12-2024

Assessing Stimulus and Cognitive Control in Adults, Preschool Children, Capuchin Monkeys (Sapajus apella), and Rhesus Macaques (Macaca mulatta) in a Conditional Discrimination Task

Brielle T. James

Georgia State University

Assessing Stimulus and Cognitive Control in Adults, Preschool Children, Capuchin Monkeys (*Sapajus apella*), and Rhesus Macaques (*Macaca mulatta*) in a Conditional Discrimination Task

by

Brielle James

Under the Direction of Michael J. Beran, PhD

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2024

ABSTRACT

Cognitive control has been defined as the use of various executive functions (such as inhibition, attention control, and task shifting) to execute goal-directed behavior. For this reason, cognitive control is considered essential for complex cognition and flexible behavior, especially in the face of response competition or novel circumstances. However, complex behavior has also been explained by behaviorists using associative stimulus-response theories. The current study examined the role of cognitive and stimulus control in problem-solving behavior. Adult humans, preschool children, adult rhesus macaques, and adult capuchin monkeys completed a computerized conditional discrimination task. Participants completed various psychomotor tasks using a specialized cursor to do so. The cursors differed in speed, size, and their ability to move through walls in a way that made each cursor beneficial for two specific tasks. Participants completed three of the six tasks with the specialized cursors to learn through associative experience which cursor was optimal for each task. Those that could not discern this on their own experienced a correction procedure to facilitate learning. Participants who were able to make correct conditional discriminations were then presented with the remaining three tasks to use with the specialized cursors for the first time. With the aid of correction experience, adults learned how to discriminate the cursors based on their functionality during the primary tasks. They also showed optimal use of the cursors during the first three trials of the transfer tasks, suggesting cognitive control over behavior and a conceptual understanding of the tools. Children showed partial success and monkeys showed minimal success in learning to discriminate the cursors based on their functionality during the primary tasks. Neither children nor monkeys showed generalization of choice behavior to the transfer tasks. However, both groups did show limited associative learning of the cursors' optimality during these tasks. The present results highlight the role that both cognitive and stimulus control play in successful choice behavior in a novel conditional discrimination task. As well as their influence on flexible, goal-driven behavior across species and age ranges.

INDEX WORDS: Cognitive control, Discrimination, Capuchin monkeys, Rhesus macaques

Assessing Stimulus and Cognitive Control in Adults, Preschool Children, Capuchin Monkeys (*Sapajus apella*), and Rhesus Macaques (*Macaca mulatta*) in a Conditional Discrimination Task

by

Brielle James

Committee Chair: Michael Beran

Committee: Sarah Barber

Sarah Brosnan

Robert Hampton

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

December 2024

DEDICATION

This dissertation is dedicated to my mother and father – my two biggest fans! They have supported me in every way possible and always believed I was capable of more than any of us could imagine. This achievement would not have been possible without their unyielding love, guidance, faith, encouragement, and pride in me.

In Loving Memory of

Larry Donell James

June 4, 1954 – February 3, 2023

ACKNOWLEDGEMENTS

First, I would like to acknowledge my husband, Jamahri Gaines. He has loved, supported, and cared for me in countless ways as I navigated graduate school. Without his encouragement and support, I would not be where I am today. He walked alongside me as a partner, friend, confidant, helping hand, sounding board, and voice of reason. From late nights to upending life events over the years, he has stood beside me in ways only he can, to bring us through to the end of this milestone.

I also need to thank my advisor, Dr. Michael Beran, for his dedication to mentorship throughout our time working together. Without his guidance, advisement, and understanding, I would not have achieved all that I have as an academic. I would also like to thank the members of my dissertation committee, Dr. Sarah Barber, Dr. Sarah Brosnan, and Dr. Robert Hampton, for their feedback, support, and understanding throughout this dissertation process.

I want to acknowledge, as well, my colleagues and friends at Georgia State University's Language Research Center. Without their commitment to research, along with animal welfare, I would not have been able to conduct the studies that established my academic career. I also would not have had as much fun over the years without our many gatherings and lighthearted moments together. I am also thankful for the post-docs, post-bacs, and graduate students of the Comparative Intelligence and Cognition Lab for their support through data collection and life as a graduate student.

Over the course of my graduate career, I have formed relationships with fellow students that have turned into genuine friendships that extend well beyond the constraints of school. They have supported me through both my darkest and happiest days in ways that I will never be able to say thank you enough for. They are my found family in all the ways that matter. I must also acknowledge my family, friends, and sorority sisters that have been on this journey with me before and during graduate school. They have always had faith in me, especially on the days when I did not have faith in myself. These are people that have kept me grounded, kept life in perspective, and brought joy to my life in countless ways.

Thank you all! I am beyond grateful for each of you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 COGNITIVE CONTROL

For more than half a century, researchers have offered many different theories and definitions of cognitive control. As a construct, cognitive control has been a foundational and long-standing principle of cognitive psychology even if theorists have not converged on a singular definition (Cohen, 2017). No matter the specific definition of what cognitive control is, it is accepted that the ability to regulate one's internal cognitive state is required for complex cognition, such as critical thinking, reasoning, planning, problem-solving, and decision-making, especially in the face of different or new circumstances (e.g., Diamond, 2013; Miller & Cohen, 2001; Nigg, 2017). Research has shown that the complex cognitive functions that constitute cognitive control rely on the prefrontal cortex (PFC) of the brain (e.g., Diamond, 2002; Engle & Kane, 2004; Kane & Engle, 2002; Miller, 2000; Miller & Cohen, 2001; Miyake et al., 2000; Nigg, 2017), as well as the anterior cingulate cortex (Engle & Kane, 2004; Nigg, 2017; for a review, see Jurado & Rosselli, 2007).

Unquestionably, complex cognition is important for daily functioning (Diamond, 2013; Miyake & Friedman, 2012), and cognitive control guides adaptive behavior in everyday situations (Duverne & Koechlin, 2017). The utility of this capacity is evident in the ability to solve a mental math problem, follow directions, or complete a long-term project at school or work. Indeed, self-regulation of cognition is related to numerous life outcomes, including intelligence (e.g., Kane & Engle, 2002; Shipstead et al., 2016), mental health, physical health, quality of life and relationships, school success, occupational success, and development (for reviews, see Diamond, 2013; Meier & Kane, 2017; Miyake & Friedman, 2012; Nigg, 2017). Interestingly, it is the capacity for these high-level cognitive processes that some people believe makes humans unique from other species (e.g., Cohen, 2017; Diamond, 2013; Duverne & Koechlin, 2017; Nigg, 2017).

A consensus in the definition of cognitive control as a construct in psychology is imperative given cognitive control's influence on a wide range of domains across human functioning. As research in different areas grows, understanding what each other means when talking about "cognitive control" is important to ensure we are interpreting information about the appropriate aspect(s) of this capacity. A better understanding of the mechanisms that are responsible for complex cognition will facilitate a better understanding of intelligent, goal-directed behavior, and improve interventions to circumvent behavior's susceptibility to irrational influences and failure (Cohen, 2017).

1.1 History of Cognitive Control

Given the importance of cognitive control, the historical development of cognitive control theories provides an important perspective when conducting research on this topic. Gardner and colleagues (e.g., Gardner, 1961; Gardner et al., 1968; Gardner & Long, 1960) identified several principles of cognitive control: conceptual differentiation in categorizing, field articulation, scanning, and level-sharpening. Conceptual differentiation is the ability to form abstract concepts (Gardner et al., 1968). Field articulation is one's ability to manage relevant and irrelevant cues through attention control (Gardner, 1961; Gardner & Long, 1960). Scanning, which is relevant for perceptual learning, is reflected in how many looks in a free scanning task a person habitually uses to make a perceptual judgment (Gardner & Long, 1960). The extensiveness of scanning varies across individuals and was conceptualized as the scanning of external fields (cues) and internal fields (ideas). Lastly, Gardner and Long describe level-sharpening as the degree to which related percepts and memory traces begin to assimilate. As a result of this assimilation, memories and perceptual experiences become undifferentiated.

Conceptual differentiation has been tested using free categorizing tests of objects, persons, and behavior, showing that the number of groups that subjects can form differs based on their ability for conceptual differentiation (Gardner et al., 1968). Those individuals who are considered low in conceptual differentiation in categorizing have more abstract concepts compared to those high in conceptual differentiation. The latter group can produce more categories for objects than for people. Individuals that are high in field articulation showed less susceptibility to cognitive illusions (Gardner, 1961; Gardner & Long, 1961), less retroactive interference (Gardner & Long, 1960), and better selective attention during a free association test (Gardner, 1961) allowing them to produce words more related to the stimulus words than subjects low in field articulation. The latter group produced fewer overall associations and reported frequent "blocking," which suggested that they were more distractable. Gardner and Long (1960) observed differences between extensive and limited scanners in attention control (measured by recording eye movements) when viewing optical illusions and during a size judgment task. More specifically, when viewing a succession of 150 squares that gradually increased in size, "sharpeners" were able to keep pace with the increasing size and maintain high accuracy in size judgments and rankings. "Levelers" struggled to do the same, presumably because they were less able to differentiate the successive squares as a result of new squares assimilating with an aggregate memory trace of the previously shown smaller squares. Once again, Gardner and Long saw differences in retroactive interference and intrusion errors between those high and low on level-sharpening.

Cognitive control, according to Gardner and colleagues (e.g., Gardner, 1961; Gardner et al., 1968; Gardner & Long, 1960), is a person's use of these principles to guide their attention appropriately to produce adaptive behavioral responses based on the situation. They asserted that an individual's tendency for certain strategies are stable components, vary across individuals, and comprise the "adaptive style" or "personality" of their cognitive functioning.

Also interested in individual differences in cognition, Hammond and Summers (1972) identified cognitive control as a separate and independent process from knowledge acquisition. According to Hammond and Summers, cognitive control is how knowledge is applied and used in various situations, particularly during learning, judgment, and interpersonal interactions. They argued that due to the independence between knowledge and control, poor task performance can result from having either a perfectly controlled cognitive system but not enough knowledge or perfect knowledge acquisition processes but imperfect control. In other words, two people can present with identical performances but for different mechanistic reasons. Using a multiple-cue probability learning task, Hammond and Summers demonstrated empirical dissociation of knowledge and control. Participants completed either a hard (nonlinear cue relationship) or easy (linear cue relationship) inference task. The researchers provided the participants with the knowledge needed for accurate performance in the tasks (i.e., they explained the cue relationships) to equate knowledge acquisition between the two tasks. Given the difference in difficulty (and therefore also the ease with which participants would be able to exercise cognitive control), participants showed lower performance in the hard task compared to the easy task, despite knowledge being the same. Interestingly, by the end of testing, control performance was equal to knowledge for participants completing the easy task, but performance was lower than knowledge for participants completing the hard task. This dissociation between knowledge and control demonstrates that even when knowledge is complete, imperfect cognitive control can prevent high achievement.

The concept of cognitive control was also heavily influenced by the dual-process concept of attention put forth by Posner and Snyder (1975), who made the distinction between automatic and conscious processing. Posner and Snyder argued that automatic processes are the automatic activation of a previously learned response. From their viewpoint, automatic processes occur (1) without intention, (2) in the absence of conscious awareness, and (3) without interfering with other concurrent mental processes. It is important to note that the latter point is referring to effects on processing, not behavioral responses as a result of that processing. Posner and Snyder referenced studies of the Stroop effect as primary evidence of the lack of intention required for automatic processing. The Stroop task (Stroop, 1935) creates interference between automatic processing and controlled processing. Traditionally, this task requires participants to say the meaning of color words (i.e., read the word) or name the color that the word is printed in. Typically, participants are easily able to read a list of color words, regardless of the color of the font (e.g., GREEN printed in red, where the correct response would be "green"). When reading the same list, participants make slower verbal responses when naming the color of the letters if the color of the font is incongruous with the color word (e.g., GREEN printed in red color, where the correct response would be "red"). This difficulty is not seen, however, when participants name the color of shapes or symbols, so color naming is not the challenge. These results demonstrate that reading is a highly practiced and automatic response, and therefore incongruous trials in the reading task (but not the color naming task) require controlled processing of attention to override automatic processing. During incongruent trials, automatic responding would lead to incorrect responses, so this task design puts automatic and controlled responding in competition with each other. The Stroop effect, as it has become known, is the slower or incorrect responding to incongruous stimuli due to interference caused by irrelevant and incongruous, but prepotent, stimuli features. For this reason, the Stroop task is commonly considered to be a measure of inhibition.

Using the Stroop effect as an example, Posner and Snyder (1975) identified automatic processes as those that subjects cannot avoid engaging in even if they desire to ignore that aspect of the input. Other tasks also generate evidence of automatic processes. Dichotic listening studies (e.g., Moray, 1959), which require participants to attend to one ear and respond aloud based on what is heard, while ignoring the information that is being played in the other ear, test a person's awareness of unattended information. Posner and Snyder pointed to dichotic listening studies during which words presented to the unattended ear interfered with the shadowing of items in the attended ear or led to a conditioned galvanic skin response as evidence that automatic processing occurs even when attention is focused elsewhere and does not require awareness. They also argued that automatic processes can occur in parallel with other processes, resulting in the simultaneous processing of information without conscious effort or attention to do so. They found that when making auditory or visual classifications, participants' responses were slow but still accurate when responding to an unexpected (and therefore given no or little attention) modality. As with dichotic listening studies, unattended information was still being processed and automatically activating conditioned internal representations.

Automatic processing is held in comparison to the conscious processing system (Posner & Snyder, 1975). This conscious control of attention and cognition is a limited-capacity processing mechanism. Once controlled attention is directed toward an activity there are fewer remaining attention resources available for processing other stimuli. Posner and Snyder reviewed research that supports a cost-benefit analysis to conceptualize the interaction between automatic and conscious processing. They hypothesized that the automatic processing of a stimulus will always occur, which facilitates the processing of that input. Directing conscious attention to the processing of that same stimulus will increase the processing capability for that stimulus. By combining automatic activation with conscious attention participants will demonstrate faster response times and a reduction in errors. However, the commitment of the conscious processor will lead to inhibition (the cost of which is long response times, for example) in the processing of any other signals. Therefore, new stimuli will be able to easily activate automatic processing but will struggle to initiate new conscious, nonhabitual responses because the conscious processing system is already at work. In other words, there is a capacity constraint on controlled processing that does not exist for automatic processing. This cost-benefit relationship is observed when manipulating the validity of a priming cue (including the use of neutral cues) to vary whether conscious attention is directed toward a stimulus. Participants do well (more accurate and faster responses) when their attention is directed to a stimulus (e.g., a priming stimulus cues the target stimulus) but make slower or incorrect responses to unexpected items (e.g., the prime is not a valid cue of the target). Increases in reaction time to a secondary task when attention has been actively (i.e., consciously) directed towards the primary task also support these claims. Overall, Posner and Snyder asserted that conscious processing directs participants' attention to particular stimuli or internal representations in order to reduce interference from outside stimuli and inhibit habitual responses. This maintenance of attention is effortful and of a limited capacity. However, it must be noted that conscious intentions and strategies cannot prevent the activation of automatic processes.

Posner and Synder's (1975) theory of automatic and conscious processing was built upon by Schneider and Shiffrin (1977; Shiffrin & Schneider, 1977), who also recognized automatic and controlled processing as two fundamental information processing modes that can control information management, attention, and behavioral responses. Through the development and review of numerous detection and search tasks, Schneider and Shiffrin identified automatic processing as being habitual and stimulus-driven (internal or external stimuli), while controlled processing is controlled by and through the attention of the subject. Schneider and Shiffrin, like Posner and Snyder, also identified automatic processing as operating in parallel with other processes (i.e., it does not impact the capacity limitations of the overall processing system). As a result of this parallel nature, automatic processing is largely unaffected by dual-task interference. Automatic processing is the result of learned, associative connections in long-term memory that are difficult to suppress or modify. In the context of a search task, for example, automatic processing is seen in the form of automatic detection. This is when a learned target stimulus attracts attention and initiates a particular response automatically, regardless of other stimuli or memory load. This is in contrast to controlled processing, which requires active attention. As a result of this high attention demand, controlled processing has a limited capacity and is easily affected by cognitive load. Instead of automatic detection, in a search task with little practice and/or inconsistent targets (meaning automatic processing is unavailable), controlled processing is seen through controlled searches that use a serial comparison technique to compare each possible target with each presented stimuli until the correct one is found – evidence of a much slower and more effortful strategy compared to automatic detection. Despite its attentional and capacity limitations, controlled processing has the benefit of easily being altered and applied to novel situations for which automatic responses have not been learned. It is actually controlled processing that facilitates the learning of automatic processing. Schneider and Shiffrin (1977) demonstrated that repeated practice of a task requiring controlled processes can become automatic in nature.

Automatic processing and conscious, controlled processing (Posner & Snyder, 1975; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977) are now often described in terms of bottom-up and top-down processing. That is, behavior is said to stem from two discrete levels: a base (“bottom-up”) level responsible for automatic processing and a supervisory (“top-down”) level responsible for controlled processing (e.g., Miller & Cohen, 2001; Nigg, 2017). Specifically, controlled processing acts on the goal-relevant information afforded from automatic processing (e.g., cue information, previously learned associations, priming). This distinction between top-down processing and bottom-up processing is the framework for many current definitions of cognitive control. As previously described for automatic processing, bottom-up processes (BUP) are automatic, stimulus-driven, habitual, and generally independent of mental capacity. This is in contrast to top-down processes (TDP) which are deliberate, sequential, require various cognitive processes, and are capacity limited. TDP are driven by internal goals and representations instead of sensory stimuli. TDP are especially needed in situations where habitual responses have not yet been established, such as when responding to novel problems, or when BUP would lead to incorrect responses, such as in situations that create response conflicts or coactivation. In a review of self-regulation, Nigg (2017) acknowledged that BUP are usually the targets of TDP. TDP can activate, suppress, or bias BUP (i.e., monitor and regulate automatic processes), so theorists often refer to TDP as supervisory processes or executive control. Overall, BUP and TDP are both important functions of the mind and work together to facilitate goal-oriented behavior (e.g., Evans & Stanovich, 2013).

Gardner and Hammond, along with their colleagues, were among the first to operationalize cognitive control as a construct (Gardner, 1961; Gardner & Long, 1960; Gardner et al., 1968; Hammond & Summers, 1972). Posner and Snyder (1975), along with Shiffrin and Schneider (1977; Schneider & Shiffrin, 1977) proposed the distinction between automatic and controlled processing, namely that controlled processing requires (1) effortful attention, (2) is vulnerable to interference from automatic processing, and (3) demonstrates capacity constraints in processing, that pushed cognitive control to the center of research in cognitive psychology (Cohen, 2017). More recently, these concepts of cognitive control have been referred to as bottom-up and top-down processes (Nigg, 2017). Research on cognitive control has made significant progress in identifying the more precise and mechanistic components of this capacity. However, this has led to various theories of what specifically defines cognitive control.

1.2 Current Theories of Cognitive Control

Cognitive control has been defined more recently in several, at times conflicting, ways (see Bugg & Crump, 2012; Nigg, 2017, for a review). Diamond (e.g., 2011, 2013) discussed cognitive control as an umbrella term for three specific cognitive processes. Importantly, instead of "cognitive control," Diamond referred to the crucial top-down processes needed for complex cognition as executive functions. From this perspective, executive functions, executive control, and cognitive control are all synonymous. Diamond (2013) specifically defined executive functions as top-down processes needed to resist habitual responding when purposeful, planned, and goal-directed behavior is needed. Like all the aforementioned controlled processes, executive functions require effortful processing and make flexible behavior possible. Diamond (e.g., 2013) identified three core executive functions that underly complex cognition, which are inhibition, working memory, and cognitive flexibility.

Diamond (2013) stated that inhibitory control consists of response inhibition (i.e., self-control) and interference control (i.e., selective attention and cognitive inhibition). Response inhibition (e.g., Mischel et al., 1989; Rachlin et al., 1991) is defined as maintaining control over behavior in the face of temptation, so as not to act impulsively. Self-control relies on the inhibition of action and prepotent responses. Diamond characterized selective attention (which has also been referred to as focused, controlled, or executive attention) in the same way as Posner (Posner & DiGirolamo, 1998; Posner & Snyder, 1975): as interference control at the level of attention through the intentional, top-down suppression of attention to irrelevant stimuli. As previously described, selective attention exists in direct contrast to bottom-up attention, which is involuntary, automatic, and driven by external stimuli. Lastly, cognitive inhibition is the suppression of prepotent mental representations. Cognitive inhibition is essential for intentional forgetting (e.g., Anderson & Levy, 2009) and resisting proactive and retroactive interference (e.g., Hedden & Park, 2003; Postle et al., 2004).

Diamond (2013) also viewed working memory as an essential cognitive process for problem-solving, and she defined it as the ability to hold and manipulate information in your mind. In other words, working memory includes mentally manipulating information that is no longer perceptually available. Working memory and inhibitory control support each other for optimal functioning (Diamond, 2013). The relationship between the two processes is obvious when you consider that working memory is needed to hold task goals in mind, which will inform what information is relevant and what information needs to be inhibited. This theoretical perspective is supported by a study with 4- and 5-year-old children, in which Bodrova and Leong (2007) found that visual reminders of a previously stated task goal improved children's self-control during a reading activity. At the same time, keeping a task goal in mind requires the inhibition of thoughts and memories (i.e., cognitive inhibition; Diamond, 2013), especially previous thought patterns when trying to find a novel solution to a problem. Mind-wandering is a great example of this aspect of the relationship between inhibitory control and working memory (e.g., Kane et al., 2007). Unsurprisingly, research indicates that working memory and inhibitory control rely on the same limited cognitive capacities (e.g., Wais & Gazzaley, 2011).

The last of the core executive functions according to Diamond (2013) is cognitive flexibility. This is the ability to change or adjust your thinking, especially in the face of new task rules or demands. Cognitive flexibility allows organisms to switch between tasks (i.e., task sets) and various ways of thinking about stimuli or problems (i.e., mental sets). The low-level executive functions (working memory and inhibitory control) combine, resulting in cognitive flexibility, an intermediate-level executive function (Diamond, 2013; Nigg, 2017). Cognitive flexibility requires inhibition to disengage attention from a previous perspective and works closely with working memory to take advantage of the new approach. Interestingly, Diamond argued that these three core executive functions then combine themselves to facilitate the high-level executive functions of reasoning, problem-solving, and planning. According to Diamond (2013), the correlation of individual executive functions with intelligence (e.g., Duncan et al., 2008; Roca et al., 2010) results from fluid intelligence being synonymous with reasoning and problem-solving executive functions.

Researchers such as Friedman and Miyake (e.g., Friedman & Miyake, 2017; Miyake & Friedman, 2012; Miyake et al., 2000) also view cognitive control to be a specific set of closely related, but independent, executive functions that support goal-directed behavior. Miyake et al. (2000) had participants complete various general (e.g., Wisconsin Card Sorting Test, Tower of Hanoi, operation span) and targeted (e.g., local-global task, Stroop task) executive function tasks.

