5-4-2023

Can explicit processes support implicit category learning?: The effect of relevant rule-oriented selective attention on implicit learning

Andres Sanchez

Georgia State University

Can explicit processes support implicit category learning?: The effect of relevant rule-oriented selective attention on implicit learning

by

Andres Sanchez

Under the Direction of J. David Smith, PhD

A Thesis submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2023

ABSTRACT

Categorization is a crucial component of human cognition. Multiple systems theories suggest categories can be learned by explicit or implicit processes/systems depending on the type of category (e.g., Ashby & Valentin, 2017). Research examining the interaction between these systems found that explicit learning impairs implicit performance (Ashby & Crossley, 2010; Crossley & Ashby, 2015; Sanchez et al., 2020). The nature of this impairment remains unclear. The current study examined the effect of selective attention to rule dimensions that were either relevant or irrelevant to a later implicit categorization task to better understand how this impairment occurs. The results suggested that attention to relevant dimensions is crucial for implicit learning. Both systems can learn in parallel as long as the relevant category information is attended. This suggests the primary mechanism of implicit impairment by the explicit system may be drawing attention away from relevant information rather than rule-based strategy perseveration.

INDEX WORDS: Categorization, Multiple systems, COVIS, Explicit learning, Implicit learning

Can explicit processes support implicit category learning?: The effect of relevant rule-oriented selective attention on implicit learning

by

Andres Sanchez

Committee Chair: J. David Smith

Committee: Michael Beran

Sarah Barber

Barbara Church

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

May 2023

DEDICATION

I dedicate this thesis to my family. They have provided constant support that has made this endeavor possible. They have allowed me to pursue my interests and dreams and for that I am truly appreciative. I also dedicate this thesis to Alexandria Reeves. Thank you for your time and lending an ear whenever I talked about this project and my interests in cognitive psychology generally. Thank you for your support and words of affirmation during times that I felt inadequate. This thesis would not have been possible without you.

ACKNOWLEDGEMENTS

First, I would like to acknowledge my advisor and committee chair, Dr. J. David Smith. Thank you for your support and assistance throughout this thesis. Your dedication for categorization research, among other topics, has and will continue to inspire me. Second, I want to acknowledge my committee members, Dr. Michael Beran, Dr. Sarah Barber, and Dr. Barbara Church, for their willingness to serve on my committee, for their support, and for their comments. Third, I want to acknowledge fellow graduate student, Brooke Jackson, who always made herself available any time I had questions. I want to especially acknowledge and thank, Dr. Barbara Church. Thank you for reading an endless number of drafts and for challenging me. Without your guidance, this thesis would not have come to fruition in the way that it did. All of those acknowledged here, have without a doubt made me a better researcher, critical thinker, and writer. For that I am endlessly grateful.

TABLE OF CONTENTS

LIST OF FIGURES

1 INTRODUCTION

Categorization is the ability to psychologically organize objects and/or ideas using cognitive grouping mechanisms in mental space (e.g., Smith et al., 2022). These organized mental groupings can be based on strict perceptual similarity or completely abstract relations. In effect, category learning creates functional groupings by which meaningful interactions with the world can be experienced (e.g., Seger & Miller, 2010). For instance, categorization allows us to discern whether we are facing a friend or foe, food is edible or expired, or whether experimental evidence supports a theory or not. Once we have correctly determined category membership, we can take an appropriate course of action. Due to the adaptive value of effective categorization, research on this topic has been extensive (e.g., Ell & Ashby, 2006; Jacoby & Brooks, 1984; Medin & Schaffer, 1978; Rosch, 1973; Smith et al., 2012, 2015; Unger & Sloutsky, 2021). However, the particular cognitive mechanisms and grouping principles that humans use to form categories is still a matter for theoretical debate (for reviews see Ashby & Maddox, 2005; Poldrack & Foerde, 2008).

The categorization literature has considered whether a single system or multiple category-learning systems are needed to understand category learning (e.g., Ashby et al., 1998; Love et. al., 2004; Nosofsky & Zaki, 1998). This debate in categorization is reminiscent of the debate about whether memory is composed of a single system or multiple systems (e.g., Brooks & Baddeley, 1976; Cohen & Squire, 1980; Knowlton et al., 1994; Roediger, 1990; Tulving, 1972; 1985). In the last 25 years, numerous multiple system (or multiple process) accounts have theorized that different category learning systems/processes govern different types of category learning (e.g., Ashby et al., 2011; Erickson & Kruschke, 1998; Smith & Church, 2021). A dominant multiple systems theory is the COmpetition between Verbal and Implicit Systems

(COVIS) approach. COVIS hypothesizes distinctions between an explicit-declarative system that learns quickly through hypothesis testing of rules and an implicit-procedural system that learns by unconscious stimulus-response associations.

With this distinction in mind, multiple-systems theorists have explored how these systems may interact and/or impair each other during category learning and category decision-making (Ashby & Crossley, 2010; Crossley & Ashby 2015; Sanchez et al., 2020). This research has generally found that explicit learning impairs implicit processes. The current study seeks to further clarify the mechanism involved in the impairment of implicit category learning by the explicit system. Specifically, it examines the role that top-down processes, like selective attention, have on the impairment of implicit learning seen during interactions between the explicit and implicit system. It also examines the effect that selective attention may have on the perceptual and memory processes necessary for category learning.

To foreshadow the layout of my thesis, I begin by detailing the primary theoretical ideas about how humans categorize, explaining their approaches, assumptions, and limitations. I will then characterize single versus multiple systems theories focusing on the exemplar comparison theory versus implicit/explicit approaches. I will explain their main assumptions, differences, and the evidence for each. Next, I will lay out the types of tasks and stimuli that are normally used to study dissociations of the proposed explicit and implicit systems of category learning. I will then go into further detail examining the issue of interaction vs competition between those systems. I will then explain the role of selective attention and strategy selection in category learning and decision-making generally and under the implicit/explicit framework. I then present previous research investigating manipulations of relevant dimensions and instructions, and fully lay out the logic and design of this study. Finally, I present the results and discuss them in terms of their theoretical implications to the multiple systems perspective. Future research is suggested as well.

2 THEORIES OF CATEGORIZATION

There has been a long-standing debate regarding which theoretical perspective best accounts for category learning. Philosophers have been interested in categorization since at least the time of the ancient Greek philosopher Aristotle. Experimental psychology's study of category learning can be traced back to Hull's (1920) early behaviorist experiments. Over time, as the cognitive perspective came to dominate the study of human categorization, three classes of theories emerged as the primary competitors in the study of category learning (for review see Goldstone et al., 2018). Classical theory, prototype comparison theory, and exemplar comparison theory have been in competition for over 50 years, but in the last few decades, newer hybrid models that incorporate multiples of these proposed categorization processes have come to the forefront.

Classical theory is exemplified by Bruner et al.'s (1956) experiments illustrating the hypothesis testing of rules. There are three main assumptions of classical theory (e.g., Smith & Medin, 1981). The first assumption is that a summary representation of a category is abstracted to represent the entire category. The second, and most critical, assumption is that an object is defined as belonging to a particular category because it has the necessary and sufficient features of the category in question. In other words, a feature that represents a category must be present in all objects in the category (necessary) and every object with all the defining features is a member of the category (sufficient). To illustrate this assumption further, consider the case of an equilateral triangle. An equilateral tringle's defining features are its three equal sides and being a closed figure. These features are necessary and jointly sufficient to call all objects with these features an equilateral triangle. Classical theory assumes that clear rules, such as if it has three equal sides and is a closed figure then it is an equilateral triangle, allow category learning. The third assumption is that subcategories of a concept must have the defining criteria of the larger concept. For instance, squares as a subset of a quadrilateral must have the defining features of a quadrilateral to be a subset. With these assumptions in mind, proponents of the classical theory (e.g., Bruner et al. 1956; Restle, 1962) suggested that category rules are the leading and perhaps only process by which categories are represented and learned (e.g., Levine, 1975). However, subsequent research has weakened classical theory because it has become clear that rule formation cannot explain all category learning (e.g., Medin & Schaffer 1978; Rosch, 1973a).

There are three general criticisms/limitations of the classical theory. First, classical theory cannot account for disjunctive concepts (e.g., Smith & Medin, 1981). Disjunctive concepts are concepts that do not necessarily have to have all the features in question to belong to the category. Disjunctive concepts have category members that have either one or the other of the features needed for category membership but exemplars with both features are not members (e.g., Snow & Rabinovitch, 1969). For example, a concept could include an object being green or a circle to belong to the category but not a green circle. Second, classical theory erroneously assumes that subsets of categories are easily identifiable or agreed upon by all persons, but research has shown that this is not the case (e.g., Smith & Medin, 1981). Third, many (if not most) categories do not have clear defining features but rather are best learned through understanding what characteristics are typical for the category members (e.g., Wittgenstein, 1953). Wittgenstein famously employed the concept of games to illustrate his argument. Wittgenstein asked, what would the defining feature of games be that would apply to all games? If one understands games to be football, Ping-Pong, tennis, solitaire, etc., it becomes quite difficult to determine the defining criteria. Instead, he argued that categories can best be learned through understanding what characteristics are typical for the category members (for review see, Goldstone et al., 2018). He called this similarity relation among category members family resemblances.

Prototype comparison theory was developed to account for classical theory's shortcomings and to explain how family resemblance categories are learned (e.g., Rosch & Mervis, 1975). Prototype comparison theory assumes that while learning a category, a prototype representation (central tendency) of typical features of that category forms (e.g., Rosch, 1973). New potential category members are compared to the prototype and category decisions are made based on similarity. Early research suggested that this categorization process of comparison to the central tendency or prototype could explain performance with natural categories such as color and shape (Rosch, 1973a), artificial categories (Posner & Keele, 1968; Reed, 1970), and semantic categories (Rosch, 1973b). Prototype comparison theory also assumes that the prototypical examples are more likely to be recalled first (Mervis et al., 1976) and prototypical examples are learned first by children (Rosch, 1973b). The assumption of faster recall was supported by research showing that the more typical members of a category are categorized faster than less typical members (Rosch, 1973b). For example, semantic categorization tasks found that typical birds like robins were categorized faster than less typical members such as chickens (Rosch, 1973b). Unlike the classical theory's focus on defining criteria for membership, prototype comparison theory does not require that a category have necessary and sufficient features. This allows it to explain a wider range of categorization performance. In fact, prototype comparison may explain most non-human animals' categorization performance (for review see Smith et al., 2022).

Despite prototype comparison theory's ability to account for a wide range of categorization phenomena, it is still ill-equipped to explain the ability to learn atypical category members, devise clear cut category boundaries, or create ad hoc categories (Smith et al., 2022). For example, prototype comparison theory's main assumption of comparing to a central tendency cannot easily explain how humans learn to consistently categorize dolphins as mammals even though they have more visual similarity to fish. Second, prototype comparison theory is often criticized for its unclear or fuzzy boundaries. For instance, on a continuum of color, where exactly does the red start and end and when does orange or pink begin (e.g., Geeraerts, 1989)? In other words, the category boundaries are not sharply defined when a similarity comparison is made to a single prototype. Crucial category information such as the category range is lost. Finally, prototype comparison theory is unable to account for categorization processes that rely on logic, ad hoc groupings, or randomness (e.g., Smith et al., 2022). For instance, in an ad hoc grouping (Little et al., 2006) of what to take in the event of your house burning (i.e., children, documents, pictures, money, etc.), a prototype comparison is not useful for forming such a disparate category.

Exemplar comparison theory addresses some of the limitations of prototype comparison theory (e.g., Medin & Schaffer, 1978; Nosofsky, 1987). Exemplar comparison theory assumes that all experienced category examples are singly and separately stored in memory rather than consolidated into a representation of central tendency. Judgements about category membership are made by similarity comparisons with these stored examples in memory (Medin & Schaffer, 1978; Smith & Medin, 1981). Medin and Schaffer (1978) argued that exemplars are learned and stored in memory and grouped as categories depending on the learning context. Under this context theory of categorization, probe stimulus features and the context in which they occur together act as a cue for memory retrieval of similar exemplars (Medin & Schaffer, 1978). If similarity is sufficient with the stored exemplars, the new stimulus is likely to join previously learned exemplars and be placed into the category. Nosofsky (1986, 1987) extended context theory to include multidimensional continuous stimuli and developed the generalized context model (GCM). The GCM uses a multidimensional scaling approach to model similarity. Exemplars are assumed to represent points in a multidimensional psychological space. Here, similarity is a decreasing function of distance between exemplars in that psychological space. Selective attention is also factored into the GCM. Selective attention serves to optimize performance by systematically changing the psychological space of the categories based on attention to the dimensions that produce the best average performance (Nosofsky, 1986).

Another exemplar model is an exemplar connectionist model, ALCOVE (Kruschke, 1992). ALCOVE uses an error correcting learning process to instantiate exemplar category learning processes. ALCOVE assumes an attentional and similarity weight that learns through error to guide dimensional attention and association between exemplar nodes and category nodes (Kruschke, 1992; Kurtz, 2007). Because there is a direct comparison to previously stored examples, exemplar comparison theory can account for learning atypical examples and ad hoc categories through memory activation of exemplars. For instance, within the category of couches, once an atypical member like a geometrical cushion couch is stored in memory as a category member, its memory representation can simply be retrieved. However, exemplar comparison theory is not free from criticism.

In order to account for atypical examples and ad hoc categories, exemplar comparison theory proposes unlimited numbers of stored exemplars at the expense of cognitive economy (e.g., Smith & Medin, 1981; Smith et al., 2022). As with GCM, ALCOVE also faces the issue of storage demands of the hidden nodes (the exemplars) and lack of category constraints (Kurtz, 2007). The plausibility of a memory system that can hold thousands of exemplars of a single category let alone of multiple different categories is questionable (e.g., Smith et al., 2022). It is unlikely that a memory system of such capacity would develop. Even if it did, some have argued that a composite of the exemplars must be formed in working memory for similarity comparison (Hintzman,1986). In such a case, then the comparison process is more akin to prototype comparison theory (Smith et al., 2021). Another criticism of exemplar comparison theory is its inability to fit categorization data that prototype theory can fit when category size, structure, and stimulus dimensions are manipulated (Minda & Smith, 2001, 2002). Across four experiments, Minda and Smith (2001) demonstrated that a prototype model fit the data better than the exemplar model under a number of circumstances. The prototype model was particularly good at accounting for large categories and complex stimuli sets. This suggests that exemplar comparison theory may be best suited for categories that are small, poorly structured, and less complex. Further, Smith (2002) demonstrated that exemplar comparison theory's predicted typicality gradient produces an inadequate fit to the typicality gradient of the dot distortion categorization task. Rather, the task was best fit by the typicality gradient predicted by prototype comparison theory. These results and others put into question exemplar comparison theory's ability to fully explain all categorization processes both empirically and psychologically (e.g., Smith, 2002).

