ comparator. Since it is a RB_Y rule the x-value of the stimulus is ignroed. Thus the new rule will have the form:

IF stimulus_Y ≤ 0.7: 'B' ELSE: 'A'.

And in this specific case: IF 0.7 ≤ 0.7: 'B' ELSE: 'A' resulting in a label of 'B'.

This model captures the idea that participants generate a rule based on their most recent experience and use it until it is not "good enough". They then replace the rule with another one and repeat the process until a rule that meets their standard is found. We allow the model to explore a much wider range of rules, most of which inaccurate, in the hopes of finding an appropriate one that will meet its satisficing criteria. As the old saying goes: "There must be a pony in here somewhere".

Satisficing Anchored Model with Bias

During the post-experimental interviews, participants nearly all report that they considered thickness more important than orientation during the first block. This is unsurprising both because the stimuli they see lend themselves to a thickness rule, and that most participants do very well on that block. To capture this tendency, as well as to improve the fit of the model, the fifth model we consider is identical to the previous one except that it has a bias towards generating thickness rules over orientation rules. All other mechanisms, as well as the base-rule satisficing threshold were the same. During model optimizations, a RBx bias of 0.65 was found to be the optimal value across all trial orders.

Hierarchical Rule Based Models

All the models described thus far have been single step models capable of testing one-dimensional rules based on thickness or orientation. As discussed later, these models are not able to perform as well as the best learners in our experiments. We therefore need models that use more sophisticated rules in order to produce performance in line with participant behavior. The space of possible rule complexity we could add to the model is vast. Given that our guiding principle in model development is parsimony and interpretability, we restricted the set of possible rules based on participant verbal report. When interviewed, the best learners articulate a base rule "If it's thick, then it's an 'A'" as well as an exception to this rule "unless it's also very tilty, then it's a 'B'". We thus restricted our hierarchical models to sequential versions and limited them to a two-step process where the second step takes the form of an exception to the base-rule. In the one-step rule evaluation, the rule takes the form:

IF base_rule: *answer*; ELSE: *alternate-answer*.

Where *answer* and *alternate-answer* may be 'A' and 'B' (or vice versa).

The process is similar in the two-step evaluation except that instead of evaluating the base-rule first, an exception is first checked. If the exception does not account for the stimulus, the base-rule is then evaluated.

IF exception: *answer*; ELSE: base_rule (as above)

(i.e. if the *exception* doesn't apply, check the *base_rule*)

This feature was added to the *incremental RBx learner* and the *Satisficing Anchored model with Bias*. It was not implemented in the parallel *incremental RBx & RBy learner* since such a system would require much more complexity in order to function. For example, each of the hundreds of potential hierarchical rules would need to be tracked and would compete on each trial for behavior. This is both highly complex, and unlikely to reflect the mental state of participants.

Though these models are capable of generating a two-step rule, they only begin generating them based on necessity. This was done because participants have no reason to invent exceptions to a simple thick | thin rule in the Bait phase since they never see any stimuli that are not accounted for by that rule. Each model initially attempts to discover simple rules drawing from either RBx in the case of the *incremental RBx learner*, or both RBx and RBy for the *satisficing anchored model with bias*. The first time a rule is modified from one- to two-steps is when a one-step rule whose accuracy exceeds the base-rule satisficing threshold subsequently drops below that threshold. This is meant to capture the idea that “I’ve found a really good (one-step) rule, but it’s not as good as it used to be. Perhaps there’s something extra I need to add to this rule to make it better.”

In the *satisficing anchored model*s, when a one-step rule falls below the satisficing threshold it is replaced. We take a similar approach to the second step in a two-step rule. Once a second step (exception) is generated both the *incremental RBx learner* and the *satisficing anchored* models can replace their exceptions when they drop below an *exception satisficing threshold*. This threshold is a separate parameter from the *base rule satisficing threshold* and is necessarily set to a higher value. During optimization of the *hierarchical incremental learner* a value of 80% was found to produce the optimal fit across all trial orders. For the *hierarchical* *satisficing anchored* model, values of 78% and 82% for the *base rule satisficing threshold* and *exception satisficing threshold* were found to provide the best fits.

Figure 4.7 depicts an example of a hierarchical base rule and exception. In the case of the *hierarchical satisficing anchored* model, if a simple rule is found that achieves greater than 78% accuracy and then drops below that threshold the first exception is generated. This only happens from Block 2 onwards since any rule that achieves such a high level of accuracy in block 1 is used throughout the block. Once the exception is generated, each time the model is incorrect it first evaluates the exception. If the exception is not good enough, it is replaced. Then the base rule is evaluated and if it is not good enough, it is replaced.

The process for generating an exception is identical to that of a base-rule except that both vertical and horizontal boundaries are generated. The steps are:

1) Both horizontal (orientation) and vertical (thickness) boundaries are created centered on the current stimulus. This divides the space into quadrants.
2) For each boundary a comparator is randomly selected. This isolates a specific quadrant.
3) That quadrant is assigned the label of the current stimulus.

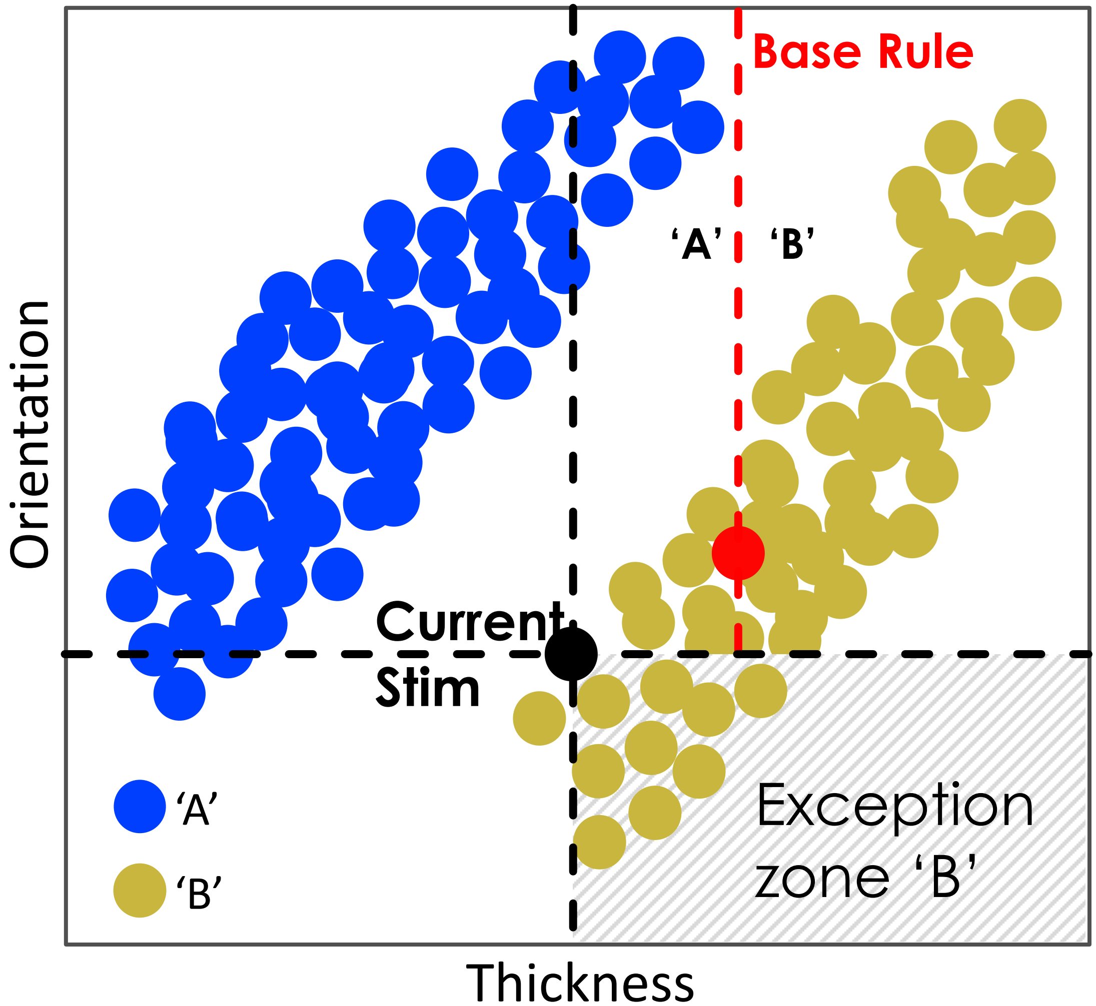


Figure 4.7: A cartoon representation of a hierarchical rule. If the model exceeds, then drops below the base-rule satisficing threshold, it generates and begins to consider quadrant exceptions to its base rules (thickness | orientation). If the exception accuracy falls below the exception satisficing threshold it is replaced.

In the case above, the red dashed line represents the current base-rule decision boundary anchored to the stimulus (red dot) used to generate it. In this case the rule is of the form:

IF stimulus_X ≤ Boundary: 'A' ELSE 'B'.

which is a fairly good rule. Notice that it only misses the bottom left of the orange (B) distribution including the current stimulus.

A new quadrant exception rule is generated with both vertical and horizontal decision boundaries anchored to the current stimulus. For each boundary a comparator is selected at random. In this case the vertical boundary has the ≥ operator and the horizontal boundary has the ≤ operator. Since the stimulus is a 'B', the quadrant denoted by the conjoint comparators is assigned the label 'B'. Thus, the rule is of the form:

$$\left.\begin{array}{c}\text{IF stimulus\_X} \geq \text{vertical\_boundary} \\ \& \\ \text{stimulus\_Y} \leq \text{horizontal\_boundary: 'B'}\end{array}\right\} \text{Exception}$$

ELSE: check base rule.

For example, let us consider a case where an existing RBx base-rule with a boundary of 0.7, comparator of ≤ and a label of 'A' (i.e. IF stimulus_X ≥ 0.7: 'B' ELSE: 'A') already exists and we want to generate an exception. If the current stimulus coordinates are (0.5, 0.3) and the label is 'B', vertical and horizontal boundaries are set at 0.5 and 0.3 respectively. This divides the space into four quadrants. Then for the vertical boundary, the comparator ≥ is randomly selected and for the horizontal boundary the comparator ≤ is randomly selected. This isolates the bottom-right quadrant. That quadrant is given the label 'B'. On the next trial the model evaluates the stimulus based on the following order:

1) Check the exception. If applicable, assign the label. If not:
2) Check the base rule.

If the stimulus on the next trial has the coordinates (0.6, 0.4) the order of operations would be:

1. Check the exception:
    a. Is 0.6 ≥ 0.5 (vertical boundary)? True / False
    b. IF True: is 0.4 ≤ 0.3 (horizontal boundary)? True / False
        i. IF True: Assign exception label: 'B'.
2. Check the base-rule (IF either 1 or 1a are False):
    a. IF 0.6 ≥ 0.7: 'B' ELSE: 'A'

In this case, the stimulus would be assigned the label 'A'. It fails the second exception check (1b) and thus is not in the appropriate quadrant. It then fails the base rule check (2a) and is thus assigned label 'A'.

To test whether hierarchical rules were a better description of participant behavior than a diagonal rule as implemented in PINNACLE 2.0, we added an incremental diagonal strategy to the hierarchical *satisficing anchored model with bias*. Since no participant ever articulated a diagonal strategy (i.e. IF thickness is greater than orientation: ‘B’ ELSE ‘A’) a model that predicted an abundance of diagonal strategy use could be taken as evidence for implicit strategy use rather than explicit II. In this model, if the base rule fell below the *base-rule satisficing threshold* the model was able to randomly select between thickness, orientation, or diagonal rules. Diagonal rules were restricted to one-step and had no exceptions attached to them. They were only modified while in use using an incremental learning style to approximate implicit learning similar to PINNACLE 2.0.

Criteria for Model Success

Our dataset includes several characteristics that are important to account for if we are to consider a model to be a good approximation of human performance. These criteria are:

1. The model must match participant performance on Block 1.
2. The model must show a significant reduction in accuracy in Block 2.
3. The model must be capable of producing variability in performance similar to participants in Blocks 3-5.