Using latent variables, shifting, updating, and inhibition were identified as the key executive functions that comprised cognitive control. Set shifting is one's ability to allocate and regulate attention in order to shift flexibly between task and/or mental sets, which is crucial for problem solving. This ability is synonymous with attention switching and task switching, and it overlaps with cognitive flexibility as described by Diamond (2013). Set shifting requires disengaging from irrelevant task sets or mental sets in order to actively engage in the appropriate ones. Information updating and monitoring (which Diamond simply referred to as working memory) is the monitoring of information in working memory for relevance to the current task and the replacement of any outdated information with newer, more relevant information. Lastly, the inhibition of prepotent responses, which Diamond referred to as inhibitory control (more specifically, response inhibition), involves deliberately suppressing dominant and/or automatic responses when necessary to allow for controlled responses. Confirmatory factor analyses revealed that shifting, updating, and inhibition are highly correlated with each other but still distinct and separable factors. This seminal three-factor model demonstrates both unity and diversity among executive functions (Friedman & Miyake, 2017; Miyake & Friedman, 2012; Miyake et al., 2000). In other words, these three functions are clearly separable (i.e., diversity), as evidenced by structural equation modeling which revealed that the functions contribute to performance to different extents depending on the executive task. However, at the same time, they must share common underlying mechanisms (i.e., unity) resulting in high intercorrelations among them.

Other researchers have also used statistical modeling to uncover the executive functions driving complex behavior, resulting in slightly different theoretical perspectives. Karr et al. (2018) found that a bifactor model (without inhibition) was a better fitting model for executive functions in adults after reanalyzing several studies' confirmatory factor analyses. Karr et al. (2022) later used network models to find that shifting mediates the correlation between inhibition and updating. Despite the differences, they argued that this still highlights unity and diversity among executive functions. Cirino et al. (2018) identified a best-fitting bifactor model that included a common executive function factor and five specific factors (working memory span/manipulation and planning, working memory updating, generative fluency, self-regulated learning, and metacognition). Overall, there seem to be general commonalities across theoretical frameworks using these statistical methods. A review of the literature revealed that the three most common factors included in these models are inhibition, working memory, and shifting (Karr et al., 2018).

Researchers studying executive functions have identified working memory as an important subcomponent of cognitive control. However, there is another theory of working memory and its capacity that equates executive functioning and cognitive control with working memory (Diamond, 2013). Shipstead et al. (2016) defined working memory as the ability to temporarily maintain and manipulate information. As in the previous discussion of controlled processes, working memory is limited in capacity and resource-demanding (Engle & Kane, 2004; Meier & Kane, 2017). Asserting an executive attention account, a person's working memory capacity is determined by their ability to focus attention on goals and critical task-relevant information, as well as resist distraction to attention (Conway & Engle, 1994; Engle & Kane, 2004; Engle, Kane et al., 1999; Engle, Tuholski et al., 1999; Kane & Engle, 2000, 2002; Shipstead et al., 2014). In other words, attention control is crucial to maintaining relevant information and task goals during complex cognition (Meier & Kane, 2017; Shipstead et al., 2015). According to this theory (e.g., Engle, 2002; Engle & Kane, 2004; Kane et al., 2007;

Shipstead et al., 2016), cognitive control should be studied through the lens of working memory, which can be reduced to executive attention (i.e., attention control by a higher-order executive function). Specifically, this two-factor theory of cognitive control consists of maintenance, the ability to maintain attention on critical information, and disengagement, the ability to withdraw attention from irrelevant information or block attention to outdated information. Together, maintenance and disengagement make up executive attention. Variation in working memory capacity is, therefore, due to variation in people's ability and efficiency (i.e., general processing speed) to maintain information (i.e., attentional focus) and resolve response conflict (i.e., attentional inhibition, Engle & Kane, 2004; Kane et al., 2007; Shipstead et al., 2016), especially when prepotent or habitual behavior creates response competition with task-appropriate behavior. The importance of executive attention is highlighted by Miyake et al. (2000) and Miyake and Friedman (2012), who explained that the three core executive functions from their theory (shifting, updating, and inhibition) may be highly correlated latent factors because of a shared controlled attention component (Engle, Kane et al., 1999; Engle, Tuholski et al., 1999) amongst all controlled processes that allows for goal and information maintenance and use in processing. Regardless of the task and, therefore, regardless of the specific executive function, controlled attention is required and responsible for the unity they describe in their model. Diamond (2013) asserted that inhibitory control (specifically, the selective attention component of interference control) is the same as executive attention as it is talked about here. Like the correlations Diamond (2013) reported, working memory, as defined by Engle and colleagues, has been found to correlate strongly with fluid intelligence (e.g., Conway et al., 2003; Kane & Engle, 2002) because they both rely on executive attention to organize cognitive processing in a goal-oriented manner for strategy selection and problem monitoring (Shipstead et al., 2014;

2016). Like Miyake et al. (2000), Shipstead et al. (2014) used structural models to identify key mediating factors (primary memory, secondary memory, and attention control) between working memory and fluid intelligence.

In summary, while Diamond refers to selective attention as a component of inhibition, Miyake views attention control through the lens of "shifting" between mental and task sets, and Engle views executive attention as the sole mechanism of cognitive control. Regardless of the terminology, however, all three theorists recognize the importance of attentional control, as did earlier definitions of cognitive control (Posner & Snyder, 1975; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Interestingly, working memory is also a common thread among theories. However, Miyake (who referred to this capacity as updating) and Diamond both view working memory as a subcomponent of cognitive control, while Engle would argue that working memory is the entirety of cognitive control. Despite these similarities across theories, however, each theorist offers different reasoning as to why cognitive control is so highly correlated to general intelligence. Diamond argued that the three core executive functions (inhibition, working memory, and cognitive flexibility) combine to create the higher-order executive functions (reasoning, problem-solving, and planning) responsible for intelligence. Miyake credited a shared controlled attention component for the strong correlations among executive functions. Engle argued that general intelligence and working memory share a common mechanism (executive attention) to explain their correlation.

1.3 Development of Cognitive Control in Children

The developmental progression of cognitive control offers significant value in further understanding how humans come to use this complex mode of processing. Whether cognitive control is driven by one central mechanism (executive attention) as Engle would suggest or if cognitive control is hierarchical in nature as Diamond would imply, is still up for debate. Understanding the development of executive functioning may offer insight into the components that make up these skills. Development of cognitive control has been primarily studied in children between the preschool years and late adolescence although it has been suggested that executive functions should be studied through late adulthood and across the lifespan to better understand the full development of these abilities (e.g., Best et al., 2009; Best & Miller, 2010). Especially considering that the maturation of cognitive control is closely related to the development of the prefrontal cortex (e.g., Brydges et al., 2013; Bunge et al., 2002; Dempster, 1992) throughout childhood and adolescence.

Bunge et al. (2002) found that children (8- to 12-year-olds) demonstrated less attentional control and response inhibition compared to adults during intermixed trials of Eriksen flanker and Go/No-Go tasks. Brydges et al. (2013) found the same result when comparing 8- to 11-year-old children to 18-year-olds on a Go/No-Go task. Similarly, Bub et al. (2006) tested elementary school children (7 to 11 years old) on the Stroop task and found that the younger children (less than 9 years old) struggled to maintain attention on the appropriate task set (word reading vs. color naming) compared to the older children (9 years old and older). In a review of the literature on executive functioning in children across a wide age range, Best et al. (2009) and Best and Miller (2010) concluded that the broad developmental milestones of inhibition, working memory, planning, and shifting develop at different rates.

Inhibition has the most noticeable improvement during the preschool years (Best et al., 2009; Best & Miller, 2010). Age-related differences in performance on the Dimensional Change Card Sort (DCCS) task demonstrate the changes in children’s inhibition during this time. This task requires children to initially sort a deck of cards along one of two dimensions (e.g., shape or color). However, later, the task then switches to require them to sort along the other dimension, requiring them to inhibit the now prepotent response to sort based on the first dimension. By age four, children are generally able to complete this complex response inhibition task with significantly less perseveration than younger children. These large, early improvements between the ages of three and five are then followed by more subtle, slower improvements through adolescence. Best and Miller differentiate these early conceptual gains in inhibition from the refinements in accuracy (likely through improved efficiency at blocking unwanted responses) that occur through later childhood and adolescence. This theory is supported by the findings of Davidson et al. (2006) that adults showed response time effects in their performance on three spatial incompatibility tasks while children showed accuracy effects. According to Davidson et al., the adults likely slowed down their responses on difficult trials to maintain their accuracy, while the impulsivity of the children resulted in consistent response times across trials but low accuracy on difficult trials.

However, Zelazo and colleagues (e.g., Zelazo et al., 2003; Zelazo & Fryer, 1998) argued that the DCCS task is not about inhibition. Instead, they stated that 3-year-old children continue sorting the cards by the first dimension because, even though they understand both pairs of rules, they are unable to integrate them into a rule system that follows a hierarchical structure. Based on the cognitive complexity and control theory, children's perseveration is a function of complexity. This theory states that the formal relationships among rules follow a hierarchical tree structure such that one rule can be embedded under another and controlled by it (for example, in the DCCS task, the rules could be "if sorting by color, then if red car, …here, and if blue flower, …there, but if sorting by shape, then if red car, …there, and if blue flower, …here). Therefore, the developmental transitions that are observed in children's performance are because 2-year-olds can only represent a single rule at a time, while 3-year-olds can consider a single pair of conditional if-then rules, and 5-year-olds can represent higher order rules and navigate between two conflicting pairs of rules. With age, children gain the ability to handle an additional level of complexity which affords them better control over their behavior. Zelazo and Fryer defended that perseverance on the DCCS task is not an issue of inhibition because children fail to sort by the second dimension after sorting by the first dimension for only one trial (i.e., before the initial response was able to become habitual). Additionally, they are able to switch to a different, incompatible rule when only one pair of rules is involved (so, it is not about being unable to change previous ways of responding). Finally, Zelazo and Fryer noted that discerning between two pairs of rules is not an issue of working memory capacity because children can sort based on four simultaneous rules that are not embedded.

Working memory and shifting also emerge during the preschool years (Best et al., 2009; Best & Miller, 2010). However, unlike the burst of improvement observed in inhibition, these executive functions improve in a gradual, linear manner through adolescence. Given that inhibition and working memory processes are important for successful shifting (a response set must be maintained in working memory and alternative sets inhibited from activation in order to successfully shift), shifting naturally has a longer developmental progression. While preschool-aged children can shift between simple tasks with reduced inhibition demands, improved metacognition during adolescence facilitates continued development of shifting (Best & Miller, 2010). Best and Miller reported that task switching between complex tasks during middle adolescence has been reported to match adult-like levels of this ability. However, Davidson et al. (2006) reported that cognitive flexibility measured in 13-year-olds was below adult levels. Finally, planning sees the largest improvement in later childhood or adolescence (Best et al., 2009). However, it must be noted that planning processes are typically measured using complex tasks that may not be suitable for assessing this ability at younger ages.

Building on the framework of cognitive control by Miyake et al. (2000), which considers executive functioning to be a unitary construct with three dissociable components (working memory, response inhibition, and shifting), Garon et al. (2008) found that factor analyses indicated that measures of executive functioning in 2- to 5-year-old children also clustered into distinct factors. Therefore, by the end of preschool development, executive functions have already been organized into partially distinct components. Using exploratory and confirmatory factor analyses, Lehto et al. (2003) also found three interrelated factors (working memory, inhibition, and shifting) when examining the performance of 8- to 13-year-old children on a battery of executive function tasks that is in line with the three-factor model of Miyake et al.

1.4 Defining Cognitive Control in Nonhuman Primates

In addition to developmental studies, comparative data is crucial to understanding human cognition because it elucidates the evolutionary progression of complex cognition. Studying whether humans' and nonhuman primates' capacity for cognitive control is similar allows for a better understanding of the evolutionary origins and phylogeny of cognition, providing insight into the fundamental aspects of the human mind and the conditions for which complex cognition and behavior are required. Comparative psychology informs theory about the evolution of cognition and plays a crucial role in connecting the natural and social sciences (such as ecology and psychology) within and between disciplines. Additionally, extensive research has shown the advanced cognitive abilities of nonhuman primates and their continuity with those of humans, making nonhuman primates an ideal model to compare and generalize results across species. Studying nonhuman primate cognition provides an informative reference point for humans' cognitive abilities, as well as reveals the fascinating nature of animal minds more generally (Kaufman et al., 2021; Maestripieri, 2003; Roitblat et al., 1984; Shettleworth, 2009; Tomasello & Call, 1997; Zentall & Wassermann, 2012).

In an exploration of attention control in nonhuman primates, Washburn (1994, 2016) tested rhesus monkeys on a Stroop-like numerical task. Rhesus monkeys were taught to make a two-choice discrimination between two arrays consisting of a different number of letters and then indicate which array had the most items. During some trials, monkeys were shown two arrays of Arabic numerals. The monkeys had previously been taught the relative values of each of the Arabic numerals used. Congruent trials were those in which the more numerous arrays also contained the numerals of larger value (e.g., five 4s vs. two 2s, for which the correct response would be the five 4s). On incongruent trials, the more numerous arrays consisted of the smaller valued numbers (e.g., three 2s vs. seven 0s). Washburn found that the monkeys responded fastest to congruent trials and slowest to incongruent trials, compared to baseline trials (arrays of letters). Monkeys' accuracy was also poorer for incongruent trials. Overall, this Stroop-like interference and facilitation (which also partially matched that of human participants who completed the same task) was indicative that the monkeys' automatic processing of symbol meaning created response competition in the primary task (quantity discriminations), despite its irrelevance to the task. Washburn concluded that the mechanisms of automatic versus controlled attention must be similar in humans and monkeys and that competition for attentional control occurs in nonhuman primates.

Interestingly, Washburn (2016) identified three forms of cognitive control. That is, the focusing of attention (and subsequent control of behavior) can be determined by stimuli (environmental constraints), experience with stimuli (experiential constraints), and conceptual processing of stimuli (executive constraints). Stimulus control, which is characterized as stimulus-driven and automatic processing, can be the result of either experiential or environmental constraints. Experiential constraints are conditioned responses that are elicited automatically by stimuli. Environmental constraints are more purely stimulus-driven, reflexive, and bottom-up, such as attentional capture. Executive constraints result in goal-directed, effortful processing that is controlled and supervisory. Executive control, also referred to as voluntary control, allows for effective responding when automatic responding would lead to errors. Washburn argued that the Stroop effect should be discussed as a competition between experiential cues (from the symbol meanings) and controlled processing. In summary, Washburn characterized automatic processing as being driven by stimulus-response associations (which may be innately prepotent [environmental] or learned through experience [experiential]), and controlled processing as being voluntary and effortful and driven by executive attention. In other words, cognitive control should be conceptualized within the context of attention, which is subject to stimulus control and executive control.

In contrast to this focused definition of cognitive control, other researchers view the term as an umbrella term referring to multiple processes. Beran (2015) identified cognitive control as the various regulatory processes used for complex cognition. Beran retains the same crucial distinctions between controlled processes and automatic processes held by others (e.g., Schneider & Shiffrin, 1977). That is, controlled processes require selective and effortful attention and have a limited capacity. Beran, Menzel et al. (2016) defined cognitive control processes as the use of deliberate, effortful responses to inhibit habitual, automatic, associative responses when adaptive and flexible behavior is required. In other words, cognitive control is required to overcome stimulus control of behavior, the former relying sometimes on inhibitory control to do so.

According to this definition, all executive and regulatory mental processes are cognitive control processes, making cognitive control a label for a plethora of cognitive capacities, such as metacognition (the monitoring and controlling of cognitive processes to assess levels of uncertainty and confidence, as well as seek additional information if needed, during responding, Flavell, 1979; Nelson & Narens, 1994), self-control (self-regulation to decline a readily available reward of lower value for a more valuable, but harder to obtain, reward, Mischel et al., 1989), executive attention (selective attention to specific stimuli and/or stimuli features, e.g., Engle & Kane, 2004; Rueda et al., 2005, 2004), recall of episodic memory (autobiographical memories, e.g., Tulving, 1993), and prospective memory (forming and retaining intentions in order to execute an intended response in the future, e.g., Marsh & Hicks, 1998). Executive attention, for example, is the control of attention, while episodic and prospective memory require controlled accessing of memories. As such, each of these processes (among others) are a form of cognitive control (Beran, 2015; Beran, Menzel et al., 2016). Therefore, the various nonhuman primate studies demonstrating metacognition (Basile et al., 2009; Beran et al., 2013, 2015; Beran, Perdue, Church, et al., 2016; Beran & Smith, 2011; Brady & Hampton, 2021; Brown et al., 2019; Hampton, 2001; Hampton et al., 2004, 2020; Kornell et al., 2007; Shields et al., 1997, 2005; Smith et al., 1997, 1998, 2003, 2006, 2013; Templer & Hampton, 2012; Templer et al., 2018; Tu et al., 2015; Zakrzewski et al., 2014), self-control (Addessi et al., 2011; Beran & Evans, 2006; Beran, Perdue, Rossettie, et al., 2016; Beran, Rossettie, et al. 2016; Bramlett et al., 2012; Evans & Beran, 2007; Evans et al., 2012; Evans, Perdue, Parrish et al., 2014; Evans & Westergaard, 2006; Judge & Essler, 2013; Parrish et al., 2018, 2013; Perdue et al., 2015; Tobin et al., 1996), executive attention (Beran et al., 2007; French et al., 2018; Hassett & Hampton, 2022; Herrmann & Tomasello, 2015; Smith et al., 2013), episodic memory (Basile & Hampton, 2017; Hampton et al., 2005; Menzel, 1999; Sayers & Menzel, 2012; Schwartz, 2005; Schwartz & Evans, 2001; Schwartz et al., 2002), prospective memory (Beran et al., 2012; Evans & Beran, 2012; Evans, Perdue, & Beran, 2014; Perdue et al., 2014), and working memory (Basile & Hampton, 2013a; Brady & Hampton, 2018a, 2018b; Brown & Hampton, 2020; Templer et al., 2019; Tu & Hampton, 2014) are all examples of cognitive control in these nonhuman animals.

The dual-process model of recollection and familiarity processes in memory, which are relatively slow and fast processes, respectively (Basile & Hampton, 2013b), offers an interesting parallel to the automatic and controlled processes reviewed thus far. The difference in processing time in rhesus monkeys' memory in this model of recognition is evidence of the cognitive control of memory during recollection and the automatic processing of memory when using familiarity. Nonhuman primates' comparable performance to human subjects highlights the continuity of these capacities between animals and humans. Interestingly, Smith et al. (2013) found a dissociation in interference from an ongoing task on uncertainty monitoring and discrimination responses. They presented rhesus monkeys with a pixelated square that had to be discriminated as "sparse" or "dense." Across experiments, monkeys had the option to either indicate their uncertainty before responding or classify the stimulus into a third "middle" category. Under cognitive load from a concurrent matching-to-sample task, the monkeys' uncertainty responding decreased. However, their perceptual responses were unaffected. This interference is evidence of a limited-capacity system, a hallmark characteristic of controlled processes.

2 ASSOCIATIVE LEARNING

2.1 History of Stimulus-Response Theory

Cognitive control, as studied in nonhuman primates, children, and theoretically, offers a complex view of behavior that requires higher-order, supervisory processes that control lower-level inputs in order to facilitate flexible and goal-oriented behavior. However, as a field of study, behaviorism offers an alternative explanation for complex behavior that is much more simplistic. Theories of behaviorism, which have also evolved over the years, focus on the environment as the key determinant of behavior. In other words, cognitive control may also be explained associatively. There is, therefore, a debate between cognitivists and behaviorists about the mechanisms of behavior.

In stark contrast to the mentalistic explanation of cognitive control in human and nonhuman primates, there have been several theorists that believe experience is the primary driver of behavior. Watson (1913) proposed methodological behaviorism as an alternative means to explain complex human and nonhuman animal behavior. As the first behaviorist, Watson rejected the study of mental processes and their relation to consciousness and instead focused on the relationship between the environment and behavior. Methodological behaviorism (also referred to as classical behaviorism) focuses on the environmental causes of behavior by looking at the relationship between stimuli and responses (Moore, 2017). As a stimulus-response (S-R) theory, Watson rejected cognitive explanations of behavior in favor of behavioral, physiological, and environmental variables that triggered innate and/or acquired responses in an organism (Moore, 1999, 2017). Watson relied heavily on Pavlov’s theory of classical conditioning (Skinner, 1974). He thus viewed observable behavior as being triggered by automatic reflexes in response to the environmental events that preceded it. Flexible behavior, therefore, was the association of a specific response to a new stimulus (i.e., conditioned reflexes). This new association is formed when the presentation of the new stimulus aligns with the presentation of the original stimulus. For this reason, Watson considered the environment (a stimulus) as the direct cause of behavior (a response), and not thoughts, feelings, drives, and motives (Watson, 1913). With behavior being the result of connections between specific stimuli and responses, complex behavior, including language, was considered the result of very long chains of S-R associations (Moore, 2017). As a behaviorist, Watson rejected feelings, states of the mind, and mental events as causes of behavior. Instead, behavior was considered solely as it related to an individual's prior environmental history, predicted by observing that history, and controlled by manipulating it.

Building on the work of Watson, Hull (1943a, 1943b) proposed neobehaviorism. For Hull, physiological activity (such as stimulus reception, afferent neural interaction, and drive reduction) was the intervening variable between environment and behavior (Smith, 1990). According to Hull's law of stimulus generalization (Hull, 1939; Ziafar & Namaziandost, 2019), flexible behavior (that is, a new response to a stimulus) will occur when that stimulus is associated temporarily, spatially, and/or in physical characteristics with the original stimulus that triggered that response. Therefore, any stimulus that is similar to the original one can elicit the same response. As a result, a stimulus can evoke a reaction that was never associated with it directly. Hull (1939) argued that generalization is the result of stimulus equivalence, primary generalization, or secondary generalization. Stimulus equivalence can be established when the conditioned stimulus has close physical similarity to another stimulus. Primary or physiological generalization is observed when a conditioned reaction occurs in response to a new stimulus because it is from the same stimulus continuum as the original conditioned stimulus. This generalized reaction will decrease in intensity along a gradient as the difference between the new and conditioned stimuli increases. Hull also argued that when various points on the stimulus continuum have been conditioned, all stimuli along the continuum, whether conditioned or not, will also elicit a reaction. Because the degrees of separation along the continuum are no longer a factor, Hull refers to this as secondary (or indirect) generalization. Interestingly, Hull believed that these three forms of stimulus equivalence can be combined to various extents in different situations.

However, unlike Watson, Hull did not believe that behavior is the result of a chain of events. Instead, behavior was described in terms of a shaping sequence of related stimuli and their respective responses (Ziafar & Namaziandost, 2019). These learned sequences of S-R connections are conditioned habits that become the basic mechanism of behavior (Hull, 1930, 1943a; Smith, 1990). Ultimately, like all behaviorists, Hull maintained that the environment is the ultimate cause of observable behavior and not the cognitive, controlled processes of the mind.