Due to the shortcomings of classical theory, prototype comparison, and exemplar comparison theory on their own, connectionist and hybrid models that incorporate assumptions of multiple theories have developed in the last three decades. Below I will briefly characterize categorization models that have been developed in response to the limitations of the three theoretical positions mentioned above. For instance, I will describe hybrid models such as SUSTAIN (Love et al., 2004), RULEX (Nosofsky et al., 1994), ATRIUM (Erickson & Kruschke, 1998), a prototype plus exemplar comparison hybrid (Smith & Minda, 1998), and COVIS (Ashby et al., 1998).

Connectionist models generally depict cognitive processes and representations as interconnected neural networks consisting of nodes. These nodes are information processing units, that are activated in parallel with other nodes (Gluck & Bower, 1988). Here, different patterns of activation represent different percepts or concepts. In a connectionist model, category knowledge does not exist in just one particular node but rests on the connection strength between nodes and how they produce activation of output units (category responses). These models rely on weighted activation to guide categorization, and error correcting rules adjust the weights to create category learning (e.g., Gluck & Bower, 1988). SUSTAIN, a hybrid clustering connectionist model, initially assumes a simple approach to categorization (Love et al., 2004). Simple rules are first assumed by focusing on the dimensions that are most predictive of category membership. If unsuccessful, then SUSTAIN incrementally adds more clusters to account for categorization as necessary. The added clusters may take on a prototype comparison category structure or exemplar comparison category structure. Or, if needed, SUSTAIN may use both prototype and exemplar clusters to facilitate categorization. Category judgments are accomplished through a similarity judgment of input information and category representation clusters (Kurtz, 2015; Love et al., 2004).

Hybrid models assume multiple comparison processes and/or category systems for different types of category learning. For example, RULEX, a hybrid of classical theory and exemplar comparison theory, assumes the initial use of simple rules. However, if rules are unable to account for all members, memory for exceptions may develop (Nosofsky et al., 1994). The benefit of RULEX is that it can handle continuous stimulus dimensions (Nosofsky & Palmeri, 1998). Unlike single exemplar comparison models, RULEX, through rules, creates category boundaries that help to represent the categories psychologically. However, its explanatory power is limited to learning two-choice, mutually exclusive categories (Kurtz, 2007; Love et al., 2004). ATRIUM is a hybrid connectionist model that assumes the interaction of rule and exemplar modes of category learning through a gating mechanism that links the two representations (Erickson & Kruschke, 1998). It is thought that all stimuli are concurrently processed by the rule and exemplar parts (modules). The gating mechanism serves to control and push out the final output that is equal to the relative proportion that the rule and exemplar module each contributed to category learning. ATRIUM's use of rule and exemplar modules accounted for categorization data better than rule and exemplar representations alone.

Smith and Minda (1998) proposed a hybrid model of categorization that stressed the importance of prototype and exemplar comparison at different stages of category learning and with different category structures. For instance, at early stages of category learning prototype strategies were employed that then later progressed to exemplar comparisons. Additionally, the hybrid model with both prototype and exemplar parameters provided a better explanation of the data than either model alone (Smith & Minda, 1998). The authors argued that the reason for this prototype to exemplar pathway is that participants can quickly abstract a prototype from the relevant features, but adequately learning exemplars and their association to the categories takes longer, especially in large categories and with complex stimuli.

Another important hybrid model is COVIS. COVIS is a neuropsychological categorization theory that assumes two systems that compete to underlie category learning. It assumes an explicit-declarative system that creates, selects, and tests hypothesis to find verbal rules and an implicit-procedural system that lacks verbal rules but learns incrementally through stimulus-response associations (e.g., Ashby et al., 1998; Ashby & Valentin, 2017). Ashby and colleagues (1998) proposed that the explicit system uses logical reasoning to develop rules. Rules developed by the explicit system are easily verbalizable for humans. Under COVIS, rule creation, selection, testing, maintenance, and switching are thought to involve the prefrontal cortex, working memory, and executive functioning. The assumed selection of rules by the explicit system is as follows. First, all possible rules are formed based on the dimensions of the stimuli. COVIS assumes initial privilege to easier rules for hypothesis testing. Therefore, there is a higher probability of selecting unidimensional rules and rules the learner has experienced before rather than new conjunctive or disjunctive rules (Ashby et al., 1998). Once decision rules are active and held in working memory, they are used for category membership testing. For example, a decision rule may be, if the stimuli are more blue then pick A, otherwise pick B for more Red. If rules are unsuccessful, a rule switch is likely to occur. COVIS assumes that rule switching, maintenance, and selection are independent processes mediated by the explicit system.

The implicit-procedural system, on the other hand, learns slowly through dopamine-mediated reinforcement learning (e.g., Ashby & Valentin, 2017). This system learns by holistically integrating dimensional features unconsciously. COVIS assumes implicit learning to include motor mapping of stimulus-response associations. The implicit-procedural system is thought to involve the basal ganglia, specifically, the tail of the caudate nucleus, and the putamen (e.g., Ashby & Valentin, 2017). Implicit learning involving these brain structures is as follows: reward or positive feedback causes dopamine to release into the tail of the caudate nucleus. The dopamine signal then strengthens recently activated synapses through the caudate's medium spiny cells (e.g., Ashby & Valentin, 2017; Smith & Church, 2018). This process produces learning in the implicit procedural system.

A strength of COVIS is its focus on the implicit learning system's role in categorization. This role has been largely ignored in the human categorization literature (Ashby et al., 1998). As a result of including implicit learning in categorization theory, it more closely aligns the human and non-human animal literature and research in neuroscience and cognitive psychology. A potential shortcoming of this approach includes the acknowledgement that verbal rules and implicit-associative learning are probably not the only ways in which categories are learned. However, Ashby et al., (1998) make clear that they do not exclude other systems that are exemplar or perceptual in nature also being involved in categorization processes.

3 SINGLE VS. MULTIPLE SYSTEMS

The single versus multiple systems debate has been influential in both the memory and categorization domains (e.g., Ashby et al., 1998; Cohen & Squire, 1980; Poldrack & Foerde, 2008; Tulving, 1985). The debate centers on whether cognitive abilities such as categorization and memory involve a single process or multiple distinct processes. A single system approach assumes a single cognitive process or mechanism that underlies performance (e.g., Ashby & Ell, 2002). For example, a prominent single system model of memory proposed that the degree of memory retention is a function of depth or level of processing (Craik & Lockhart, 1972; Craik & Tulving, 1975), rather than assuming an attentional filtration process or multiple stores of memory to explain differences in retention (e.g., Atkinson & Shiffrin, 1968; Broadbent, 1958).

Comparable to single system memory models, single system categorization models often assume one process that determines category allocation (e.g., Hintzman, 1986; Newell et al., 2011; Nosofsky, 1986; see Section 2 for a review of the early dominant theories of categorization). For example, one categorization model uses an explicit memory model, Minerva 2, to explain category learning (Hintzman, 1986). Here, a single memory system is assumed to include multiple episodic memory traces that are all activated based on similarity to the cue. Retrieved information is then combined in working memory into an aggregate of activated episodic memory traces. Here, category knowledge is represented by a prototype created in working memory based on the retrieved information from the activated exemplar traces (Hintzman,1986).

As mentioned previously, a single system theory assumes one process or mechanism that mediates all category learning. Single system explanations are often justified by the notion of scientific parsimony (Nosofsky & Johansen, 2000). However, proponents of multiple systems argue that a theory is not as parsimonious if it must keep adding new assumptions or parameters to account for the data. Rather, multiple systems accounts may better explain categorization and memory data in total (Poldrack & Foerde, 2008).

Unlike single system theories, a multiple systems approach assumes that there are multiple independent and dissociable processes that are involved in memory or categorization (e.g., Ashby & Ell, 2002; Smith & Grossman, 2008). The multiple systems approach in categorization can be traced back to Brooks' (1978) discussion of analytic and nonanalytic forms of category learning. Analytic and nonanalytic categorization are defined by how the stimulus is the perceived. Analytic categorization involves abstracting the relevant features from the irrelevant features of a stimulus (Jacoby & Brooks, 1984). Analytic categorization often involves single-dimensional rules (Kemler Nelson, 1984) and is intentional (Brooks, 1978). Conversely, nonanalytic categorization involves a more holistic approach to the perception of the stimulus and does not break the stimulus into stable relevant and irrelevant components (Jacoby & Brooks, 1984). Rather, nonanalytic categorization depends on the integration of both relevant and irrelevant features. Nonanalytical categorization is described as incidental (Brooks, 1978) and based on overall similarity comparisons (Kemler Nelson, 1984). COVIS theory may be seen as an extension of Brooks' (1978) distinction between analytic and nonanalytical categorization.

COVIS assumes independent explicit and implicit systems (Ashby et al., 1998). Considerable evidence supporting COVIS comes from both behavioral and neuroscientific research. Behavioral research has focused on dissociations between how different variables affect performance when a task requires implicit versus explicit learning (for review see Ashby & Valentin, 2017). In some instances, manipulations hurt explicit learning performance in tasks thought to require working memory, but not performance thought to require the more automatic implicit learning. For example, research has found that performance thought to require the explicit system is impaired by a concurrent working memory task, while performance thought to require the implicit system is not (e.g., Waldron & Ashby, 2001; Zeithamova & Maddox, 2006). Also, research found that increasing the number of categories (2 versus 4) negatively affected performance thought to require explicit learning while performance thought to require implicit learning was equal in the two and four category conditions (Maddox, Filoteo et al., 2004). COVIS makes the a priori prediction that explicit learning is impaired as rules increase in quantity and complexity (conjunctive rules) because of its reliance on working memory, but implicit learning should be unaffected by this manipulation (Maddox, Filoteo et al., 2004).

Reducing feedback processing time also adversely affects the performance thought to be reliant on the explicit system more than the performance thought to be reliant on the implicit system (Maddox, Ashby et al., 2004). These researchers found that if you interrupt feedback processing by requiring participants to immediately attend to another task then explicit learning is disrupted. However, implicit learning was left intact. These findings are consistent with the COVIS theory's assumption that the explicit system recruits working memory and attention to process feedback, while the implicit system does so automatically without the need of working memory and attention (Maddox, Ashby et al., 2004).

Physiological research has also shown that there are variables that affect explicit more than the implicit reliant performance. For example, stress (Ell et al., 2011) and sleep deprivation (Maddox et al., 2009) were found to negatively affect explicit more than the implicit learning. On the other hand, positive mood was found to enhance the explicit system more than the implicit system (Nadler et al., 2010). Stress (Lupien et al., 2007) and sleep deprivation (Herscovitch et al., 1980) are known to generally have adverse effects on executive functioning and positive mood is known to enhance cognitive flexibility (Ashby et al., 1999). These findings are consistent with the COVIS theory's assumption that the explicit system is reliant on executive functioning while the implicit system is not. Experimental manipulations that affect the explicit system more than the implicit system help characterize the explicit system as one that learns rules, relying on working memory and selective attention.

If all the dissociations consistent with predictions from COVIS showed that performance thought to rely on explicit learning could be affected by variable manipulation, but implicit learning could not, there might be simpler explanations of these dissociations (floor/ceiling effects). However, there are also many studies that found dissociations where manipulations affected the performance thought to rely on the implicit system more than that of the explicit system. Behavioral research has shown that the timing and form of feedback delivery affects the performance thought to rely on implicit system more than the explicit system (Maddox et al., 2003; Smith et al., 2014). For instance, feedback delays of 2.5, 5 and 10 seconds were found to disrupt implicit learning, while explicit learning remained intact (Maddox et al., 2003). Ashby et al. (1999) found that unsupervised learning (no feedback) caused the explicit system to apply unidimensional rules to both unidimensional contrasting categories and diagonal contrasting categories. These dissociations support COVIS' assumptions that the explicit system can learn without feedback, but the implicit system relies on direct reinforcement and that reinforcement is time contingent (Ashby et al., 1998).

As mentioned in Section 2, implicit learning is hypothesized to be heavily reliant on learning of associations between perceptual stimuli and specific motor response, while rule learning is more abstract. COVIS predicts that changes in motor responses should affect implicit learning more than explicit learning. Testing this prediction, Ashby et al. (2003) investigated how switching hands and response keys at test affected explicit and implicit learning. They found that neither hand switching, nor key response switching affected performance thought to rely on explicit learning. However, key response switching did negatively affect performance thought to rely on implicit learning.

Another dissociation is the degree to which learning can generalize or transfer to an indirect categorization task. Helie and Ashby (2012) examined how learning thought to rely on the explicit and implicit systems could be applied to an indirect same-different categorization task of the same category structure. They found learning only occurred from the indirect task or could be transferred to the indirect task when performance was thought to rely on explicit processes. These results are consistent with some of the assumptions of the COVIS theory. It is consistent with the assumption that explicit learning creates category representations that are transferable and abstract. It is also consistent with the idea that implicit learning creates specific stimulus response associations (Helie & Ashby, 2012).

Possible double dissociations have also been demonstrated in further support for multiple systems theories. For example, differences in working memory capacity inversely affect performance on tasks thought to be reliant on the explicit or implicit system. DeCaro et al. (2008) found that participants with a higher working memory capacity had quicker learning in tasks thought to be reliant on explicit learning but performed worse in tasks thought to be reliant on implicit learning. Individuals with lower working memory capacity showed the opposite pattern of performance. These results suggested that a high working memory capacity may help explicit learning but may lead to explicit strategy perseveration even with category structures better learned through implicit processes. This leads to suboptimal performance. Individuals with a lower working memory capacity may be more willing to switch to implicit learning. A multitude of dissociations that affect one system more than the other or provide double dissociations have been found in the last 25 years supporting the multiple systems approaches over single system explanations (for review see Ashby et al., 2017, however, see Newell et al., 2011, for single system arguments).

Despite these various dissociations, single system proponents continue to argue for the parsimony of a single system approach. Recently, researchers have proposed that task complexity or difficulty differences can explain the many dissociations observed (Le Pelley et al., 2019; Zaki & Kleinschmidt, 2014). However, because different variables can impact tasks thought to rely on implicit versus explicit learning in opposite ways, difficulty differences cannot explain all the findings (see Ashby et al., 2020, for a refutation of the difficulty argument).

4 TYPES OF TASKS USED TO DISSOCIATE THE EXPLICIT AND IMPLICIT SYSTEMS

In the previous section, I discussed how research examined dissociations between implicit versus explicit learning through experimental variable manipulations. However, what kind of tasks are needed to determine whether implicit or explicit learning is needed? In this literature, rule-based (RB) tasks and information-integration (II) tasks are used to engage the explicit and implicit system, respectively (for review see Ashby & Maddox, 2005, 2011). They are psychophysical tasks thought to differentially promote the use of each system. In fact, much evidence for a multiple systems approach is indebted to the use of RB and II tasks (e.g., Maddox et al., 2003; Maddox et al., 2010; Zeithamova & Maddox, 2007).