During the Bait phase (block 1), nearly all participants achieve a high degree of accuracy (m=0.83, sd = 0.10). In addition, they achieve this performance by quickly (approx. 15 trials) identifying an appropriate (but not perfect) rule and continuing to use it throughout the block. It is therefore important that any model we select be capable of this initial, rapid and successful performance.

In Switch phase (block 2), all participants experience a decrease in accuracy (m=0.56, sd=0.07) due to the changing nature of the task. Thus, it is important that the model not have an objectively accurate rule before block 2 (such as a diagonal rule) which is why it was not considered.

To capture both the variability and range of performance on our task, we split our data into quartiles based on Block 2 performance (Figure 4.8: top). Participant data split this way shows rank order stability among the quartiles and thus seems to accurately describe the variance in performance. Participants who do well in block 2 continue to do well throughout the experiment whereas if they did poorly, they do not seem to recover. In addition, we see that the top quartile of participants performs much better than the bottom 3 quartiles, and the bottom two quartiles do not exceed 60% accuracy.

Of particular interest are the top and bottom quartiles. The top quartile exhibits performance between 71% in block 2 and 77% in block 5 which demonstrates they discovered a reasonably successful strategy. In addition, these participants all verbally report an accurate rule that involves a rule and an exception (i.e. If thick then A, unless tilty, then B). The bottom quartile, on the other hand, shows consistent performance around chance with 47% in Block 2 and 57% in Block 5.

Any model that we believe accurately describes our participants must account for both the variability and range of performance seen in our data. That is, it must be capable of succeeding like our participants succeed and failing like our participants fail. We are interested in developing a model of a general strategy we believe our participants used during this task rather than the idiosyncratic strategies each individual may have applied. In this case distinguishing between a similarity-based strategy, an incremental rule modification strategy, and a satisficing-based rule replacement strategy. To test whether such a general strategy is capable of accounting for both the variability and range of our dataset we run each model on every trial order seen by participants. This produces a model trace for each participant for a total of 78 model generated behavioral traces. Since each trial order participants saw is different, this produces a variety of results per model much like participants. We then split the data into quartiles based on Block 2 performance, like the participant data.

To formally assess each model in the competitive model fitting process, we constructed an error formula that is sensitive to both the variability and range of our data according to:

$$cumulative\ error = \sum_{\substack{b=1,\\ q=1}}^{5,4} \left|P_{b,q} - M_{b,q}\right| + \sum_{b=1}^{5} \left|\left(P_{b,q1} - P_{b,q5}\right) - \left(M_{b,q1} - M_{b,q5}\right)\right| \tag{4.1}$$

The first part of the equation calculates the summed difference between average quartile performance within and across blocks. The second part of the equation calculates the difference between the top and bottom average quartile performance between the model and participants within and across blocks. Figure 4.8 (bottom) shows a cartoon of these measures.

Here we first calculate the summed distance between the model's performance across trial orders and participants. To that, we add the summed difference between the top and bottom quartiles for model's performance and participants on each block since this feature is so salient.

Thus, an error score of 0.5 is better than a score of 0.75 since it indicates that the former provided a closer account of both quartile differences, and the importance of the gap between the top and bottom quartiles.

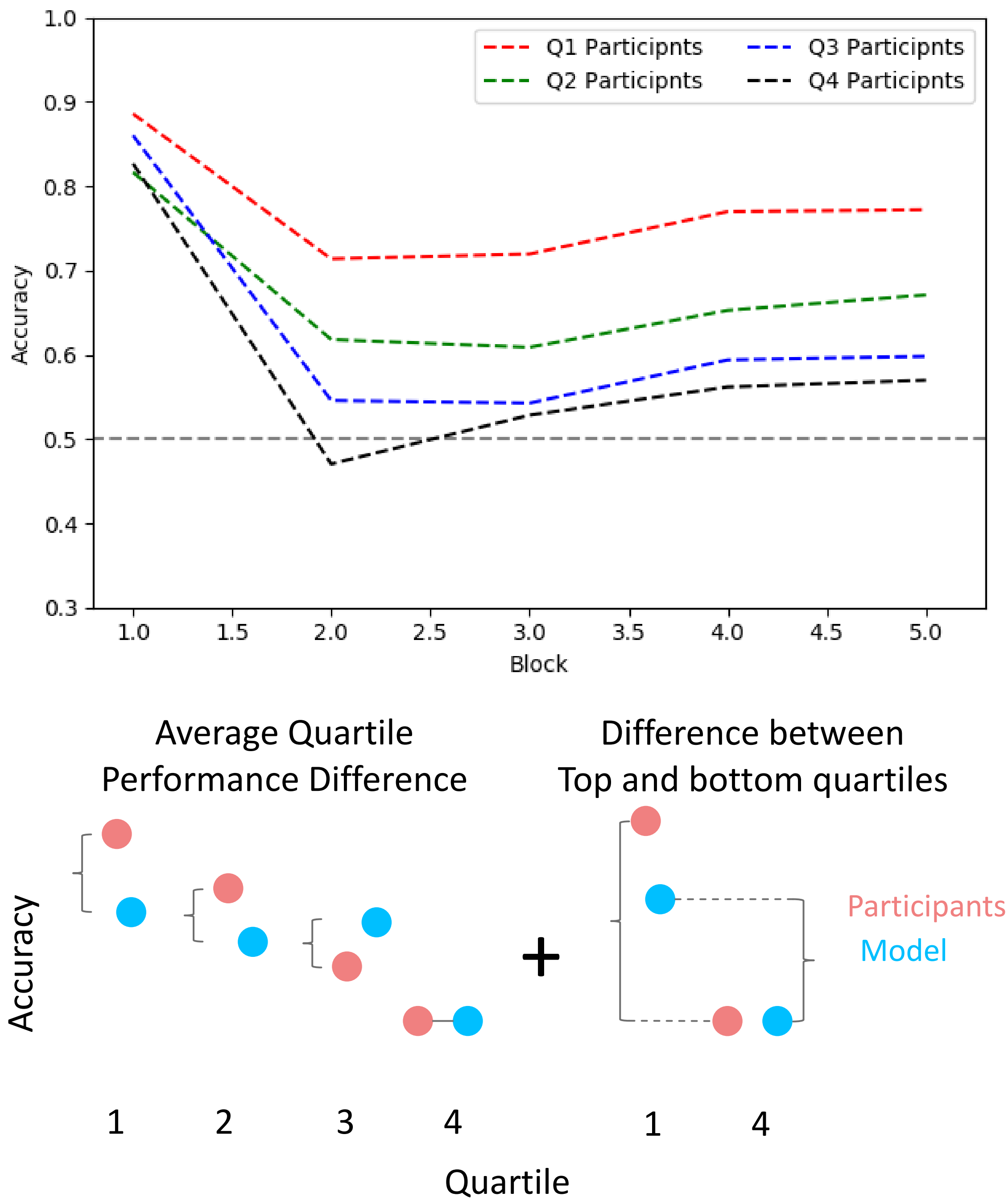


Figure 4.8. Top: Modeling dataset split into quartiles based on Block 2 performance. Participants ranked this way show rank order stability through the rest of the experiment.

Bottom: Depiction of the error function used to evaluate model fitness. Accuracy is split into quartiles based on block 2 performance. The difference between participant and model accuracy is calculated and aggregated for each block. In addition, the difference between the first and fourth quartiles for participants and the model is calculated. These error scores are then summed for a cumulative score indicating how different the two are.

Model Fitting

Given the potential complexity of modeling human behavior, we take parsimony as a guiding principle in our model development. We refrain from adding parameters and mechanisms to a model until we are satisfied that we cannot provide a better fit without them. In the following sections we consider the candidate models described earlier in an attempt to account for the current dataset as the experiment progresses. We begin with simple models that rely on a single rule without exceptions: exemplar, RB without learning, RB with incremental learning, satisficing anchored model with and without bias. At each stage we motivate the addition of a mechanism (i.e. learning, and bias respectively) based on limitations of the model without those mechanisms.

Importantly in competitive model fitting, models can only be evaluated with respect to other models under the same metrics. Since there is no objective measure of what a "good" model is, we rely on the metric to tell us which of our models is the best account of behavior. For instance, in some cases a model that does only 5% above chance is considered good since the alternatives all do worse. In other cases, a model that is 99.99% accurate is considered a bad model since the alternatives are all 99.99999% accurate. Thus, in the end we select the model with the lowest error score as the currently best fitting model, patiently awaiting the next better version to dethrone it.

The Bait: Block 1

The first criterion for our models is matching Bait phase performance. Here participants quickly discover a simple strategy that produces a high level of accuracy as well as a wide range of performance. An appropriate model should also be capable of this performance. Figure 4.9 depicts boxplots for performance on block 1 for both participants, and the first five models (those that do not include exceptions).

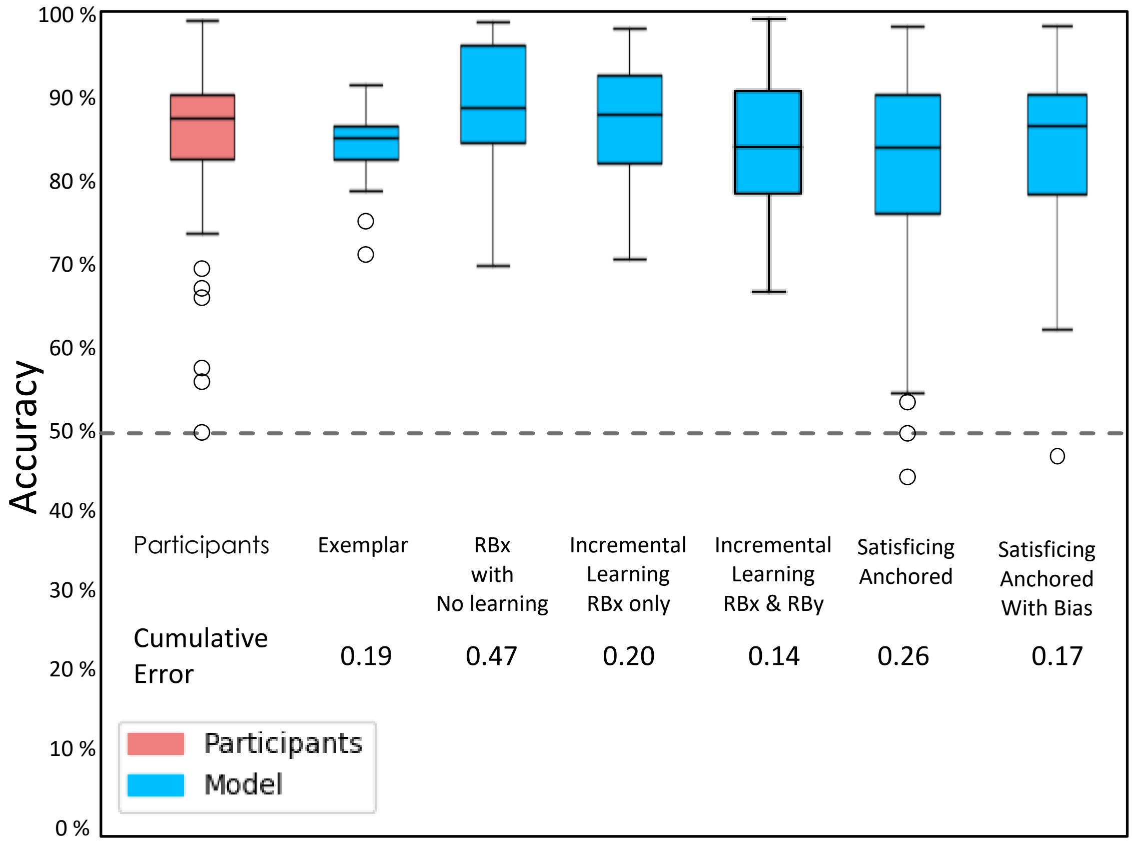


Figure 4.9: Boxplots of block 1 performance across six different models (blue) compared to participants (red). Cumulative error is reported below each boxplot indicating the discrepancy between participant and model performance on Block 1.