As a strong critic of Hull, Tolman proposed cognitive (or purposive) behaviorism, which viewed behavior as goal-directed (Tolman, 1932, 1933, 1934). Tolman and Honzik (1930) stated that behavior does not have to be rewarded for learning to occur. They found that rats that were reinforced every time they reached a designated goal box in a large maze gradually (across trials) traveled faster to that location than rats that explored with no reinforcement. The nonrewarded rats were later given food whenever they reached the goal box, like the rats that had experienced continuous reinforcement, and they also began to quickly travel to the goal box. However, their behavior changed, across trials, at a much faster rate than would be expected if they were gradually learning the goal location. Therefore, Tolman and Honzik concluded that the initially nonrewarded rats did not start learning only when reinforcement started (as would be suggested by a traditional S-R theory) but had instead been learning previously as well, which is why they were able to so quickly improve their performance. These results (referred to as latent learning) take behavior beyond being simple reinforced associations between stimuli and responses because learning had occurred in the absence of reward. Instead, behavior is purposive and a part of a larger pattern of behavior driven by the expectation of a particular outcome (Tolman, 1933; Ziafar & Namaziandost, 2019). Tolman (1948) theorized that rats were able to engage in latent learning because they had formed a cognitive map, a mental representation of their surrounding environment. Cognitive maps can, therefore, drive adaptive and flexible behavior (Wang & Hayden, 2021). As a result of this research, Tolman considered the intervening variables that mediated the relationship between the environment and behavior to be expectations and hypotheses about what behaviors will result in desired goals (Tolman 1932, 1933, 1934). While still a behaviorist, Tolman offered a more mentalistic theory of learning that incorporated cognitive processes, unlike Watson and Hull.

In opposition to this mentalistic approach, Skinner's radical behaviorism considered subjective thoughts and feelings (which were referred to as "private behaviors") as behaviors that could be explained by the environment in the same way observable behavior could be (Skinner, 1953, 1974, 1984). This consideration of overt and covert behaviors was a departure from Watson's theory. However, like Watson, Skinner viewed internal states as an inadequate explanation for behavior (Skinner, 1974). Introspective observations, such as consciousness and the mental processes of the mind, are determined by one's environmental conditions and considered "behavior from within" (Skinner, 1984). "Knowing," for example, was identified as a behavior in and of itself (Skinner, 1974). He defined self-knowledge as simply describing one's public or private environment because mentalistic processes do not have causal influence over observable behavior. Instead, control is held by the environment and is the ultimate cause of both behavior and mentalistic expressions (as opposed to mental processes mediating behavior and the environment).

Skinner argued that the basic principles of behavior are explained by the principles of operant conditioning (Skinner, 1963). Building on Thorndike's Law of Effect (Thorndike, 1933), Skinner characterized behavior based on its utility to result in a certain consequence (Skinner, 1963, 1981). For example, gaining food or escaping dangerous stimulation are both reinforcing consequences. If a contingency is created between the occurrence of a behavior and a consequence, behaviors that lead to reinforcement increase in frequency, while those that do not decrease. Eliminating the reinforcing consequence will also lead to a decrease in response rate (i.e., extinction). When stimuli are incorporated into the response-reinforcement contingency, by reinforcing a response in the presence of that stimulus, the behavior becomes more likely to be emitted in the presence of said stimulus. This contingency between stimulus, response, and outcome is an effect known as stimulus control of an operant (Skinner, 1963).

According to Skinner (1963, 1974), behavior is defined in terms of its contingencies of reinforcement. The characteristics of the response and/or its complexity (e.g., force or frequency required to press a lever) can be manipulated by changing the contingencies. It is the changes in the contingencies of reinforcement within the environment that, therefore, change behavior. Thus, operant conditioning can be a means of adapting to changes in the environment. Through this process, individuals are able to acquire learned behavior patterns that allow them to navigate their environment quickly and successfully (Skinner, 1981). Including behaviors outside of their innate repertoire, resulting in new forms of behavior that can be shaped and maintained. Problem solving can therefore be reduced to behavior controlled by complex terminal contingencies, that can be sequentially or concurrently arranged, involving multiple stimuli and responses (Skinner, 1963). Differences in problem-solving ability may, therefore, lie in whether an individual can respond correctly under the problem's contingencies without requiring a long series of intervening steps.

While not a traditional stimulus-response theory, Timberlake hypothesized that behavior systems also influence behavioral expression and learning during conditioning (Timberlake, 1993; Timberlake & Lucas, 1989). That is, all organisms have biological predispositions that influence conditioning, and learning is the result of an interaction between biological traits and conditioning mechanisms (Lucas, 2019). Behavior systems theory puts behavior in an evolutionary context instead of reducing it to links between stimuli, responses, and reinforcement outcomes (Timberlake, 1993, 1994). Instead, behavior fits into a systematic hierarchy of activities that are organized around various functions (e.g., feeding, reproduction, predation). At the lowest level of organization are perceptual-motor actions. These are various specific behaviors, such as crawling, sniffing, pawing, etc., that are the components of responding. Actions are organized into behavioral modules (e.g., locomotion refers to various types of physical movements through space). Modules are then categorized into modal groups (e.g., general search). Modes include multiple modules and modules can fall under multiple modes. Finally, these modes fit within a larger goal-directed subsystem (e.g., predation). Therefore, this four-level model organizes behavior into behavior systems that are called upon to navigate the environment, such that the activation of modes activates multiple modules of behavior (Lucas, 2019). In other words, organisms have innate patterns of behavior that are activated by environmental cues.

These systems are pre-organized material for learning (Timberlake, 1993). The functional systems are adapted to the organism's environment (Timberlake & Lucas, 1989) and determine what behaviors are available for responding. As a result, learning engages subsystems that recruit many innate actions, not just the activities that precede rewards (Lucas, 2019). Behavior systems theory, therefore, characterizes behaviors as flexible and opportunistically engaged by relevant environmental cues. Studies of learning must therefore be mindful of an organism's ecology, since these ecological mechanisms drive the behaviors can be observed (Timberlake, 1993). Reinforcement must be consistent with the natural behaviors of the organism, since how compatible conditioned stimuli are with an animal's biological predispositions affect how that animal will react to the stimuli, and subsequently experimenters' observations and conclusions about behavior (Zentall et al., 2019). For example, rats have easily adapted to running in mazes because wild rats live in underground burrows that consist of numerous tunnels and passages, so the mazes tap into preestablished dispositions for trail following, searching, and exploring (Lucas, 2019). Zentall et al. observed species differences between rats and pigeons in their suboptimal choice behavior. However, when behavior systems theory was used to explain these differences, they found that the results originated from differences in how the animals responded to the conditions used to study the effect and that these differences were based on the species' ecology. Species differences were differentiated from learning differences when the search modes elicited by stimuli were matched between the species and, as a result, similar behavior was observed across species. For this reason, it is important that species-fair methodology is used in the comparative study of behavior.

2.2 Stimulus Generalization and Discrimination

Most relevant to the present study is stimulus control as it relates to stimulus generalization and discrimination. Stimulus generalization occurs when a response conditioned to one stimulus occurs in response to another stimulus (Shepard, 1958). The magnitude of this tendency to respond is based on the similarity between the second stimulus and the stimulus that was originally conditioned. For example, Guttman (1963) found that a pigeon reinforced to peck at a circular key will peck keys of similar shape, color, and/or size. Because generalization gradients are a factor of stimulus similarity (Hull, 1939, 1943a), the response rate of this behavior to the new stimulus will decrease in correlation with the difference in physical properties from the originally reinforced stimulus.

When the response is differentially reinforced based on a certain property, stimulus control in the form of discrimination will be observed. That is, differential reinforcement along the stimulus continuum will result in discrimination of a stimulus despite the previous generalized response to that same stimulus (Hull, 1947). Interestingly, in a study of pigeon pecking behavior to a response key, Hanson (1959) found that when a stimulus along a gradient of wavelength was trained after the initial training of a single stimulus (e.g., discrimination training of a 560 mμ stimulus after initial training at 550 mμ), post-discrimination generalization gradients shifted to a location along the wavelength continuum that was between both stimuli.

2.3 Conditional Discrimination in Nonhuman Primates

One method of studying flexible behavior in the face of multiple learning problems is through use of conditional discrimination tasks. Harlow (1949) studied the formation of learning sets to identify how individuals adapt to changing environments beyond trial and error. Learning set studies (e.g., Fobes & King, 1982; Harlow 1949, 1951; Meador et al., 1988; Washburn et al., 1989a, 1989b; Washburn & Rumbaugh, 1991) present subjects with a series of novel, two-choice discrimination problems for which they must identify (across a block of at least six trials per problem) which stimulus has been designated as the S+. However, there are no clues about which stimulus will be rewarded until the first choice is made. The outcome of the first trial of each problem is the first learning opportunity. Harlow (1949) tested rhesus macaques with over 300 object-quality discrimination learning problems, using different pairs of stimuli each time. He found that while the learning curves of the earlier problems showed a gradual increase in performance that was suggestive of trial-and-error learning, the learning curves for the last 56 problems reached near perfect performance by trial two of each problem. By assessing the percentage of correct responses on trial two of each discrimination problem, Harlow measured how well the animals were able to learn from the first training trial. The transition from gradual trial-by-trial learning to efficient one-trial learning represents the formation of a learning set. That is, across problems, the animals were learning to learn, such that they could identify the S+ from a single trial instead of requiring extended experience with the problem. Learning sets transformed a problem that was initially ambiguous to immediately solvable because a single training trial provided the problem's solution.

In addition to object-quality discrimination problems, Harlow (1949) also interposed right-position discriminations within the larger series of problems presented to the monkeys. Despite the changing reinforcement contingencies, by the end of the experiment the monkeys showed a high level of performance for both problems, suggesting that they were able to form and use two independent learning sets (right-position and object-quality). Similarly, monkeys were able to form discrimination reversal learning sets when Harlow would reverse the reinforcement contingencies after a set number of trials for each problem. He measured their percentage of correct responses on the first trial after the reversal (i.e., reversal trial 2) and found that their performance over the final reversal problems was as efficient as trial two of the original discrimination. He also reported that 2- to 5-year old children presented with blocks of discrimination and discrimination reversal problems also showed the formation of a learning set across problems, evidence that the same basic mechanisms operate in both human and nonhuman primates. Overall, Harlow considered this learning process (the formation of a discrimination learning set that culminated in the ability for one-trial learning) as the gradual learning of insight. Learning set is the generalized ability to learn future discrimination problems easily and is therefore an important mechanism for problem solving and adaptive behavior. Instead of several hundred problems worth of reflexive responses to specific stimuli, the formation of learning sets increases one's capacity to adapt and solve problems with minimal errors.

A modification and extension of the learning set paradigm was developed by Rumbaugh (1969, 1970) called the Transfer Index. In addition to testing for learning set formation, Transfer Index testing (e.g., De Lillo & Visalberghi, 1994; Kinoshita et al., 1997; Rumbaugh & Pate, 1984; Rumbaugh et al., 1972; Washburn & Rumbaugh, 1991; Washburn et al., 1989b) is also an assessment of transfer of learning. Instead of a set number of trials per two-choice discrimination problem, measuring transfer index requires a subject to meet a certain criterion level of performance on each problem in the series before the discrimination reversal occurs. That is, the learning set is formed that facilitates a certain high level of performance on two-choice visual discrimination problems and then the reinforcement cue is switched. How quickly the subject meets the criterion performance they had pre-reversal when doing the reversal problems is then calculated as a ratio of correct responses on each trial type across blocks of problems. The cue reversal requires subjects to alter their stimuli choices in accordance with the new reinforcement contingencies, in direct opposition to their prior experience. A transfer of learning would allow for an easier adjustment to the new testing conditions, as opposed to new learning. The mastery of each problem to a criterion level of performance assesses sustained attention and discrimination learning in a way that traditional learning set studies cannot because of a small, fixed number of trials per problem (Rumbaugh et al., 1972). Transfer index has been studied across the primate order, including humans (e.g., Rumbaugh & Gill, 1973; Rumbaugh & Pate, 1984).

In addition to learning set and Transfer Index, primates' ability to solve complex discrimination problems has also been assessed through conditional discrimination studies. Conditional discrimination learning requires subjects to make discriminations based on changing cues (i.e., "if-then" relations). Depending on the cue (e.g., background color, stimulus orientation), different discrimination responses are required. Unlike cue-reversal during learning set and transfer index studies, the cues can change from trial to trial. For example, a simple conditional discrimination example would be, when presented with a circle and a triangle, the subject chooses the circle if the background of the test area is white but chooses the triangle if the background is black (Thomas & Kerr, 1976). Schrier (1970) tested rhesus macaques in a conditional discrimination study where the object pairs differed in either form or size and the conditional cue was the color of the test tray or the color of the object pair. Other research has included multiple conditional cues (e.g., Spaet & Harlow, 1943), including Nissen (1951) who taught chimpanzees to solve conditional discrimination problems with 16 concurrent cues. Warren (1960, 1964) gave rhesus monkeys conditional cues of object color and object orientation for two-choice discrimination problems. Warren (1964) found an additive effect of cueing, in that the monkeys' problem solving was improved when both color and orientation were available as cues compared to when only one cue was used. Conditional discrimination studies with other primates, such as capuchin monkeys (e.g., D'Amato, Salmon, & Colombo, 1985; D'Amato, Salmon, Loukas et al., 1985) and squirrel monkeys (e.g., Barge & Thomas, 1969; Burdyn & Thomas 1984; Thomas & Kerr, 1976), have also been conducted.

In opposition to stimulus-response theories, some have argued that conditional discrimination tasks can also require conceptual processes in monkeys by using an oddity task for the discrimination problems (Thomas & Kerr, 1976) or conceptual discrimination problems and conditional cues (Burdyn & Thomas, 1984). The latter, for example, required monkeys to make conditional discriminations of concept exemplars (for "sameness" and "difference") based on concept exemplar cues (of "triangularity" and "heptagonality," respectively).

2.4 Conditional Discrimination in Children

Conditional discrimination has also been studied in young children and, at times, has revealed interesting age effects among preschool children. Like nonhuman primates, complex behavior such as cognitive control can be evaluated through an associative lens. For example, the DCCS task (reviewed previously, e.g., Zelazo et al., 2003) is similar to conditional discrimination tasks, in that they require a sort of task switching. Golin and Liss (1962) studied 3.5 to 6-year-olds' ability to solve conditional discrimination problems. Children were required to modulate their selection of either a circle or a triangle based on the objects' background. Initially, children learned to select one of the objects presented on the same background. Later, the stimuli were presented on the other background and the reinforcement contingencies were reversed. After reaching the learning criterion for both the discrimination and reversal learning phases, the conditional discrimination phase involved alternation of the two backgrounds when presenting the stimuli. Children across this age range were able to learn the initial discrimination.

The reversal phase was the hardest (as defined by the number of reversal errors that were made) for the 3.5 to 4-year-old children. However, approximately 10% of the 3.5 to 4-year-olds, 17% of the 4.5 to 5-year-olds, and 54% of the 5.5 to 6-year-olds met the learning criterion for the conditional discrimination phase. In general, the conditional discrimination phase was difficult for all of the children. Golin and Liss argued that the children's difficulty with keeping track of which reinforcement contingences were paired with which backgrounds contributed to their poor performance.

Golin (1965) presented children with conditional discrimination problems, as well. However, the children no longer had experience with the initial discrimination or the reversal problems beforehand. Instead, they were trained on both conditional problems simultaneously. Almost half (47%) of the older children (5.5 to 6 years old) performed above chance levels, while only 13% of the children in both the younger (3.5 to 4 years old) and middle (4.5 to 5 years old) age groups performed significantly above chance. Rudy et al. (1993) used a similar procedure as Golin and Liss (1962) to study conditional discrimination in 4- and 5-year-old children. Children had much more success in the study by Rudy et al., in that all the children passed the discrimination and reversal phases. However, the conditional discrimination phase still presented a challenge. Only 11% of the 4-year-olds, but 89% of the 5-year-olds met the learning criterion. Boelens et al. (1989) also reported successful conditional discrimination learning by five-year-old children. However, Hill (1962) found that the ability to solve conditional-oddity problems did not appear until age six. Hill compared the ability of 9 to 13 mo (M = 1 year), 16 to 24 mo (M = 1.7 years), 35 to 61 mo (M = 3.9 years), and 71 to 105 mo (M = 6.7 years) children and found that the three youngest age groups performed at chance levels after 100 trials while the oldest children reached a 65% success rate. Overall, it appears that conditional discrimination is more difficult for younger preschool children compared to older children (Golin, 1965; Golin & Liss, 1962; Hill, 1962; Rudy et al., 1993). However, Jordan et al. (2001) reported successful conditional discrimination learning (with three A/B contingencies) in 2- to 4-year-old children. Additionally, Martínez et al. (2009) observed simple conditional discrimination learning in children ranging from four to six years old.

Andrews et al. (2012) studied 4- to 6-year-old children on their ability to make conditional discriminations when the background color was the distinguishing cue. However, they found that less than 50% of 4- and 5-year-olds were able to meet criterion for a conditional discrimination task. However, this number increased dramatically, to nearly all children reaching criterion for conditional learning, when they were given simpler reversal problems first, similar to the three-phase method used by Rudy et al. (1993) and Golin and Liss (1962). Andrews et al. found an interesting age effect in children's ability to solve conditional discrimination problems. That is, older children relied on relational processing (a cognitive process based on the theories of relational complexity and cognitive complexity and control) to solve conditional discriminations more than younger children, while younger children were more likely to rely on associative processing.

While age effects were not reported, Pérez-González and Serna (2003) found that children were able to generalize contextual cues when solving conditional discrimination problems. Children ranging in age from 10 to 17 years old, were trained on A/B conditional discriminations (i.e., A1/B1 and A2/B2) using a matching-to-sample task with arbitrary visual symbols. They were then given new stimuli (X1 and X2) as a contextual cue of the A/B reinforcement contingencies. When new conditional discriminations were taught with E/F stimuli, the children transferred the relations indicated by the contextual cues X1/X2 without direct training to do so. In other words, the contextual stimuli were able to control discriminations of novel relations. Similar results of stimulus control generalization of conditional cues were found previously in adults and an 11-year-old child by Pérez-González (1994).

2.5 Conclusion

While research in psychology has not led to a single definition of cognitive control, there are key commonalities across current theories of cognitive control. The idea that working memory and inhibition are both required to engage and disengage from task-relevant information is a central idea across theories. Cognitive flexibility and executive attention have also been identified as important aspects of cognitive control. There is also an agreement that cognitive control is an executive and effortful process needed to resist habitual responding when flexible, goal-directed behavior is needed. It has been argued that cognitive control allows for the resolution of response competition when prepotent responses are not ideal, especially in the face of novel or ambiguous circumstances that require controlled responses. Majority of the executive functions that make up cognitive control first appear during the preschool years in children. However, the developmental milestones of inhibition, working memory, planning, and shifting occur at different rates. Additionally, primate species exhibit numerous behaviors that suggest the roots of cognitive control exist across humans' phylogenetic history.

In contrast to cognitive control, another theoretical perspective to explain behavior control is that of associative learning. While the history of behaviorism includes various stimulus-response theories, behaviorists have agreed that behaviors are automatic responses to external stimuli. From this perspective, the locus of control can be found in the environment (as opposed to an executive process of the mind). Navigation of ambiguous and novel situations was accredited to generalization and discrimination, which changes according to the degree of similarity between the conditions and new stimuli. In fact, Abrahamse et al. (2016) argued that cognitive control is actually just an extension of associative learning and can be explained by learning-based principles. For example, flexible behavior often attributed to cognitive control may instead be conditional discrimination and stimulus generalization. From a developmental and comparative perspective, young children and nonhuman primates have also exhibited the capacity for these processes.

In addition to the dichotomy between stimulus and cognitive control, the interplay between associative learning and cognitive control must also be further explored, since they are both important for meaningful and advantageous behavior. For example, attention control is required to focus on goal-relevant information when carrying out complex tasks. However, attentional capture (automatic, involuntary, and stimulus-driven redirection of attention; e.g., Theeuwes, 1994; Yantis & Jonides, 1984, 1990) ensures survival in potentially threatening situations. Several theories have been proposed that suggest a joint control of behavior between stimulus- and cognitive-driven mechanisms. Toates (1998) emphasized the relative weighting of external stimuli and internal cognitions and goals on their influence on behavior, and that both learning mechanisms and cognitive processes are always both present. Therefore, it is the change in these relative weightings that is responsible for flexible and adaptive behavior. According to Toates, there are three influences on behavior – stimuli, cognition, and physiological states. There is a direct link between stimuli and responses, as is supported by stimulus-response theories. Cognition can modulate the S-R connection, while motivation from physiological states can modulate both the S-R connection and cognition. These connections receive feedback from changes or consequences in the external and internal world. Conditional discrimination, according to Toates, is the result of differential reinforcement establishing the connection between specific stimulus-response links to specific cognitive sets. For example, the link between S1 and R1 is specific to a cognitive set, while the link between S2 and R2 is tied to a different cognitive set. As the cognitive set varies in response to changing environmental factors (in complex situations this variation would happen rapidly), modulation of the S-R links changes accordingly. In other words, the discriminative stimulus modulates the S-R connections via cognitive links (e.g., an assessment of the current task goals).

According to this theory, in novel situations, when there are no stimuli that have particularly strong links to behavior available, weight will be given to cognitive processing to establish cognitive maps, goals, expectations, etc., based on sensory input (Toates, 1998). Conversely, when a cognitive solution is unknown, weighting is switched to the S-R links. There is no longer a dichotomy between automatic and controlled processing, but rather a difference in the weighting between them occurs, essentially creating a spectrum of control. Developmental effects are seen as the result of changes in the point of control over behavior.

Washburn's (2016) identification of environmental, experiential, and executive constraints on behavior also suggest multiple sources of behavioral control. While Barge and Ferguson (2000) suggested that complex social behavior, such as motivation and goal setting, various social behaviors, and social judgment, can occur automatically. They argued that executive control processes (e.g., working memory), to a certain extent, are guided automatically by the environment just like processes that do not require conscious guidance. They resisted the traditional dichotomy supported by many cognitivists and behaviorists between the cause of conscious and nonconscious processes (i.e., internal vs. external and controlled vs. automatic, respectively). That is, environmental determinants of behavior should also be applied to controlled processes and complex behavior can be habitual and subconscious.

It is important to distinguish between associate and cognitive explanations for behavior so we can better understand the mechanisms that control behavior. Distinguishing stimulus and cognitive control, including identifying the testing conditions that evoke one or both of these processes, is important for understanding and predicting the behavior of humans, animals, and the interactions between them. By understanding how behavior works, it can be manipulated to improve functioning in various capacities important for everyday life, such as children rearing, training animals, personal and professional relationships, and performance at work and/or school. Identifying the mechanisms of behavior may also lead to mitigation of errors in ideal responding (e.g., failure of self-control when attempting to remain on a diet or stop smoking) and the development of interventions that would be useful to treat various behavioral disorders, for example, and train service animals for populations that need them.