RB tasks, in their simplest form, are tasks featuring categories that require participants to selectively attend to one dimension that is diagnostic of category membership. This is a process that is thought to be accomplished by explicit reasoning. In more complex forms, conjunctive and disjunctive rules may also be used for rule-based categorization. RB tasks are thought to recruit the explicit system because they foster verbalizable rules, selective attention, and rule selection (Ashby et al., 1998). Overall, the explicit system privileges easily verbalizable rules. Research has found that these rule selection processes involve the anterior cingulate and prefrontal cortices (Posner & Petersen, 1990). COVIS assumes that a cingulate-prefrontal cortex network learns rules and facilities explicit rule selection (Ashby et al., 1998). Consistent with this view, as described in Section 3, evidence has demonstrated that RB tasks target brain regions known to be involved in declarative memory, working memory, and executive functioning (e.g., Ashby et al., 1998; Ashby & Valentin, 2017).

RB categories may be represented in the following way. Assuming a two-dimensional perceptual space to include an x and y dimension that can be graphed on an x and y axis, explicit unidimensional rules partition the space either vertically or horizontally (see Figure 1, rule-based task perceptual space). To describe this with verbal rules, it may be as follows, “respond A to the small objects and B to big objects.” Thus, clearly defined boundaries can be constructed that allow easily verbalizable rules.

On the other hand, II tasks require participants to combine both dimensions for optimal performance and a mapping of stimuli to motor responses. II tasks require procedural learning. Learning of II tasks may take the form of nonanalytical processing or of weighted probabilistic category information learning (e.g., Ashby & Valentin, 2017). Due to the II task demand for holistic processing the implicit system is thought to take over categorization learning and response (Ashby et al. 1998). When simple rules can no longer provide useful categorization information, the implicit system is thought to take over categorization judgements. II tasks are well suited to recruit the implicit system because they require the integration of multiple dimensions in ways that are difficult to verbalize for optimal categorization performance. Within the same two-dimensional perceptual space, categories that integrate multiple dimensions may be represented by partitioning the space with diagonal lines. II categories are very difficult to verbalize. However, these categories can be characterized as respond A to stimuli that are more dimension X than dimension Y and respond B to stimuli that are more dimension Y than dimension X (see Figure 1, information-integration perceptual space).

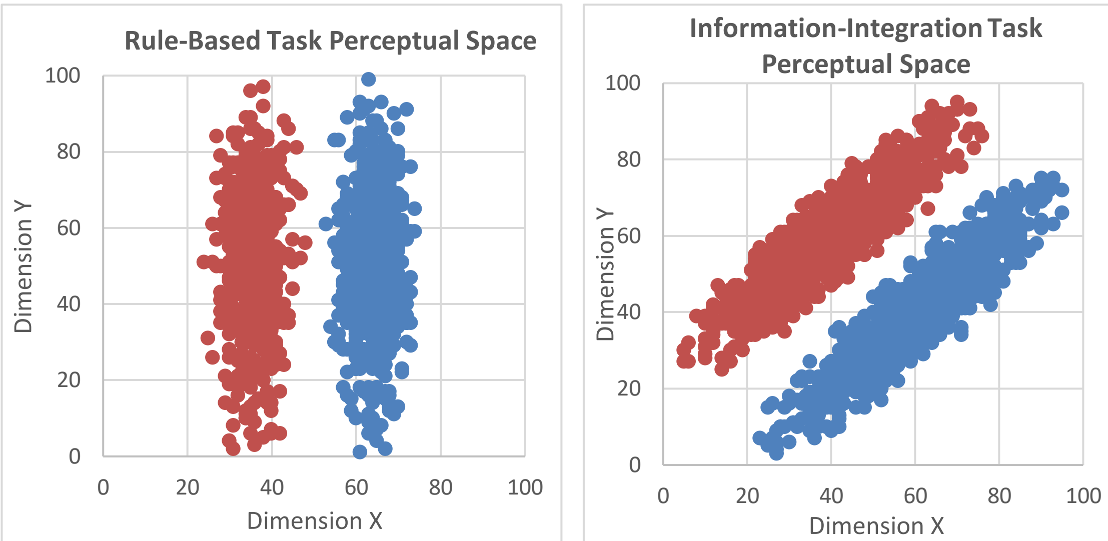

Figure 1 Visual representations of a rule-based category structure and an information-integration

Note. Category A stimuli shown by red dots; Category B stimuli by blue dots.

5 EXPLICIT AND IMPLICIT SYSTEM INTERACTION

If a multiple systems approach is assumed, a crucial next question is how the systems in question interact or compete to produce successful categorization performance. Several questions along these lines have been advanced by researchers (Ashby & Crossley, 2010; Ashby & Maddox, 2011). For example, are the explicit and implicit systems fully independent? Do they learn in parallel? Do they compete in category learning? If so, how, and when does competition occur? The current section will highlight key studies that attempt to answer such questions.

Neuroscience studies suggest that declarative (explicit) and procedural (implicit) memory systems interact in a competitive fashion (Poldrack et al., 1999; Schroeder et al., 2002). Research examining implicit versus explicit category learning has come to similar conclusions (e.g., Nomura et al., 2007). One behavioral study examined how the explicit and implicit systems interact using a hybrid category structure task that was assumed to activate both systems (Ashby & Crossley, 2010). The hybrid category structure task meant that some stimuli could be categorized using explicit learning and some only using implicit learning. The researchers hypothesized that if the explicit and implicit systems operated independently then the hybrid task should be learned as easily as an information-integration task, if not easier. Further, Ashby and Crossley reasoned that facilitating interactions between the two systems would be supported if performance was better in the hybrid than the II task. Conversely, competition between the two systems would be implicated if the hybrid task was more difficult. Across both experiments the results showed that only two out of 53 participants used the optimal hybrid strategy. Instead, a simple unidimensional rule strategy was used by most participants. After ruling out difficulty and task switching cost hypotheses, Ashby and Crossley argued that the reason for the suboptimal performance by the participants was a competitive relationship between the explicit and implicit systems. They discussed two possible forms of competition. First, the explicit system, as the psychologically privileged system, stops implicit learning all together whenever rules are salient. Second, the two systems may learn in parallel but competition at response/output stages may occur with the most confident system dominating.

A subsequent study seeking to disentangle which form of competition was operating examined whether implicit learning occurs at all when the explicit system is used (Crossley & Ashby, 2015). The researchers emphasized the neuroscientific evidence for continued striatal activation (involved in implicit learning) when the explicit system was in control, suggesting both systems may learn in parallel (Foerde et al., 2006). The study further explored the hypothesis of a competition for control over category responses by examining whether learning still occurred in both systems. The researchers developed a behavioral paradigm that examined category response control and actual learning by each system in a single experiment. In the study, researchers included four different conditions to examine underlying implicit learning during rule learning tasks by looking at transfer performance to a later II task. Two conditions featured rule-based training and two featured implicit learning during training. One condition of each training type switched categories at test. In the conditions that switched, Category As became Category Bs and vice versa. The other two conditions featured congruent categories at test (the category membership remained the same). The researchers hypothesized that if implicit learning occurred during rule learning, then transfer from rule-based training to a congruent II task should be better. The results showed that participants in the congruent RB training condition performed significantly better than the rotated condition. This was interpreted as evidence of parallel learning. Crossley and Ashby (2015) concluded that the explicit and implicit systems learn in parallel, and that competition is at the point of motor response control and not during learning. However, several important questions remained regarding what the implicit system learns while the explicit system is in control and whether learning is done independently or is reliant on perceptual and attention processes that also drive explicit learning. The current study seeks to examine these remaining questions about how the explicit and implicit system interact.

6 SELECTIVE ATTENTION AND STRATEGY SELECTION

Though Crossley and Ashby (2015) concluded that implicit learning proceeds normally during rule learning, other research has suggested that selective attention processes employed during rule learning may actually impair the learning of a fully integrated perceptual representation of both dimensions (Sanchez et al., 2020). This section will discuss what selective attention and strategy selection is and past research on the role of selective attention and strategy selection in categorization. Further, it will discuss selective attention and strategy selection in relation to COVIS.

In *Principles of Psychology* (1890/1950), James defined (selective) attention as "… taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration of consciousness is of its essence. It implies withdrawal from some things in order to deal effectively with others…". Here James aptly describes selective attention as a process of singling out specific stimuli or features for deliberate examination. Generally, selective attention facilitates the reduction and simplification of our experiences and can be manipulated based on our current goals and desires (Unger & Sloutsky, 2021). In other words, selective attention helps us focus on what is relevant to the experience at hand and filters out irrelevant information that may clutter cognitive processing.

Selective attention and learning how to use it optimally is a crucial component of categorization (Blair et al., 2009). Since Shepard et al. (1961), there has been an emphasis on the importance of selective attention to category learning. Category models like ALCOVE (Kruschke, 1992), ATRIUM (Erickson & Kruschke, 1998) and RULEX (Nosofsky et al., 1994) operate explicitly on an attentional mechanism that guides attention optimally. COVIS assumes that selective attention to relevant or diagnostic features is critical for rule formation and testing (Ashby et al., 1998). One study examining selective attention with eye-tracking technology has shown that humans, by the end of the task, can learn to allocate attention optimally by focusing on the relevant dimension or dimensions (Rehder & Hoffman, 2005a). Humans, therefore, have a flexible attentional system that allows for optimal focus during category learning.

The categorization literature has also raised questions about strategy use and strategy selection. Questions about how strategy selection impacts categorization, and research investigating what type of strategies participants employ has been crucial to understanding categorization data within a multiple process framework (Helie et al., 2017). Restle (1962) defined strategy as “a particular pattern of responses to stimuli.” Restle (1962) identified three different potential forms of strategy selection. One form involves the consideration and use of a random single strategy from the first trial and continued use if correct throughout and does not consider another strategy until needed. The second form involves the consideration of all strategies available simultaneously then testing those readily in memory and/or elicited by the task. The third form involves a random sampling of all different strategies and then a resampling when error pushes for new strategies.

More recent theoretical models assume specific learning strategies that are employed by participants in a categorization task. Two dominant hypothesized learning strategies, similarity comparison and rule/hypothesis testing strategies, have often been examined together to explain the relative role each strategy has in categorization performance. For instance, Allen and Brooks (1991), using 5-dimensional animal figures, found support for an exemplar similarity strategy that interrupted a diagnostic rule-based strategy. Some studies have found that early in learning participants employ simple rule testing (e.g., Johansen & Palmeri, 2002). However, other studies have concluded that early learning involves the use of multiple strategies before a diagnostic rule is found if available (Rehder & Hoffman, 2005a).

Further research has suggested there may be a rule testing module and an exemplar (similarity) module both operating during categorization (Rehder & Hoffman, 2005b). Still, other researchers have suggested a preference for rule strategies. For instance, Mathy and Feldman

(2009) found that categories were learned significantly quicker and better with a rule-based order presentation than a similarity order presentation.

COVIS assumes a rule based and a holistic approach that are employed by the explicit and implicit system, respectively. As discussed in Section 5, explicit learning may interfere with implicit learning (Ashby & Crossley, 2010). This suggests that the two systems are in competition and the explicit system may be the preferred process. Further research by Crossley and Ashby (2015) concluded that the interference between the systems is not actually an interference of learning but rather occurs at the response level. However, Church et al. (2018) found that selective attention to rules may impair the learning of the unattended dimensions relevant to implicit learning. This suggested that the explicit system with its employment of selective attention may cause direct impairment of implicit learning.

Additional research examined the role of selective attention to the relevant and irrelevant dimensions. It suggested that when attention is allocated to all dimensions during rule learning, participants can learn the II structure (implicit learning) better than when they only focus on the irrelevant rule dimension during training. However, focusing attention on all dimensions versus only the irrelevant dimension after initial learning had no impact. These findings suggested that selective attention may directly impair implicit learning by preventing the learning of the necessary dimensions (Sanchez et al., 2020). In other words, rather than impairment occurring during control of response outputs, impairment may be due to selective attention changing what is learned. The current study will extend this line of work by directly examining the effects of directing selective attention to relevant versus irrelevant dimensions during an initial rule-focused learning phase on implicit learning. It will allow us to understand the nature of the interaction between the explicit and implicit systems and pinpoint how competition can occur.

In summary, selective attention and strategy selection may be best understood as processes that aid in categorization, especially in rule generation and rule use (e.g., Ashby et al., 1998). Selective attention helps guide attentional efforts optimally to the relevant features of the category. Strategy selection guides learning and decision making to the method that is best (hopefully) for the task contingences.

7 PAST EXPERIMENTS WITH ATTENTION MANIPULATIONS

One task manipulation that has been used to examine selective attention is attentional instructions directing participants' attention to particular dimensions. This section will explore past experiments that manipulated dimensional relevancy and/or instructions to better understand the role of selective attention in categorization.

Understanding the role of attention in human cognition has been a central question in cognitive and developmental psychology (e.g., Anderson et al., 1973; Folstein et al., 2012; Kahneman, 1973; Soto & Ashby, 2015; Unger & Sloutsky, 2021). Tasks that focus on how humans and non-human animals attend and switch attention to relevant dimensions has a long history in discrimination learning and reversal shift learning tasks (e.g., Kohler, 1938; Kruschke, 1996; Lashley, 1942; Wolff, 1967). Studies using these types of tasks have shown that humans without cognitive impairments can readily shift attention to different dimensions, although the type of shift is a factor in the ease of attentional shift (Kruschke, 1996). For instance, studies have shown that intra-dimensional shifts are easier than extra-dimensional shifts. In other words, it is easier to learn when shifts to dimensional values within the same dimension occur than shifting to a completely different dimension (Kruschke, 1996).

Most relevant to the theoretical underpinnings of the current study, there are a number of studies that have manipulated attention to dimensions and attentional instructions within a

COVIS framework (e.g., Grimm & Maddox, 2013, Rosedahl et al., 2021, Sanchez et al., 2020). One study that manipulated attentional instructions to relevant or irrelevant dimensions in rule-based and information-integration categorization tasks found that rule-based categorization benefited from instructions to relevant dimensions, while information-integration category learning benefited from a focus on the irrelevant dimension. The researchers concluded that the focus on relevant dimensions aided rule-based categorization because it helped the process of selecting and testing the correct rule. Meanwhile, the focus on the irrelevant dimensions helped information-integration learning because working memory was filled with irrelevant information. The authors argued that this disengaged the explicit system because it was focused on irrelevant information, allowing the implicit system to learn the relevant dimensions (Grimm & Maddox, 2013).