All models perform similarly and predict behavior somewhat in line with participants. The exemplar model predicts very consistent behavior. The non-learning RBx only model predicts a range of performance but tends to over-perform. Both satisficing models predict the largest range of behavior and provide a reasonable account of participant performance.

The parallel *incremental learning with RBx & RBy* model has the lowest error score and this seems most likely to describe what participants did. Note that its performance is similar to the sequential incremental RBx only learner and differs in its ability to produce lower accuracy performance due to the RBy strategy.

Exemplar Model

The exemplar model calculates the distance between novel stimuli and each category of previously seen stimuli as a proxy for perceptual similarity. It assigns the stimulus to the nearest category. In Block 1, the exemplar model provides a reasonably good fit to participant performance. It is capable of rapidly increasing its performance in line with participants and does so across all trial orders. Compared to participants, the average performances are quite similar, though the exemplar model has a more restricted range compared to participants because its performance only depends on the particular trial order it encounters with no additional sources of variance. The cumulative error for this model on block 1 was 0.19. Despite not accounting for the very best and worst performers, the exemplar model provides a reasonable account of behavior for most participants on this block.

Rule Based Models

Basic RB X

The next model we consider is restricted to a single rule, based on thickness, and is unable to modify its boundary. The boundary is initialized between the first ‘A’ and first ‘B’ stimuli it sees and responds randomly before that. All stimuli to the left of the boundary are ‘A’s and those to the right are ‘B’s. While even this simple model achieves high accuracy, it tends to over-perform compared to participants. The cumulative error for this model on block 1 was 0.47. Since perseverating on the initial estimate of the boundary seems to produce better performance on average than our participants, it is unlikely they adopted this strategy.

Incremental Learning - RBx

The next model is one that can modify its decision boundary incrementally in response to feedback. When incorrect, the boundary is incremented by a fixed learning rate in the direction that would have produced the correct answer. Both the range and central tendency of this model's performance is more in line with participant performance. The cumulative error score for this model on block 1 was 0.20 which was a significant improvement over the non-learning model. It seems implausible that people strictly insist on optimizing a single rule rather than searching for new ones. Nevertheless, we see that even just a simple model produces a fair account of participant behavior which is what makes these decision bound theory models so attractive and important to rigorously test for validity.

Incremental Learning – RBx & RBy

The next model adds an additional strategy to the incremental learner in the form of a rule based on orientation. Each rule predicts category membership in the form of a *p(A)* value, and the most confident rule is selected to guide behavior. When correct, the rule becomes more confident and thus more likely to drive behavior in the future. When incorrect, the rule loses confidence and its boundary is incremented in the direction that would have produced the correct answer. This model performs very similarly to the version with only RBx strategies unsurprisingly as in this block it ends up relying almost exclusively on that strategy. The cumulative error for this model on block 1 was 0.14. The stimuli shown to participants vary on just two dimensions, both of which are represented by the model. It seems reasonable to presume that participants tested both of these strategies before converging on the appropriate one, in this case thickness.

Satisficing Anchored Model

The next model shifts the focus from modifying a set of fixed rules to generating and evaluating different rules. It considers strategies based on both thickness (RBx) and orientation (RBy) and is capable of generating four types of rules: RBx with ‘A’s on the right or left of the boundary, and RBy with ‘A’s above or below the boundary. The model uses a rule until it drops below a satisficing threshold at which point it replaces it with a new one. The cumulative error for this model on block 1 was 0.26 and importantly it produced a much wider range of behavior than previous models. That said, the abundance of underperformance means that it fails too often to find an appropriate rule in the Bait phase. It is therefore less likely that participants adopted a strategy of an unconstrained and unbiased search for an appropriate strategy.

Satisficing Anchored with Bias

The final model we consider during the Bait phase is identical to the satisficing anchored model but has a preference towards thickness rules. This bias was added to both better reflect what we know about apparent participant preference for thickness over orientation rules as well as to allow it to converge on a successful strategy more consistently in the Bait phase. The cumulative error for this model on block 1 was 0.17. It produces both an appropriate variability as well as range when compared to participants though if anything it does have a slight tendency to underperform. This model represents a plausible alternative account to the incremental RBx & RBy model of how participants navigated this initial phase.

Interim Summary

During the Bait phase the differences between the exemplar, parallel incremental RBx & RBy, and sequential satisficing anchored with bias models were quite small. Each provides a plausible account of participant behavior. Since even the worst performing model – Basic single-step RBx with no learning – is still in line with participant behavior, we retain all models as we consider the second criterion – matching behavior during the Switch phase.

The Switch (Block 2)

The second criterion for our models is matching performance during the Switch phase. Here participants are confronted with a sudden change in the experiment where we begin sampling from never-before-seen areas of the category space. These regions specifically disconfirm the previously successful thickness rule and thus pose a challenge to our participants. During this phase accuracy drops as participants adjust to the change and begin exploring various alternative strategies. An appropriate model should also exhibit this profile of performance. Figure 4.10 depicts boxplots for performance on blocks 1 and 2 for participants and block 2 performance for each model we considered.

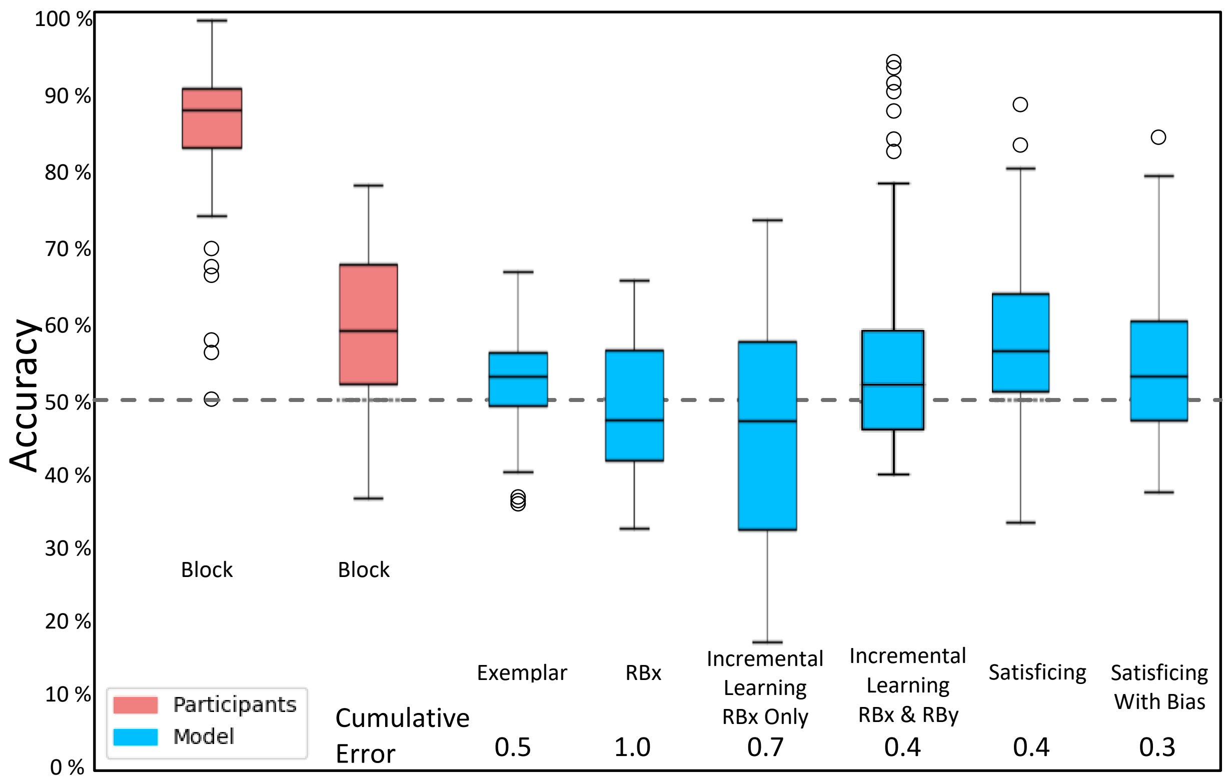


Figure 4.10: Boxplots of block 2 performance across six different models (blue) compared to participants (red). Cumulative error (block 1 + 2) is reported below each boxplot indicating the discrepancy between participant and model performance on Blocks 1 & 2.

All models here tend to under-perform compared to participants indicating that participants are doing something more sophisticated than the current set of models.

Notably the single step incremental learning RBx only model predicts a wide range of below chance performance suggesting this is not what participants were doing. The parallel incremental learning RBx & RBy model tends to over-perform suggesting that though successful, participants did not seem to adopt this strategy.

The satisficing anchored model with bias had the lowest cumulative error across all models indicating it likely represents the closest account of what participants are doing.

Exemplar Model

During the Bait phase, the exemplar model provided a reasonably good fit to participant performance. This was because it is capable of rapidly forming an accurate representation of the observed categories. In the Switch phase, we see its performance crash to hover around chance with a moderate range. Of note, it on average under-performs compared to participants and its range of possible performance is smaller than participants. That is to say, while it can fail like participants fail during the Switch, it cannot succeed like participants succeed. This is because the model depends on accumulating enough exemplars such that the average distance between new stimuli in the disconfirming areas and previously seen stimuli from the same category become smaller. Initially this is not the case, since a new stimulus might have thick lines but belong to category ‘B’ and thus be closer in space to the ‘A’ category. Until enough exemplars of these thick ‘B’s are encountered the model mistakenly attributes them to category ‘A’ (and vice versa for new ‘A’ stimuli with thin lines). If we wanted the model to do better on the switch, we would need to restrict the number of previous exemplars it has access to such that those seen in the Bait phase would have been effectively “forgotten”. While this solution would modify the central tendency of the model’s performance it would not increase its range of possible performance. Our participants perform both successfully and poorly while the model would be capable of only one or the other. The cumulative error for this model on blocks 1-2 was 0.53.

Rule Based Model

Basic RB X

The basic single-step RBx model predicts that participants should perform on average below chance during the Switch phase. This is because most stimuli shown during this phase specifically do not conform to a RBx rule though enough do to maintain near chance performance. The fact that this model cannot respond to feedback actually protects it from doing worse, as we will see with the next model. The cumulative error for this model on blocks 1-2 was 1.08. While this model predicted reasonable behavior during the Bait phase, its inability to consider an orientation rule leads it to predict worse behavior than participants display. Further, considering that it cannot learn is grounds for ruling out this model as a plausible account of how participants navigated this task. Because of this, we will no longer consider this model as we move forward.

Incremental Learning - RBx

The single-step incremental learning model displays the widest range of performances across all the models we consider. At its worst, its accuracy is as low as 15%, and at best around 70% depending on the trial order and parameter values. Given that in many cases the model is confronted with an abundance of stimuli that do not conform to a thickness rule, the model thrashes between boundary values, leading to poor performance. While the model is capable of performing as poorly (and more so) as participants, it is not capable of performing as well as the best performers due to being restricted to thickness rules. The cumulative error for this model on blocks 1-2 was 0.7. For similar reasons to the static RBx model, we will no longer consider this model as we move forward.

Incremental Learning – RBx & RBy

The parallel incremental learner that uses both RBx and RBy strategies does better than both variants of the single-step RBx only model since it is able to account for at least some of the stimuli by shifting to a RBy strategy. This is especially true for stimulus orders from experiment 3 in which the disconfirming zones are sampled at a 3:1 ratio during this phase. Since these stimuli are largely accounted for by an orientation rule this strategy leads to higher performance but misses stimuli from the areas seen in Bait phase. We thus see an increased range of possible performances predicted by this model. The cumulative error for this model on blocks 1-2 was 0.40. Though the model still predicts that participants should be performing worse than they are, the increased range of performances suggests this model better approximates participants during this phase.