3 PRESENT STUDY

As I have outlined, cognitive control has been defined in multiple ways (e.g., Diamond, 2013; Miyake et al., 2000; Shipstead et al., 2016). There is consensus that cognitive control allows for flexible, adaptive, and goal-directed behavior. When faced with challenging, ambiguous, or novel situations, it has been hypothesized that executive functions are recruited to mitigate response conflict. Alternatively, behaviorists have argued that flexible, adaptive behavior can be associative. For example, Skinner's (1974) longstanding theory of operant conditioning emphasizes that behaviors that lead to reinforcement will increase in frequency, while those that do not decrease, because all behavioral responses are associated with consequences. Based on this theory, flexible behavior is the result of changing contingencies of reinforcement within the environment. It can be considered adaptive in the sense that individuals acquire learned behavior patterns that allow them to navigate their environment quickly and successfully (Skinner, 1981). Additionally, ambiguous and novel situations can be solved via stimulus generalization (Hull, 1939).

This debate about the role of cognitive control versus stimulus control has been explored among animal researchers. Thomas and Kerr (1976) used squirrel monkeys to test their ability to complete a conceptual conditional discrimination task. They had monkeys complete an oddity task using a conditional cue of background color. In the task, the monkeys had to use the available cue (white tray or black tray) to make a conceptual discrimination ("odd" or "not odd", respectively). By requiring a conceptual solution to the discrimination problem and using new oddity problems on each trial, the monkeys were unable to associate specific stimulus configurations (e.g., white background/circle, blue background/triangle) with differential reinforcement outcomes on which to base their conditional responses. All three monkeys showed a high level of success on the conditional discrimination task, evidence of a conceptual interpretation of conditional discrimination in the absence of stimulus-response learning of stimulus configurations. Burdyn and Thomas (1984) continued to dissociate stimulus and cognitive control over conditional discrimination problems by having squirrel monkeys make conditional discriminations of concept exemplars (sameness and difference) based on concept exemplar cues (triangularity and heptagonality, respectively). All of the animals (*N* = 4) were able to remember conceptual information to make conceptual choices.

Andrews et al. (2012) also attempted to disentangle the role that cognitive control and configurational learning play in solving conditional discrimination problems by studying children four to six years old. Specifically, they were interested in whether children use relational or configural processing. Relational complexity theory (e.g., Andrews & Halford, 2002; Halford et al., 1998) describes a cognitive form of processing that requires forming representations of the relational structure between problems. That is, the solution for one problem is mapped analogically to a second problem that follows the same form but involves different components or stimuli. Relational complexity theory is closely related to the theory of cognitive complexity and control theory (Zelazo et al., 2003). Andrews et al. found that older children were more likely to rely on relational processing, whereas younger children relied on associative processing. Additionally, fluid intelligence was found to be a significant predictor of performance for relational processing but not associative processing. Therefore, it seems that conditional discrimination tasks can be solved with either associative or cognitive mechanisms, relational processing is measurably distinct from associative processing, and the role of these processes in behavior changes with age.

The present study continued the exploration of how cognitive and associative factors impact behavior across species and development using a novel conditional discrimination task. I aimed to assess whether adults (Experiment 1), children (Experiment 2), and monkeys (Experiment 3) demonstrated optimal and flexible use of computerized tools based on the specialized cursors' function or the cursors' associative history. The present study was intended to extend the research of conditional discriminations that involved the use of conceptual information (e.g., Burdyn & Thomas, 1984; Thomas & Kerr, 1976). However, the present study differed in that (1) instead of participants discriminating conceptual information based on stimulus-based features, they were required to discriminate various stimuli based on conceptual information (functionality), and (2) then generalize their knowledge of that functionality. Therefore, there were two main questions of interest. Can participants learn to differentiate stimuli based on functionality? Can participants then generalize their optimal use of stimuli to different tasks? Critically, the speed with which discrimination learning occurred would distinguish cognitive control and stimulus control within the same task. Additionally, the present study assessed the tendency for generalization in the presence of limited physical similarity between conditioned and new stimuli. Given that stimulus similarity is a key component of generalization and discrimination across stimuli according to stimulus-response theories, but not theories of cognitive control, the present study was able to potentially discern the mechanisms of cognitive and stimulus control from each other in their control over behavior. This discernment would provide important information in the debate over whether associative learning or cognitive control are responsible for flexible, adaptive behavior.

In this study, a set of computer cursors served as "tools" to aid the participants in solving various psychomotor tasks. Tasks varied in nature, and the cursors varied in appearance and movement, such that the cursors were better adapted (i.e., optimal) for specific tasks. By using the optimal cursor, the participants were able to meet the specific demands of each task more efficiently (i.e., reducing response times) and increase their success of obtaining reinforcements (i.e., increasing the speed of reinforcement). Participants' ability to optimize their tool use by learning which specialized cursors were best for three different primary tasks was assessed. For the participants that were able to do this, their ability to utilize cognitive control to generalize optimal behavior to three transfer tasks was also assessed.

Participants were first allowed to freely choose between three different cursors to complete three different tasks. This initial phase of testing (Tool Testing) investigated whether differences in efficiency were sufficient associative cues to differentiate the optimal cursor for each task. For those that were unable to demonstrate learning of the cursors' functionality on their own, a separate phase of testing (Correction) introduced a correction procedure during which differential feedback was used to further support learning. Finally, participants were given three additional tasks (Generalization Testing) that they had limited experience with to see if they were able to generalize their knowledge of the cursors' functionality to transfer tasks with similar task environments but limited perceptual similarities. Successful generalization would have appeared as immediate success during the transfer tasks, as well as suggested a conceptual understanding of each tool and cognitive control over choice behavior. Since the participants would not have had direct experience using the specialized cursors with these new tasks, automatic and habitual responses had not been established yet, so immediate success would presumably not have been learning-based. Therefore, the present study was designed to identify the extent to which associative learning was required for this conditional discrimination task, as well as whether cognitive control or associative learning support the transfer of this skill to the transfer tasks.

If participants are able to discriminate various stimuli based on the concept of functionality, they should choose to use the optimal cursor for each task. If they cannot, they would not meet the testing criterion for the Tool Testing or Correction phases. Additionally, in the absence of explicit stimulus cues to support generalization and extensive associative experience to support learning, if human behavior is predominantly controlled by executive functions rather than associative learning, then participants would show optimal performance during the earliest trials of each task. If instead, participants were relying on specific stimulus configurations to perform optimally they would require high trial counts to gradually learn this behavior through associative experience.

Given that previous research has shown that humans and nonhuman primates are capable of cognitive control, it was predicted that all three groups would learn and generalize cursor functionality to optimize their performance, suggestive of conceptual understanding and cognitive control over behavior. However, differences in the extent of success across age and species were expected. If mature executive functioning supported better understanding and performance, then there would be differences in success across participant groups. However, if they did not, equal levels of success would potentially be observed. Given that executive functioning is still developing in children and there is variability in the extent of cognitive control observed in nonhuman animals, adults were predicted to show the highest level of success (defined as the number of individuals that showed successful performance) in both learning and generalization behavior, followed by children, and then monkeys.

Lastly, adult participants were also tested with and without explicit instructions of the cursors' functionalities to compare mature adult performance to that of nonhuman primates with no language capabilities and children with developing language. If the capacity and understanding of language facilitates performance in the present study, then the adults who received explicit instructions about the tasks would choose the optimal cursor at a higher frequency than those that did not. If instructions were not an important factor for learning, then there would be no difference between these groups. It was predicted that instructions about the key features of each cursor would facilitate learning and generalization.

In the event that humans and/or monkeys did not demonstrate the ability to generalize their knowledge about the cursors' functionality, the present study would be potential evidence of their reliance on associative processing (and potentially limitations of cognitive control) to complete this conditional discrimination task. The three-phase design of the present study allowed for a comparison of learning across the different phases to examine potential changes in performance as a result of the various experiences. To the extent that the participants were unable to learn which cursor was optimal for each primary task, various factors (learning history, associative cues, conceptual generalization) were assessed to identify potential factors that prevented learning.

By including adult humans, children, and monkeys as participant groups, behavior can be examined across development and species. Humans unquestionably have the capacity for advanced cognitive control, while animal behavior is often first attributed to associative control. By including species with (presumably) opposite control mechanisms guiding behavior, interesting comparisons can be made when opposing processes are assessed using the same paradigm. Additionally, children have developing capacities for cognitive control and developing language skills, compared to adults fully developed control- and language-abilities and the absence of language abilities in monkeys altogether, making for interesting developmental comparisons.

3.1 Data Analysis Plan

To test the hypotheses, several analyses were used across all three experiments. Because there were six different tasks designed to create different task demands for the participants, task set and task type were often included as separate (and sometimes the only) independent variables during data analysis to assess participants' choice behavior on the individual tasks. To test the optimality of the cursors (defined as faster response times for each task when the specialized cursor was chosen compared to either of the other two cursors), 2 x 3 mixed ANOVAs were used to test the effects of cursor accuracy and task type, respectively, on trial response time. Additionally, the amount of time participants spent moving the transparent cursor through barriers blocking the path to the target location was approximated to assess whether this cursor was being used as it was designed to be.

Performance was always defined as the percentage of time (out of 100%) that participants chose the correct (i.e., optimal) cursor for each task in comparison to chance (33%). To determine whether it made a difference if adults received explicit or general instructions about the specialized cursors, 2 x 2 x 3 mixed ANOVAs were used to assess the effects of instructions, task set, and task type, respectively, on their performance. Similarly, to determine if there was an effect of receiving correction experience or not on performance, 2 (correction) x 2 (task set) x 3 (task type) mixed ANOVAs were used. A one-way ANOVA was used to assess for changes in monkeys' performance across time (10 blocks). Binary logistic regressions were used to assess whether receiving explicit instructions (for the adults only) and/or assigned task set were significant predictors of a participant's likelihood of needing correction experience after the first phase of testing. When relevant (e.g., to assess for potential biases in cursor selection), one-sample t-tests or binomial tests were used to compare participants' cursor choice behavior to chance levels of responding. Chi-square Goodness of Fit tests, as well as qualitative analyses, were also used to compare the frequency of participants' cursor choices across tasks.

4 EXPERIMENT 1

4.1 Participants

Undergraduate students (N = 58) at Georgia State University, ranging in age from 18 to 30 years old (M = 19.81, SD = 2.02), participated in this study in exchange for course credit. The convenience sample consisted of twenty-four participants that self-identified as women, one as queer, and the remaining as men. Participants provided informed consent for their voluntary participation and all testing complied with protocols and procedures approved by the Georgia State University Institutional Review Board.

4.2 Apparatus

Undergraduate students were tested at individual computer stations with a 17-inch laptop touchscreen computer. All testing was conducted on the campus of Georgia State University in the Department of Psychology. Participants were tested in a single 30 to 45 m testing session. They responded using joystick deflections via a USB game controller to control a cursor on the computer screen. Completing each trial resulted in a large yellow smiley face being shown on the computer screen. When relevant, incorrect responses resulted in a large yellow sad face being shown on the computer screen. All computer tasks were written in Visual Basic 6.0.

4.3 Design and Procedure

4.3.1 Task Training

Participants completed six psychomotor tasks[1] – CHASE, ERASE, FENCE, SIDE, BARRIER, and CATCH (see Figure 4-1).

---

[1] Tasks are capitalized throughout as this is a convention in our laboratory going back to the original development of computerized testing of nonhuman primates. It also aids the reader in understanding when I refer to a specific task.

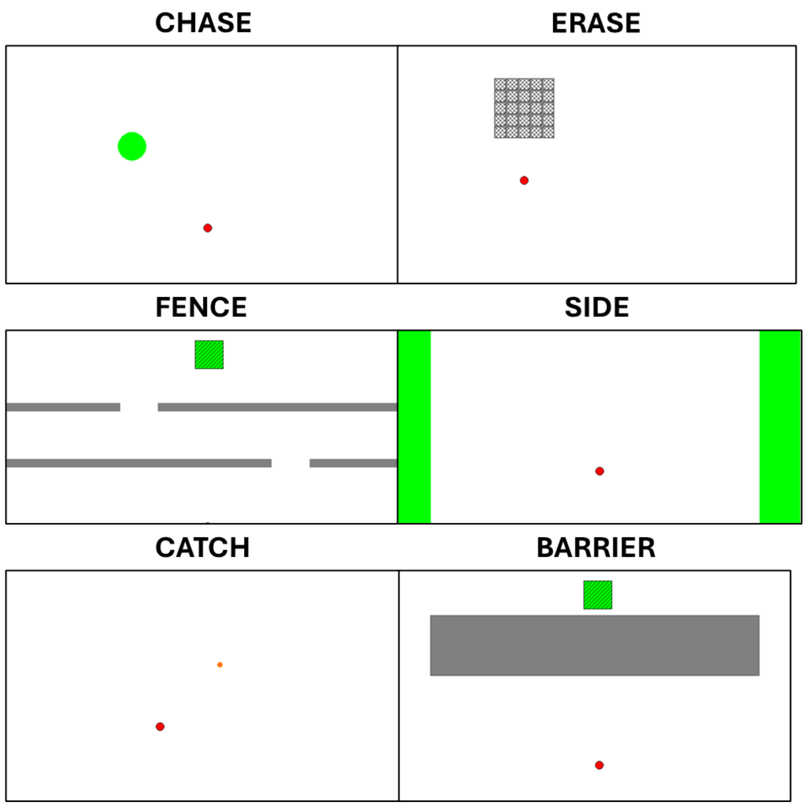

Figure 4-1 Example Trials of Psychomotor Tasks

In the CHASE task, participants had to move the cursor to contact a moving target (as described by Richardson et al., 1990). The moving target (a green circle) moved along a random course across the computer screen whenever the participants moved the cursor, until the cursor and the target collided.

In the ERASE task, participants had to move the cursor to contact several small stationary targets (a 4 x 4 or 5 x 5 grid of small squares on either the left or right side of the screen). Contact with any of the small squares resulted in that individual square disappearing from the screen. Through repeated movements of the cursor, participants had to contact each individual target until the entire grid was gone.

In the FENCE task, participants had to move the cursor through either a one-level or two-level maze to reach a target (green striped square) at the top of the screen. The maze consisted of either one or two horizontal gray bars, each of which had a single gap located along it. The target was located at the top center of the screen. The openings in the maze appeared in different possible locations determined randomly across trials. The participants had to move the cursor with a series of left, right, and upward movements to navigate through the openings in the maze to reach the target.

In the SIDE task, participants had to contact two of the side walls of the computer screen with the cursor. The two walls that had to be contacted were determined at random across trials. Target walls (either the left, right, or upper wall) were designated by a green color along the edge of the computer screen until contacted by the cursor. Participants were allowed to contact the target walls in any order, but the cursor was re-recentered after the first target was contacted.

In the BARRIER task, a large gray rectangle was displayed in the center of the computer screen. Participants needed to navigate the cursor around the large rectangle to contact a green square at the top of the computer screen. If the cursor came in contact with the gray rectangle before reaching the target, it was relocated back to the starting location in the center of the bottom of the screen. Thus, the participants needed to navigate the cursor around the rectangle without touching it to reach the green square. Critically, there were six levels of this task to facilitate training. Level 1 consisted of the largest possible gray rectangle (which was five times the length of the target square) being presented with the green target square below it on either the left or right side of the rectangle. Level 2 consisted of the green target square being presented above the grey rectangle. However, this time the gray rectangle was the same width as the target square. Levels 3 to 6 consisted of the green target square always being presented above the gray rectangle in the center of the screen, while the gray rectangle incrementally progressed in length from being two (Level 3) to five (Level 6) times the length of the target square.

In the CATCH task, either large green squares or small green rectangles moved vertically down the computer screen in a straight line from various starting locations at the top of the screen (determined at random throughout the trial). Participants had to move the cursor to catch the targets as they fell down the screen by lining the cursor up beneath them. When a target was touched by the cursor, it disappeared and the cursor recentered in the bottom center of the computer screen. If the target reached the bottom of the computer screen without contacting the cursor it disappeared and another target began to fall from the top of the screen. The targets continued to fall down the screen from various starting points until either one large green square or two small green rectangles were contacted with the cursor, depending on the testing phase.

Testing began with written instructions shown on the computer screen of the general objective of each task (specifically, “to move the cursor around to touch the green target(s) or clear the screen of the presented object.”). After confirming they had read the instructions, participants indicated their readiness to work by contacting a “start” button in the center of the computer screen with the cursor, using the joystick with their hand. Doing so centered the cursor in the middle of the bottom of the screen and started one trial of one of the described tasks. Upon completing that trial, participants contacted another start button to initiate the next trial. To gain experience with each task, participants progressed through five CHASE trials, four easy (4 x 4 grid) and four hard (5 x 5 grid) ERASE trials, four easy (one horizontal bar) and four hard (two parallel horizontal bars) FENCE trials, five SIDE trials, one trial of each level of BARRIER, and four easy (catch 1 large green square) and four hard (catch 2 small green rectangles) CATCH trials. After completing those 40 trials, participants advanced to the Tool Testing phase.

4.3.2 Tool Testing

In Tool Testing, participants learned which specialized cursor was best suited for each of the tasks (see Table 1). The cursors differed in size, speed, and ability to move through barriers. Importantly, every cursor could be used to complete every task. However, only one of the three cursors led to the most efficient and optimal performance for each task; that is, each cursor was specialized for a specific task and a corresponding generalization task. Cursor 1 moved across the screen quickly (relative to the other cursors), which was beneficial in the CHASE and SIDE tasks. However, it was small, making it a suboptimal cursor for the ERASE and CATCH tasks. Cursor 2 moved slowly but was large, making it an ideal cursor for the ERASE and CATCH tasks (since a larger cursor was able to remove the stationary stimuli from the computer screen faster and catch falling objects without being centered directly beneath them), but not the CHASE or SIDE tasks because of its slow speed. Cursor 3 was slow-moving and small, making it inefficient in the CHASE, SIDE, CATCH, and ERASE tasks. However, only Cursor 3 had the ability to move through barriers (instead of having to go around them). Completing the FENCE and BARRIER tasks with Cursor 3 was, therefore, more efficient because participants were able to travel through the center of the rectangle in the BARRIER task instead of traveling around the outer edge, as well as through the walls of the MAZE task instead of only through the designated openings.

Table 1 Cursor Features and Task Set Assignments

Tool	Key Features	Task Set 1	Task Set 2
	Speed: fast Size: small Barriers: go around	CHASE	SIDE
	Speed: slow Size: large Barriers: go around	ERASE	CATCH
	Speed: slow Size: small Barriers: go through	FENCE	BARRIER

Tasks were completed as described in Task Training, with three exceptions. First, after initiating each trial, participants had to choose one of the three cursors to complete the task with before they could continue with the trial. The cursors were presented at the bottom of the screen while the task was shown (but inaccessible until a cursor choice had been made) in the upper half of the screen (see Figure 4-2). Once a cursor had been chosen, the remaining cursors disappeared, the regular cursor was replaced by the chosen specialized cursor, and the participants were able to complete the task using the chosen cursor. This allowed them to learn the best cursor for each task based on feedback from efficient and inefficient choices.

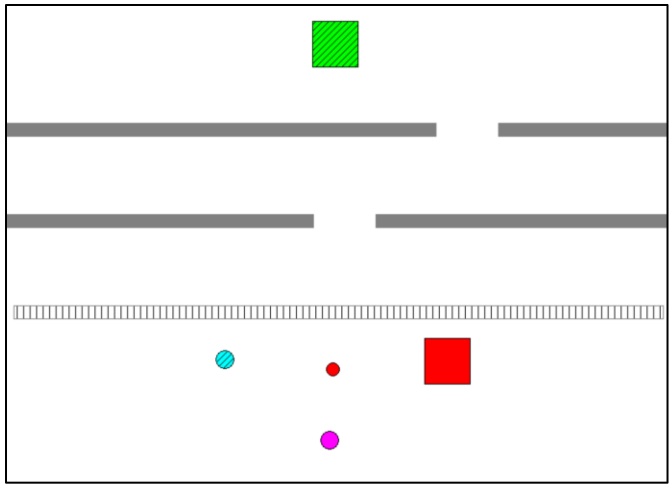

Figure 4-2 Example of Cursor Choice During a Trial
Note. The trial setup was used for Tool Testing, Correction, and Generalization Testing. The location of the cursors was randomized across trials.

The second significant difference from Task Training was that participants got experience using the cursors with only three of the six tasks. Participants were given either the FENCE, CHASE, and ERASE tasks (Task Set 1) or the CATCH, BARRIER, and SIDE tasks (Task Set 2). Finally, before participants started Tool Testing, half of them were given general instructions about the cursors:

"Each cursor is different, and you can choose whichever one you want to use each time. There is a set number of trials you have to complete before you can move on. If you figure out how to complete each trial as fast as possible, the faster you will be done."

With only general information about the cursors, participants had to learn about the cursors' unique properties through experience. The remaining participants were given explicit descriptions of the most important feature of each cursor:

"There are 3 cursors: a large red square, a fast purple circle, and a blue striped circle that can move through walls. You can choose whichever cursor you want to use each

time. There is a set number of trials you have to complete before you can move on. If you figure out how to complete each trial as fast as possible, the faster you will be done."

Within a testing session, the tasks within the assigned set were presented at random. To meet the Tool Testing criteria, participants had to choose the optimal cursor at least seven out of ten trials for each task. The participants who met this training criteria advanced to Generalization Testing (described below). Participants who did not meet these criteria in 60 trials advanced to the Correction phase.

4.3.3 Correction

During the Correction phase, trials proceeded as described in Tool Testing with one exception. If a participant did not choose the optimal cursor at the beginning of the trial, a sad face was shown on the computer screen for three seconds. They then had to complete the task with the correct cursor. Once again, tasks within the assigned set were presented at random. As in Tool Testing, participants had to choose the optimal cursor at least seven times for each task. The participants that met this training criterion within 45 total trials advanced to Generalization Testing. The participants who did not meet this criterion were discontinued from additional testing because they were unable to learn the differential functionalities of the cursors.

4.3.4 Generalization Testing

During Generalization Testing, participants were reintroduced to the three tasks from Task Training that were *not* in their assigned task set. Trials proceeded as described in Tool Testing. That is, at the beginning of each trial participants selected a cursor to complete the presented task with and then immediately began that task with the chosen cursor. However, participants were no longer completing trials for the three tasks they were shown during Tool Testing (and Correction, if needed). Instead, they completed the three tasks from the opposite task set (e.g., participants assigned to Task Set 1 completed SIDE, BARRIER, and CATCH during Generalization Testing). As shown in Table 1, each primary task of the participants' assigned task set corresponded to a transfer task in the opposite task set, such that whichever cursor was best for the primary task was also the optimal cursor for the transfer task, due to similar testing environments. Up to this point, participants did not have any experience using the specialized cursors when faced with the transfer tasks. Therefore, generalizing their knowledge of each cursors' functionality to make efficient choices in the new testing conditions would lead to more overall success and faster optimal performance, compared to learning (again) which cursor was best for the presented tasks through experience. Within a testing session, the transfer tasks for the assigned set were presented at random. Trials otherwise proceeded as previously described in Tool Testing. Participants completed 60 total trials (20 trials per task).