Another study focused on the effect that instructions had on rule-based versus information-integration categorization tasks (Rosedahl et al., 2021). Here, the researchers gave participants either explicit instructions telling them how they should categorize or no specific instructions. It was found that explicit instructions detailing the nature of the categories benefited rule-based learning compared to the no specific instructions. This contrasted with the effect on information-integration learning where there were no differences between the instruction condition and the no specific instructions condition. These results confirmed the a priori COVIS prediction that the explicit system would benefit from this abstract information, but the implicit system would not (Rosedahl et al., 2021).

Sanchez et al. (2020) also manipulated attentional instructions and, unlike Grimm and Maddox (2013) and Rosedahl et al. (2021), found that when participants were instructed to attend to all dimensions during rule-based category learning they showed significantly higher accuracy and more optimal strategy use during a later information-integration test than when not instructed to attend to all dimensions. The researchers also manipulated instructions at test that told participants that rules were no longer useful, and an intuitive based strategy should be used. Still, only those that had attentional instructions to focus on all the dimensions showed better accuracy and more optimal strategy use. This suggested that it was selective attention to irrelevant rule dimensions interfering with learning the other dimensions relevant to the implicit task rather than a perseveration of explicit rule-based response strategies at test that caused interference with implicit performance.

The current study, while similar to Grimm and Maddox (2013), and Rosedahl et al. (2021), asks a different question. It asks how the interference between the explicit and implicit system (see Section 5 for more details) occurs. The current study also attempts to resolve the contrasting results of Grimm and Maddox 2013 and Sanchez et al., 2020. Similar to Sanchez et al., (2020), it seeks to understand whether learning of the underlying II structure can occur during rule strategy use. However, the current study directly examines the role that attention to the relevant dimensions plays in whether learning takes place. If attention is crucial, it would suggest that selective attention to the irrelevant dimensions and inattention to relevant dimensions may have produced the impairment of implicit performance in past experiments. This would support the hypothesis that explicit rule focus interferes with the necessary learning for the implicit system to perform well.

This contrasts with the hypothesis that implicit impairment is the result of a rule-based strategy perseveration, despite the rule's suboptimality. According to this hypothesis, the explicit system remains in control of category decision making once a rule strategy is initially found to be optimal, no matter latter feedback (or instructions) that it is no longer optimal. The distinction between these hypotheses focuses upon the locus of the impairment of implicit performance. Impairment may occur during the category learning stage when the category learning systems are gathering information. Thus, the implicit system may not have the necessary information about the category to perform well. On the other hand, implicit category learning may occur normally with all the information it needs for success, but the continuation (perseveration) of rule strategies prevents what the implicit system has learned from being exhibited. In that case, impairment would be during response outputs (category decision making).

8 INTRODUCTION TO CURRENT EXPERIMENT

The current investigation is guided by the COVIS categorization framework. It assumes there is an explicit-declarative system that operates to learn rule-based categorization through hypothesis testing and active decision processes. It assumes there is an implicit-procedural system that integrates multi-dimensional stimuli at a pre-decision stage that incrementally learns to holistically map the stimuli to responses via the mechanism of reward (Ashby & Gott, 1988; Ashby & Maddox, 2005). Additionally, the II tasks that are typically used to engage the implicit learning system is an integral part of the design of this study. This experiment further examines the “second generation” of categorization questions using the multiple systems perspective (Ashby & Maddox, 2005). The “first generation” questions dealt with the viability of the COVIS multiple systems approach itself. Strong evidence supports these first-generation questions (for review see Ashby & Valentin, 2017). Now, the subsequent generation has the duty to understand the relationship between the systems hypothesized by this theoretical approach. In other words, how do the explicit and implicit systems interact? Does using one impair the other’s ability to learn? If so, how? During competition which system dominates category responses? Why? These and other “second generation” questions remain largely unanswered. The current study is designed to illuminate some of these questions.

Research attempting to answer these second-generation questions has generally found interference between systems (Ashby & Crossley, 2010; Crossley & Ashby, 2015; Sanchez et al., 2020). To clearly investigate this interference, it has been necessary to depart from the traditional rule-based partitioning of the perceptual space (see Figure 1). For example, Crossley and Ashby (2015) had three phases in their experiment. In the congruent condition with rule training, the 3 phases had the same II categories. The difference among the phases was how the perceptual space was divided in each phase. In the first phase, the perceptual space was divided along the vertical 50 level and was based on one of the dimensions. Category As were to the left of the vertical 50 level divider and Category Bs were to the right. In Phase 2, the perceptual space was now divided along the horizontal 50 level divider and was based on another dimension. In the first two phases, the researchers excluded a section of the perceptual space for each category to ensure that only rule-based categorization occurs. In Phase 3, the full II category range was available (see Figure 2 for a re-creation of the perceptual space used). This methodology allowed Crossley and Ashby (2015) to examine whether implicit learning occurred while the explicit system was in control. The current study uses a similar methodology to investigate whether implicit learning can occur if attention is directed to the relevant dimensions during a rule-oriented task.

This issue was investigated by comparing implicit learning when one selectively attends to rules about the dimensions that are relevant to the implicit category structure versus when one attends to rules about dimensions that are irrelevant to that structure. By manipulating instruction across three conditions that direct attention to either the two relevant dimensions, two irrelevant dimensions, or no dimensions during a training phase, we examined the primary cause of implicit performance impairment seen after explicit rule focus. We may see that selective attention to irrelevant rule dimensions impairs the learning of other dimensions or it may be that rule strategy perseveration is causing incorrect responding by the explicit system at test. If selective attention impairs implicit learning about other dimensions, then tasks that direct attention successively to rules based on the relevant dimensions will allow participants to learn the II structure significantly better than those who attend to irrelevant dimensions. This finding would suggest that the primary cause of impairment of implicit performance is not learning the relevant perceptual information because of selective attention to the wrong dimensions.

Conversely, if there is not a significant difference between the relevant and irrelevant condition this finding would suggest that incorrect responding caused by RB strategy perseveration impairs implicit performance. Additionally, the condition with no instructions directing participants' attention to specific dimensions serves as a control condition that allows for comparisons of rates of implicit learning when selective attention is clearly employed to specific dimensions versus when participants must determine their own attentional strategies.

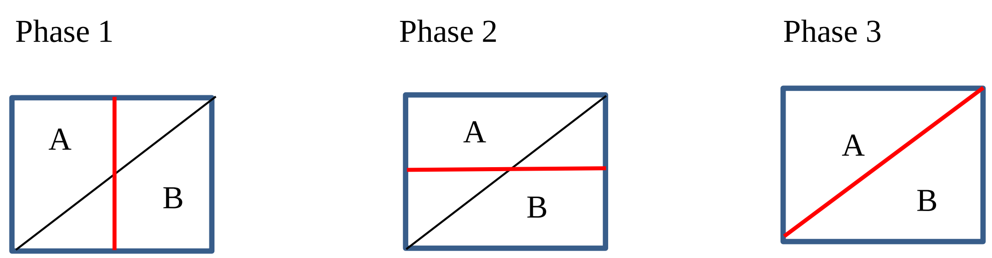

Figure 2 Re-creation of the perceptual space used in Crossley & Ashby (2015)

Note. This is the perceptual space used in each of the three phases of the congruent condition with rule training.

9 METHODS

The current experiment presents an RB category task (training phase) and an II category task (test) using rectangular stimuli that vary along 4 dimensions – proportion of pixel colors, rectangle box size, pixel density, and location of vertical line in the rectangular stimuli. Box size and pixel density are the relevant dimensions comprising the II structure. The irrelevant dimensions are the proportion of pixel colors and vertical line location.

9.1 Participants

One hundred and ten college students with normal or corrected to normal vision were recruited from Georgia State University psychology classes through SONA to participate in an online experiment for partial course credit. Participants were pseudo-randomly assigned to one of three conditions (39 assigned to the relevant, 32 assigned to the irrelevant and 39 assigned to the control condition). The average partial eta-square of relevant studies (Church et al., 2018; Sanchez et al., 2020) was used to calculate the minimum sample size estimate of 21 participants per condition to achieve 80% power using G*Power, and then the target sample size was rounded up to 30 per condition to maximize statistical power. Criteria for exclusion from analyses included not finishing the full experiment (200 training trials; 120 test trials) and showing significant side bias (greater than or equal to 75% choice of only one response). A total of 15 participants were excluded for not finishing the experiment (8 from the relevant condition; 1 from the irrelevant condition; 6 from the control condition). A total of 5 participants were excluded due to a significant side bias (1 from the relevant condition; 1 from the irrelevant condition; 3 from the control condition). A total of 20 participants were excluded from the data analysis leaving a sample of 90 participants, 30 participants in each condition.

9.2 Stimuli

The stimuli consisted of computerized pixelated rectangular boxes created in Turbo pascal 7.0. The rectangles were shown at the center-top of the screen on participants' personal computers for this online experiment. The experiment consisted of screenshots of the stimuli used to create the training and testing phase of the experiment with PsychoPy software (Peirce et al., 2019).

The rectangles varied on 4 different dimensions – the rectangle's box size, the pixel density (more or less pixels), color proportion of light red to cyan pixels and the left to right placement of a yellow vertical line in the rectangle. All 4 dimensions had 101 levels (levels 0-100). Figures 3a and 3b show examples of training stimuli and testing stimuli respectively.

A. B.

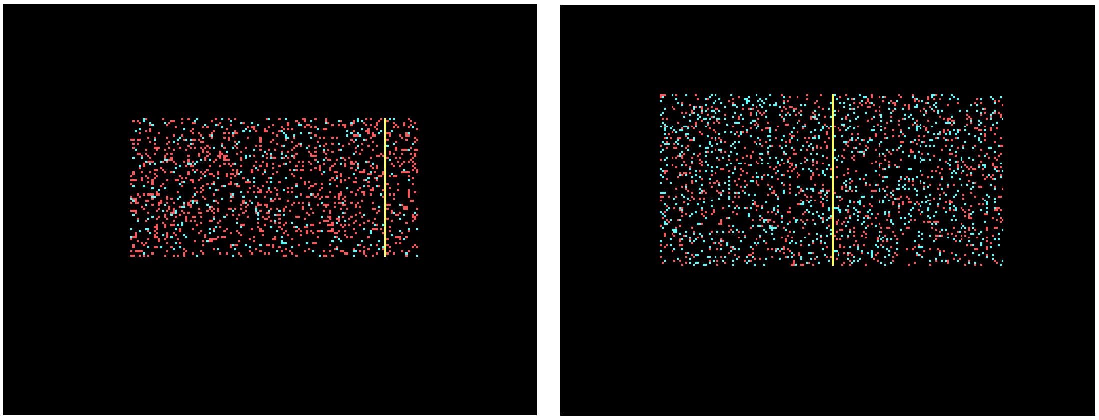

Figure 3 Examples of Training and Testing Stimuli

Note. A. An example of a training stimulus; the stimulus has level 78 pixel color mixture, level 36 rectangle size, level 64 pixel density, and level 89 vertical line location. B. An example of a testing rectangle stimulus. The stimulus has level 50 pixel color mixture, level 76 rectangle size, level 56 pixel density, and level 50 vertical line location.

9.3 Training Stimuli: Relevant Condition

The relevant condition stimuli were created so that the As and Bs for size or pixel density category rules were located within the same perceptual space as the As and Bs of the underlying diagonal II structure. Like Crossley and Ashby (2015), the rule stimuli divided into As and Bs based on their placement on either side of the vertical or horizontal 50-level line. They were relevant to the II structure because both used the same size and pixel density dimensions. This methodology allowed me to examine whether giving positive feedback for relevant dimension rules, that were consistent with the II structure, allowed implicit learning. The idea is implicit learning may occur if it receives useful information from the explicit system.

Whichever of the relevant dimensions was not the focus of the current category rule instructions could range randomly in the perceptual space, but always within the constraints of the II structure. For example, when the relevant rule instructions focused on the size, the stimulus was in category A of the size dimension and category A of the II size/pixel density category but could be in either the A or B space of the density rule. On the other hand, when the relevant rule instructions focused on pixel density, the stimulus shown would be in category A of the pixel density dimension and category A of the II size/pixel density category but could be in either the A or B space of the size rule. The two other irrelevant dimensions, pixel color proportion and vertical line position varied randomly with each having examples above and below the 50 level in the perceptual space in both relevant rule conditions. Figure 4 shows the placement in perceptual space of the relevant dimensions in the II structure, and how the correct rules fell in that perceptual space. For instance, on the X axis, the size rule categories were divided along the vertical 50-level line. Size rule stimuli that were left of the vertical 50-level line are category A (dark blue dots). Size rule stimuli that were right of the vertical 50-level line are category B (dark orange dots). Note that the X-axis rule stimuli could range from 0-50 and 50-100 for category A and B, respectively, as long as they are also within the II diagonal structure. On the Y axis, the pixel density rule categories were divided along the horizontal 50-level line. Pixel density rule stimuli that were above the horizontal 50-level line are category A (light blue dots). Stimuli that were below the horizontal 50-level line in pixel density were category B (light orange dots). Note that the Y-axis rule stimuli could range from 50-100 and 0-50 for category A and B, respectively, as long as they were also within the II diagonal structure. One hundred stimuli met the criteria described for the box size rule instructions and 100 met the criteria for the pixel density rule instructions.