Satisficing Anchored Model

The single-step satisficing anchored model performs similarly to the parallel *incremental RBx & RBy* model as it too is able to explore both thickness and orientation rules. It predicts a slightly over-performs compared to participants as they explore a wide range of strategies. In this phase, the model cannot find a rule that consistently meets its satisficing threshold because of the asymmetric stimulus sampling. This therefore predicts that participants are constantly searching for rules during this phase since no rule seems to account for all the stimuli they see. The cumulative error for this model on blocks 1-2 was 0.42. While it does provide a decent account of behavior during this phase, the model's inability to consistently match participant performance in the Bait phase led us to focus on the variant that includes a bias towards thickness rules. As such, this version of the satisficing anchored model will no longer be considered.

Satisficing Anchored Model with Bias

The single-step satisficing anchor model with a bias towards thickness rules provides the best fit to our data across the Bait, and Switch phases. While it tends to under-perform on average, it predicts the wide range of behavior seen in participants. The difference in fit values between the no-bias and bias versions of this model suggest that participants tend to retain a preference for thickness rules even though they are presented with an abundance of evidence to the contrary. This is supported by prior data as well. In designing experiment 3, the reason we elected to sample the disconfirming areas so heavily (3:1 in the Switch phase, then 2:1 in the Learning phase) was because prior small sample pilot studies showed that participants perseverated on the thickness rule from the Bait phase, apparently electing to ignore or discount evidence that it was no longer a viable strategy. The cumulative error for this model on blocks 1-2 was 0.35. The error low score and wider range produced by this model suggests that it provides the best account of the data thus far.

Interim Summary

During the Switch phase, several models demonstrated an inability to match key metrics of participant performance and were thus dropped from further consideration. Both the non-learning and incremental learning versions of the RBx only model predicted an abundance of below chance performance during the Switch phase due to their restricted strategy set. Though they made reasonable predictions during the Bait phase, it is unrealistic to think that participants only considered a single strategy during the entire experiment. We therefore ruled them out as plausible accounts of participants behavior. The parallel incremental learning with RBx & RBy model provided a good account of behavior by adaptively alternating between the two strategies. Interestingly, on several trial orders, this produced better performance than participants suggesting that participants adopted a less effective strategy during this phase. The single-step satisficing anchor model with no bias, while providing a decent account of the data, was outdone by the variant with bias both in the Bait, and Switch phases. In addition to that, the variant with bias provides a closer account of participant behavior and verbal report. We therefore elected to focus our attention on the variant with bias. This model predicts that participants failed to find a good enough rule during this phase and therefore constantly generated and replaced a variety of rules. The fact that this model does not predict an abundance of overperformance during this phase, like the parallel incremental learning RBx & RBy model, suggests this is a closer account of what participants were likely doing.

The Learning Phase (Blocks 3-5)

The third criterion for our models is matching the broad range of performance profiles observed in the Learning phase. Here participants seem to have adopted a wide range of strategies to varying degrees of success. The important aspect of these data for our modeling is creating a model whose mechanisms allow it to match behavior that range from the worst to the best learners. To capture participant success during this phase, PINNACLE 2.0 included a diagonal strategy meant to represent a two-dimensional explicit rule, or implicit strategy use. In experiment 5 few, if any, participants zeroed in on a strategy as good as a diagonal since that would lead to perfect performance. To further support this claim, we included a diagonal strategy in the *Satisficing Anchored Model with Bias and Exceptions* model. Modeling results predict that participants considered this strategy only 10% of the time. We discuss the results in the relevant section. Figure 4.11 depicts boxplots for performance on blocks 1, 2 and 5 for participants and block 5 performance for each model we considered in the Learning phase.

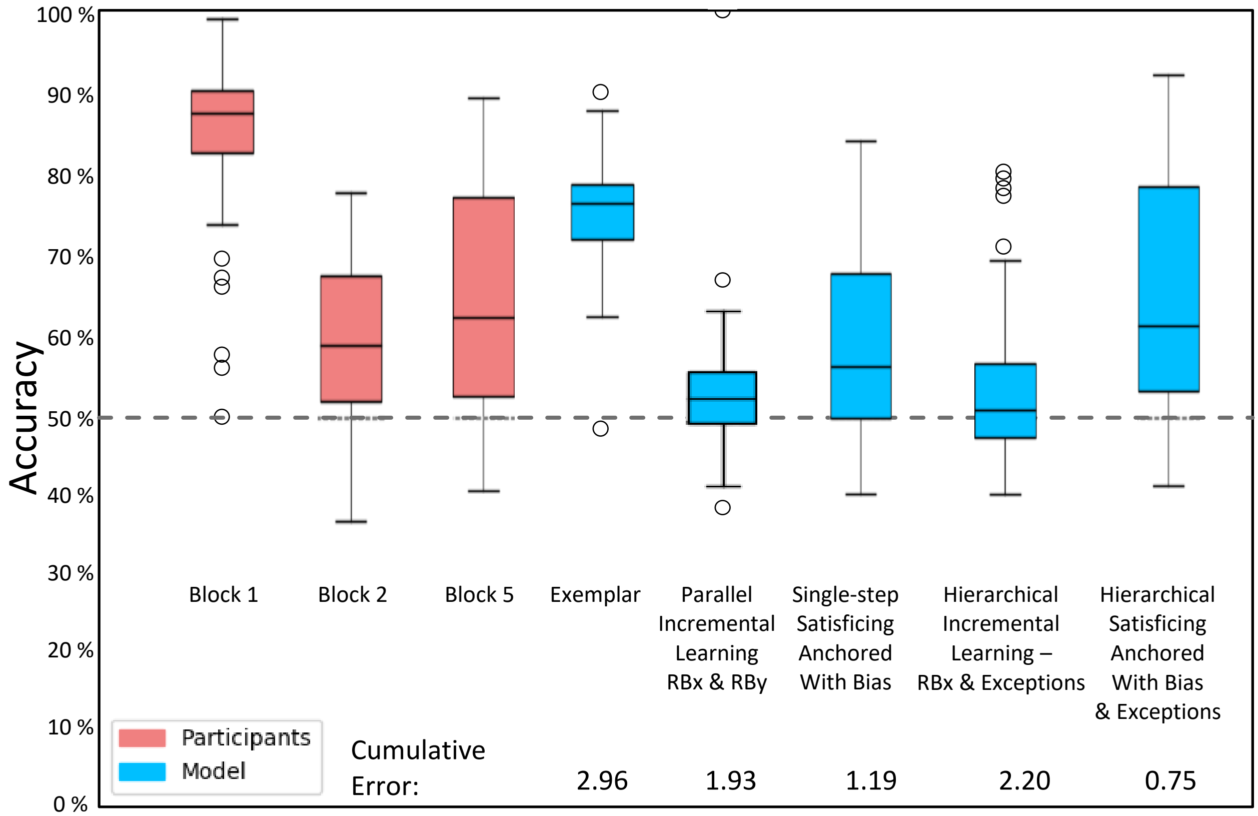


Figure 4.11: Boxplots of block 5 performance across five different models (blue) compared to participants (red). Cumulative error (blocks 1 - 5) is reported below each boxplot indicating the discrepancy between participant and model performance across blocks 1 - 5.

The exemplar model cannot account for the range of participant behavior and is restricted to matching the best learners. The parallel single-step incremental and hierarchical incremental learners also cannot account for the range of behaviors and are restricted to matching the worst learners.

The two satisficing anchored models both approximate the range and variability of performance seen in participants. Importantly, the hierarchical satisficing anchored model is able to match both the worst and best learners while the single-step version cannot achieve do as well as the best participants.

Exemplar Model

As early as block 3 (not shown) the exemplar model predicts that participants ought to perform at the level shown in Figure 4.9 and remain stable throughout the rest of the experiment. This is because by block 3 the model predicts that participants would have already experienced enough exemplars to recover from the Switch phase and perform consistently well. In fact, the model predicts performance in line with the best participants but cannot produce the range of performances that participants do. This is a fundamental limitation of the exemplar model. Versions that match the top performers cannot fail and versions that match the worst performers cannot succeed, even on block 1. The cumulative error across all blocks was 2.96, the highest of any model we considered. As such, it is highly unlikely that participants are using a decision-making process analogous to a similarity-based judgments. If they were, the models predict we would see very different profiles of behavior.

Rule Based Models

Incremental Learning – RBx & RBy

By the end of the experiment, the parallel incremental learning model capable of using both RBx & RBy strategies does not do much better than it did in the Switch phase. We see its range has reduced, indicating that both strategies have narrowed in on their optimal solutions and simply trade-off between them, but it still tends to produce performance around chance. Critically, it is incapable of matching the best learners though in the Switch phase it found some success by more consistently favoring an orientation rule. The cumulative error across all blocks was 1.93. While this model can fail like our participants it cannot succeed like them, the opposite problem of the exemplar model.

Satisficing Anchored with Bias

The single-step satisficing anchor model with a bias towards thickness rules predicts a range of behavior more indicative of participant behavior. It produces a wide range of performance and captures low and moderately performing participants. However, it struggles to match the best learners. This suggests that no thickness or horizontal rule by itself is sufficient to produce such learning (also evidenced by the parallel *incremental – RBx & RBy* model which implies that participants are using a more complex strategy). The cumulative error across all blocks was 1.19. While this model produces a relatively low error score, it cannot succeed the way participants do indicating that it is an incomplete account of human behavior. As such we turn to more complex models to create a better theory of what people are doing.

Hierarchical Models

To address the limitation of the single step *satisficing anchored* model with bias and bring the model more in line with participant verbal report, we developed a hierarchical version that incorporates quadrant exceptions. A diagonal strategy was also incorporated to test whether the observed behavior was more in line with a hierarchical rule strategy, or a single-step diagonal one. Additionally, to provide an alternative to the hierarchical satisficing framework we developed a hierarchical *incremental RBx* model as well.

Incremental Learning - RBx with Quadrant Exceptions

The first model capable of generating and testing exceptions we consider is the hierarchical incremental learner that only uses an RBx base-rule. We elected to implement exceptions within this simpler version of the incremental learner rather than the RBx & RBy version due to the added complexity inherent tracking the vast number of possible two-step rules, the *p(A)* calculations, and the decision-making associated with such a system. Moreover, the added benefit of the RBy strategy is mostly moot by the end of learning and removing it does not significantly change its performance in Block 4-5. While performance of the hierarchical model is similar to the single step version, we see an increased range of performance due to these exceptions. The cumulative error across all blocks was 2.2. Importantly, no instance of the model performed as well as the best participant indicating this is not a suitable description of what participants are doing.

Satisficing Anchor with Bias, Diagonal, and Quadrant Exceptions

The final model we considered was the hierarchical *satisficing anchored model* that had access to thickness, orientation, and diagonal strategies, a bias towards thickness, and generated quadrant exceptions to its base-rule. Though it has access to an objectively perfect strategy (diagonal), the parameters that produce the best fit predict that participants only considered diagonal strategies 10% of the time and maintained a preference towards thickness rules throughout the experiment. Looking at the boxplot performance in Figure 4.9 we immediately notice that the model predicts behavior that is similar across the variability, range, and central tendency to participants. The model is capable of producing behavior similar to both the worst and best performers. The cumulative error across all blocks was 0.75. This model provides the best fit out of all the models we tested and does so with just 3 parameters (*RBx bias*, *exception satisficing threshold*, *base rule satisficing threshold*; half the number that PINNACLE 2.0 had). The appeal of the model is not just due to its performance, but also because of its simplicity and intuitive operation. People explore different rules of varying complexity, using them until they are deemed not "good enough", and then they are replaced. Despite the possibility of happening upon an explicit diagonal strategy, or switching to rely on implicit knowledge, participants seem to persist in using two-step explicit strategies.

Summary

In this chapter we presented eight models that attempted to account for participant behavior both conceptually and through specific mechanisms. We showed why both the exemplar and several incremental learning models – highly successful models of category learning –struggle to account for the data from Dynamic Cat. We introduced a new satisficing mechanism for decision-making we had not previously considered whose focus rests on rule generation and evaluation rather than modification of a fixed set of strategies. In addition to being a simpler model, this model provides both a better account of participant behavior and a more intuitive understanding of what they are likely doing.