4.4 Results and Discussion

4.4.1 Phase Progression

Of the total number of participants (*N* = 58), half (*n* = 29) received explicit instructions about the salient feature of each cursor (i.e., speed, size, or transparency) and the other half did not. Fifteen participants from each of the instructional conditions were assigned to task set 1. The remaining 14 participants from each condition were assigned to task set 2. The total number of participants that progressed through each phase is shown in Table 2.

Table 2 Number of Participants that Progressed Through Each Phase of Experiment 1

	Total	Completed Generalization Testing	Did Not Require Correction	Passed Correction	Did Not Pass Correction
No Instructions	29	28	8	20	1
Instructions	29	26	14	12	3

4.4.2 Tool Testing

4.4.2.1 Cursor Efficiency

To determine if each of the optimal cursors was, in fact, optimal as they were designed to be (optimal was defined by faster trial completion times), a two-way ANOVA was performed. The effects of accuracy (defined as whether the optimal cursor was chosen or not) and task type on response time were examined using a 2 (accuracy) x 3 (task type) within-subject ANOVA (see Figure 5-3). All post-hoc comparisons were conducted with a Bonferroni correction. The means and standard deviations for each condition are presented in Table 3. It should be noted that due to missing cases (i.e., some participants never chose the correct and/or incorrect cursor for certain tasks), this ANOVA was conducted with a subset (n = 35) of the total sample. Of most interest to this analysis is the main effect and interaction that assess the effect of accuracy on response time, that is, whether participants completed trials faster when they were using the optimal cursor compared to when they were not. Mauchly's Test of Sphericity was statistically significant for the interaction, $\chi^2(2) = 7.88$, $p = .02$, so a Greenhouse-Geisser correction was used. The results indicated a significant main effect of accuracy, $F(1, 34) = 90.42$, $p < .001$, partial $\eta^2 = .73$. Trials completed with the optimal cursor were completed significantly faster (M = 8.06 sec) than when they were not (M = 11.38 sec), $p < .001$. There was a nonsignificant interaction between accuracy and task type, $F(1.65, 56.09) = .71$, $p = .47$.

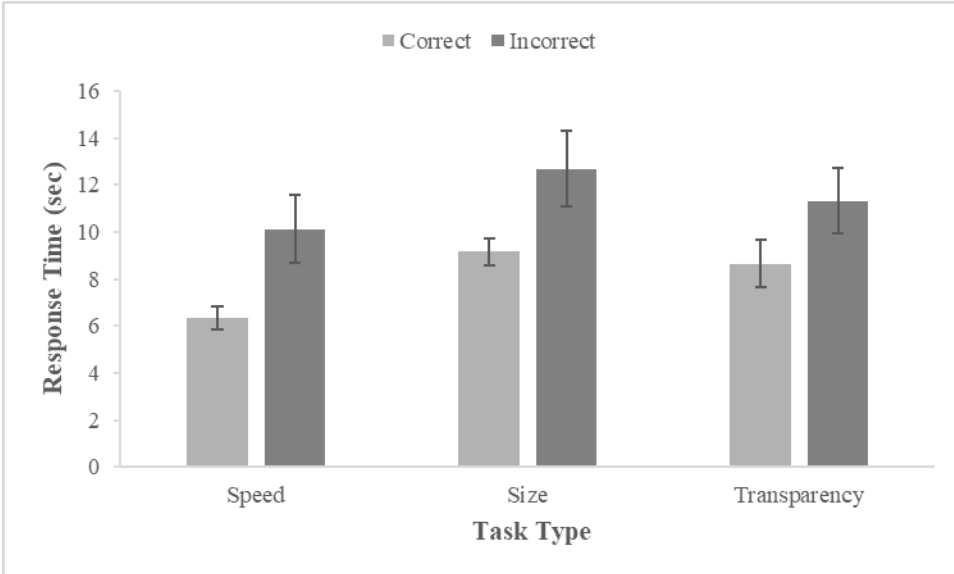

Figure 4-3 Average Response Time During Tool Testing

Note. Accuracy (recorded as correct or incorrect) was defined as whether the optimal cursor was chosen or not. The speed tasks were CHASE and SIDE, size tasks were ERASE and CATCH, and transparency tasks were FENCE and BARRIER. Data were collapsed across task set, instruction, and correction conditions. Errors bars represent 95% confidence intervals.

Table 3 Descriptive Statistics for Response Times During Tool Testing

	Correct		Incorrect	
Task Type	*M*	*SD*	*M*	*SD*
Speed	6.36	1.44	9.17	1.65
Size	9.17	1.65	12.70	4.64
Transparency	8.66	2.96	11.33	4.02

Note. Accuracy (recorded as correct or incorrect) was defined as whether the optimal cursor was chosen or not. The speed tasks were CHASE and SIDE, size tasks were ERASE and CATCH, and transparency tasks were FENCE and BARRIER. Data were collapsed across task set, instruction, and correction conditions.

In addition to differences in response time, whether the participants were using the transparent cursor as intended (i.e., moving through barriers instead of around them) was also calculated. For the BARRIER and FENCE tasks, the amount of time that participants spent moving the cursor through blockades was measured. For the BARRIER task, participants that moved the cursor in a vertical line from the bottom of the screen directly to the target stimulus at the top of the screen (i.e., the most direct route to complete the trial) would have taken approximately one second to traverse the barrier. Participants that moved from the bottom of the screen to the top of the screen during the FENCE task would have spent 1/3 or 2/3 seconds traveling through one or two walls, respectively, to reach the target. Therefore, the optimal route for the BARRIER and FENCE tasks would have produced, on average, scores of 98 and 45, according to how the computer program recorded this. For both the BARRIER (M = 83.86, SD = 41.36) and FENCE (M = 43.52, SD = 26.56) tasks, participants were using the transparent cursor to move through obstacles when completing the transparency tasks with the optimal cursor. Suggesting that, on the group level, they understood how to use the transparent cursor correctly.

4.4.2.2 Overall Performance

A three-way mixed ANOVA was performed to evaluate the effects of receiving explicit instructions or not, the assigned task set, and the task type on performance (see Figure 5-4). All post-hoc pairwise comparisons were conducted with a Bonferroni correction. The means and standard deviations for each condition of the 2 x 2 x 3 factorial design are presented in Table 4. The results indicated a significant main effect for task type, $F(2, 108) = 17.60$, $p < .001$, partial $\eta^2 = .25$, and instructions, $F(1, 54) = 9.93$, $p = .003$, partial $\eta^2 = .16$. The main effect for task set, $F(1, 54) = 2.32$, $p = .13$, was not significant.

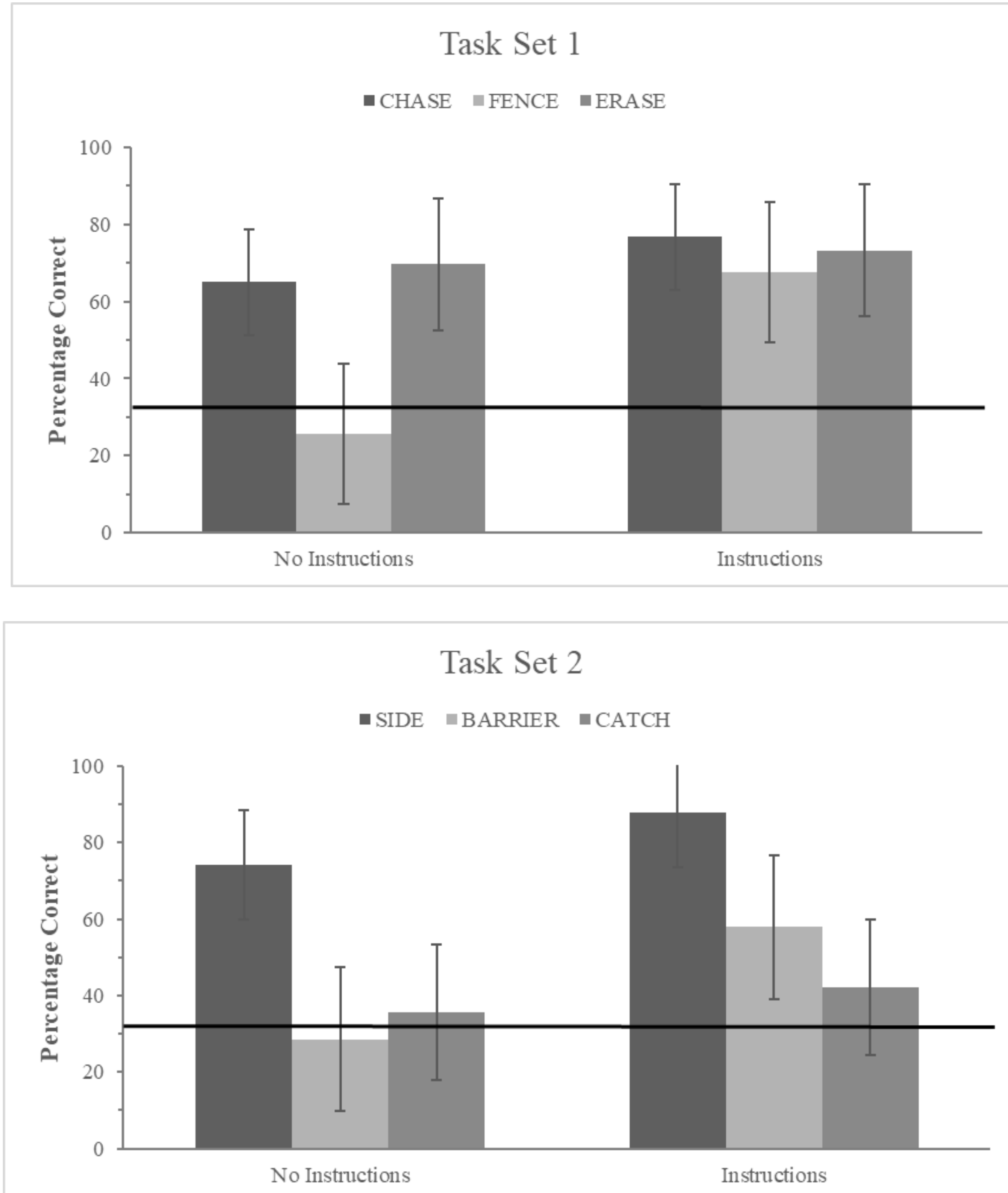

Figure 4-4 Average Performance During Tool Testing

Note. Errors bars represent 95% confidence intervals. Performance reflects the percentage of correct responses, defined as choosing the optimal cursor relative to chance. The black horizontal line represents the 33% chance level.

Table 4 Descriptive Statistics for Performance During Tool Testing

Task Set	Task	Instructions	*M*	*SD*
1	CHASE	No	65.00	31.78
		Yes	76.77	29.70
	FENCE	No	25.56	22.49
		Yes	67.54	36.98
	ERASE	No	69.67	22.71
		Yes	73.22	28.89
2	SIDE	No	74.25	26.97
		Yes	87.81	13.36
	BARRIER	No	28.54	35.05
		Yes	57.91	43.56
	CATCH	No	35.55	36.98
		Yes	42.07	41.30

Note. Performance reflects the percentage of correct responses as defined by choosing the optimal cursor. Task set 1 without instructions $n = 15$; task set 1 with instructions $n = 15$; task set 2 without instructions $n = 14$; task set 2 with instructions $n = 14$.

There was also a significant interaction between task type and task set, $F(2, 108) = 8.40$, $p < .001$, partial $\eta^2 = .14$. Participants assigned to task set 1 struggled significantly more with the FENCE task ($M = 46.55$) than the CHASE task ($M = 70.88$), $p = .001$, and the ERASE task ($M = 71.45$), $p = .006$. Their performance on the CHASE and ERASE tasks did not significantly differ, $p = 1.00$. Participants assigned to task set 2 performed significantly better on the SIDE task ($M = 70.88$) compared to both the CATCH task ($M = 38.81$), $p < .001$, and BARRIER task ($M = 43.22$), $p < .001$. Their performance on the CATCH and BARRIER tasks was not significantly different, $p = 1.00$. Overall, the participants assigned to task set 1 seemed to particularly struggle with their transparency task (FENCE), while the participants assigned to task set 2 struggled with their transparency (BARRIER) task and size (CATCH) task.

Pairwise comparisons were also used to compare the task sets to each other based on the task type (as opposed to comparisons of the individual tasks to each other within each task set). Participants in both task sets performed similarly well in choosing the optimal cursor on the speed tasks (CHASE, $M = 70.88$, and SIDE, $M = 70.88$) and showed similar difficulty on the transparency tasks (FENCE, $M = 46.55$, and BARRIER, $M = 43.22$,). In both cases the difference in performance between task sets was nonsignificant, $p = .15$ and $p = .72$, respectively. However, a significant difference was observed for the size tasks, $p < .001$. Participants assigned to task set 1 performed significantly better on their size task (ERASE, $M = 71.45$) compared to the participants assigned to task set 2 (CATCH, $M = 38.81$). In other words, participants in neither task set struggled with their speed tasks, and they both struggled with their transparency tasks. For the size tasks, the participants assigned to task set 1 (ERASE) outperformed those assigned to task set 2 (CATCH). However, given that the participants assigned to task set 2 struggled with two tasks (BARRIER and CATCH), while participants assigned to task set 1 only significantly struggled with one task (FENCE), this result was expected.

The two-way interaction between task type and instructions was also significant, $F(2, 108) = 4.46$, $p = .01$, partial $\eta^2 = .08$. This interaction indicated that the transparency tasks (BARRIER and FENCE) were the only tasks for which performance was significantly affected by receiving instructions. The participants who received instructions ($M = 62.73$) did significantly better than those who did not ($M = 27.05$), $p < .001$. However, both instructional conditions performed similarly on their speed tasks (SIDE and CHASE; Instructions $M = 82.29$, No Instructions $M =$

69.62), *p* = .08, as well as on their size tasks (CATCH and ERASE; Instructions *M* = 57.65, No Instructions *M* = 52.61), *p* = .56.

The interaction between task set and instructions, *F*(1, 54) = 0.05, *p* = .82, and the three-way interaction, *F*(2, 108) = 0.33, *p* = .72, were non-significant.

4.4.2.3 Cursor Bias

Another potential strategy to complete the present task was for participants to most frequently select the cursor that was the fastest across all tasks (i.e., satisficing responses instead of selecting the optimal cursor for each individual task). Twenty-seven participants (46.6%) selected their overall fastest cursor most often across all three tasks. The same number of participants did not. The most frequent cursor choice could not be calculated for four participants because they chose two of the three cursors equally often. Therefore, there seems to be an overall bias among approximately 50% of the participants to choose their overall fastest cursor.

4.4.3 Correction

4.4.3.1 Overall Performance

A three-way mixed ANOVA was performed to evaluate the effects of receiving explicit instructions, assigned task set, and task type on performance (see Figure 4-5). All post-hoc pairwise comparisons were conducted with a Bonferroni correction. The means and standard deviations for each condition of the 2 x 2 x 3 factorial design are presented in Table 5. The results indicated that there was no significant main effect for task type, *F*(2, 64) = 1.70, *p* = .19, and no significant main effect for instructions, *F*(1, 32) = 0.36, *p* = .56. However, there was a significant main effect for task set, *F*(1, 32) = 11.52, *p* = .002, partial η^2 = .27, such that performance was significantly better for participants assigned to task set 1 (*M* = 93.30) than task set 2 (*M* = 73.76). Therefore, it seems that one particular set of tasks (CHASE, FENCE, and

ERASE) was more amenable to the correction procedures, resulting in better performance. All two-way interactions (task type and task set: $F[2, 64] = 1.35$, $p = .27$; task type and instructions: $F[2, 64] = 1.04$, $p = .36$; task set and instructions: $F[1, 32] = 1.23$, $p = .28$) and the three-way interaction, $F(2, 64) = 1.88$, $p = .16$, were non-significant.

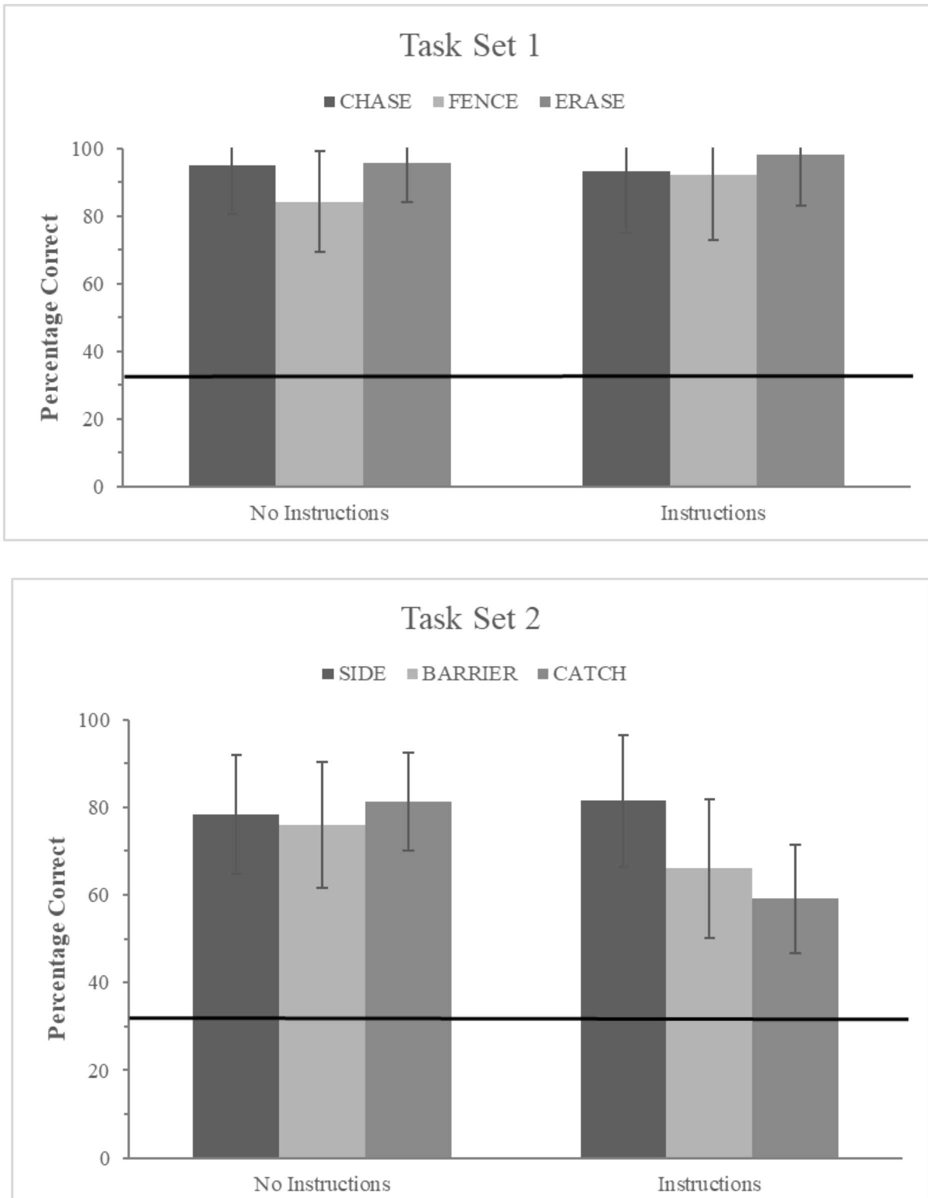

Figure 4-5 Average Performance During the Correction Phase

Note. Performance reflects the percentage of correct responses, defined as choosing the optimal cursor relative to chance. The black horizontal line represents the 33% chance level. Errors bars represent 95% confidence intervals.

Table 5 Descriptive Statistics for Performance During the Correction Phase

Task Set	Task	Instructions	*M*	*SD*
1	CHASE	No	94.89	9.79
		Yes	93.27	10.51
	FENCE	No	84.29	9.69
		Yes	92.27	14.59
	ERASE	No	95.76	10.61
		Yes	98.15	4.54
2	SIDE	No	78.35	28.75
		Yes	81.45	27.45
	BARRIER	No	76.09	26.60
		Yes	66.08	32.57
	CATCH	No	81.41	13.67
		Yes	59.19	30.94

Note. Performance reflects the percentage of correct responses as defined by choosing the optimal cursor. Task set 1 without instructions $n = 10$; task set 1 with instructions $n = 6$; task set 2 without instructions $n = 11$; task set 2 with instructions $n = 9$.

The lack of an interaction involving task type and task set is an interesting contrast to the performance observed during tool testing. During tool testing, certain tasks were significantly more difficult for each task set. However, during the Correction phase, there were no significant differences in performance between task types or individual tasks. This suggests that, even on the hardest tasks, the correction procedure was able to successfully correct the erroneous choices the participants were making. Additionally, whether participants received instructions or not also did not influence performance.

A binary logistic regression was conducted to investigate the effects of assigned task set and receiving explicit instructions on the likelihood of having to complete the Correction phase. The model was not statistically significant, $\chi^2(2) = 4.80$, $p = .09$, explaining only between 8% (Cox & Snell R Square) and 11% (Nagelkerke R Square) of the variance in Correction phase completion. These findings indicate that neither task set nor instructions are important factors in determining the likelihood of passing tool testing.

The effect of experiencing the Correction phase on performance during Generalization Testing is explored in detail below.

4.4.4 Generalization Testing

4.4.4.1 Overall Performance

A three-way mixed ANOVA was performed to evaluate the effects of receiving explicit instructions, assigned task set, and task type on performance (see Figure 4-6). All post-hoc pairwise comparisons were conducted with a Bonferroni correction. The means and standard deviations for all conditions of the 2 x 2 x 3 factorial design are presented in Table 6. The results indicated a significant main effect for task type, $F(2, 100) = 11.03$, $p < .001$, partial $\eta^2 = .18$. However, the main effects for instructions, $F(1, 50) = .13$, $p = .72$, and task set, $F(1, 50) = 2.04$, $p = .16$, were not significant. The two-way interactions between task type and instructions, $F(2, 100) = 1.73$, $p = .18$, and task set and instructions, $F(1, 50) = 0.42$, $p = .52$, were nonsignificant. The three-way interaction, $F(2, 100) = 0.23$, $p = .79$, was also nonsignificant.