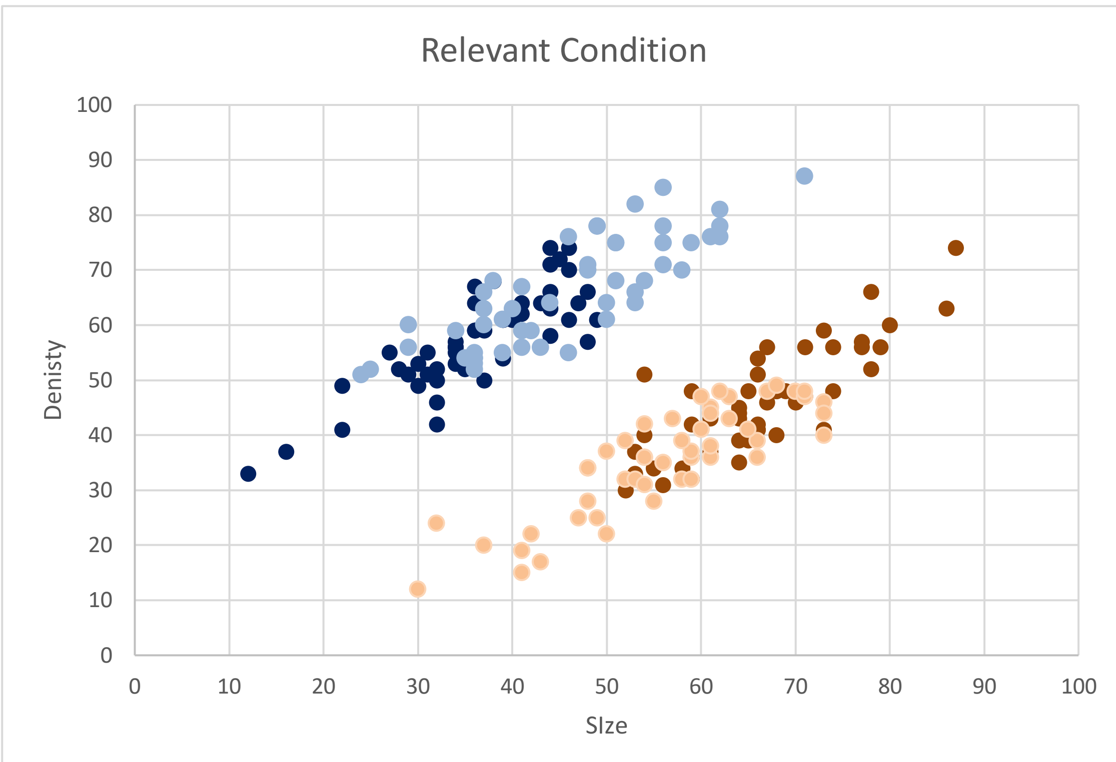

Figure 4 The perceptual space containing the Relevant Condition's stimuli

Note. Size is the X dimension and Density is the Y dimension. Dark blue dots are category A size rule stimuli within the category A II structure. Dark orange dots are category B size rule stimuli within the category B II structure. Light blue dots are category A pixel density rule stimuli within the category A II structure. Light orange dots are category B pixel density rule stimuli within the category B II structure.

9.4 Training Stimuli: Irrelevant Condition

The irrelevant condition stimuli were created so that the As and Bs for proportion of pixel color or vertical line position category rule instructions were also located within the same perceptual space as the As and Bs of the underlying II structure. Whichever of the irrelevant dimensions was not the focus of the current category rule could range randomly above and below the 50 level in the perceptual space for that dimension. For example, when the irrelevant rule instructions focused on the proportion of pixel color, the stimulus was in category A of the proportion of pixel color rule and category A of the II size/pixel density category but could be in either the A or B space of the vertical-line-position rule. On the other hand, when the irrelevant rule instructions focused on a vertical-line position, the stimulus was in category A of the vertical-line position rule and category A of the II size/pixel density category but could be in either the A or B space of the proportion of pixel color rule. Figures 5 shows a graph of the placement of the stimuli used to train rules for color proportion and vertical line in relation to the perceptual space of the II structure. The dark blue dots are the color proportion category A rule stimuli. Figure 5 shows where they were located in the II category A structure. The dark orange dots were category Bs for the color proportion rule stimuli. Figure 5 shows where they were located in the II category B structure. The light blue dots were category As for the vertical line dimension rule, and the light orange dots were Bs for the vertical line rule. Figure 5 also shows where these stimuli were located in the A and B space of II category structure. This stimuli design allowed us to see whether giving positive feedback for irrelevant dimension rules, that were consistent with the II structure, allowed for parallel learning. The rule stimuli in the irrelevant condition ranges more across the II structure simply because the rule dimensions are different dimensions from the dimensions that define the II structure. Therefore, the same vertical and horizontal line restrictions required by the relevant rule are not necessary. One hundred stimuli met the criteria described for the proportion of pixel color rule instructions, and 100 met the criteria for the vertical line position rule instructions.

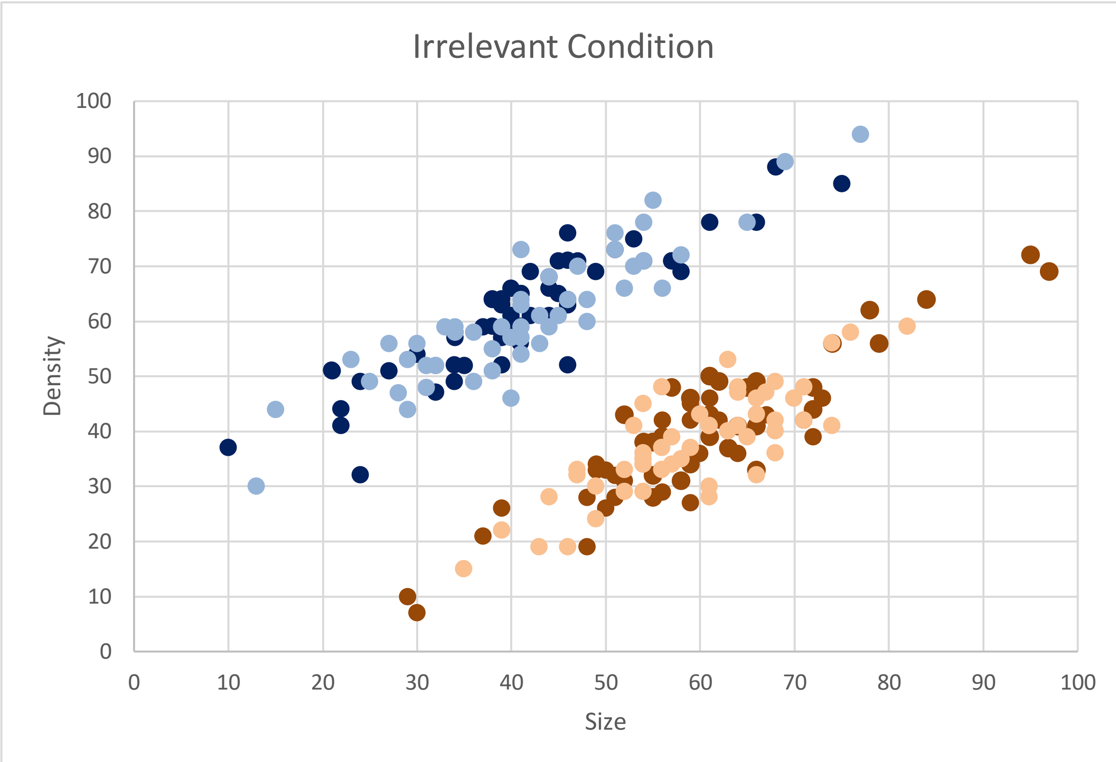

Figure 5 The perceptual space containing the Irrelevant Condition's stimuli

Note. Size is the X dimension and Density is the Y dimension. Dark blue dots are category A color proportion rule stimuli within the category A II structure. Dark orange dots are category B color proportion rule stimuli within the category B II structure. Light blue dots are category A vertical line position rule stimuli within the category A II structure. Light orange dots are category B vertical line rule stimuli within the category B II structure.

9.5 Training Stimuli: Control Condition

Two hundred stimuli were created for the training phase of the control condition as well. All stimuli obeyed a major diagonal II category structure using size and pixel density as the X and Y dimensions respectively (see Figure 6). Twenty-five of the training stimuli also had size levels well below 50. Another 25 had size levels well above 50. Twenty-five had pixel density well above 50, and another 25 had pixel density well below 50. Twenty-five had color mixture levels well above 50, and another 25 had color mixture well below 50. Finally, 25 stimuli had vertical line placements well above 50 and another 25 had vertical line placements well below 50. Half the stimuli were As within the II structure and half were Bs. The stimuli were randomly intermixed. Because simple rules cannot be successful in this presentation, an implicit strategy that learns the major diagonal II category structure based on more holistic stimulus response pairings was optimal.

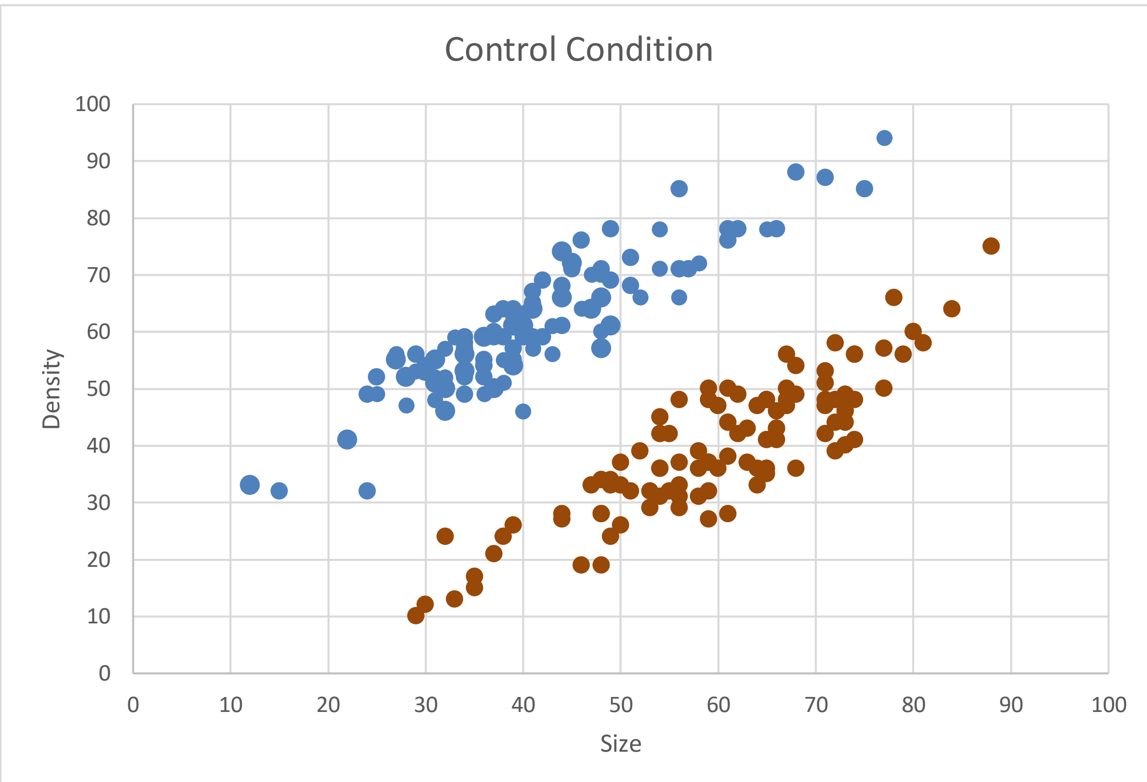

Figure 6 The perceptual space containing the Control Condition's stimuli

Note. The dimensions size and pixel density are the X and Y dimensions respectively (Category A stimuli shown by blue dots; Category B stimuli by orange dots).

9.6 Testing Stimuli

One hundred and twenty stimuli were created for the testing phase. All stimuli used the same major diagonal II structure (size and pixel density) used to create the training stimuli. The two irrelevant dimensions (color mixture and vertical line placement) were kept at a constant level of 50 to remain uninformative to categorization (see Figure 3b for an example). Half the stimuli were As within the II structure and half were Bs.

9.7 Procedure

The experiment had two phases, the training phase, and the test phase. During the training phase, participants saw the 200 stimuli described above. For the relevant and irrelevant conditions, the training phases were segmented into 10 blocks of 20 trials that told participants which dimensions to attend. Each block directly instructed the participant to focus on a single dimension to make their categorization decisions. Before each switch, an instruction screen directed participant's attention to the rule dimension for that block. In the relevant condition, participants switched between focusing on the size rule and the pixel-density rule. In the irrelevant condition, they switched between the color-mixture rule and the line-placement rule. After training, the participants completed 120 test trials. After the test, the experiment ended, and participants were debriefed.

On each trial, the rectangular stimulus appeared on the center-top of the screen. The category options, A and B, were presented below and were aligned to the left and right side of the stimulus, respectively. In between the A and B, a small white cross was presented. Participants were told to press the "S" key on the keyboard for the A category and the "L" key on the keyboard for the B category. Participants received immediate feedback after each trial. Correct responses were followed by green text saying "Correct!" and white text indicated their total accumulated points for the experiment. After incorrect responses, red text said "Incorrect.", and white text indicated total accumulated points. When incorrect, participants received a 4 second timeout before the next trial. Correct responses gained 1 point and incorrect responses lost 1 point.

9.8 Design

A 3x(6) mixed factorial design was used for this experiment. The dependent variable was the proportion of correct category responses in the testing phase. The between-participants independent variable was the attention condition with three levels (relevant, irrelevant, and control), and the within-participant independent variable was block with six levels (Blocks1-6). Each block contained 20 trials.

9.9 Instructions

Instructions before the training phase in the relevant and irrelevant conditions were as follows: “You will see an object and will need to categorize it as either A or B. To help you learn we will tell you which dimension is important during this particular block of trials. At first you will have to guess what belongs in the A or B category but will learn with practice. Press the S key for A responses and the L key for B responses. You will gain 1 point for a correct response and lose a point and receive a 4 second timeout for an incorrect response. Incorrect responses will cost you time to earn points and lengthen the experiment. If you have read and understood the instructions, press y to start the experiment." At the start of each 20-trial block switch the instructions appeared as follows: "Pay attention to the INSERT HERE dimension for optimal performance. Press Y to continue.”

Instructions before the training phase in the control condition were as follows: “You will see an object and will need to categorize it as either A or B. There will not be any simple rules you can use to categorize the objects. At first you will have to guess what belongs in the A or B category but will learn with practice. Press the S key for A responses and the L key for B responses. You will gain 1 point for a correct response and lose a point for an incorrect response along with a 4 second timeout. Incorrect responses could cost you time to earn points and could lengthen the experiment. If you have read and understood the instructions, press y to start the experiment."

At the testing phase, the instructions in all conditions were as follows: “Now we are in the second phase. You will see the same kind of objects as before and you will keep categorizing objects as A or B. The As and Bs still feel like As and Bs. Categorize based on what feels more like an A and what feels more like a B. It is okay to use your intuition. You will gain 1 point for a correct response and lose a point for an incorrect response along with a 4 second timeout. Incorrect responses could cost you time to earn points and could lengthen the experiment. If you have read and understood the instructions, press y to start the experiment.”

9.10 Data Analysis: Accuracy and Learning

A 3x(6) general linear model (GLM) using proportion correct during the testing phase as the dependent measure with Attention Condition (relevant, irrelevant, and control) as the between-subjects independent variable and Block (6 blocks of 20 trials each) as the within-subject independent variable was conducted. This overall analysis was used to examine any main effects and interactions between the variables. Subsequent planned comparison independent t-tests were conducted to examine any significant differences among the three conditions. All statistical tests were two-tailed with an α of .05.