Though the hierarchical *satisficing anchored with exceptions* model accounts for both the range and variability of the data, it is important to temper success with realism. Whenever we are confronted with such high variability in behavior, it is almost certainly the case that participants adopt a wide range of behaviors absent in our models. We do not claim this is the best account of our data, only that it is the best account of our data so far. The fact that such a simple model seems to predict behavior similar to participants is certainly an encouraging start.

To further understand how our participants navigated this task we turn to the neuroimaging data collected during experiment 5. In the next chapter we describe results from preliminary neuroimaging analyses that identify neural correlates of successful vs. unsuccessful learning as well as how these results inform further development of the PINNACLE model.

Chapter 5: Neuroimaging Analysis

To navigate the complex and dynamic decision-making task in Dynamic Cat participants relied on a broad range of strategies, both successful and not. The work discussed thus far has focused on characterizing specific mechanisms that provide a plausible account of this behavior culminating in PINNACLE 2.1. A crucial insight from the modeling work was the importance of generating and evaluating a wide variety of rules during the decision-making process for successful learning. Rather than perseverate on a specific strategy or explore a restricted set, participants displayed flexible behavior that was responsive to task demands which we captured in the satisficing and rule-generation mechanisms. In this chapter we describe preliminary evidence of the neural correlates of these processes.

A critical moment in our experiment is the Switch phase (block 2) where participants are confronted with a significant and abrupt change in task demands. Participants had to adapt to maintain performance by searching for new appropriate strategies. This process began with the sudden change in the Switch phase and continued throughout the subsequent trials in the Learning phase. Such behavior requires engaging variety of functions including decision-making, evidence accumulation and integration, and working memory. The more these processes are engaged the more likely successful solutions to the problem are to be found. Indeed, we see that the top learners display more activity right dorsolateral prefrontal cortex (BA-9), right inferior parietal lobule and right precuneus, areas previously linked with these processes, across the Switch and Learning phases (Blocks 2-5) compared to the worst learners based on a median split. Additionally, we find evidence for differential processing of correct versus incorrect feedback among these groups during the Switch phase (block 2). Throughout this phase left dorsolateral prefrontal cortex (BA-8, BA-9), left precuneus, and left superior parietal lobule showed increased activity associated with incorrect trials for Learners compared to Non-Learners. Considering these results in light of PINNACLE 2.1 implies that learners were less likely to satisfice with their rules and were thus more likely to engage in the rule generation and replacement process. Doing so reliably led to increased chance of discovering better solutions and thus better outcomes.

Methods and Materials & Procedure

As described in chapter 4, thirty-five participants were recruited from the community and performed the Dynamic Cat task in experiment 5 while in a fMRI scanner. In sum, seven participants were excluded as described in experiment 5's Exclusions section.

Our neuroimaging protocol used a Siemens 3T fMRI scanner. During the experiment, we ran a T2*-weighted echo planar (EPI) scan covering 35 interleaved axial slices (TR = 2200 ms, TE = 21 ms, multiband = 2, voxel size = 2×2×2 mm) for 218 volumes in each of six scans. For anatomical localization, high-resolution, 3D MP-RAGE T1-weighted scans (voxel size = 1×1×1 mm; 128 axial slices) were collected for each participant following the functional runs.

All data were preprocessed using AFNI version 19.1.6. All anatomical scans were aligned and normalized to the MNI-152 atlas. EPI data were motion corrected, aligned to anatomical data and smoothed with a 6mm Gaussian kernel. A first level general linear model (GLM) analysis for each participant was run using the 3dDeconvolve function in AFNI. For all analyses each event was modeled as a 1.5s block beginning at trial onset. Second level group analyses were done using the 3dttest++ function in AFNI with two groups (top half vs bottom half of learners).

Results

To identify the neural correlates of successfully navigating the dynamic decision-making process we began by contrasting all trials for the top and bottom half of learners based on a median split during the Switch phase (block 2). Learners showed increased activity in right dorsolateral prefrontal cortex (BA-9; peak voxel MNI-coordinates: 36, 52, 8; 51 voxels) across all trials when compared with non-learners ($t=3.4$, $p<0.00$, minimal cluster size of 40 contiguous voxels ($320mm^3$)). Figure 5.1a shows activity for the top versus bottom learners on all trials in block 2.

A main prediction of PINNACLE 2.1 is that on incorrect trials, participants evaluate their rule and consider whether it should be replaced. If so, they generate a new rule and update the current one under consideration. Based on this prediction, we contrasted correct versus incorrect trials for learners versus non-learners in the Switch phase where this behavior was most likely to be observed. Four areas including left dorsolateral prefrontal cortex (BA-8 peak voxel MNI-coordinates: -52, 20, 34; 141 voxels, BA-9 peak voxel MNI-coordinates: -30, 54, 18; 66 voxels), left precuneus (peak voxel MNI-coordinates: -28, -50, 42; 71 voxels), and right superior parietal lobule (peak voxel MNI-coordinates: 26, -70, 54; 48 voxels) reflected increased activity for incorrect trials in learners ($t=3.75$, $p<0.001$, minimal cluster size of 40 contiguous voxels ($320mm^3$)). Figure 5.2 shows activation regions and ROI activations for these regions.

During the Switch phase, participants were confronted with the difficult task of discovering new adaptive strategies to the task. According to PINNACLE 2.1 this process involves several steps: 1) evaluate the current rule, 2) if necessary, reject the rule, 3) generate a new rule, 4) replace the current rule with the new one. In the human brain, these operations are plausibly associated with the regions showing increased activity just described. Dorsolateral prefrontal cortex has been identified in a variety of both decision-making and working memory tasks (DM: Barraclough et al., 2004; Van't Wout et al., 2005; Heekeren et al., 2008; Rangel, et al., 2008; Nomura & Reber, 2012. WM: Curtis & D'Esposito, 2003; Donahue & Lee, 2015; Cieslik et al., 2012). The precuneus has been associated with episodic memory tasks (Cabeza and Nyberg, 2000; Rugg and Henson, 2003) as well as memory retrieval (Buckner et al., 1996, Fletcher et al., 1998). Superior parietal lobule has been identified as active in sensorimotor integration and visuospatial processing and working memory (Wolpert, Goodbody & Husain, 1998; Corbetta et al., 1995; Koenings et al., 2009) tasks. The increased activity observed in these areas associated with incorrect trials in learners suggests that these areas likely reflect activity related to processes analogous to rule evaluation, generation, and replacement. Simply put, learners are less satisfied with their performance during the Switch and are therefore less likely to satisfice (consider their rule 'good enough'). They are therefore more likely to generate and evaluate rules when wrong than learners thus enabling them to eventually find a successful one.

To further characterize differences between successful learners and non-learners we contrasted all trials from both the Switch and Learning phases (blocks 2-5). The right inferior parietal lobule (peak voxel MNI-coordinates: 48, -40, 46; 82 voxels) and right precuneus (peak voxel MNI-coordinates: 8, -66, 38; 67 voxels) both showed increased activity for learners compared to non-learners ($t=3.4$, $p<0.001$, cluster size of 40 contiguous voxels ($320mm^3$)). These areas have been found in previous studies of category learning (Forstmann et al., 2008; Braunlich, Gomez-Lavin, Seger, 2015) and perceptual decision-making (Wenzlaff et al., 2011) indicating that even after the critical Switch phase, learners do better by engaging areas that allow them to integrate information about the stimuli in meaningful ways and use that information to guide better rule generation resulting in successful categorization. Table 5.1 summarizes the neuroimaging findings.

Regions	Side	*t*	*voxels*	*x*	*y*	*z*
Block 2: Learners > Non-Learners, All Trials						
DLPFC – BA 9	R	3.4	51	36	52	8
Block 2: Learners > Non-Learners, Incorrect > Correct						
DLPFC – BA 8	L	3.75	126	-52	20	34
DLPFC – BA 9	L	3.75	66	-30	54	18
Precuneus	L	3.75	71	-28	-50	42
Superior Parietal Lobule	R	3.75	48	26	-70	54
Blocks 2-5: Learners > Non-Learners, All Trials						
Inferior Parietal Lobule	R	3.4	82	48	-40	46
Precuneus	R	3.4	67	8	-66	38

Table 5.1

Areas of activity identified across conditions. *t* indicates the t-value for two-sampled, two-tailed t-test with a top vs bottom median split of learners based on block 2 performance. Voxels indicates the size of the cluster of activity after surviving a minimal cluster size of 40 contiguous voxels equating 320mm^3 of BOLD signal. XYZ coordinates denote peak voxel activity in MNI space.

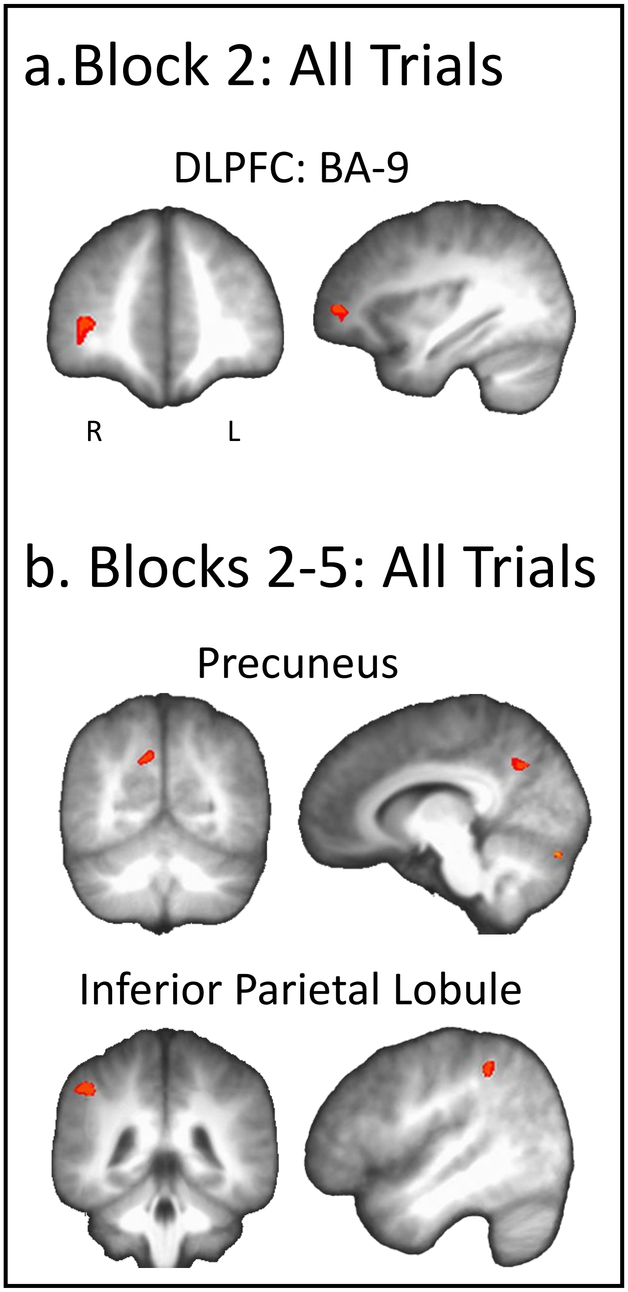


Figure 5.1:

a. Neural activity associated with all trials for the top half of participants based on a median split of performance in block 2. Learners show greater activity in right dorsolateral prefrontal cortex (area BA-9) which has been previously shown to be associated with decision-making and working memory.

b. Neural activity associated with all trials for the top half of participants based on a median split of performance in block 2. Learners show increased activity in right precuneus and right inferior parietal lobule.

These areas have been found in previous studies of category learning and perceptual decision-making indicating that even after the critical Switch phase, learners do better by engaging areas that allow them to integrate information about the stimuli in meaningful ways and use that information to guide better rule generation resulting in successful categorization.