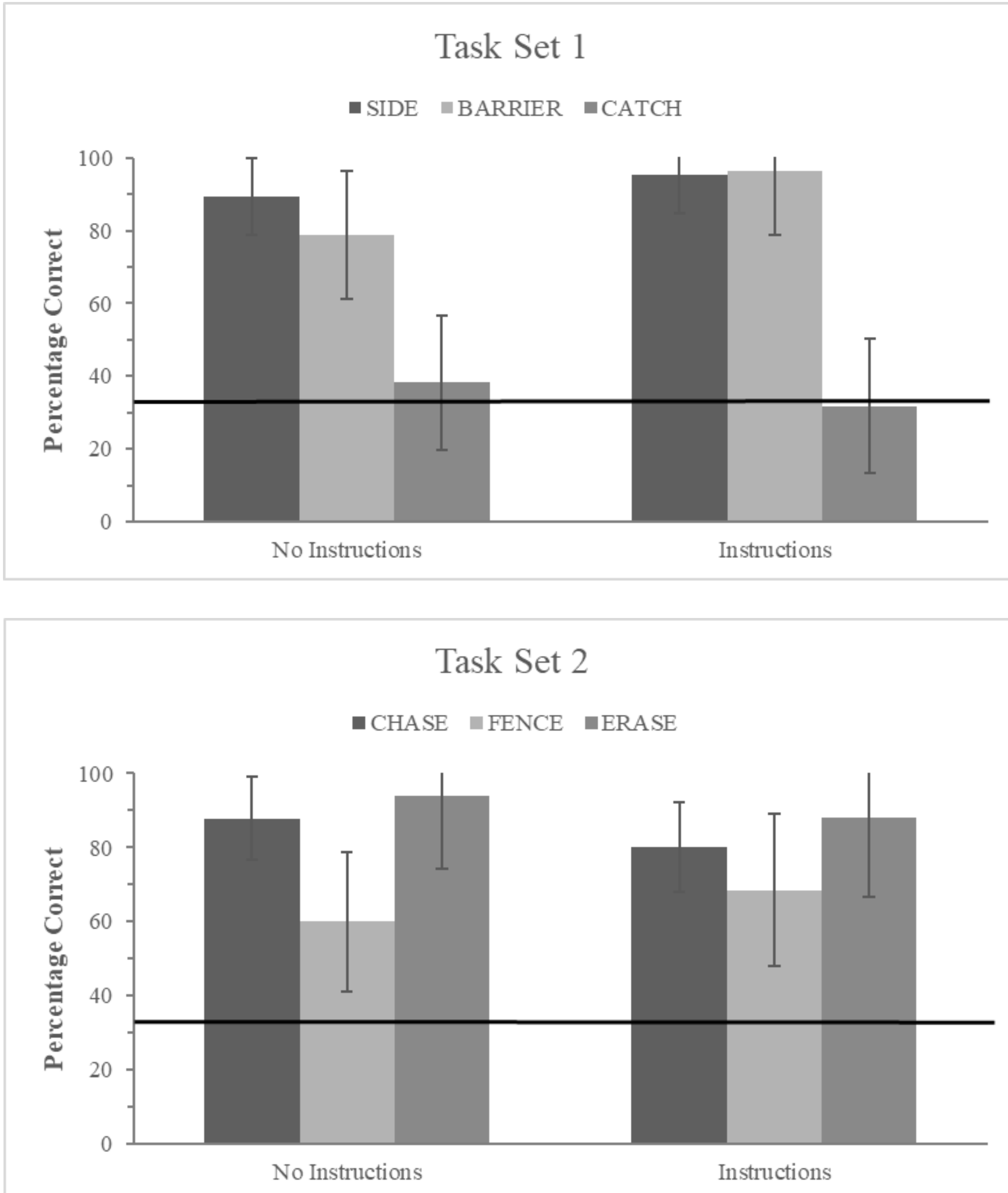

Figure 4-6 Average Performance During Generalization Testing

Note. Performance reflects the percentage of correct responses, defined as choosing the optimal cursor relative to chance. The black horizontal line represents the 33% chance level. Data were collapsed across correction conditions. Errors bars represent 95% confidence intervals.

Table 6 Descriptive Statistics for Performance During Generalization Testing

Task Set	Task	Instructions	*M*	*SD*
1	SIDE	No	89.39	15.73
		Yes	95.32	5.36
	BARRIER	No	78.73	37.68
		Yes	96.39	4.33
	CATCH	No	38.24	43.60
		Yes	31.80	43.99
2	CHASE	No	87.83	18.78
		Yes	80.16	34.92
	FENCE	No	59.95	42.89
		Yes	68.51	38.71
	ERASE	No	93.97	6.41
		Yes	88.17	29.43

Note. Performance reflects the percentage of correct responses as defined by choosing the optimal cursor. Data were collapsed across correction conditions. Task set 1 without instructions $n = 15$; task set 1 with instructions $n = 15$; task set 2 without instructions $n = 13$; task set 2 with instructions $n = 11$.

There was a significant interaction between task type and task set, $F(2, 100) = 31.07$, $p < .001$, partial $\eta^2 = .38$. Participants assigned to task set 1 struggled with the CATCH task ($M = 35.02$) significantly more than the SIDE ($M = 92.36$), $p < .001$, and BARRIER ($M = 87.56$) tasks, $p < .001$. However, performance for these participants on the SIDE and BARRIER tasks did not differ significantly, $p = 1.00$. Participants assigned to task set 2 struggled significantly more with FENCE ($M = 64.23$) than they did ERASE ($M = 91.07$), $p = .02$, as well as CHASE

(*M* = 83.99), $p = .03$. Their performance did not differ significantly between the CHASE and ERASE tasks, $p = .96$.

In other words, when collapsed across instructional conditions, participants did similarly well on two of the assigned tasks but struggled with the remaining task. Interestingly, however, the difficult task for each task set was ideal for different cursors. Participants assigned to task set 1 had a harder time completing the size task during Generalization Testing (CATCH), so much so that they were performing at chance levels (33%). While participants assigned to task set 2 had the most trouble with the transparency task (FENCE). The remaining two tasks were performed similarly well for each task set and both were significantly better than the hardest task within that task set. This is interesting because participants assigned to task set 1 struggled the most with their transparency task during tool testing (FENCE, *M* = 46.55), not the size task. However, participants assigned to task set 2 struggled with both the size task (CATCH, *M* = 38.81) and the transparency task (BARRIER, *M* = 43.22) during tool testing, but only the transparency task during Generalization Testing.

Based on the results thus far, it was clear that receiving instructions did not significantly affect participants' performance during Generalization Testing (unlike tool testing). More notably, the two-way interaction between task set and task type revealed that each task set found a particular task significantly more difficult than the others, specifically one of each of the size (task set 1 - CATCH) and transparency (task set 2 - FENCE) tasks. However, both task sets performed similarly well on their speed tasks (SIDE and CHASE).

To explore the importance of correction on performance during Generalization Testing a three-way mixed ANOVA was performed. Since there was no main effect or interactions for receiving instructions in the previous analysis, participants' data were collapsed across instructional conditions and a 2 x 2 x 3 ANOVA was used to examine the effects of assigned task set, whether the participant did or did not complete the Correction phase, and task type on performance. All post-hoc pairwise comparisons were conducted with a Bonferroni correction. Mauchly's Test of Sphericity was statistically significant, $\chi^2(2) = 6.49$, $p = .04$, so a Greenhouse-Geisser correction was used. Consistent with the previous analysis, there was a significant main effect for task type, $F(1.78, 88.97) = 10.39$, $p < .001$, partial $\eta^2 = .17$.

There was also a significant interaction between task type and task set, $F(1.78, 88.97) = 28.26$, $p < .001$, partial $\eta^2 = .36$. This interaction indicated that participants that were assigned to task set 1 did significantly worse on CATCH ($M = 34.98$) than they did on SIDE ($M = 92.38$) and BARRIER ($M = 88.17$), both $p < .001$. However, there was no difference in performance between SIDE and BARRIER, $p = 1.00$. These pairwise comparisons reflect the same relationships between task type and task set as the previous 2 x 2 x 3 ANOVA (instruction x task set x task type). For the participants assigned to task set 2, none of the pairwise comparisons between the tasks (CHASE, $M = 84.67$; FENCE, $M = 68.74$; ERASE, $M = 92.26$) showed a significant difference in performance, all $p > .05$. Compared to the previous 2 x 2 x 3 analysis, performance on the FENCE task for participants assigned to task set 2 improved when the participants' data were collapsed across correction conditions ($M = 68.74$; compared to instructional conditions, $M = 64.23$). This improvement likely resulted in the lack of a significant difference between tasks for this analysis compared to the previous one (for which performance on FENCE was significantly lower than the other two tasks). However, the same pattern of results for participants assigned to task set 2 remained. That is, participants performed best on the ERASE task, followed by the CHASE task, and then the FENCE task. As a result, FENCE was the lowest performing task for both analyses.

The remaining main effects (task set: $F[1, 50] = 3.05$, $p = .09$; correction: $F[1, 50] = 2.51$, $p = .12$), two-way interactions (task type and correction: $F[1.78, 88.97] = 2.62$, $p = .08$; task set and correction: $F[1, 50] = 0.31$, $p = .58$), and three-way interaction, $F(1.78, 88.97) = 0.10$, $p = .89$, were all non-significant.

Overall, including whether the participants required the Correction phase after completing tool testing as a factor in the mixed ANOVA was not relevant to the relationship between task type and task set. Instead, the results matched the results of the previous 2 x 2 x 3 ANOVA that the specific task being completed predicted performance. However, getting experience with the Correction phase did not. Participants that were able to correctly discern the differential functionalities of the cursors on their own (without need of the Correction phase) did not show better performance during Generalization Testing than those who needed the Correction phase.

4.4.4.2 Generalization

One-sample t-tests indicated that participants' performance (defined as the percentage of optimal cursor choices) on all the tasks except for CATCH, $t(29) = 0.21$, $p = .83$, was significantly better than chance levels (33%) of responding (CHASE, $t[23] = 9.25$, $p < .001$; ERASE, $t[23] = 14.08$, $p < .001$; FENCE, $t[23] = 3.71$, $p = .001$; SIDE, $t[29] = 27.10$, $p < .001$; BARRIER, $t[29] = 10.67$, $p < .001$). Therefore, apart from one task, participants were able to successfully identify the optimal cursor for each of the transfer tasks. To assess whether participants' performance was due to generalization, as opposed to learning through experience which cursor worked best for the transfer tasks, performance was assessed for the first three trials of each task. By looking at performance across only the earliest trials, participants would not have yet had enough experience to associatively learn the correct cursor for each task, but instead would more likely be relying on their knowledge of the cursors' functionality when making cursor selections during their first few choices of each task. One-sample t-tests indicated that participants' performance during the first three trials of each task on all tasks except for CATCH, $t(29)$, = -0.45, $p = .66$, was significantly better than chance levels (CHASE, t[23] = 8.31, $p < .001$; ERASE, t[23] = 10.82, $p < .001$; FENCE, t[23] = 3.41, $p = .002$; SIDE, t[29] = 8.75, $p < .001$; BARRIER, t[29] = 11.38, $p < .001$). Therefore, participants chose the correct cursor at statistically significant rates for five of the six tasks within the first three trials, suggesting successful generalization. These results also match those of the initial binomial tests when performance was assessed across all trials.

To compare performance during the first three trials on each of these tasks to each other, a 2 x 3 ANOVA was used to examine the effects of assigned task set and task type on performance (see Figure 4-7). Because previous analyses indicated that receiving instructions, as well as completing the Correction phase, did not significantly affect performance, participants' data were collapsed across these conditions. The results indicated a significant main effect for task type, $F(2, 104) = 7.09$, $p = .001$, partial $\eta^2 = .12$. There was no significant main effect of task set, $F(1, 52) = 2.09$, $p = .15$. There was also a significant interaction between task type and task set, $F(2, 104) = 25.92$, $p < .001$, partial $\eta^2 = .33$. This pattern of results reflects the same results of the previous 2 x 2 x 3 ANOVAs that assessed overall performance. Therefore, when performance is examined for only the earliest trials of Generalization Testing, the results are the same as when performance is examined overall. This suggests that the optimal performance observed across all trials was established during participants' first three choices for each task.

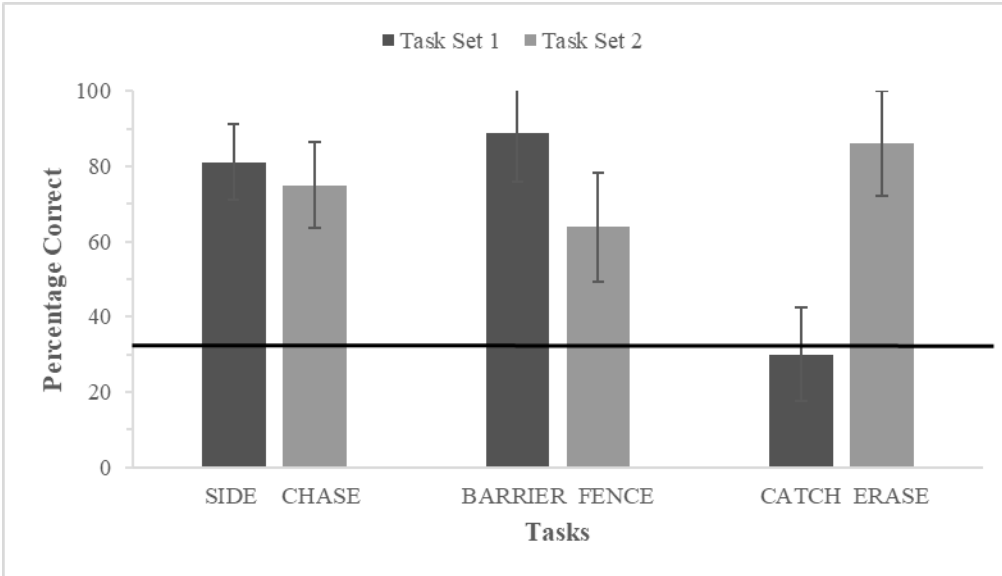

Figure 4-7 Average Performance During the First Three Trials of Each Task During Generalization Testing

Note. Performance reflects the percentage of correct responses, defined as choosing the optimal cursor relative to chance. The black horizontal line represents the 33% chance level. Data were collapsed across instruction and correction conditions. Errors bars represent 95% confidence intervals.

Overall, the results indicated that neither correction nor instruction affected performance during Generalization Testing. Instead, apart from CATCH, participants were able to correctly choose the optimal cursor for the transfer tasks. Additionally, they did so during even the earliest trials of each task, suggesting successful generalization of the cursors' functionalities. Generalization supported optimal performance without requiring extended experience of using the different cursors with the transfer tasks; instead, participants were able to correctly use the cursors with the transfer tasks immediately. Investigation of the relationship between task set and task type on the group level revealed that participants assigned to task set 1 found CATCH to be the hardest task (performing at 33% chance levels), while participants assigned to task set 2 performed the worst on FENCE (although usually not significantly so). This is consistent with initial binomial tests on the group level that indicated that performance was significantly better than chance levels for all the tasks (all $p < .001$), except for CATCH ($p = .14$).

4.4.4.3 Performance Level Comparisons

Task performance was also analyzed based on individual success (defined as performing above 33% chance on one, two, or three tasks) to see if there were any differences in performance for the highest performers compared to those who were less successful. Analyzing the data in this way is important because the strongest performing participants may have overshadowed other participants when analyzed on a group level. On average, out of three tasks, participants chose the correct cursor significantly above chance levels for 2.33 (*SD* = 0.70) tasks. One participant (1.9%) chose correctly for zero tasks, four participants (7.4%) chose correctly for only one task, 25 participants (46.3%) chose correctly for two tasks, and 24 participants (44.4%) chose correctly for all three tasks. Therefore, strong (3 tasks) and average (2 tasks) performers were represented at equal rates. The means and standard deviations for performance on each task, as a function of how many tasks the participants chose the correct cursor for significantly more often than chance, is shown in Table 7. For the participants only performing above chance levels for one task (*n* = 4), the transparency tasks (BARRIER and FENCE; *M* = 27.76, *SD* = 45.60) and the size tasks (CATCH and ERASE; *M* = 25.00, *SD* = 50.00) were the hardest. This is consistent with the previous analyses that consistently showed that CATCH and FENCE were the hardest tasks for the participants assigned to task set 1 and task set 2, respectively. For the participants who performed above chance levels for two tasks (*n* = 25), the size tasks (CATCH and ERASE; *M* = 35.48, *SD* = 43.51) were the hardest. This result is likely due to participants' performance on CATCH being at chance levels.

Table 7 Descriptive Statistics for Performance During Generalization Testing as a Function of the Number of Tasks Participants Performed Above Chance Levels

Number of Tasks	Task	*M*	*SD*
1 (*n* = 4)	SIDE/CHASE	66.25	28.69
	BARRIER/FENCE	27.76	45.60
	CATCH/ERASE	25.00	50.00
2 (*n* = 25)	SIDE/CHASE	89.58	17.87
	BARRIER/FENCE	72.92	39.29
	CATCH/ERASE	35.48	43.51
3 (*n* = 24)	SIDE/CHASE	95.41	5.32
	BARRIER/FENCE	92.74	9.72
	CATCH/ERASE	93.96	8.02

Note. Performance reflects the percentage of correct responses, defined as choosing the optimal cursor relative to chance (33%). Data were collapsed across task set, instruction, and correction conditions. One participant did not choose the correct cursor significantly above chance for any task.

A qualitative comparison of strong and average performing participants shows clear differences in dispersion and variability between the groups. The participants that scored above chance for three tasks showed high performance (≤ 92%) for each task, while the participants that scored above chance for only two tasks showed a gradual increase in performance across tasks (CATCH/ERASE *M* = 35.48, BARRIER/FENCE *M* = 72.92, SIDE/CHASE *M* = 89.58). This can also be seen in the difference in range for the average task performance for each group (range: three tasks = 2.67, two tasks = 54.10). Additionally, the strongest performers also showed much less variability in their performance for all task types compared to the participants with average performance (as defined by the standard deviation for each condition, see Table 7).

Overall, almost half of the participants (44.4%) did well generalizing the cursors to the transfer tasks.

4.4.4.4 Cursor Bias

The analyses of participants' overall performance consistently found CATCH (task set 1) and FENCE (task set 2) tasks to be the hardest tasks for participants. Therefore, identifying whether there was any consistency in which cursor was being chosen incorrectly for these tasks was necessary. For each participant, the most frequent cursor chosen (big, fast, or transparent) when completing the hardest task was identified. The total number of participants for each cursor (big: $n = 1$, fast: $n = 22$, and transparent: $n = 2$) was analyzed using a chi-square goodness-of-fit test, $\chi^2(2, N = 25) = 33.68$, $p < .001$. This analysis revealed that participants erroneously chose the fast cursor significantly more often than expected and the big and transparent cursors both less often than expected. Thus, participants were most likely to (erroneously) select the fast cursor when completing the CATCH and FENCE tasks.

5 EXPERIMENT 2

5.1 Participants

Preschool children (*N* = 27, 14 females), ranging in age from three to five years old (*M* = 56.3 months, *SD* = 6.67 months), were tested at two daycares in Atlanta, Georgia, and one daycare in Decatur, Georgia. Participants were recruited using convenience sampling. Due to attrition (*n* = 2) and experimenter error (*n* = 1), three children did not fully complete Experiment 2, but their data has been included for the phase that they did complete. Children voluntarily worked with experimenters during normal school hours several times per week. These children had experience completing computerized and manual tasks in exchange for a small toy or sticker at the end of each test session, which typically lasted for 10 to 15 minutes. All testing complied with protocols and procedures approved by the Georgia State University Institutional Review Board, including parental informed consent and children's assent at the start of each testing session.

5.2 Apparatus

Participants were tested on a 17-inch laptop computer. They responded using joystick deflections via a USB game controller to control a cursor on the computer screen. Completing each trial resulted in a large yellow smiley face being shown on the computer screen. Incorrect responses, when relevant, resulted in a large yellow sad face being shown on the computer screen. All computer tasks were written in Visual Basic 6.0.

5.3 Design and Procedure

Children experienced the same experimental design and procedures that were used for the adults in Experiment 1, with three exceptions. First, during Task Training, instead of reading general instructions on the computer screen at the beginning of the testing session, experimenters explained to the children how to complete each task using simple instructions during the first trial of each task (e.g., when explaining the CHASE task the experimenter would say “your job is to use this joystick to move the cursor around and catch the green circle”). Second, children were able to complete the experiment across multiple 10 to 15 m testing sessions (as opposed to the adult participants that completed the entire experiment in one 30 to 45 m session). Therefore, they were told or reminded of the task objective(s) at the beginning of every session they completed for Task Training, Tool Testing, Correction, and Generalization Testing. Third, all of the children were given simple and explicit descriptions of the most important feature of each cursor (e.g., “the purple circle can move really fast, the blue striped circle can move through walls, and the red square is really big!”). The children were told or reminded of the special features for each cursor at the beginning of every session they completed for Tool Testing, Correction, and Generalization Testing. This is a stark difference from Experiment 1, for which only half of the participants were given explicit instructions and they only received them at the beginning of Tool Testing, because of the repeated testing sessions that were required for the children. Adults completed Experiment 1 in one testing session. Therefore, the participants that were only given general instructions were still able to rely on relatively recent task experience to inform their cursor choices. Since children completed Experiment 2 across multiple testing sessions, explicit instructions were given to them in order to reduce the memory demands when completing the present study.

Children completed Task Training, Tool Testing, Correction, and Generalization Testing as previously described in Experiment 1. All of the same tasks, cursors, and task sets were used. Participants learned how to complete all six tasks with a regular cursor in Task Training, they then gained experience using the specialized cursors with the three tasks in their assigned task set during Tool Testing and Correction (if required), and then used the specialized cursors for the first time to complete the three transfer tasks from the opposite task set during Generalization Testing. All training criteria, testing criteria, and trial counts were also the same.

5.4 Results and Discussion

5.4.1 Phase Progression

Of the total number of participants (N = 27), fifteen participants were assigned to task set 1. The remaining 12 participants were assigned to task set 2. The total number of participants that progressed through each phase is shown in Table 8. Due to attrition and experimenter error three participants only completed the Tool Testing phase and 24 participants completed the experiment in its entirety. Data from the three participants that did not finish the study was included for the phase that they completed. All three of these participants would have required correction training if they had progressed to the next phase.

Table 8 Number of Participants that Progressed Through Each Phase of Experiment 2

	Total	Completed Tool Testing Phase	Completed Correction Phase	Completed Generalization Testing Phase	Did Not Require Correction	Passed Correction	Did Not Pass Correction
n	27	27	22	12	2	10	12

Note. Three participants only completed the Tool Testing phase due to attrition and experimenter error; all of which would have progressed to the Correction phase if they finished the study. Twenty-four participants completed the experiment in its entirety.

5.4.2 Tool Testing

5.4.2.1 Cursor Efficiency

To determine whether participants completed trials faster when they were using the optimal cursor compared to when they were not, a two-way ANOVA was performed. The effects of cursor accuracy and task type on trial response time were examined using 2 x 3 within-subject

ANOVA (see Figure 5-1). All post-hoc comparisons were conducted with a Bonferroni correction. The means and standard deviations for each condition are presented in Table 9. Due to missing cases, a subset (*n* = 24) of the total sample was used for this analysis. Mauchly's Test of Sphericity was statistically significant for the interaction, $\chi^2(2) = 11.04$, $p = .004$, so a Greenhouse-Geisser correction was used. The results indicated a significant main effect of accuracy, $F(1, 23) = 130.13$, $p < .001$, partial $\eta^2 = .85$. There was also a significant interaction between accuracy and task type, $F(1.43, 32.98) = 9.45$, $p = .002$, partial $\eta^2 = .29$. The optimal cursor was more efficient for the size (ERASE and CATCH) and transparency (FENCE and BARRIER) tasks, both $p < .001$. However, there was no significant difference in response time between trials completed with the correct cursor compared to the incorrect cursors for the speed tasks (CHASE and SIDE), $p = .11$.