9.11 Data analysis: Strategy Modeling

Strategy modeling was used to determine the strategy used by each participant during the test phase. This was accomplished by comparing the individual’s performance to predicted performance from four different types of strategy models, X-rule, Y-rule, GLC (a diagonal boundary), or guessing. The rule-based models assume that participants apply a categorization decision criterion on a single dimension. For example, participants using a rule-based strategy may set their decision criterion as respond Category A to stimuli that are a small box size and Category B to stimuli that are a large box size. The rule-based models produce either a horizontal or vertical line that best fit the decision criterion set by the participants. This model features a perceptual noise free parameter and a decision criterion free parameter. The GLC model assumes that participants partition the perceptual stimuli space using a linear decision bound. The GLC model uses slope and intercept to create the diagonal linear line that best partitions the participant's category responses. This model features a perceptual noise free parameter, and the slope and intercept of the line as free parameters. The Guessing model assumes the participants randomly guess so both category responses are equally likely. The best fit to each of the participant's overall strategy responses was determined based on the smallest Bayesian Information Criterion (BIC; Schwarz, 1978). This determined how many participants used either a rule-based, an II, or a guessing strategy in each condition. A chi square analysis was used to see whether there were significant differences in the type of strategies used most often by the participants in the different conditions.

9.12 Results: Accuracy and Learning

The 3x(6) GLM examining performance across testing blocks in the three conditions found no significant within-subject main effect of block, $F\ (4.434, 385.742) = 1.405$, $p = .227$, $\eta p^2 = .016$, and no significant interaction between the attention and block conditions, $F\ (8.868, 385.742) = .610$, $p = .786$, $\eta p^2 = .014$, suggesting that participants' performance did not change significantly during the testing phase, and this was true for all the attention conditions. However, the analysis did find a significant between-subjects main effect of attention condition, $F\ (2, 87) = 9.368$, $p < .001$, $\eta p^2 = .177$, reflecting the fact that performance differed based on the attentional manipulation. To fully understand this main effect, planned comparisons were done comparing the following conditions: the relevant versus irrelevant, irrelevant versus control, and relevant versus control. The comparison of relevant versus irrelevant found significantly better performance in the relevant condition t (58) = 4.159, p <.001, d = 1.074 (relevant condition, M = .619, SD = .088: irrelevant condition, M = .535, SD = .066). The comparison of the irrelevant condition versus the control condition (M = .577, SD = .068), found significantly better performance in the control condition, t (58) = 2.436, p =.018, d = .629. The comparison of the relevant condition and the control condition found significantly higher performance in the relevant condition, t (58) = 2.032, p = .047, d = .525. Figure 7 presents the average performance during each of the six 20 trial blocks for the three conditions.

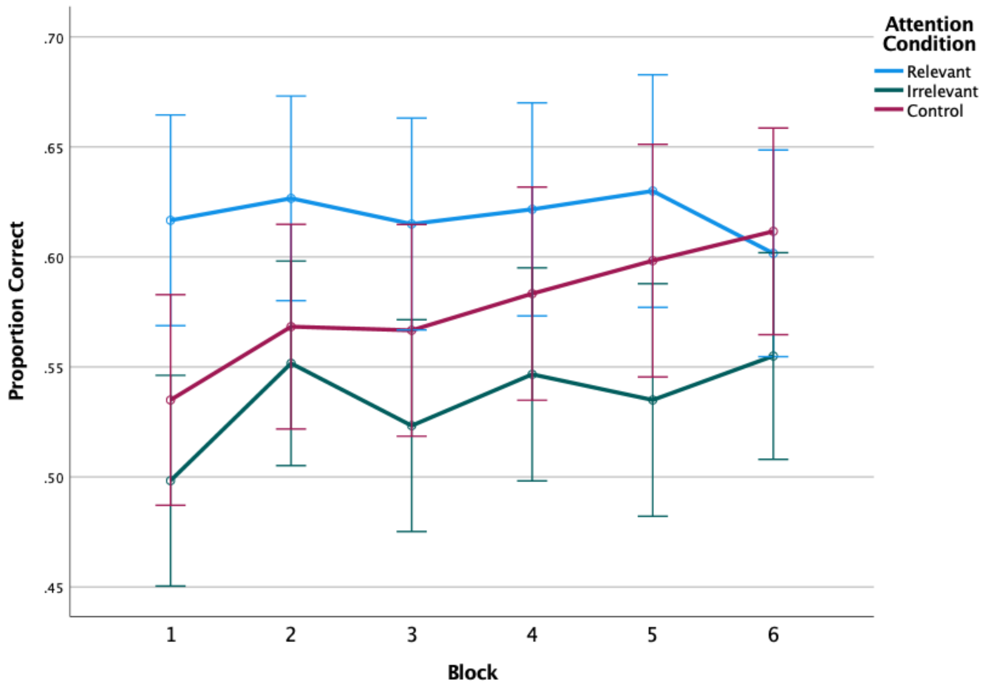

Figure 7 Proportion correct across block for the three attention conditions

Note. Error bars reflect 95% confidence intervals.

9.13 Results: Strategy Modeling

To further understand why the participants in the relevant condition performed better, strategy modeling was used to determine the number of participants in each condition who based their decisions on a rule strategy, an II strategy, or guessing. This analysis helped us to understand whether participants are in fact learning the correct II decisions bounds during explicit system control. It also produced lines depicting the strategy that best partitioned the participant's decision space in each condition. This can be seen in Figure 8. Horizontal and vertical lines suggest rule strategies. Diagonal lines depict II strategies with diagonal lines proceeding from the bottom left corner to the top right corner depicting the major-diagonal boundary that correctly partitioned the categories.

In the relevant condition, eleven participants were best fit by rule strategies, ten were best fit by a guessing strategy and nine were best fit by an II strategy. Out of the nine II strategists, eight showed the appropriate diagonal line. In the irrelevant condition, nine participants were best fit by rule strategies, nineteen were best fit by a guessing strategy and two were best fit by an II strategy. Out of the two II strategists, one showed the appropriate diagonal line. In the control condition, thirteen participants were best fit by rule strategies, fifteen were best fit by a guessing strategy, and two were best fit by an II strategy. Out of the two II strategists, one showed the appropriate diagonal line.

A 3x3 chi square analysis was used to see if there were differences between the attention conditions in the number of participants using the 3 different strategies. The analysis found a significant difference between the conditions X^2 (4, N = 90) = 11.038, p = .026, w = .350. To investigate exactly where the significant differences lay, we examined the adjusted standardized residual for each cell and converted them to probability values. Bonferroni correction was applied to the probabilities to help reduce the risk of Type I error. This analysis revealed that the only cell that was significantly different across the conditions was the II strategy in the relevant condition, p = .002. These results indicated that participants in the relevant condition used significantly more II strategies than participants in the irrelevant or control conditions. More participants who paid attention to the relevant dimensions, even though they were doing so during a rule task, learned the correct II decision boundary than participants who exclusively paid attention to the irrelevant dimensions or had no help directing attention.

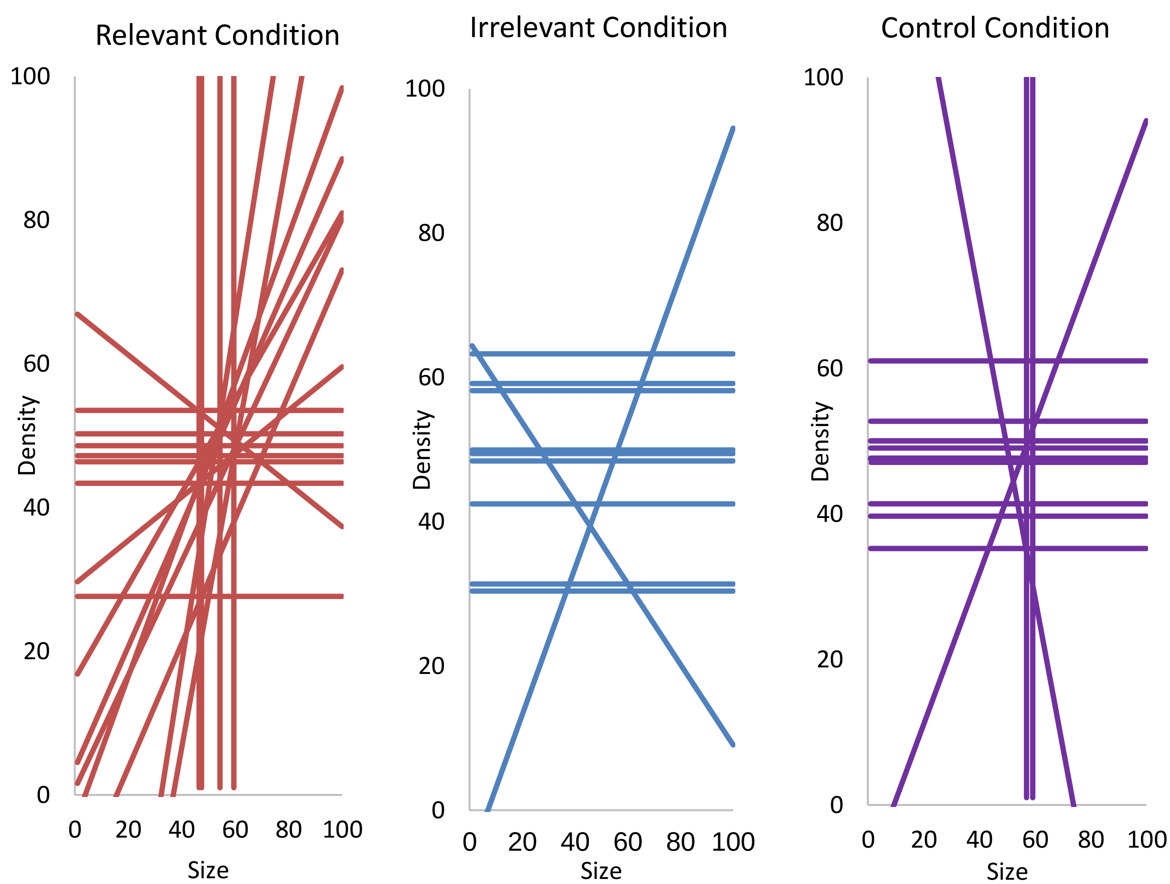

Figure 8 Strategy modeling graphs for each condition

Note. The graphs represent the decision bounds by each participant in each condition. Guessing strategists are not depicted.

10 DISCUSSION

The accuracy data analysis did not find a main effect of block or interaction between the attention conditions and block. This suggests that there was little or no further learning during the testing phase, and this was true in all conditions. However, as predicted by the attentional hypothesis, there was a main effect of condition. Specifically, participants in the relevant condition performed significantly better than both the irrelevant and control conditions. In other words, the relevant condition participants learned the underlying implicit information during the rule-focused learning phase when the explicit system should have been operating. The strategy modeling further confirmed that when participants are directed to pay attention to relevant rule dimensions, they are significantly more likely to learn the correct II decision boundary and use the appropriate II strategy at test than in either of the other conditions.

These findings suggest that, when participants' attention is directed to all the relevant dimensions for later implicit testing, explicit rule use interferes less with implicit learning. As long as attention is directed to the necessary dimensions for the tasks, systems learn in parallel. The significantly worse performance in the irrelevant condition, when compared to both the control and relevant conditions suggest that when selective attention is directed to irrelevant dimensions it significantly harms learning about the relevant dimensions. This is consistent with the hypothesis that selective attention draws attention away from the necessary dimensions for correct implicit learning, and this is the primary mechanism of interference with implicit performance after explicit focus.

Also, because performance in the control condition was actually significantly lower than performance in the relevant condition, rule-strategy perseveration can be ruled out as the primary explanation of the implicit impairment found in this experiment. The rule-strategy perseveration hypothesis clearly predicts that the control condition should have the highest performance, because they were never directed to use rules and their stimuli could never be successfully partitioned by using rules. Therefore, they should be less likely to have a rule strategy to perseverate. This finding clearly allows us to reject the strategy perseveration hypothesis and instead, the results are consistent with the hypothesis that selective attention to inappropriate information for implicit learning is what interferes with implicit performance when people focus on explicit rules.

The significantly better performance in the relevant condition compared to the control condition is also theoretically interesting. Many multiple systems theorists would assume that the participants in the control condition should perform better because it would be most likely to foster an implicit strategy during training that could follow through to the test. However, that is clearly not the case. One possible explanation is that the explicit system may have been used during the control condition to search for rules, despite attempts to foster a more implicit strategy, and the strategy modeling indicated some support for that idea. It may be that because they had no direction about where to attend, they spent a lot of time attending to the irrelevant dimensions. They still performed better than those who spent all their time attending to the Irrelevant dimensions, but not as well as those who always attended to the relevant information even though they did so in explicit rule tasks.

The goal of this study was to determine the primary reason explicit focus during training interferes with later implicit performance. Previous research has produced competing hypotheses about how this occurs (Crossley & Ashby, 2015; Sanchez et al., 2020). One hypothesis assumes that appropriate implicit and explicit learning about the stimuli proceed in parallel but previous successful rule-use causes people to perseverate on a rule strategy. Therefore, they do not use their implicit learning to respond during the test. The other hypothesis assumes that because selective attention during the rule task is directing attention to irrelevant or incomplete information the necessary implicit learning about the stimuli cannot occur. The current study uses rule-oriented training tasks that focused participants' attention to dimensions that were relevant or irrelevant to a later II test. It compared them to each other and to a learning task that tried to focus participants on implicit learning. These comparisons were used to tease these hypotheses apart. The strategy perseveration hypothesis made the strong prediction that the control condition should be better than either rule task, and it should produce more optimal II strategy use. This prediction was clearly falsified. Instead, consistent with the attentional hypothesis the condition focusing participant's attention to both relevant dimensions produced significantly better II test performance and more optimal II strategy use. The fact that the relevant condition generated significantly better performance than both the irrelevant and control conditions indicates that even when the explicit system is operating implicit learning occurs as long as the relevant dimensions to the II category structure are fully attended.

10.1 Significance of Findings

These findings indicate that when the explicit system is in use and is selectively attending to category information irrelevant to the implicit system, the explicit system interferes with implicit learning. In this case, selective attention is the primary mechanism of implicit learning interference. This interference takes place during learning and not during response outputs/decision making. This conclusion is substantiated by the finding that both the relevant and the control condition performed significantly better than the irrelevant condition. If selective attention did not play a role in interfering with later implicit performance, then only the control condition should have generated significantly better performance than the irrelevant condition.

However, because the relevant condition was better than the irrelevant and the control conditions, the category learning systems must be learning in parallel. Further, because the relevant condition generated significantly better performance at test, it is possible that explicit system processes, like selective attention, directly impact shared perceptual inputs that affect both explicit and implicit learning. Not only does this research suggest that the implicit system interference that has been seen in previous research is unlikely to have been caused by strategy perseveration, but it also suggests an important role for attention in both types of learning.