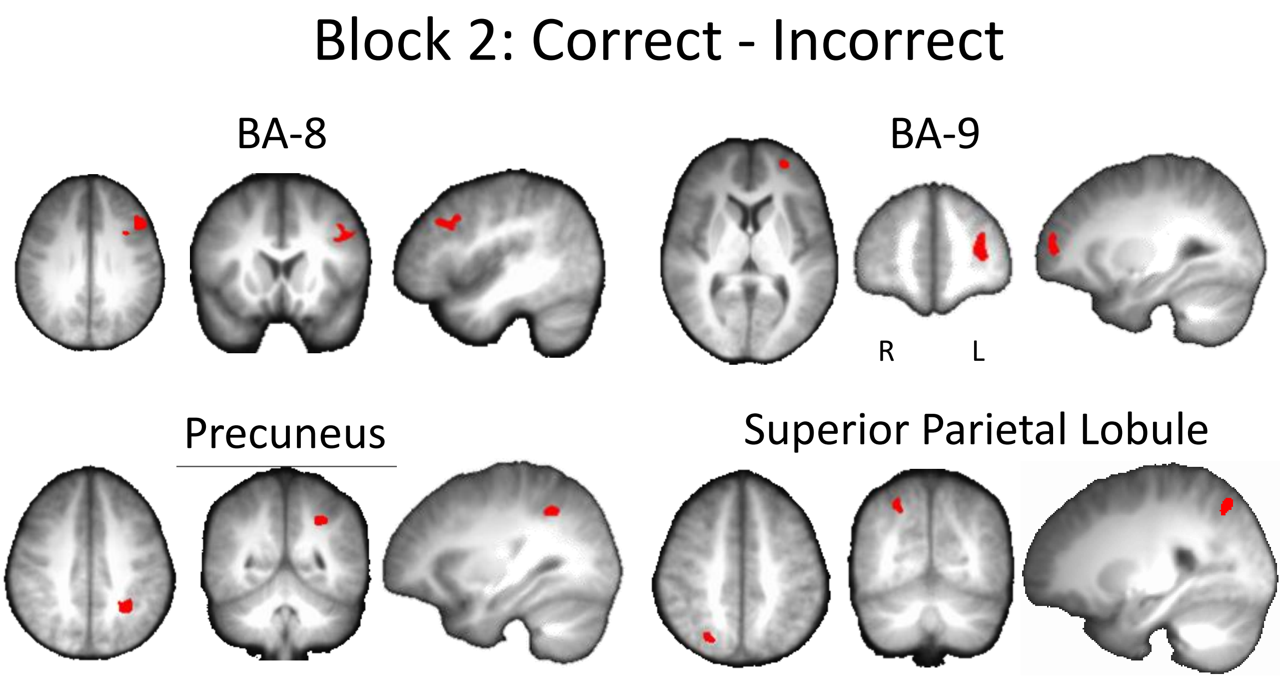
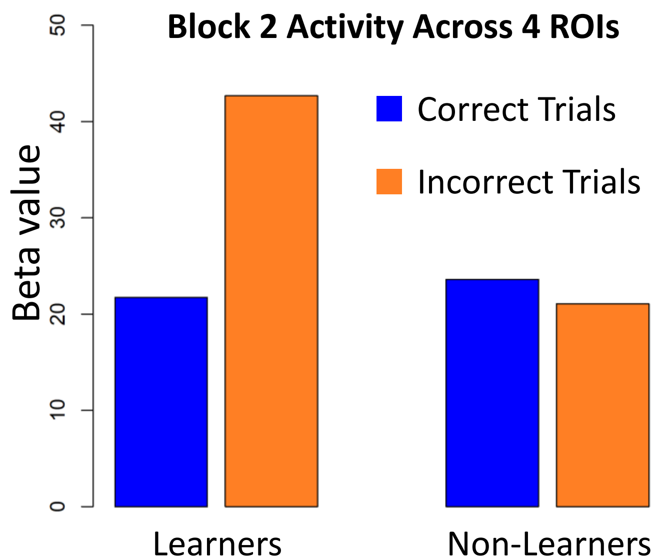


Figure 5.2:

Block 2 Correct vs Incorrect trials

TOP: Activations for 4 ROIs more active for Incorrect trials in Learners versus Non-learners.

Bottom: Average beta values across all ROIs. Individual ROIs all show the same pattern of increased activity for Incorrect trials in Learners compared to the other conditions.



Summary

Taken together these results suggest that people who did better on this task engaged more often in a process of assessing their current rule and if necessary, replacing it. This was most pronounced in response to incorrect feedback during block 2. Activity in the dorsolateral prefrontal cortex suggests an increased working-memory load associated with generating a new rule, as well as processes associated with goal-directed decision-making such as comparative evaluations (i.e. “Is this rule good enough?”). Precuneus involvement suggests a process of accumulating and updating relevant information as people try to make sense of all they have experienced thus far, as well as memory retrieval as they access and modify current rule representation. Superior and inferior parietal lobules suggest that people may be reevaluating stimuli (both current and previously seen) in light of new information (i.e. “perhaps the rules have changed and all ‘A’s are now ‘B’s!”).

The observed results also help rule out accounts of behavior that do not predict differential processing between learners and non-learners as well as between correct and incorrect trials. For example, the exemplar model predicts the same activity on each trial regardless of feedback. The stimulus is stored in long-term memory with a label based on the response and feedback. There is no additional work to be done whether correct or incorrect. As we discuss in the following section, these data also rule out the current implementation of PINNACLE 2.1.

PINNACLE 2.1a

The differences in neural activity between learners and non-learners in block 2 indicate that learners responded differently to incorrect trials during this task. This pattern of results poses a challenge to PINNACLE 2.1. The current model predicts that both learners and non-learners have the same *base-rule* and *exception satisficing thresholds* and thus should be equally likely to engage in the rule evaluation, generation, and replacement processes. What distinguishes the two groups is effectively luck; some people happen upon a successful rule while others keep searching. To account for these data, the model would need to systematically produce more instances of strategy switching in learners than non-learners. This is possible if learners have a higher *base-rule* and *exception satisficing thresholds* than non-learners. In that case, learners have higher standards as to when they satisfice. Non-learners, on the other hand, satisfice at a lower accuracy and thus do not engage in as much rule evaluation, generation and replacement.

We incorporated these requirements into PINNACLE 2.1a leading to a new variant that provides a better fit than PINNACLE 2.1 and offers a plausible account for both the observed behavior and neural activity. Figure 5.3 shows participant and model performance across all five blocks of Dynamic Cat. The cumulative error for the model is 0.62. Optimization identified a base-rule threshold value of 72% and an exception rule satisficing threshold value of 92% for learners while non-learners are assigned thresholds of 52% and 57% respectively. These thresholds align with the observed mean accuracies on blocks 2-5 for the top and bottom half of learners in experiment 5 (Learners: accuracy ranged from 46% to 91%; Non-Learners: accuracy ranged from 47% to 63% with one participant in that group achieving up to 85% accuracy on block 5).

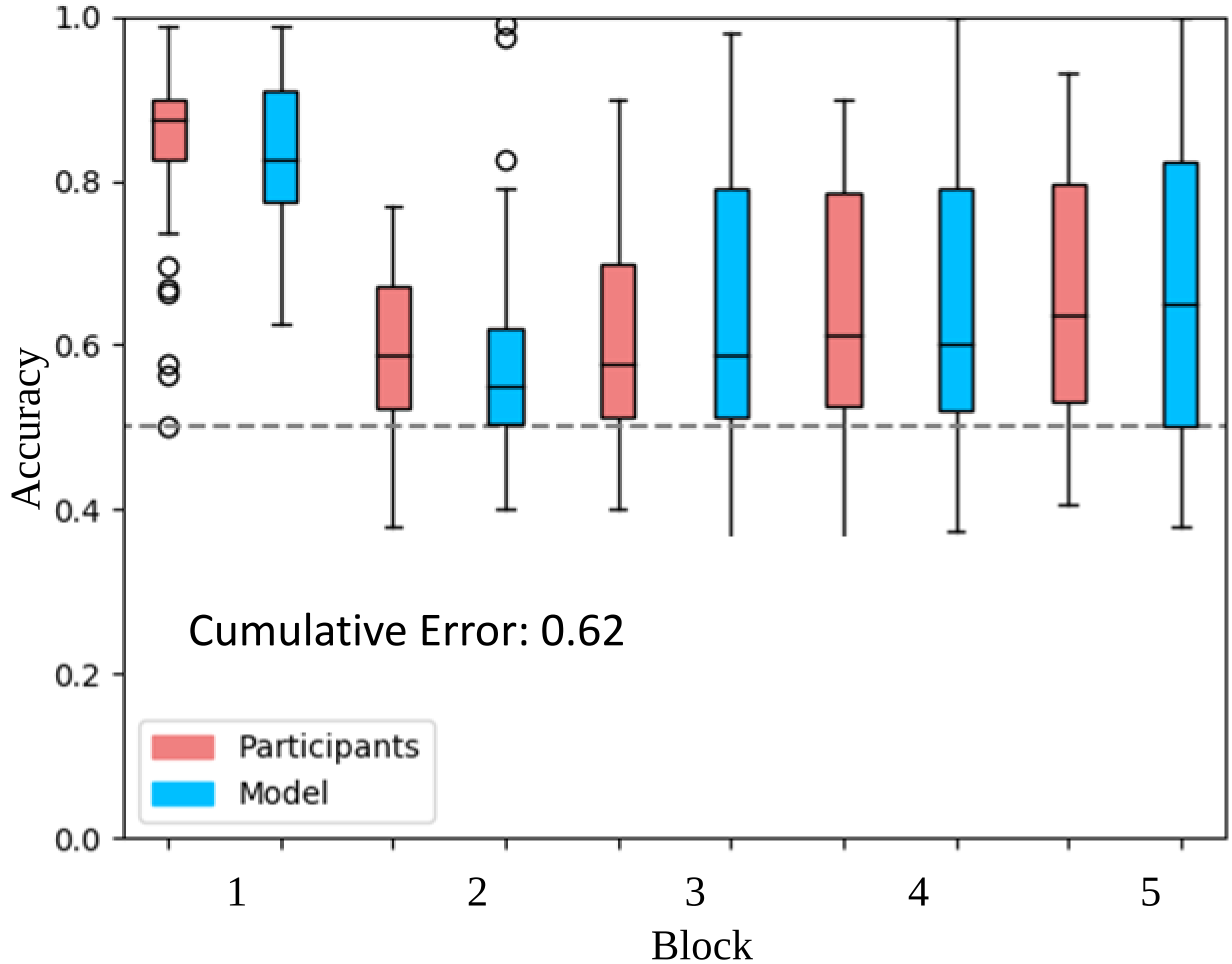


Figure 5.3:

Boxplots for blocks 1-5 of participants (red) and PINNACLE 2.1a (blue). In this model the top half of participants based on a median split of block 2 performance were assigned a higher base-rule and exception satisficing threshold than the bottom half. This leads to learners engaging in more rule evaluation, generation, and replacement (i.e. not satisficing) compared to non-learners who satisficed at a lower accuracy level. This behavior provides a plausible account for both all behavioral and neuroimaging data on the Dynamic Cat task.

Chapter 6: Discussion

In this study we challenged participants, across five experiments, to solve a deceptively complex decision-making task that forced them to dynamically explore a variety of strategies during category learning. Understanding these processes from both a psychological and neurological perspective is challenging due to the inability to directly measure them through introspection or observation. In many cases the outcome of a decision-making task is a button-press or other physical response from which we cannot infer the driving cognitive mechanisms. This ambiguity has led to a variety of different explanations for how people acquire categories such as exemplar and decision-bound theory. Do people make category judgements based on a similarity comparison, or by applying a rule? From just the single button press available to us it is impossible to tell, and as my advisor says: *"there's more than one way to skin a cat-egory"*. While a complete understanding of how humans navigate this problem will require much more work, our results help rule out several previously successful accounts of visual category learning and suggest a plausible framework for understanding the dynamic decision-making process.