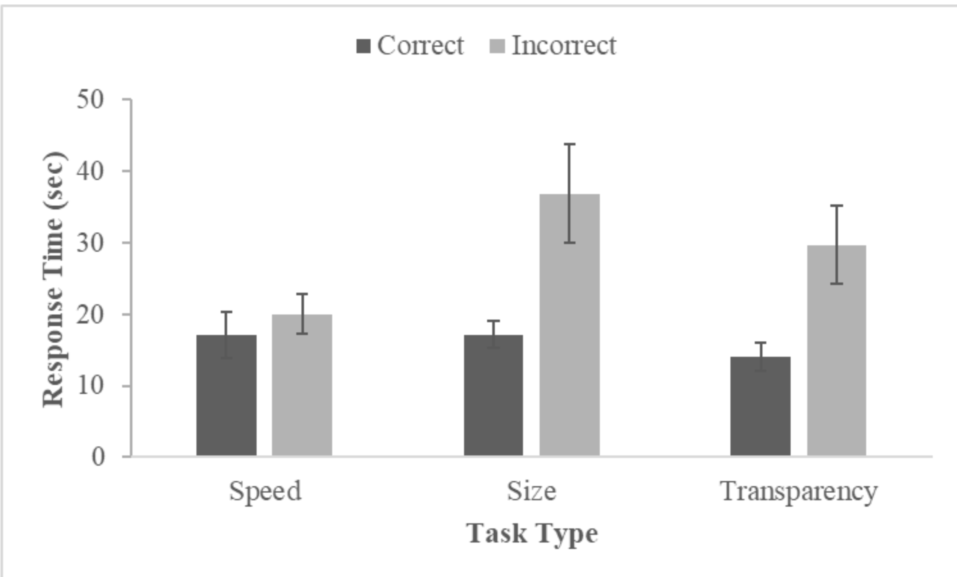

Figure 5-1 Average Response Time During Tool Testing

Note. Accuracy (recorded as correct or incorrect) was defined as whether the optimal cursor was chosen or not. The speed tasks were CHASE and SIDE, size tasks were ERASE and CATCH, and transparency tasks were FENCE and BARRIER. Data were collapsed across task set and correction conditions. Errors bars represent 95% confidence intervals.

Table 9 Descriptive Statistics for Response Times During Tool Testing

	Correct		Incorrect	
Task Type	*M*	*SD*	*M*	*SD*
Speed	17.10	7.79	20.05	6.53
Size	17.16	4.35	36.86	16.38
Transparency	14.05	4.47	29.71	12.85

Note. Accuracy (recorded as correct or incorrect) was defined as whether the optimal cursor was chosen or not. The speed tasks were CHASE and SIDE, size tasks were ERASE and CATCH, and transparency tasks were FENCE and BARRIER. Data were collapsed across task set and correction conditions.

Along with differences in response time, participants' behavior when using the transparent cursor was also assessed. To determine if the cursor was being used to move through barriers (instead of around them) in the BARRIER and FENCE tasks, the duration of time that participants were inside the barrier or fence walls with the cursor was calculated. When the most direct path to complete the trial is taken for either task (moving from the bottom of the screen to the top of the screen in a vertical line), participants would have been in the central barrier for about one second in the BARRIER task and about 1/3 (one wall) or 2/3 seconds (two walls) in the FENCE task. Therefore, the optimal route for the BARRIER and FENCE tasks would have produced, on average, scores of 98 and 45 given how this was measured by the computer program. For both the BARRIER (M = 111.98, SD = 79.10) and FENCE (M = 53.94, SD = 29.72) tasks, participants were using the transparent cursor to move through obstacles when completing the transparency tasks with the optimal cursor. Suggesting that, on the group level, they understood how to use the transparent cursor correctly.

5.4.2.2 Overall Performance

A two-way mixed ANOVA was performed to evaluate the effects of assigned task set and task type on performance (see Figure 5-2). All post-hoc comparisons were conducted with a Bonferroni correction. The means and standard deviations for each condition of the 2 x 3 factorial design are presented in Table 10. There were no significant main effects for task set, $F(1, 25) = 0.08$, $p = .78$, or task type, $F(2, 50) = 0.43$, $p = .65$. The interaction, however, was significant, $F(2, 50) = 4.01$, $p = .02$, partial $\eta^2 = .14$.

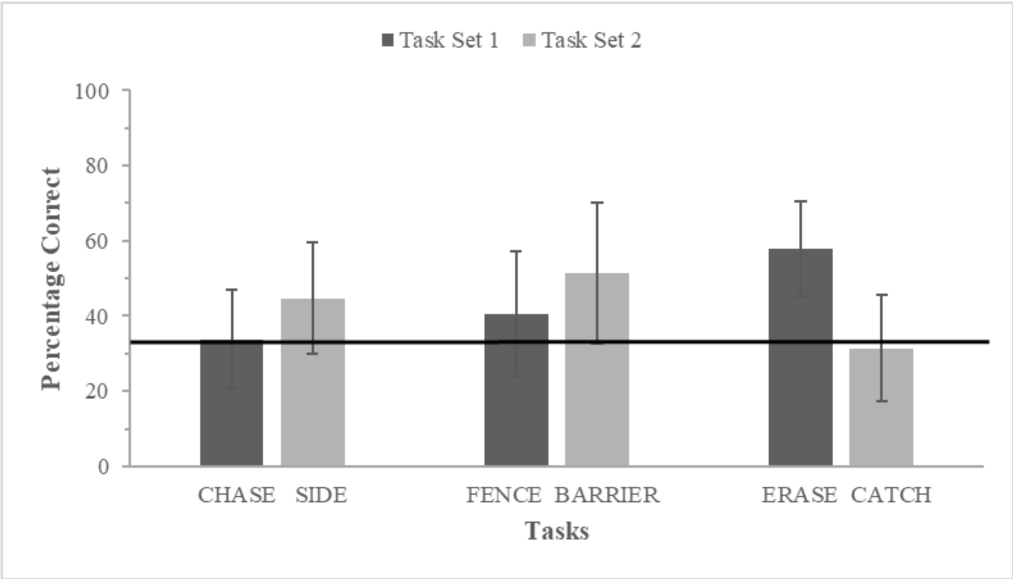

Figure 5-2 Average Performance During Tool Testing

Note. Errors bars represent 95% confidence intervals. Performance reflects the percentage of correct responses, defined as choosing the optimal cursor relative to chance. The black horizontal line represents the 33% chance level.

Table 10 Descriptive Statistics for Performance During Tool Testing

Task Set	Task	*M*	*SD*
Task Set 1	CHASE	33.81	24.95
n = 15	FENCE	40.52	30.53
	ERASE	58.00	25.12
Task Set 2	SIDE	44.76	24.47
n = 12	BARRIER	51.48	33.07
	CATCH	31.37	21.90

Note. Performance reflects the percentage of correct responses as defined by choosing the optimal cursor.

The interaction revealed that the participants assigned to task set 1 did significantly better on their size task (ERASE, *M* = 58.00) than those assigned to task set 2 (CATCH, *M* = 31.37), *p* = .01. The difference in performance between task sets was nonsignificant for the speed tasks (CHASE, *M* = 33.81, and SIDE, *M* = 44.76), *p* = .26, as well as the transparency tasks (FENCE, *M* = 40.52, and BARRIER, *M* = 51.48), *p* = .38. Given that the participants assigned to task set 2 performed below chance levels (33%) for CATCH, this result is not surprising. Interestingly, this pattern of results (i.e., significantly different performance for the size tasks only) mirrors those found for the adult participants (Experiment 1) during tool testing.

Pairwise comparisons also revealed that participants assigned to task set 1 performed significantly worse on CHASE (*M* = 33.81) than they did on ERASE (*M* = 58.00), *p* = .04. However, their performance between CHASE and FENCE (*M* = 40.52), *p* = 1.00, as well as ERASE and FENCE, *p* = .39, did not differ significantly. For the participants assigned to task set 2, none of the pairwise comparisons between the tasks (SIDE, *M* = 44.76; BARRIER, *M* = 51.48; CATCH, *M* = 31.37) showed a significant difference in performance, all *p* > .05.

5.4.2.3 Cursor Bias

Based on the experimenter's observations during testing, it seemed that the children had a bias to choose the same cursor that was chosen on the previous trial. To investigate this potential bias, a one-sample t-test was used to compare the percentage of trials for which the chosen cursor matched the cursor that was chosen in the previous trial to chance levels (33%) of responding. The results indicated that participants chose the cursor that was chosen on trial *n* -1 (M = 49.97, SD = 15.48) significantly more often than chance, $t(26) = 5.58$, $p < .001$. As a result of this inefficient strategy, the majority of the participants (25 of 27) required correction training after completing the Tool Testing phase.

One possible reason why participants' cursor choices matched across trials was that they were choosing the cursor that was the fastest across all tasks. Ten participants (37.0%) selected their overall fastest cursor most often during Tool Testing. The remaining participants (n = 17; 63.0%) did not show this same bias. Therefore, approximately one-third of the children used this alternative strategy, potentially as a form of satisficing.

5.4.3 Correction

5.4.3.1 Overall Performance

A two-way mixed ANOVA was performed to evaluate the effects of assigned task set and task type on performance (see Figure 5-3). The means and standard deviations for each condition of the 2 x 3 factorial design are presented in Table 11. There were no significant main effects for task set, $F(1, 20) = 0.20$, $p = .89$, or task type, $F(2, 40) = 0.02$, $p = .99$. The interaction between task set and task type was also nonsignificant, $F(2, 40) = 1.82$, $p = .18$. These results indicated that participants performed similarly across all tasks and both task sets.

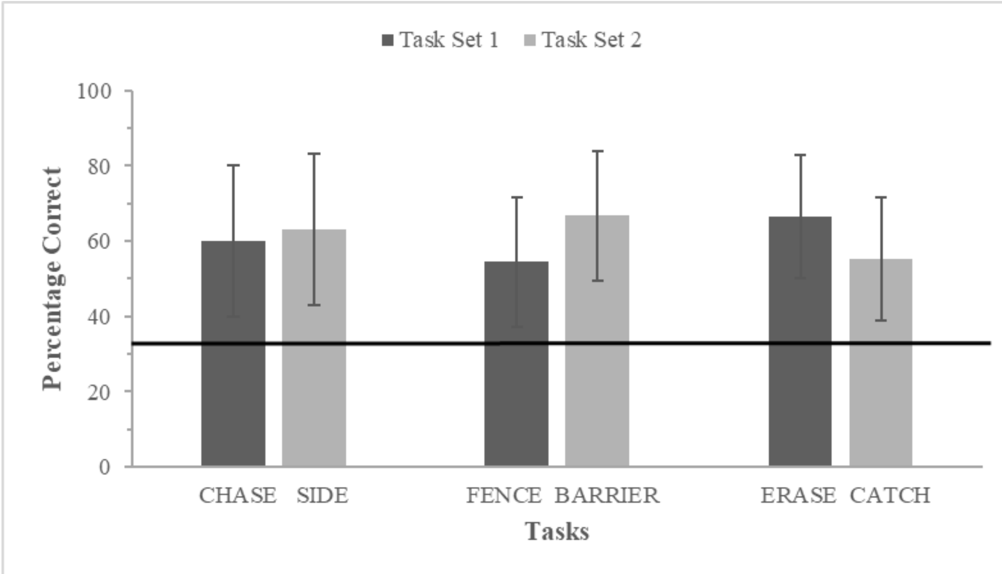

Figure 5-3 Average Performance During the Correction Phase

Note. Errors bars represent 95% confidence intervals. Performance reflects the percentage of correct responses, defined as choosing the optimal cursor relative to chance. The black horizontal line represents the 33% chance level.

Table 11 Descriptive Statistics for Performance During the Correction Phase

Task Set	Task	*M*	*SD*
Task Set 1	CHASE	60.05	31.74
n = 11	FENCE	54.43	24.27
	ERASE	66.49	18.06
Task Set 2	SIDE	63.21	32.37
n = 11	BARRIER	66.75	29.89
	CATCH	55.21	32.20

Note. Performance reflects the percentage of correct responses as defined by choosing the optimal cursor.

A binary logistic regression was conducted to investigate the effects of assigned task set on the likelihood of having to complete the Correction phase. The model was not statistically significant, $\chi^2(1) = 2.48$, $p = .12$, explaining only between 9% (Cox & Snell R Square) and 21%

(Nagelkerke R Square) of the variance in Correction phase completion. These findings indicate that task set was not an important factor in determining the likelihood of passing tool testing.

5.4.3.2 Cursor Bias

As shown in Table 8, 22 participants completed the Correction phase. Twelve of these participants did not meet the criteria to move on to the Generalization Testing phase. To explore whether participants were displaying the same cursor bias that was observed during Tool Testing (that is, choosing the cursor that matched the cursor freely chosen on the previous trial), a one-sample t-test was used to compare the percentage of trials for which the chosen cursor matched the cursor chosen on trial $n - 1$ to chance levels of responding (33%). This analysis was conducted with the subset of participants that did not pass the Correction phase ($n = 12$). Consistent with the Tool Testing phase, the percentage of time participants chose the cursor that matched the previous trial ($M = 44.05$, $SD = 10.31$) was significantly greater than chance, $t(11) = 3.60$, $p = .004$. Therefore, these participants likely did not pass this phase due to their persistent use of the inefficient strategy that they adopted during the Tool Testing phase.

5.4.4 Generalization Testing

5.4.4.1 Overall Performance

A two-way mixed ANOVA was performed to evaluate the effects of assigned task set and task type on performance (see Figure 5-4). The means and standard deviations for each condition of the 2 x 3 factorial design are presented in Table 12. The results indicated nonsignificant main effects for task set, $F(1, 10) = 0.73$, $p = .41$, and task type, $F(2, 20) = 1.76$, $p = .20$. The interaction was also nonsignificant, $F(2, 20) = 0.41$, $p = .67$. Therefore, like the Correction phase, regardless of the task or assigned task set, participants' performance was similar.

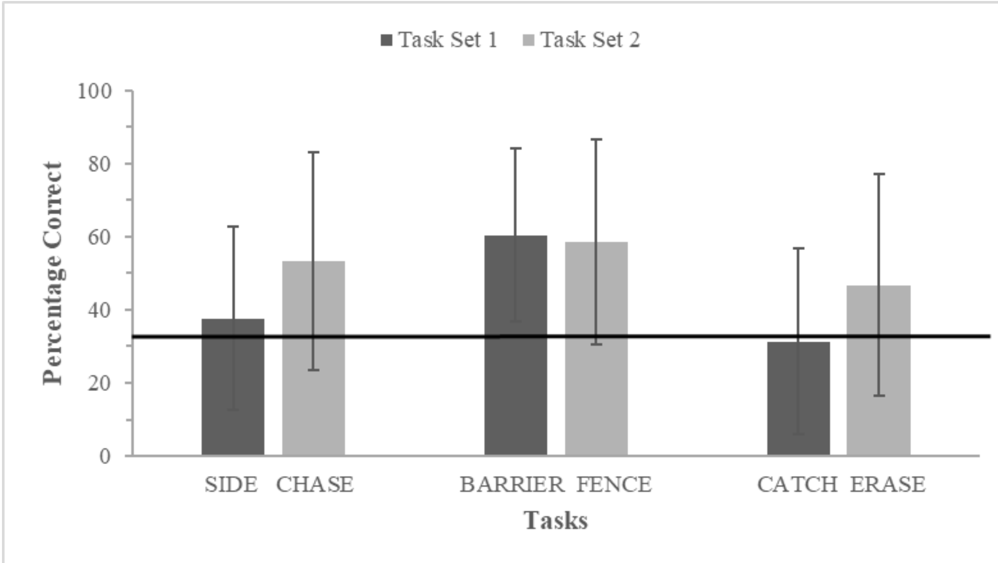

Figure 5-4 Average Performance During Generalization Testing

Note. Errors bars represent 95% confidence intervals. Performance reflects the percentage of correct responses, defined as choosing the optimal cursor relative to chance. The black horizontal line represents the 33% chance level. Data were collapsed across correction conditions.

Table 12 Descriptive Statistics for Performance During Generalization Testing

Task Set	Task	*M*	*SD*
Task Set 1	SIDE	37.58	21.90
n = 7	BARRIER	60.32	25.16
	CATCH	31.32	26.01
Task Set 2	CHASE	53.26	38.65
n = 5	FENCE	58.54	31.98
	ERASE	46.79	35.76

Note. Performance reflects the percentage of correct responses as defined by choosing the optimal cursor. Data were collapsed across correction conditions.

5.4.4.2 Correction Experience

To explore the importance of the Correction phase on performance during Generalization Testing a three-way mixed ANOVA was performed. A 2 x 2 x 3 ANOVA was used to examine the effects of assigned task set, whether the participant did or did not complete the Correction phase, and task type on performance. See Table 13 for the means and standard deviations of each condition. Consistent with the previous analysis there were no significant main effects (task set, $F[1, 9] = 2.36$, $p = .16$; correction, $F[1, 9] = 3.11$, $p = .11$; task type, $F[2, 18] = 1.23$, $p = .32$). The two-way interaction between task type and task set was nonsignificant, $F(2, 18) = 0.35$, $p = .71$, as well as the two-way interaction between task type and correction, $F(2, 18) = 0.01$, $p = .99$. Since there were only two participants that did not experience the Correction phase, the two-way interaction between task set and correction, as well as the three-way interaction, could not be calculated because the degrees of freedom for these effects was zero. Once again, regardless of the task or assigned task set, participants' performance did not significantly differ. These results also suggest that the same is true for whether participants required the Correction phase after completing tool testing. However, given that only two participants were able to correctly discern the differential functionalities of the cursors on their own, any potential effects of correction experience on performance during Generalization Testing likely was not detectable.

Table 13 Descriptive Statistics for Performance During Generalization Testing as a Function of Correction Experience

Task Set	Task	Correction	*M*	*SD*
1	SIDE	No	55.95	8.42
		Yes	30.23	21.57
	BARRIER	No	77.50	3.54
		Yes	53.45	27.19
	CATCH	No	52.11	45.40
		Yes	23.01	14.03
2	CHASE	Yes	53.26	38.65
	FENCE	Yes	58.54	31.98
	ERASE	Yes	46.79	35.76

Note. Performance reflects the percentage of correct responses as defined by choosing the optimal cursor. Task set 1 without correction $n = 2$; task set 1 with correction $n = 5$; task set 2 with correction $n = 5$. There were no participants without correction experience that were assigned to task set 2.

5.4.4.3 Generalization

One-sample t-tests indicated that participants' performance was significantly above chance levels of responding (33%) for the BARRIER task, $t(6) = 2.84$, $p = .03$. However, performance did not significantly differ from chance for any of the other tasks (SIDE, $t[6] = 0.51$, $p = .63$; CATCH, $t[6] = -0.20$, $p = .85$; CHASE, $t[4] = 1.15$, $p = .31$; FENCE, $t[4] = 1.76$, $p = .15$; ERASE, $t[4] = 0.84$, $p = .45$). For the majority of the transfer tasks, participants did not successfully identify the optimal cursor. To assess whether participants' performance in the BARRIER task was due to generalization, performance was assessed for the first three trials of that task. This would allow an assessment of performance before the participants had gained extended experience using the different cursors with the transfer task. A one-sample t-test indicated that participants' performance during the first three trials of the BARRIER task was not significantly better than chance levels, $t(6) = 0.42$, $p = .69$. This suggests that participants learned associatively which cursor was optimal for BARRIER. Unlike the adult participants (Experiment 1), children did not generalize their knowledge of the cursors' functionality to the transfer tasks. Instead, for BARRIER, their success was established through experience over the course of testing.

5.4.4.4 Performance Level Comparisons

Task performance was also analyzed based on individual success, defined as the number of tasks for which the optimal cursor was selected significantly more often than chance (33%). On average, participants chose the correct cursor significantly above chance levels for 1.25 (*SD* = 1.06) tasks. Three participants (25.0%) chose correctly for zero tasks, five participants (41.7%) chose correctly for only one task, two participants (16.7%) chose correctly for two tasks, and two participants (16.7%) chose correctly for all three tasks. The means and standard deviations for performance on each task, as a function of how many tasks the participants chose the correct cursor for significantly above chance levels of responding, is shown in Table 14. The majority of participants performed above chance for only one task, specifically the transparency tasks (BARRIER and FENCE). This is likely due to the aforementioned significantly above chance performance on the group level for the BARRIER task.

Table 14 Descriptive Statistics for Performance During Generalization Testing as a Function of the Number of Tasks Participants Performed Above Chance Levels

Number of Tasks	Task	*M*	*SD*
0 (*n* = 3)	SIDE/CHASE	26.07	6.44
	BARRIER/FENCE	25.87	5.95
	CATCH/ERASE	37.57	13.50
1 (*n* = 5)	SIDE/CHASE	40.04	34.07
	BARRIER/FENCE	65.94	26.80
	CATCH/ERASE	16.18	13.23
2 (*n* = 2)	SIDE/CHASE	47.13	29.77
	BARRIER/FENCE	73.81	10.10
	CATCH/ERASE	40.07	31.72
3 (*n* = 2)	SIDE/CHASE	78.32	23.22
	BARRIER/FENCE	80.00	0.00
	CATCH/ERASE	89.72	7.80

Note. Performance reflects the percentage of correct responses, defined as choosing the optimal cursor relative to chance (33%). Data were collapsed across task set and correction conditions.

5.4.4.5 Cursor Bias

Given participants' bias during the Tool Testing and Correction phases to choose the cursor that matched the cursor chosen in the previous trial, determining whether participants' poor performance during Generalization Testing was due to this same strategy was important. Interestingly, participants did not choose the same cursor as the previous trial ($M = 40.54$, $SD = 14.22$) at a percentage significantly different from chance, $t(11) = 1.76$, $p = .11$. It should be noted, however, that the absence of this bias during Generalization Testing could have been due to exclusion of the children that did not pass the Correction phase (i.e., the previous bias was driven by the low performing children that did not meet the Correction phase criterion). Therefore, the percentage of time the participants chose the same cursor that was used on the previous trial was re-analyzed for the Tool Testing phase but only for the subset of children (n = 12) that advanced to Generalization Testing, to determine if these participants showed the original bias. On average (M = 49.02, SD = 15.72), the children that completed the Generalization Testing phase had shown a previous bias during the Tool Testing phase to choose the same cursor as trial n -1 at a percentage significantly above chance levels, $t(11) = 3.46$, $p < .01$. Therefore, the absence of this bias during Generalization Testing was the result of a change in the participants' choice behavior. When presented with the transfer tasks, despite their poor performance, the participants that progressed to the generalization phase did not fall back on the strategy that they used during both of the previous phases.