However, Grimm and Maddox (2013) found that directing attention to irrelevant dimensions seemed to allow for better implicit learning and suggested that it minimized explicit system interference. In their study, they examined whether directing attention to a relevant rule dimension or an irrelevant rule dimension affected explicit and implicit learning differently. They gave instructions before a category learning task that told participants a dimension they should attend to. In their first experiment, the optimal category boundary was determined by a conjunctive rule, and in a second experiment it was an II category structure. Grimm & Maddox (2013) argued that an RB strategy would interrupt implicit learning, but this would only occur when the explicit system was focused on a dimension needed for the implicit task (relevant). Their findings seemed to indicate that this idea might be correct. However, the experiment presented here found the opposite result. Specifically, it found that focusing attention on irrelevant dimensions directly interfered with implicit learning. Why the seemingly contradictory findings?

The study presented here used a separate training and testing phase. During the training phase, participants always saw stimuli that were consistent with their rules and the II structure, and only in the test phase were the rules no longer correct. The effects examined in this study happened during this test phase after participants had learned what they would learn about the categories (as indicated by the fact that there is no more significant learning during the test phase). Grimm and Maddox (2013) directed attention to dimensions in a category situation where simple rules were not correct from the start, and when participants tried to use them, they got substantial feedback telling them they were incorrect. Also, the effects in their II experiment happened during the first four blocks of learning after which everyone reached asymptote. This suggests that early in learning directing attention to single dimensional rules based on one of the relevant dimensions for the II structure harmed learning more than directing attention to and punishing responses based on an irrelevant dimension. This punishment for attending to relevant dimensions likely prevented implicit learning from occurring. This is very different from having participants learn rules about relevant versus irrelevant dimensions with an underlying II structure that is receiving positive feedback as well. The current study is better suited to examining whether explicit and implicit learning can operate in parallel because the stimuli allow for both types of learning to receive positive feedback at the same time. The finding from this study that attention to the relevant dimensions is vital for implicit learning to occur suggest a possible alternative explanation for Grimm and Maddox's (2013) finding. Perhaps, directing attention to only one dimension and then punishing performance based on focusing solely on that dimension made people less likely to attend to that dimension, and that interfered with learning compared to when they directed attention away from an irrelevant dimension.

Interestingly, in this study the relevant condition produced significantly better performance than the control condition. This finding may be because of selective attention's role in facilitating implicit learning. Because selective attention is thought to be a function of the explicit system, this raises interesting questions about when and how the explicit system can help implicit learning.

Along a similar idea, Paul and Ashby (2013) assessed how the explicit and implicit systems interacted by simulating II and hybrid categorization tasks (see Ashby & Crossley, 2010). They hoped that by changing different parameters in the categorization model they could tell where and how interaction occurred. They asked whether the two systems were entirely encapsulated from one another and learned from two separate feedback signals or whether they both learned from the same feedback signal. They also asked if the two systems could switch control on a trial-by-trial basis (soft switch) or if once a switch happens the system in question remained in control for an extended period of time (hard switch).

In their first set of simulations, their first model, Model 0, featured full system independence, two separate feedback signals, and the soft switching capabilities that COVIS theory originally assumed. Model 0 was ruled out because it produced good performance in both II and hybrid tasks. However, past research has shown that humans can rarely learn hybrid categories (Ashby & Crossley, 2010).

Their second model, Model 1 featured a shared feedback signal and a hard switching architecture. Here, implicit learning was interrupted until the hard switch occurred and the implicit system could control responses. Paul and Ashby (2013) argued that this model was also not viable because it consistently got to around 75% accuracy when the explicit system was in control during an II task but then dropped to chance after the switch. They reported zero II studies in which this data pattern has occurred. Their third model, Model 2, also featured a single feedback source but had a soft switching architecture. Here the model accounted for learning a lot better as long as the implicit system could control most of the responses. Again, they questioned the viability of this model for humans because there is little evidence that humans can engage easily in soft switching (Paul & Ashby, 2013).

Paul and Ashby (2013) speculated that perhaps the explicit system can actually "train/bootstrap" the implicit system because of the shared feedback source. Therefore, in simulation set 2, they made additional changes to the models' parameters. In these models it was assumed that the feedback information being used by the explicit system is also made available to the implicit system, and it continues to learn. After this change in models 1 and 2 was made, the data patterns more closely resembled that of human patterns. The main takeaway from Paul and Ashby (2013), was that the implicit system could learn while the explicit system was in use because it still had access to the feedback.

The current experiment, to the author's knowledge, provides the only data that seems to suggest that explicit system processes, like selective attention, can directly benefit implicit learning by manipulating its inputs as well. The qualification is that the explicit system must be directing attention to all the relevant input information that is necessary for the implicit system to learn the appropriate stimulus response pairings. This research suggests that the explicit and implicit category systems share not only feedback but also the perceptual inputs that are impacted by selective attention. When the explicit system employs selective attention on irrelevant information, the perceptual inputs to both systems reflect primarily the information of the irrelevant dimensions. However, when attention causes the perceptual inputs to reflect information that is relevant to both systems, the implicit system can learn from this information and use it later.

In summary, modifications from the original hypothesis of COVIS theory seem necessary. The learning systems may be independent, but they clearly share feedback signals and inputs providing many points of interaction. Selective attention can affect both systems by affecting those shared perceptual inputs. The control of selective attention may explain why explicit focus can have such a profound impact on implicit learning, and, as this study shows, that impact can be negative or positive depending on where attention is directed. Here and in Paul and Ashby's (2013) simulations, the implicit system seems to learn from tasks that focus on the explicit system as long as information relevant to the implicit system's later test is attended and appropriate shared feedback is given.

10.2 Future directions

The issue of how category systems interact and the role that selective attention plays could benefit from an evolutionary perspective. Just like memory systems (e.g., Sherry & Schacter, 1987), categorization systems may have an evolutionary function and trajectory, and evolution may tell us something about these systems. A comparative approach could allow one to examine whether selective attention plays a role in associative learning and what differences there may be between species.

Particularly relevant to this, there is evidence that monkeys show the same basic learning patterns in rule-based and information-integration tasks as humans, giving further evidence to the multiple systems perspective. In fact, monkeys (Old World and New World monkeys) show a preference for RB tasks and may be using rule-like processes (e.g., Smith et al., 2012). This is in sharp contrast with other animals such as pigeons (e.g., Castro et al., 2020; Qadri et al., 2019; Smith et al., 2011) and rats (e.g., Broschard et al., 2019) that seem to lack a preference for rule-based categories and learn both RB and II categories at equal rates. Given the approach and method laid out by the current study and past research, it can be assumed that the species that seem to have distinct category systems (monkeys) may show the same results as humans. Their explicit processes (selective attention) may affect their implicit processes (i.e., associative learning). Further, animals (rats and pigeons) that have not shown distinct category systems and few signs of selective attention may not show any differences between having relevant/irrelevant rules. Rats and pigeons may not be able to employ the analytic strategies available to an explicit category system, rather they may approach the stimuli wholistically and learn the II tasks at equal rates regardless of presence of relevant/irrelevant rule focus during pre-training.

Future work should also want to examine other manipulations/variables that may affect the interaction between the category systems. For instance, low and high working memory capacity has been found to affect explicit and implicit process differently (e.g., DeCaro et al., 2008). Could a low or high working memory capacity affect how much implicit learning can occur while the explicit system is in use? Could feedback delays also have an effect on how the two systems can learn in parallel? Past studies have found that feedback delays disrupt implicit learning but not explicit learning (Maddox et al., 2003; Maddox & Ing, 2005; Smith et al., 2014). Could a feedback delay during rule-based training negatively affect the underlying implicit learning even when the explicit system is attending to relevant information for implicit system? Hopefully the current study can provide inspiration for the next steps and questions that are needed to further understand how category learning takes place in humans and animals.

REFERENCES

Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General*, *120*(1), 3–19. https://doi.org/10.1037/0096-3445.120.1.3

Anderson, D. R., Kemler, D. G., & Shepp, B. E. (1973). Selective attention and dimensional learning: A logical analysis of two-stage attention theories. *Bulletin of the Psychonomic Society*, *2*(5), 273–275. https://doi.org/10.3758/bf03329269

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*(3), 442–481. https://doi.org/10.1037/0033-295X.105.3.442

Ashby, F. G., & Crossley, M. J. (2010). Interactions between declarative and procedural-learning categorization systems. *Neurobiology of Learning and Memory*, *94*(1), 1–12. https://doi.org/10.1016/j.nlm.2010.03.001

Ashby, F. G., & Ell, S. W. (2002). Single versus multiple systems of learning and memory. In H. Pashler & J. Wixted (Eds.), *Stevens' handbook of experimental psychology: Methodology in experimental psychology* (pp. 655–691). John Wiley & Sons Inc.

Ashby, F. G., Ell, S. W., & Waldron, E. M. (2003). Procedural learning in perceptual categorization. *Memory & Cognition*, *31*(7), 1114–1125. https://doi.org/10.3758/BF03196132

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), 33–53. https://doi.org/10.1037/0278-7393.14.1.33

Ashby, F. G., Isen, A. M., & Turken, A. U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review*, 106(3), 529–550. https://doi.org/10.1037/0033-295X.106.3.529

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, *56*(1), 149–178. https://doi.org/10.1146/annurev.psych.56.091103.070217

Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences*, *1224*(1), 147–161. https://doi.org/10.1111/j.1749-6632.2010.05874.x

Ashby, F. G., Smith, J. D., & Rosedahl, L. A. (2020). Dissociations between rule-based and information-integration categorization are not caused by differences in task difficulty. *Memory & Cognition*, *48*(4), 541–552. https://doi.org/10.3758/s13421-019-00988-4

Ashby, F. G., Queller, S., & Berretty, P. M. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Perception & Psychophysics*, *61*(6), 1178–1199. https://doi.org/10.3758/BF03207622

Ashby, F. G., & Valentin, V. V. (2017). Multiple systems of perceptual category learning: Theory and cognitive tests. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 157–188). Elsevier Academic Press. https://doi.org/10.1016/B978-0-08-101107-2.00007-5

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: II*. Academic Press. https://doi.org/10.1016/S0079-7421(08)60422-3

Blair, M. R., Watson, M. R., Walshe, R. C., & Maj, F. (2009). Extremely selective attention: Eye-tracking studies of the dynamic allocation of attention to stimulus features in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(5), 1196–1206. https://doi.org/10.1037/a0016272

Broschard, M. B., Kim, J., Love, B. C., Wasserman, E. A., & Freeman, J. H. (2019). Selective attention in rat visual category learning. *Learning & Memory*, *26*(3), 84–92. https://doi.org/10.1101/lm.048942.118

Broadbent, D. E. (1958). *Perception and communicatio*n. Pergamon Press. https://doi.org/10.1037/10037-000

Brooks, D. N., & Baddeley, A. D. (1976). What can amnesic patients learn? *Neuropsychologia*, *14*(1), 111–122. https://doi.org/10.1016/0028-3932(76)90012-9

Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch, & B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–211). Erlbaum Associates.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking.* John Wiley and Sons.

Castro, L., Savic, O., Navarro, V., Sloutsky, V. M., & Wasserman, E. A. (2020). Selective and distributed attention in human and pigeon category learning. *Cognition*, *204*, 104350. https://doi.org/10.1016/j.cognition.2020.104350

Church, B. A. Valdez, G. E., Boomer, & J. & Smith, J. D. (November 2018). *Implicit and explicit category learning: Independence or competition?* Poster presented at the 59th Annual Meeting of the Psychonomic Society. New Orleans, LA.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, *210*(4466), 207–210. https://doi.org/10.1126/science.7414331

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294. https://doi.org/10.1037/0096-3445.104.3.268

Crossley, M. J., & Ashby, F. G. (2015). Procedural learning during declarative control. *Journal of Experimental Psychology: Learning, memory, and cognition*, *41*(5), 1388–1403. https://doi.org/10.1037/a0038853

DeCaro, M. S., Thomas, R. D., & Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition*, *107*(1), 284–294. https://doi.org/10.1016/j.cognition.2007.07.001

Ell, S. W., & Ashby, F. G. (2006). The effects of category overlap on information-integration and rule-based category learning. *Perception & Psychophysics*, *68*(6), 1013–1026. https://doi.org/10.3758/bf03193362

Ell, S. W., Cosley, B., & McCoy, S. K. (2011). When bad stress goes good: Increased threat reactivity predicts improved category learning performance. *Psychonomic Bulletin & Review*, *18*(1), 96–102. https://doi.org/10.3758/s13423-010-0018-0

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*(2), 107–140. https://doi.org/10.1037/0096-3445.127.2.107

Foerde, K., Knowlton, B. J., Poldrack, R. A., & Smith, E. E. (2006). Modulation of competing memory systems by distraction. *PNAS Proceedings of the National Academy of Sciences* *of the United States of America*, *103*(31), 11778–11783. https://doi.org/10.1073/pnas.0602659103

Folstein, J. R., Palmeri, T. J., & Gauthier, I. (2013). Category learning increases discriminability of relevant object dimensions in visual cortex. *Cerebral Cortex*, *23*(4), 814–823. https://doi.org/10.1093/cercor/bhs067

Geeraerts, D. (1989). Introduction: Prospects and problems of prototype theory. *Linguistics*, *27*(4), 587-612. https://doi.org/10.1515/ling.1989.27.4.587

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, *117*(3), 227–247. https://doi.org/10.1037/0096-3445.117.3.227

Goldstone, R. L., Kersten, A., & Carvalho, P. F. (2018). Categorization and concepts. In S. Thompson-Schil (Ed.), *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th ed., pp. 275-317). Wiley. https://doi.org/10.1002/9781119170174.epcn308

Grimm, L. R., & Maddox, W. T. (2013). Differential impact of relevant and irrelevant dimension primes on rule-based and information-integration category learning. *Acta Psychologica*, *144*(3), 530–537. https://doi.org/10.1016/j.actpsy.2013.09.005

Helie, S., & Ashby, F. G. (2012). Learning and transfer of category knowledge in an indirect categorization task. *Psychological Research*, *76*(3), 292–303. https://doi.org/10.1007/s00426-011-0348-1

Hélie, S., Turner, B. O., Crossley, M. J., Ell, S. W., & Ashby, F. G. (2017). Trial-by-trial identification of categorization strategy using iterative decision-bound modeling.