Exemplar models based on the Nosofsky general context theory suggest that people categorize novel stimuli by comparing them to previously seen exemplars of each category and associating them with the more similar one. Thus, a chihuahua seen for the first time is more likely to be called a dog than a cat (or rat) even though it may share some similarities with cats (or rats). While successful at describing a range of category learning data, this framework was insufficient for describing how people performed the Dynamic Cat task. Exemplar models can be tuned in a variety of ways that allow them to match a range of learning rates reflective of successful or unsuccessful learners. However, participants performing the Dynamic Cat experiment showed profiles of being both successful and unsuccessful learners in the same trace as evidenced by the rapid initial success in the Bait phase followed by a range of behaviors from gradual increases in success to prolonged stable near-chance performance. Our modeling work rules out the possibility that participants primarily relied on this type of strategy. Were they to do so, the theory predicts that they would have recovered from the Switch phase far more rapidly than they did. Furthermore, the exemplar model learns by storing stimuli in long-term memory, to be retrieved the next time a comparison is needed. This operation happens regardless of whether the response elicited positive or negative feedback. Feedback is used to inform which category the stimulus should be associated with. This architecture thus predicts equal neural activity on correct and incorrect trials and would likely be less associated with areas involved in working-memory and evidence accumulation. The asymmetric neural activity on incorrect trials for learners we observed suggests a different process was at play.

Another equally successful theory of how people learn categories is the Ashby decision-bound theory. According to this theory, people use something analogous to a linear decision-boundary that partitions a perceptual space into two categories and acts as a criterion for evaluating stimulus category membership. In its most common formulation, a rule-based model maintains a single decision-boundary and modifies it based on feedback. Conceptually, this implies that a single rule is maintained and tweaked until it produces the best performance it can. While it does not seem very plausible that people settle on a particular rule (in this case one based on thickness) and endlessly tweak it in search of success it has nonetheless been a successful model of category learning. Results of attempting to fit such a model to our data shows that while this approach reasonably predicted participant performance in the Bait, and Switch phases, it was unable to produce the range of behaviors seen by the end of learning.

Specifically, it was unable to succeed like participants succeeded and produce high performance by the end of the experiment. Ruling out this hypothesis implies that participants were engaging in a more sophisticated strategy while navigating this task.

In considering more sophisticated accounts of participant behavior, we tested a parallel incremental learning model in which two one-dimensional rules (in this case thickness and orientation rules) compete for control of behavior on each trial. This model proved more successful than the single rule variant but was similarly unable to match the best learners in the dataset. This inability rules out the hypothesis that participants relied solely on simple one-dimensional rules based on the stimulus features that learn incrementally and compete with each other for behavior. Indeed, in post-experimental interviews participants reported attempting a wide range of strategies as well as more elaborate hierarchical ones though their ability to describe the exact nature of these rules were limited due in part to the nature of post-experimental reflection of strategy use.

Clearly, a more elaborate framework was necessary to understand how participants navigated this task. To that end, the PINNACLE architecture was designed based on our understanding of the systems neuroscience of multiple memory systems, reward processing, and decision-making. It provides a theoretical framework for understanding people's behavior using multiple competitive interacting strategies. According to the theory, for a given decision, people can rely on either explicit or implicit strategies and constantly incorporate feedback to improve those strategies. PINNACLE 1.0 was the first version of this model and included both explicit (vertical: thickness) and implicit (diagonal: thickness & orientation) representations with the ability to select between them based on relative confidence. It had previously been shown to account for learning of thickness- and diagonal-based categories. It was not, however, capable of accounting for the Dynamic Cat data primarily because the two rules it used were, in a sense, too good. In this paradigm, a thickness rule is appropriate for the Bait, and Switch phases while the diagonal rule was appropriate for the learning phase. This led to over-performance compared with participants as the model predicted that participants would naturally transition between one appropriate strategy to the other. As a note, this account has been shown to be generally accurate when learning diagonal categories: participants start with simple one-dimensional rules and gradually transition to more complex ones based on evidence. Ruling out this hypothesis means that even though just two simple rules were appropriate to accomplishing the task, most participants did not seem to use this strategy and instead were engaged in a broader and longer strategy search.

We therefore improved on the theory, in the form of PINNACLE 2.0, to address these shortcomings by allowing for more explicit strategies (thickness- and orientation-based rules) as well as an increased ability to switch among competing strategies. It also introduced the idea of replacing explicit rules rather than incrementally modifying them based on feedback and concept of anchoring these rules to the current stimulus. The anchoring approach seems far more plausible in terms of what people may actually be doing. Rather than thinking “Oh, the thickness rule should be 0.02 units less thick than it is now.” it is more akin to “Oh, this stimulus represents the boundary condition – everything thicker than this is an ‘A’.” This model showed initial success at providing a plausible account for how participants solved this task. Similar to PINNACLE 1.0 it predicted that participants began with simple explicit rules and then transitioned to more complex two-dimensional rules, some of which were inferred to be implicit. Unlike version 1.0, the updated model allowed for more transitions between one- and two-dimensional strategies before ultimately adopting the more successful two-dimensional strategy.

PINNACLE 2.0 was capable of matching Bait, and Switch phases as well as successful learners. However, it was still unable to account for how participants managed to stay at near chance performance for extended periods of time. This is primarily due to two factors: 1) both one- and two-dimensional rules learned on each trial, and 2) the two-dimensional diagonal boundary was an objectively optimal solution. Together, these two mechanisms ensured that over time the model would find an optimal solution and never remain at near-chance performance for long. Ruling this hypothesis out meant that what participants were doing extended beyond transitioning between one- and two-dimensional rules based on stimulus dimensions. While it provided a fair account of how participants could succeed in the task, it could not explain how participants failed since failure in this circumstance did not mean random guessing, and did not rely on just rapid rule thrashing.

Across all the rule-based accounts of human behavior described thus far, each one has assumed that whatever the strategy in use might be whether exemplar, or one and two-dimensional rules, they reflected the true category space. Thickness rules always assumed category ‘A’ was to the left of the boundary and ‘B’ to the right, diagonal boundaries always assigned stimuli above the diagonal to category ‘A’ and those below to category ‘B’ while exemplar models saved stimuli to the appropriate true category. However, there is no a priori reason to assume participants always mapped space this way. For example, it is entirely plausible that a participant might think that we swapped stimulus labels “Everything I called ‘A’ is now ‘B’ and vice versa.”

Under this formulation, participants are presumed to generate a variety of different rules, some objectively appropriate and some inappropriate to the task. Since participants have no way of knowing what the true category structure is, they guess. In cases where the rule seems to work, they keep using it and in cases where it does not, they replace it. Given such an approach, the decision-making solution, as well as the parallel structure implemented in PINNACLE 2.0 was no longer appropriate to arbitrating between different strategies. This was primarily due to the complexity associated with the number of potential options. To do so would require keeping track of every possible strategy and comparing across them on each trial; an unlikely account of what people do. To capture a simpler and more intuitive mechanism of rule arbitration we implemented a well-established satisficing framework in which participants use strategies that are ‘good enough’ based on a potentially latent internal satisficing threshold and replace ones that are not.

Essentially, the new version of PINNACLE, now 2.1, hypothesizes that participants begin the experiment by generating and using a simple one-dimensional rule. They used this rule so long as its cumulative accuracy stays above a given threshold. If the rule fails and drops below the satisficing threshold a new rule is generated to replace the old one. When generating a new rule, the model attempts to capture a state of uncertainty from the participant’s point of view. In these cases, whatever they thought might be happening has just been called into question. Perhaps the circumstances of the experiment have changed (again). This may mean that any previous information about the task could now be irrelevant or even misleading. The only thing participants can be sure of is the current stimulus. Thus, the new rule they generate must account for the current stimulus’s label and is therefore anchored to it. Practically, this means that when generating a vertical (thickness) rule, the boundary is set equal to the current stimulus’ thickness value (and similarly for orientation rules). This ensures that the current stimulus is categorized correctly by the new rule without necessarily making sense of any previously accumulated information.

A new rule can take on several forms: it could be based on either thickness, orientation, or a combination of both. It could swap which side of the boundary each category is on such that, for example, the left side of a vertical boundary could be either category ‘A’ or ‘B’. In addition, the model includes a limited hierarchical two step strategy in which quadrant exceptions are evaluated before linear decision boundaries. This capability was added based on participant verbal report. Another insight based on participant behavioral and interview data was that a bias towards thickness rules was necessary to account for the Bait phase. That is, people seem to be predisposed to focusing on thickness over orientation, an interesting topic for discussion in its own right. Taken together, this model predicts that people generate and test a variety of both objectively appropriate and inappropriate rules with varying degrees of complexity that make sense of the current stimulus until one is found that they deem ‘good enough’.

Finally, we turned to neuroimaging data to further understand how participants approached this complex task. Preliminary results of neuroimaging analyses found increased activity in bilateral dorsolateral prefrontal cortex, bilateral precuneus, and right superior and inferior parietal lobules for top half of learners compared to the bottom half. These results were associated with activity on all trials as well as on correct versus incorrect trials. These areas have previously been shown to be involved in working-memory, decision-making, evidence accumulation, and sensory integration. The observed results pose a challenge for PINNACLE 2.1 as they suggest that learners likely engaged in more rule evaluation, generation and replacement in response to incorrect trials than non-learners. This, however, is not predicted by the model since everyone shares the same satisficing thresholds and thus there should not be a systematic bias towards less satisficing in learners versus non-learners.

To make sense of these data in light of the success of PINNACLE 2.1, we incorporated this insight into our model leading to PINNACLE 2.1a which hypothesizes that learners have a more stringent threshold for what they consider to be a “good” rule compared to non-learners who may be willing to settle for lower rates of success. This new model therefore predicts more rule evaluation, generation and replacement following incorrect trials for learners than non-learners.

Conclusions

In this study we attempted to provide a plausible account of how people navigated a complex and dynamic decision-making task through behavioral, computational and neuroimaging methods. Developing a better understanding of this process is crucial to a broad range fields including education, medical, retail, and financial sectors as well and interpersonal relationships. Across a series of models, we ruled out several plausible hypotheses of how people navigated this process. Ultimately the best account of behavior thus far involved one in which people search for both simple and complex solutions that are *good enough* rather than *the best*. Different people have different standards of what good enough is with some people being satisfied with moderate performance while other strive towards perfection. Clearly, more work needs to be done to characterize and understand the rich and complicated nature of decision-making under uncertainty. The success of a relatively simple model in accounting for these processes is encouraging and provides a concrete step towards a fuller understanding of the human mind. Such a framework can be used to predict how people will behave in a broad range of circumstances to better inform optimal ways of making people’s lives better.

“*We have learned something about an important human activity.” – Herbert Simon*

WORKS CITED

Andersen, R. A., & Cui, H. (2009). Intention, action planning, and decision making in parietal-frontal circuits. *Neuron*, *63*(5), 568-583.

Ashby, F.G. (2013). Human category learning, neural basis. The encyclopedia of the mind. 130–134.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. Psychological Review, 105, 442-481.

Ashby, F. G., & Crossley, M. J. (2010). Interactions between declarative and procedural-learning categorization systems. *Neurobiology of learning and memory*, *94*(1), 1-12.

Ashby, F. G., & Ell, S. W. (2002). Single versus multiple systems of category learning: Reply to Nosofsky and Kruschke (2002). Psychonomic Bulletin & Review, 9(1), 175-180.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14(1), 33.

Ashby, F. G., & Maddox, W. T. (1992). Complex decision rules in categorization: Contrasting novice and experienced performance. Journal of Experimental Psychology: Human Perception and Performance, 18(1), 50.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. Annu. Rev. Psychol., 56, 149-178.

Ashby, F. G., & Valentin, V. V. (2017). Multiple systems of perceptual category learning: Theory and cognitive tests. In Handbook of Categorization in Cognitive Science (Second Edition) (pp. 157-188).

Ashby, F. G., & Waldron, E. M. (1999). On the nature of implicit categorization. Psychonomic Bulletin & Review, 6(3), 363-378.

Barbey, A. K., & Barsalou, L. W. (2009). Reasoning and problem solving: models. *Encyclopedia of neuroscience*, *8*, 35-43.

Braunlich, K., Gomez-Lavin, J. and Seger, C.A., 2015. Frontoparietal networks involved in categorization and item working memory. *NeuroImage*, *107*, pp.146-162.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. Psychological review, 108(3), 624.

Buckner, R. L., Raichle, M. E., Miezin, F. M., & Petersen, S. E. (1996). Functional anatomic studies of memory retrieval for auditory words and visual pictures. *Journal of Neuroscience*, *16*(19), 6219-6235.

Bunge, S. A., Hazeltine, E., Scanlon, M. D., Rosen, A. C., & Gabrieli, J. D. E. (2002). Dissociable contributions of prefrontal and parietal cortices to response selection. *Neuroimage*, *17*(3), 1562-1571.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of cognitive neuroscience*, *12*(1), 1-47.

Carter, C. S., & Van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. Cognitive, Affective, & Behavioral Neuroscience, 7(4), 367-379.

Corbetta, M., Shulman, G. L., Miezin, F. M., & Petersen, S. E. (1995). Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science*, *270*(5237), 802-805.

Cohen, J. D. (2008, April). The role of dopamine in reinforcement learning and cognitive control. In Biological Psychiatry (Vol. 63, No. 7, pp. 35S-35S). 360 PARK AVE SOUTH, NEW YORK, NY 10010-1710 USA: ELSEVIER SCIENCE INC.

Crossley, M. J., & Ashby, F. G. (2015). Procedural learning during declarative control. Journal of Experimental Psychology: Learning, Memory, and Cognition., 41, 1388-1403.

Curtis, C. E. (2006). Prefrontal and parietal contributions to spatial working memory. *Neuroscience*, *139*(1), 173-180.

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. Trends in cognitive sciences, 7(9), 415-423.

Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A., ... & Schlaggar, B. L. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences*, *104*(26), 11073-11078.

Duncan, J. An adaptive coding model of neural function in prefrontal cortex. Nat. Rev. Neurosci. 2001, 2, 820–829.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. Journal of Experimental Psychology: General, 127(2), 107.

Farrell, S., & Lewandowsky, S. (2010). Computational models as aids to better reasoning in psychology. *Current Directions in Psychological Science*, *19*(5), 329-335.

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., Von Cramon, D. Y., Ridderinkhof, K. R., & Wagenmakers, E. J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences*, *105*(45), 17538-17542.

Fletcher, P. C., Shallice, T., & Dolan, R. J. (1998). The functional roles of prefrontal cortex in episodic memory. I. Encoding. *Brain: a journal of neurology*, *121*(7), 1239-1248.

Fu, W. T., & Anderson, J. R. (2008). Solving the credit assignment problem: Explicit and implicit learning of action sequences with probabilistic outcomes. Psychological research, 72(3), 321-330.

Gershman, S. J., & Daw, N. D. (2017). Reinforcement learning and episodic memory in humans and animals: an integrative framework. Annual review of psychology, 68, 101-128.

Goel, V., & Dolan, R. J. (2004). Differential involvement of left prefrontal cortexin inductive and deductive reasoning. *Cognition*, *93*(3), B109-B121.

Gold, J. J., & Squire, L. R. (2006). The anatomy of amnesia: neurohistological analysis of three new cases. Learning & Memory, 13(6), 000-000.

Heekeren, H. R., Marrett, S., Ruff, D. A., Bandettini, P. A., & Ungerleider, L. G. (2006). Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. *Proceedings of the National Academy of Sciences*, *103*(26), 10023-10028.

Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. Nature reviews neuroscience, 9(6), 467.

Hintzman, D.L. (1991). Why are formal models useful in psychology? In W.E. Hockley & S. Lewandowsky (Eds.). Relating theory and data: Essays on human memory in honor of Bennet B. Murdock.

(pp. 39–56). Hillsdale, NJ: Erlbaum.

Insausti, R., Annese, J., Amaral, D. G., & Squire, L. R. (2013). Human amnesia and the medial temporal lobe illuminated by neuropsychological and neurohistological findings for patient E.P. Proceedings of the National Academy of Sciences of the United States of America, 110(21), E1953–E1962. http://doi.org/10.1073/pnas.1306244110.

Joel, D., & Weiner, I. (1997). The connections of the primate subthalamic nucleus: Indirect pathways and the open-interconnected scheme of basal ganglia– thalamocortical circuitry. Brain Research Reviews, 23, 62–78

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. Science, 273(5280), 1399-1402.

Koenigs, M., Barbey, A. K., Postle, B. R., & Grafman, J. (2009). Superior parietal cortex is critical for the manipulation of information in working memory. *Journal of Neuroscience*, *29*(47), 14980-14986.

Kruschke, J. K. (1992). ALCOVE: an exemplar-based connectionist model of category learning. Psychological review, 99(1), 22.

Kruschke, J. K., & Erickson, M. A. (1994). Learning of rules that have high-frequency exceptions: New empirical data and a hybrid connectionist model. In Proceedings of the sixteenth annual conference of the Cognitive Science Society (pp. 514-519). Lawrence Erlbaum Associates.

Love, B. C., & Medin, D. L. (1998). Modeling item and category learning. In Proceedings of the 20th Annual Conference of the Cognitive Science Society (pp. 639-644).

Love, B. C., & Medin, D. L. (1998, July). SUSTAIN: A model of human category learning. In AAAI/IAAI (pp. 671-676).

Love, B. C., Markman, A. B., & Yamauchi, T. (2000, September). Modeling classification and inference learning. In AAAI/IAAI (pp. 136-141).

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: a network model of category learning. Psychological review, 111(2), 309.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. Perception & psychophysics, 53(1), 49-70.

Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information-integration category learning. *Journal of experimental psychology: learning, memory, and cognition*, *29*(4), 650.

Maia, T. V. (2009). Reinforcement learning, conditioning, and the brain: Successes and challenges. Cognitive, Affective, & Behavioral Neuroscience, 9(4), 343-364.

Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. Nature neuroscience, 14(2), 154.

McClelland, J. L., McNaughton, B. L., & O'reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychological review, 102(3), 419.

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In Psychology of learning and motivation (Vol. 24, pp. 109-165). Academic Press.

McKinley, S. C., & Nosofsky, R. M. (1996). Selective attention and the formation of linear decision boundaries. Journal of Experimental Psychology: Human perception and performance, 22(2), 294.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. Psychological review, 85(3), 207.

Miller, E.K.; Cohen, J.D. An integrative theory of prefrontal cortex function. Annu. Rev. Neurosci. 2001, 24, 167–202.

Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. Neurobiology of learning and memory, 65-77.

Nomura, E. M., Maddox, W. T., Filoteo, J. V., Ing, A. D., Gitelman, D. R., Parrish, T. B., ... & Reber, P. J. (2006). Neural correlates of rule-based and information-integration visual category learning. Cerebral Cortex, 17(1), 37-43.

Nomura, E. M., & Reber, P. J. (2008). A review of medial temporal lobe and caudate contributions to visual category learning. Neuroscience & Biobehavioral Reviews, 32(2), 279-291.

Nomura, E. M., & Reber, P. J. (2012). Combining computational modeling and neuroimaging to examine multiple category learning systems in the brain. Brain sciences, 2(2), 176-202.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. Journal of Experimental Psychology: Learning, memory, and cognition, 10(1), 104.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. Journal of experimental psychology: General, 115(1), 39.

Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13(1), 87.

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. Journal of experimental psychology: human perception and performance, 17(1), 3.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. Psychological review, 101(1), 53.

Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of "multiple-system" phenomena in perceptual categorization. Psychonomic Bulletin and Review, 7(3), 375-402.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature neuroscience*, *7*(1), 75-79.

Parent, A., & Hazrati, L.-N. (1995). Functional anatomy of the basal ganglia. II. The place of the subthalamic nucleus and external pallidum in basal ganglia circuitry. Brain Research Reviews, 20, 128–154.

Pascual, B., Masdeu, J. C., Hollenbeck, M., Makris, N., Insausti, R., Ding, S. L., & Dickerson, B. C. (2013). Large-scale brain networks of the human left temporal pole: a functional connectivity MRI study. Cerebral cortex, 25(3), 680-702.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. Journal of neuroscience methods, 162(1-2), 8-13.

Penfield, W., & Milner, B. (1958). Memory deficit produced by bilateral lesions in the hippocampal zone. AMA Archives of Neurology & Psychiatry, 79(5), 475-497.

Poldrack, Russell A., Jill Clark, E. J. Pare-Blagoev, Daphna Shohamy, J. Creso Moyano, Catherine Myers, and Mark A. Gluck. "Interactive memory systems in the human brain." Nature 414, no. 6863 (2001): 546.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. Nature reviews neuroscience, 9(7), 545.

Reber, P. J. (2008). Cognitive neuroscience of declarative and nondeclarative memory. Advances in Psychology, 139, 113-123.

Reber, P. J. (2013). The neural basis of implicit learning and memory: a review of neuropsychological and neuroimaging research. Neuropsychologia, 51(10), 2026-2042.

Reber, P. J., Beeman, M., & Paller, K. A. (2013, July). Human memory systems: A framework for understanding the neurocognitive foundations of intuition. In International Conference on Augmented Cognition (pp. 474-483). Springer, Berlin, Heidelberg.

Rempel-Clower, N. L., Zola, S. M., Squire, L. R., & Amaral, D. G. (1996). Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. Journal of Neuroscience, 16(16), 5233-5255.

Rugg, M. D., & Henson, R. N. (2005). Episodic memory retrieval: an (eventrelated) functional neuroimaging perspective. *The cognitive neuroscience of memory encoding and retrieval*, 707-714.

Saint-Cyr, J. A., Taylor, A. E., & Lang, A. E. (1988). Procedural learning and neostyrial dysfunction in man. Brain, 111(4), 941-960.

Schultz, W., Stauffer, W. R., & Lak, A. (2017). The phasic dopamine signal maturing: from reward via behavioural activation to formal economic utility. Current opinion in neurobiology, 43, 139-148.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. Journal of neurology, neurosurgery, and psychiatry, 20(1), 11.

Smith, E. E., & Grossman, M. (2008). Multiple systems of category learning. Neuroscience & Biobehavioral Reviews, 32(2), 249-264.

Smith, E. E., Patalano, A. L., & Jonides, J. (1998). Alternative strategies of categorization. Cognition, 65(2-3), 167-196.

Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. Journal of cognitive neuroscience, 4(3), 232-243.

Squire, L. R., & Dede, A. J. (2015). Conscious and unconscious memory systems. Cold Spring Harbor perspectives in biology, 7(3), a021667.

Sutton, R. S. (1984). Temporal credit assignment in reinforcement learning.

Tekin, S., & Cummings, J. L. (2002). Frontal–subcortical neuronal circuits and clinical neuropsychiatry: an update. *Journal of psychosomatic research*, *53*(2), 647-654.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. science, 185(4157), 1124-1131.

van't Wout, M., Kahn, R. S., Sanfey, A. G., & Aleman, A. (2005). Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex affects strategic decision-making. Neuroreport, 16(16), 1849-1852.

Victor, M., & Agamanolis, D. (1990). Amnesia due to lesions confined to the hippocampus: A clinical-pathologic study. Journal of Cognitive Neuroscience, 2(3), 246-257.

Wenzlaff, H., Bauer, M., Maess, B., & Heekeren, H. R. (2011). Neural characterization of the speed–accuracy tradeoff in a perceptual decision-making task. *Journal of Neuroscience*, *31*(4), 1254-1266.

Wolpert, D. M., Goodbody, S. J., & Husain, M. (1998). Maintaining internal representations: the role of the human superior parietal lobe. *Nature neuroscience*, *1*(6), 529-533.

Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. Psychological review, 105(3), 558.

White, P.A. (2008). Beliefs about interactions between factors in the natural environment: A causal network study. Applied Cognitive Psychology, 22, 559–572.

Zola-Morgan, S., Squire, L. R., & Amaral, D. G. (1986). Human amnesia and the medial temporal region: enduring memory impairment following a bilateral lesion limited to field CA1 of the hippocampus. Journal of Neuroscience, 6(10), 2950-2967.