Behavior Research Methods, 49(3), 1146–1162. https://doi.org/10.3758/s13428-016-0774-5

Herscovitch, J., Stuss, D., & Broughton, R. (1980). Changes in cognitive processing following short-term cumulative partial sleep deprivation and recovery oversleeping. *Journal of Clinical Neuropsychology, 2*(4), 301–319. https://doi.org/10.1080/01688638008403802

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93*(4), 411–428. https://doi.org/10.1037/0033-295X.93.4.411

Hull, C. L. (1920). Quantitative aspects of evolution of concepts: An experimental study. *Psychological Monographs, 28*(1), i–86. https://doi.org/10.1037/h0093130

Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In Gordon H. Bower (Eds.), *Psychology of Learning and Motivation*, Vol 18, (pp. 1–47). Academic Press https://doi.org/10.1016/s0079-7421(08)60358-8

James, W. (1890/1950). *The principles of psychology*. Dover.

Johansen, M. K., & Palmeri, T. J. (2002). Are there representational shifts during category learning? *Cognitive Psychology*, *45*(4), 482–553. https://doi.org/10.1016/S0010-0285(02)00505-4

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Kemler Nelson, D. G. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior, 23*(6), 734–759. https://doi.org/10.1016/S0022-5371(84)90442-0

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning & Memory, 1*(2), 106–120. https://doi.org/10.1101/lm.1.2.106

Köhler, W. (1938). Simple structural functions in the chimpanzee and in the chicken. In W. D. Ellis (Eds.), *A source book of Gestalt psychology* (pp. 217–227). Kegan Paul, Trench, Trubner & Company. https://doi.org/10.1037/11496-018

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*(1), 22–44. https://doi.org/10.1037/0033-295X.99.1.22

Kruschke, J. K. (1996). Dimensional relevance shifts in category learning. *Connection Science, 8(*2), 225–247. https://doi.org/10.1080/095400996116893

Kurtz, K. J. (2007). The divergent autoencoder (DIVA) model of category learning. *Psychonomic Bulletin & Review*, *14*(4), 560–576. https://doi.org/10.3758/BF03196806

Kurtz, K. J. (2015). Human category learning: Toward a broader explanatory account. In B. H. Ross (Eds), *Psychology of Learning and Motivation* Vol 63, (pp.77–114). Academic Press. https://doi.org/10.1016/bs.plm.2015.03.001

Lashley, K. S. (1942). An examination of the "continuity theory" as applied to discriminative learning. *Journal of General Psychology*, *26*, 241–265. https://doi.org/10.1080/00221309.1942.10545168

Le Pelley, M. E., Newell, B. R., & Nosofsky, R. M. (2019). Deferred feedback does not dissociate implicit and explicit category-learning systems: Commentary on Smith et al (2014). *Psychological Science, 30*(9), 1403–1409. https://doi.org/10.1177/0956797619841264

Levine, M. (1975). *A cognitive theory of learning: Research on hypothesis testing.* Lawrence Erlbaum.

Little, D. R., Lewandowsky, S., & Heit, E. (2006). Ad hoc category restructuring. *Memory & Cognition, 34*(7), 1398–1413. https://doi.org/10.3758/BF03195905

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, *111*(2), 309–332. https://doi.org/10.1037/0033-295X.111.2.309

Lupien, S. J., Maheu, F., Tu, M., Fiocco, A., & Schramek, T. E. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and Cognition*, *65*(3), 209–237. https://doi.org/10.1016/j.bandc.2007.02.007

Maddox, W. T., Ashby, F. G., Ing, A. D., & Pickering, A. D. (2004). Disrupting feedback processing interferes with rule-based but not information-integration category learning. *Memory & Cognition*, *32*(4), 582–591. https://doi.org/10.3758/BF03195849

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), 650–662. https://doi.org/10.1037/0278-7393.29.4.650

Maddox, W. T., Filoteo, J. V., Hejl, K. D., & Ing, A. D. (2004). Category Number Impacts Rule-Based but Not Information-Integration Category Learning: Further evidence for dissociable category-learning systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 227–245. https://doi.org/10.1037/0278-7393.30.1.227

Maddox, W. T., Glass, B. D., Wolosin, S. M., Savarie, Z. R., Bowen, C., Matthews, M. D., & Schnyer, D. M. (2009). The effects of sleep deprivation on information-integration categorization performance. *Sleep: Journal of Sleep and Sleep Disorders Research, 32*(11), 1439–1448. https://doi.org/10.1093/sleep/32.11.1439

Maddox, W. T., Glass, B. D., O'Brien, J. B., Filoteo, J. V., & Ashby, F. G. (2010). Category label and response location shifts in category learning. *Psychological Research, 74*(2), 219–236. https://doi.org/10.1007/s00426-009-0245-z

Maddox, W. T., & Ing, A. D. (2005). Delayed feedback disrupts the procedural-learning system but not the hypothesis-testing system in perceptual category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 100–107. https://doi.org/10.1037/0278-7393.31.1.100

Mathy, F., & Feldman, J. (2009). A rule-based presentation order facilitates category learning. *Psychonomic Bulletin & Review*, *16*(6), 1050–1057. https://doi.org/10.3758/PBR.16.6.1050

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*(3), 207–238. https://doi.org/10.1037/0033-295x.85.3.207

Mervis, C. B., Catlin, J., & Rosch, E. (1976). Relationships among goodness-of-example, category norms, and word frequency. *Bulletin of the Psychonomic Society, 7*(3), 283–284. https://doi.org/10.3758/BF03337190

Minda, J. P., & Smith, J. D. (2001). Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(3), 775–799. https://doi.org/10.1037/0278-7393.27.3.775

Minda, J. P., & Smith, J. D. (2002). Comparing prototype-based and exemplar-based accounts of category learning and attentional allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28(*2), 275–292. https://doi.org/10.1037/027-87393.28.2.275

Nadler, R. T., Rabi, R., & Minda, J. P. (2010). Better mood and better performance: Learning rule-described categories is enhanced by positive mood. *Psychological Science*, *21*(12), 1770–1776. https://doi.org/10.1177/0956797610387441

Newell, B. R., Dunn, J. C., & Kalish, M. (2011). Systems of category learning: Fact or fantasy? In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 167–215). Elsevier Academic Press.

Nomura, E. M., Maddox, W. T., Filoteo, J. V., Ing, A. D., Gitelman, D. R., Parrish, T. B., Mesulam, M.-M., & Reber, P. J. (2007). Neural Correlates of Rule-Based and Information-Integration Visual Category Learning. *Cerebral Cortex, 17*(1), 37–43. https://doi.org/10.1093/cercor/bhj122

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115*(1), 39–57. https://doi.org/10.1037/0096-3445.115.1.39

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), 87–108. https://doi.org/10.1037/0278-7393.13.1.87

Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of "multiple-system" phenomena in perceptual categorization. *Psychonomic Bulletin & Review*, *7(*3), 375–402. https://doi.org/10.1007/BF03543066

Nosofsky, R. M., & Palmeri, T. J. (1998). A rule-plus-exception model for classifying objects in continuous-dimension spaces. *Psychonomic Bulletin & Review, 5(*3), 345–369. https://doi.org/10.3758/BF03208813

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*(1), 53–79. https://doi.org/10.1037/0033-295X.101.1.53

Nosofsky, R. M., & Zaki, S. R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science, 9*(4), 247–255. https://doi.org/10.1111/1467-9280.00051

Paul, E. J., & Ashby, F. G. (2013). A neurocomputational theory of how explicit learning bootstraps early procedural learning. *Frontiers in Computational Neuroscience*, 7, 177. https://doi.org/10.3389/fncom.2013.00177

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Poldrack, R. A., & Foerde, K. (2008). Category learning and the memory systems debate. *Neuroscience & Biobehavioral Reviews, 32*(2), 197–205. https://doi.org/10.1016/j.neubiorev.2007.07.007

Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. E. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology, 13*(4), 564–574. https://doi.org/10.1037/0894-4105.13.4.564

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*(3), 353–363. https://doi.org/10.1037/h0025953

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325

Qadri, M. A. J., Ashby, F. G., Smith, J. D., & Cook, R. G. (2019). Testing analogical rule transfer in pigeons (*Columba livia*). *Cognition*, *183*, 256–268. https://doi.org/10.1016/j.cognition.2018.11.011

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology, 3*(3), 382–407. https://doi.org/10.1016/0010-0285(72)90014-X

Rehder, B., & Hoffman, A. B. (2005a). Eyetracking and selective attention in category learning. *Cognitive Psychology, 51*(1), 1–41. https://doi.org/10.1016/j.cogpsych.2004.11.001

Rehder, B., & Hoffman, A. B. (2005b). Thirty-something categorization results explained: Selective attention, eyetracking, and models of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(5), 811–829. https://doi.org/10.1037/0278-7393.31.5.811

Restle, F. (1962). The selection of strategies in cue learning. *Psychological Review, 69*(4), 329–343. https://doi.org/10.1037/h0044672

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist*, *45*(9), 1043–1056. https://doi.org/10.1037/0003-066X.45.9.1043

Rosch, E. H. (1973a). Natural categories. *Cognitive Psychology, 4*(3), 328–350. https://doi.org/10.1016/0010-0285(73)90017-0

Rosch, E. H. (1973b). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development, and the acquisition of language* (pp. 111-144). Academic Press.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*(4), 573–605. https://doi.org/10.1016/0010-0285(75)90024-9

Rosedahl, L. A., Serota, R., & Ashby, F. G. (2021). When instructions don't help: Knowing the optimal strategy facilitates rule-based but not information-integration category learning. *Journal of Experimental Psychology: Human Perception and Performance, 47*(9), 1226–1236. https://doi.org/10.1037/xhp0000940

Sanchez, A., Church, B., Boomer, J. Zakrzewski, A., & Smith, J. (November 2020*). Why does the presence of a rule suppress learning of correlated information-integration structures: The role of selective attention.* Poster presented at the 61st Annual Meeting of Psychonomic Society, Virtual Conference.

Schroeder, J. P., Wingard, J. C., & Packard, M. G. (2002). Post-training reversible inactivation of hippocampus reveals interference between memory systems. *Hippocampus, 12*(2), 280–284. https://doi.org/10.1002/hipo.10024

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464. http://www.jstor.org/stable/2958889

Seger, C. A., & Miller, E. K. (2010). Category learning in the brain. *Annual Review of Neuroscience, 33*(1), 203–219. https://doi.org/10.1146/annurev.neuro.051508.135546

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied, 75*(13), 1–42. https://doi.org/10.1037/h0093825

Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems. *Psychological Review, 94*(4), 439–454. https://doi.org/10.1037/0033-295X.94.4.439

Smith, E. E., & Grossman, M. (2008). Multiple systems of category learning. *Neuroscience and Biobehavioral Reviews, 32*(2), 249–264. https://doi.org/10.1016/j.neubiorev.2007.07.009

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Harvard University Press.

Smith, J. D. (2002). Exemplar theory's predicted typicality gradient can be tested and disconfirmed. *Psychological Science, 13*(5), 437–442. https://doi.org/10.1111/1467-9280.00477

Smith, J. D., Ashby, F. G., Berg, M. E., Murphy, M. S., Spiering, B., Cook, R. G., & Grace, R. C. (2011). Pigeons' categorization may be exclusively nonanalytic. *Psychonomic Bulletin & Review, 18*(2), 414–421. https://doi.org/10.3758/s13423-010-0047-8

Smith, J. D., Berg, M. E., Cook, R. G., Boomer, J. Crossley, M. J., Murphy, M. S., Spiering, B., Beran, M. J., Church, B. A., Ashby, F. G., & Grace. R. C. (2012). Implicit and explicit categorization: A tale of four species. *Neuroscience and Biobehavioral Reviews*, *36*(10), 2355-2369. https://psycnet.apa.org/doi/10.1016/j.neubiorev.2012.09.003

Smith, J. D., Boomer, J., Zakrzewski, A. C., Roeder, J. L., Church, B. A., & Ashby, F. G. (2014). Deferred feedback sharply dissociates implicit and explicit category learning. *Psychological Science, 25*(2), 447–457. https://doi.org/10.1177/0956797613509112

Smith, J. D., & Church, B. A. (2018). Dissociable learning processes in comparative psychology. *Psychonomic Bulletin & Review, 25*(5), 1565–1584. https://doi.org/10.3758/s13423-017-1353-1

Smith, J. D., & Church, B. A. (2021). A dissociative framework for understanding same-different conceptualization. *Current Opinion in Behavioral Sciences*, *37*, 13–18. https://doi.org/10.1016/j.cobeha.2020.06.004

Smith, J.D., Church B.A., Jackson B.N., & Sanchez A.F. (2022). Categorization. In J. Vonk & T. Shackelford (Eds.) *Encyclopedia of Animal Cognition and Behavior*. (pp. 1083-1103). Springer, Cham. https://doi.org/10.1007/978-3-319-47829-6_1746-1

Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(6), 1411–1436. https://doi.org/10.1037/0278-7393.24.6.1411

Smith, J. D., Zakrzewski, A. C., Johnston, J. J. R., Roeder, J., Boomer, J., Ashby, F. G., & Church, B. A. (2015). Generalization of category knowledge and dimensional categorization in humans (*Homo sapiens*) and nonhuman primates (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Learning and Cognition*, *41*, 322-325. https://psycnet.apa.org/doi/10.1037/xan0000071

Snow, C. E., & Rabinovitch, M. S. (1969). Conjunctive and disjunctive thinking in children. *Journal of Experimental Child Psychology, 7*(1), 1–9. https://doi.org/10.1016/0022-0965(69)90080-0

Soto, F. A., & Ashby, F. G. (2015). Categorization training increases the perceptual separability of novel dimensions. *Cognition*, *139*, 105–129. https://doi.org/10.1016/j.cognition.2015.02.006

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381-403). Academic Press.

Tulving, E. (1985). How many memory systems are there? *American Psychologist, 40*(4), 385–398. https://doi.org/10.1037/0003-066X.40.4.385

Unger, L. & Sloutksy, V. M. (2021). *Category learning is shaped by the multifaceted development of selective attention*. Proceedings of the 43rd Annual Meeting of the Cognitive Science Society, Virtual/Vienna, Austria.

Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin & Review, 8(*1), 168–176. https://doi.org/10.3758/BF03196154

Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans.). Macmillan.

Wolff, J. L. (1967). Concept-shift and discrimination-reversal learning in humans. *Psychological Bulletin, 68*(6), 369–408. https://doi.org/10.1037/h0025189

Zaki, S. R., & Kleinschmidt, D. F. (2014). Procedural memory effects in categorization: Evidence for multiple systems or task complexity? *Memory & Cognition*, *42*(3), 508–524. https://doi.org/10.3758/s13421-013-0375-9

Zeithamova, D., & Maddox, W. T. (2006). Dual-task interference in perceptual category learning. *Memory & Cognition, 34*(2), 387–398. https://doi.org/10.3758/BF03193416

Zeithamova, D., & Maddox, W. T. (2007). The role of visuospatial and verbal working memory in perceptual category learning. *Memory & Cognition*, *35*(6), 1380–1398. https://doi.org/10.3758/BF03193609