To explore whether a general bias for the transparent cursor across tasks was the reason participants' performance was above chance for the BARRIER task, the overall percentage choice of each cursor as a function of task (Table 15) was calculated and compared. The transparency tasks were the only tasks for which the children chose the transparent cursor most often (BARRIER – 60.0%, FENCE – 59.4%). In contrast, the fast cursor was chosen the most often for both speed tasks (SIDE – 37.9%, CHASE – 54.0%) and CATCH (49.6%). Participants assigned to Task Set 2 chose the big cursor the most often for ERASE (46.5%). Therefore, a general bias for the transparent cursor was not the reason that performance in the BARRIER task was significantly above chance levels of responding. Instead, the participants assigned to task set 1 seemed to have a bias for the fast cursor, choosing it most frequently for the SIDE and CATCH tasks. The participants assigned to task set 2 chose the correct cursor most often for all of their tasks (CHASE, ERASE, FENCE) but not enough to significantly exceed 33% chance levels of performance.

Table 15 Percentage Choice of Each Cursor as a Function of Task During Generalization Testing

	Task	Cursor		
		Big	Fast	Transparent
Task Set 1	SIDE	33.1	**37.9**	29.0
	BARRIER	27.9	12.1	**60.0**
	CATCH	31.1	**49.6**	19.3
Task Set 2	CHASE	39.0	**54.0**	7.0
	FENCE	13.9	26.7	**59.4**
	ERASE	**46.5**	44.4	9.1

Note. Bolded percentages indicate the optimal cursor for each task.

6 EXPERIMENT 3

6.1 Participants

Adult male rhesus macaques (*Macaca mulatta*; *N* = 3) and adult capuchin monkeys (*Sapajus apella*; *N* = 19; 14 females) housed at Georgia State University's Language Research Center participated in the study. Participant sampling was based on convenience. Due to lack of progress and/or task engagement, 11 capuchin monkeys were removed from the study during testing, leaving a total of three rhesus monkeys and eight capuchin monkeys (seven females) that completed the study in its entirety. Each monkey was tested individually in their home enclosure but had constant visual and auditory access to nearby monkeys, as well as access to a compatible social partner or group and/or an outdoor play yard area multiple times per week, during which time no testing occurred. Food and water deprivation were not used. Instead, all the monkeys worked as they chose for food rewards (fruit-flavored chow pellets). The monkeys were fed a daily diet of primate chow biscuits and various fruits and vegetables regardless of their performance on the tasks, as well as provided with continuous access to water. All research procedures followed guidelines for working with nonhuman primates and were approved by the GSU Institutional Animal Care and Use Committee. In addition, GSU is accredited by the Association for Assessment and Accreditation of Laboratory Animal Care.

6.2 Apparatus

The monkeys were tested using the Language Research Center's Computerized Test System, which consists of a personal computer, digital joystick, color monitor, and pellet dispenser (Evans et al., 2008; Richardson et al., 1990). The monkeys used their hands to manipulate the joystick in order to control a cursor on the computer screen, and all monkeys have extensive experience doing so for a variety of computerized cognitive tasks. Completing each trial resulted in a brief melodic chime and the delivery of a fruit-flavored chow pellet. All computer tasks were written in Visual Basic 6.0.

6.3 Design and Procedure

6.3.1 Task Training

The monkeys completed the same six psychomotor tasks as the human participants, as described in previous experiments. Participants indicated their readiness to work at the beginning of each trial by contacting a button with a regular cursor and then proceeded to complete the respective trial.

Task training consisted of two parts. To complete part 1, monkeys progressed through 15 CHASE trials, 10 easy (4 x 4 grid) and 15 hard (5 x 5 grid) ERASE trials, 10 easy (1 horizontal bar) and 15 hard (2 parallel horizontal bars) FENCE trials, 15 SIDE trials, 5 trials of each level of BARRIER, and 10 easy (catch 1 large green square) and 15 hard (catch 2 small green rectangles) CATCH trials. If a monkey did not complete all of the training trials in one session, they started the next session of task training at the beginning of the block for the task that they had not yet completed and then proceeded to progress through the remaining tasks as described above. The second part of task training required monkeys to complete 10 consecutive trials of each task (60 total trials) in one testing session. After completing two 60-trial sessions in this manner, monkeys advanced to the next phase.

6.3.2 Tool Testing

As in Experiments 1 and 2, the monkeys completed Tool Testing with tasks from only one of the two task sets. The specialized cursors, tasks, and trial procedures were the same as those used by the human participants. However, instead of a yellow smiley face being shown on the computer screen at the completion of a trial, monkeys received a fruit-flavored chow pellet.

Within a testing session, tasks within the assigned set were presented at random. To meet the Tool Testing criteria, monkeys had to choose the optimal cursor on at least 20 out of 30 trials for each task in a single session. The animals that met this training criterion advanced to Generalization Testing (described below). Monkeys that did not meet this criterion in 900 trials advanced to the Correction phase.

6.3.3 Correction

Trials proceeded as described in Tool Testing with one exception. If a monkey did not choose the optimal cursor at the beginning of the trial, they experienced a 20-second time-out. After the time-out, the correct cursor was presented on the screen with the same task from that trial and they had to complete the task with the correct cursor. Once again, within a testing session, tasks within the assigned set were presented at random. As in Tool Testing, monkeys had to choose the optimal cursor during at least 20 out of 30 trials for each task. The animals that meet this training criterion advanced to Generalization Testing. The monkeys that did not meet this criterion in 900 trials were discontinued in the study because they were unable to learn the differential functionalities of the cursors.

6.3.4 Generalization Testing

Monkeys were reintroduced to the three tasks from Task Training that were *not* in their assigned task set. As with the human participants, instead of completing trials for the three tasks they had experience with during Tool Testing and Correction (if needed), the monkeys used the specialized cursors for the first time to complete the three transfer tasks from the opposite task set for which they were assigned. Within a testing session, the transfer tasks were presented at random. Trials otherwise proceeded as previously described in Tool Testing. The monkeys completed 900 total trials (300 trials per task).

6.4 Results and Discussion

6.4.1 Phase Progression

Given the small sample size, capuchin monkeys and rhesus monkeys were analyzed as one group for all analyses. Of the total number of monkeys that were included in the analyses (*N* = 11), five of them were assigned to task set 1. The remaining six monkeys were assigned to task set 2. The total number of individuals that progressed through each phase is shown in Table 16. Due to low task engagement, three monkeys did not finish the Correction phase. Due to testing errors, two monkeys (one during the Correction phase and one during Tool Testing) did not complete the correct number of trials. Therefore, there were six monkeys with complete data sets and five monkeys with partial data sets. Data from the monkeys that did not complete the whole study were included only for the phases that they fully completed.

Table 16 Number of Participants that Progressed Through Each Phase of Experiment 3

	Total	Completed Tool Testing Phase	Completed Correction Phase	Completed Generalization Testing Phase	Did Not Require Correction	Passed Correction	Did Not Pass Correction
n	11	10	7	1	0	1	6

Note. Three monkeys were excluded from the Correction phase due to low task engagement. One monkey was excluded from the Correction phase because of a testing error. One monkey was excluded from the Tool Testing phase because of a testing error. Six monkeys completed the experiment in its entirety.

6.4.2 Tool Testing

6.4.2.1 Cursor Efficiency

The effect of cursor accuracy (defined as whether the optimal cursor was chosen or not) and task type on trial response time was examined using a 2 x 3 within-subject ANOVA (see Figure 6-1). The means and standard deviations for each condition are presented in Table 17. Due to a missing case, a subset (n = 9) of the total sample was used for this analysis. There was a significant main effect of accuracy, $F(1, 8) = 10.79$, $p = .01$, partial $\eta^2 = .57$. As expected, trials completed with the optimal cursor were finished significantly faster (*M* = 11.24) than trials when the incorrect cursor was used (*M* = 18.02), *p* = .01. There was no interaction between accuracy and task type, *F*(2, 16) = 2.41, *p* = .12, suggesting that regardless of the task using the optimal cursor led to faster trial completion times.

Table 17 Descriptive Statistics for Response Times During Tool Testing

	Correct		Incorrect	
Task Type	*M*	*SD*	*M*	*SD*
Speed	6.99	1.29	12.35	2.69
Size	16.90	10.11	19.13	6.37
Transparency	9.82	6.67	22.56	12.44

Note. Accuracy was defined as whether the optimal cursor was chosen or not. The speed tasks were CHASE and SIDE, size tasks were ERASE and CATCH, and transparency tasks were FENCE and BARRIER. Data were collapsed across task sets.

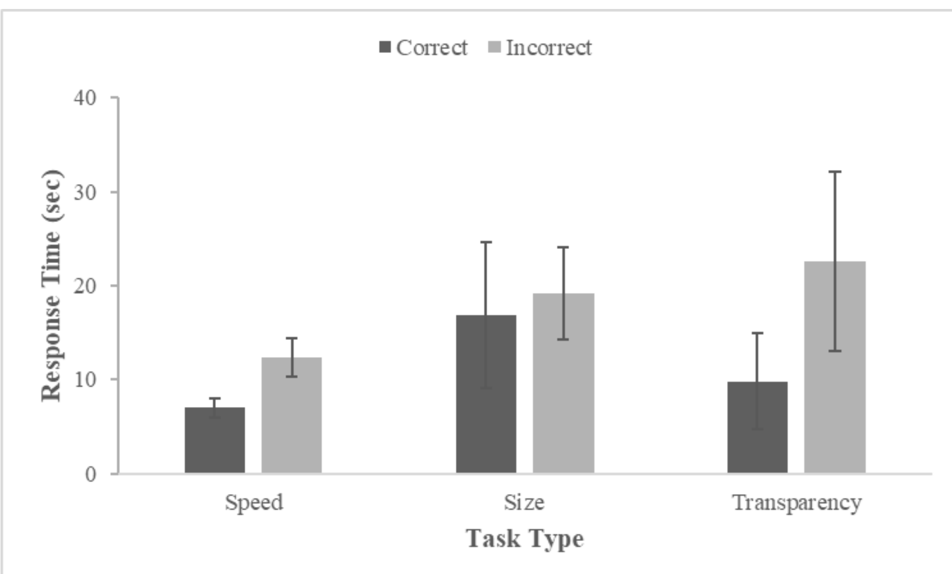

Figure 6-1 Average Response Time During Tool Testing

Note. Accuracy was defined as whether the optimal cursor was chosen or not. The speed tasks were CHASE and SIDE, size tasks were ERASE and CATCH, and transparency tasks were FENCE and BARRIER. Data were collapsed across task sets. Errors bars represent 95% confidence intervals.

Participants' use of the transparent cursor was also assessed to see if they were using the cursor to move through barriers (as opposed to around them) when completing the BARRIER and FENCE tasks. The most direct path to complete one trial of each task involved moving through barriers for approximately one second (BARRIER) or either 1/3 or 2/3 seconds (FENCE, one wall or two walls, respectively). Therefore, the optimal route for the BARRIER and FENCE tasks would have produced, on average, scores of 98 and 45 according to how the computer program recorded this. For both the BARRIER (M = 101.21, SD = 50.61) and FENCE (M = 56.23, SD = 56.63) tasks, monkeys seem to be generally using the transparent cursor to move through obstacles when completing the transparency tasks with the optimal cursor. Suggesting that, on the group level, they understood how to use the transparent cursor correctly.

6.4.2.2 Overall Performance

A two-way mixed ANOVA was performed to evaluate the effects of assigned task set and task type on performance (see Figure 6-2). All post-hoc comparisons were conducted with a Bonferroni correction. The means and standard deviations for each condition of the 2 x 3 factorial design are presented in Table 18. Mauchly's Test of Sphericity was statistically significant for the interaction, $\chi^2(2) = 6.89$, $p = .03$, so a Greenhouse-Geisser correction was used. There were no significant main effects for task type, $F(1.23, 9.84) = 0.46$, $p = .55$, or task set, $F(1, 8) = 4.25$, $p = .07$. However, the interaction was significant, $F(1.23, 9.84) = 10.03$, $p = .01$, partial $\eta^2 = .56$.

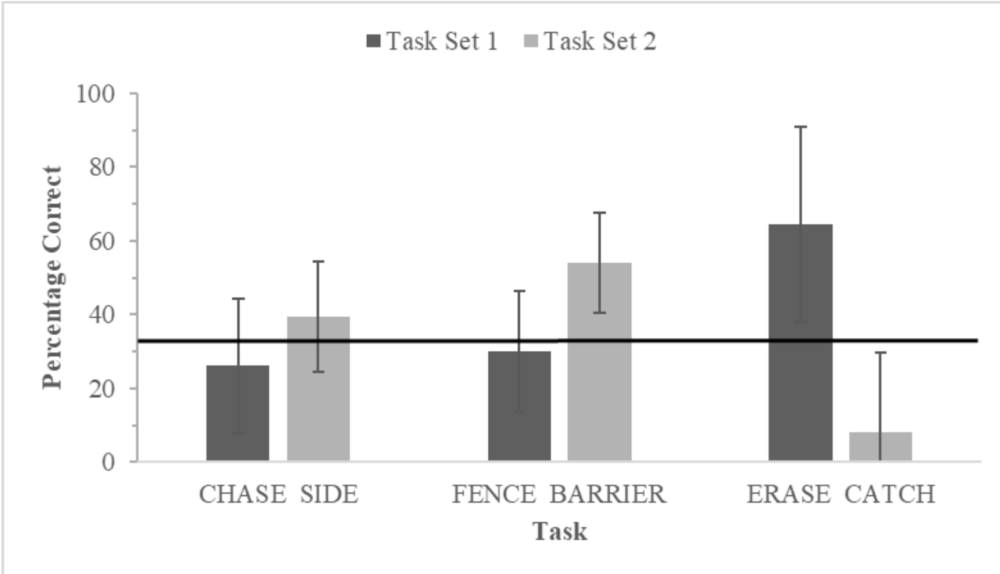

Figure 6-2 Average Performance During Tool Testing

Note. Errors bars represent 95% confidence intervals. Performance reflects the percentage of correct responses, defined as choosing the optimal cursor relative to chance. The black horizontal line represents the 33% chance level.

Table 18 Descriptive Statistics for Performance During Tool Testing

Task Set	Task	*M*	*SD*
Task Set 1	CHASE	26.08	22.98
n = 4	FENCE	30.09	22.52
	ERASE	64.50	32.79
Task Set 2	SIDE	39.46	9.17
n = 6	BARRIER	54.05	4.68
	CATCH	8.23	13.81

Note. Performance reflects the percentage of correct responses as defined by choosing the optimal cursor.

The interaction revealed that the monkeys assigned to task set 1 did significantly better on their size task (ERASE, $M = 64.51$) than those assigned to task set 2 (CATCH, $M = 8.23$), $p = .01$. The difference in monkeys' performance on the transparency tasks showed the opposite pattern. Those assigned to task set 2 (BARRIER, *M* = 54.05) significantly outperformed the animals assigned to task set 1 (FENCE, *M* = 30.09), *p* = .03. The difference in performance between task sets was nonsignificant for the speed tasks (CHASE, *M* = 26.08, and SIDE, *M* = 39.47), *p* = .23. Given that ERASE (task set 1) and BARRIER (task set 2) were the tasks with the highest percentage of correct performance for the monkeys, it was unsurprising that performance on those tasks was significantly better than performance on the respective transfer tasks.

For the monkeys assigned to task set 1, none of the pairwise comparisons between the tasks (CHASE, *M* = 26.08; FENCE, *M* = 30.09; ERASE, *M* = 64.51) showed a significant difference in performance in terms of choosing the optimal cursor, all $p > .05$. Pairwise comparisons revealed that monkeys assigned to task set 2 performed significantly better on BARRIER (*M* = 54.05) than they did on CATCH (*M* = 8.23), *p* = .03. However, their performance between SIDE (*M* = 39.47) and BARRIER, *p* = .14, as well as SIDE and CATCH, *p* = .24, did not differ significantly. Given the extremely low performance on CATCH in terms of choosing the optimal cursor, for the participants assigned to task set 2, this result was not surprising.

6.4.2.3 Cursor Bias

As was done for the previous experiment with children, a one-sample t-test was conducted to explore whether monkeys were relying on the strategy to choose the same cursor that was chosen on the previous trial. To do this, the percentage of trials for which the chosen cursor matched the cursor that was chosen in the previous trial was compared to chance levels (33%) of responding. The results indicated that the monkeys chose the cursor that was chosen on trial *n* -1 (*M* = 50.66, *SD* = 17.46) significantly more often than chance, $t(9) = 3.14$, *p* = .01. As a result of this generally ineffective strategy, all the monkeys required correction training after completing the Tool Testing phase.

One possible reason why participants' cursor choices were the same across consecutive trials was that they were choosing the cursor that was the fastest across all tasks as a form of satisficing rather than maximizing optimality. Four monkeys (40%) selected their overall fastest cursor most often during Tool Testing. The remaining six monkeys (60%) did not. Therefore, the majority of the monkeys were not using this alternative optimizing strategy.

6.4.3 Correction

6.4.3.1 Overall Performance

A two-way mixed ANOVA was performed to evaluate the effects of assigned task set and task type on performance (see Figure 6-3). The means and standard deviations for each condition of the 2 x 3 factorial design are presented in Table 19. There were no significant main effects for task set, $F(1, 5) = 1.31$, $p = .30$, or task type, $F(2, 10) = 0.13$, $p = .88$. The interaction between task set and task type was also nonsignificant, $F(2, 10) = 4.04$, $p = .052$. The results indicated that the monkeys performed similarly across all tasks and both task sets.

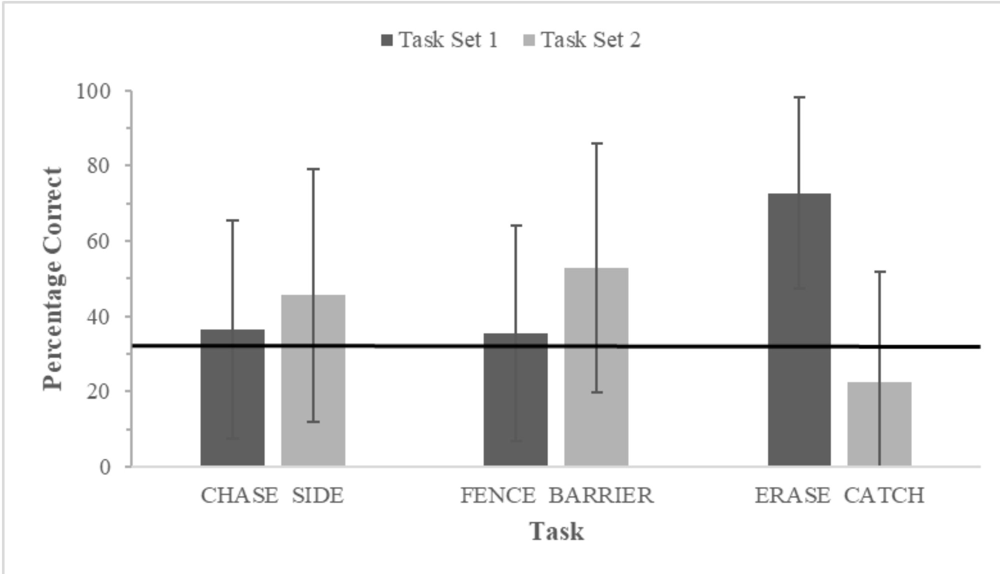

Figure 6-3 Average Performance During the Correction Phase

Note. Errors bars represent 95% confidence intervals. Performance reflects the percentage of correct responses, defined as choosing the optimal cursor relative to chance. The black horizontal line represents the 33% chance level.

Table 19 Descriptive Statistics for Performance During the Correction Phase

Task Set	Task	*M*	*SD*
Task Set 1	CHASE	36.40	28.47
n = 4	FENCE	35.50	27.53
	ERASE	72.77	22.73
Task Set 2	SIDE	45.59	7.56
n = 3	BARRIER	52.91	10.31
	CATCH	22.52	14.44

Note. Performance reflects the percentage of correct responses as defined by choosing the optimal cursor.

6.4.3.2 Performance Over Time

The animals' cursor choices were organized into 10 blocks in order to assess performance over the course of the Correction phase. For the monkeys that did not pass Correction, and therefore completed all 900 trials of this phase, ten blocks of 90 trials each were analyzed. For one capuchin monkey (Paddy), the only monkey to meet the Correction criterion (across 450 total trials), ten blocks of 45 trials were analyzed. Group and individual performances are shown in Figure 6-4 and Figure 6-5, respectively. At the group level, there was a significant effect of testing block, $F(9, 54) = 2.08$, $p = .048$. However, none of the pairwise comparisons between blocks were significant, all $p > .05$. Visual observation of Figure 6-4 shows an increase in the group's overall performance in Blocks 9 and 10. On an individual level, even in Block 1, Paddy was choosing the optimal cursor more often than the other monkeys. Therefore, it is likely that the increase in performance that the monkeys experienced overall in the final blocks of testing was sufficient improvement for Paddy because her performance was already higher than the other monkeys. However, her performance decreased during Blocks 5 through 7 to levels more consistent with group performance. None of the other monkeys ever showed significant improvement in their performance.

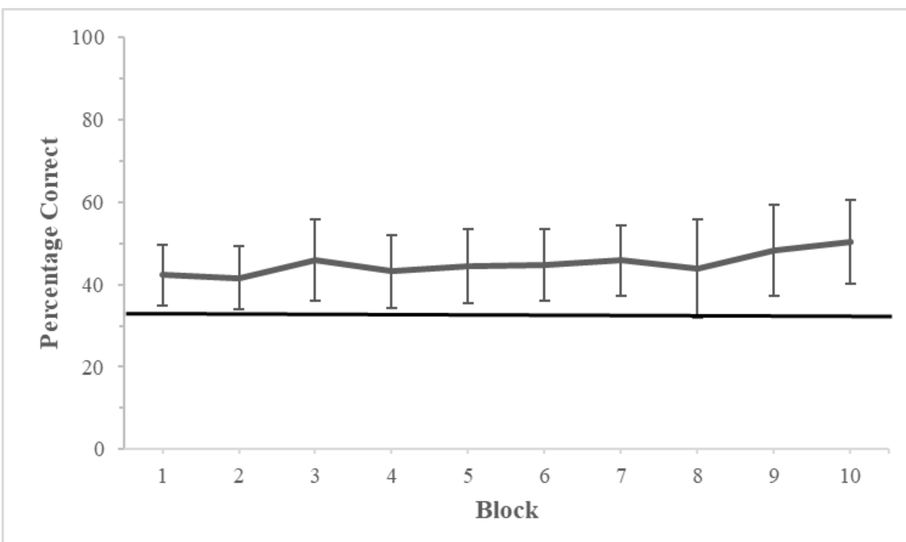

Figure 6-4 Average Performance During the Correction Phase as a Function of Block

Note. Errors bars represent 95% confidence intervals. Performance reflects the percentage of correct responses, defined as choosing the optimal cursor relative to chance. The black horizontal line represents the 33% chance level. Data were collapsed across task set and task type.

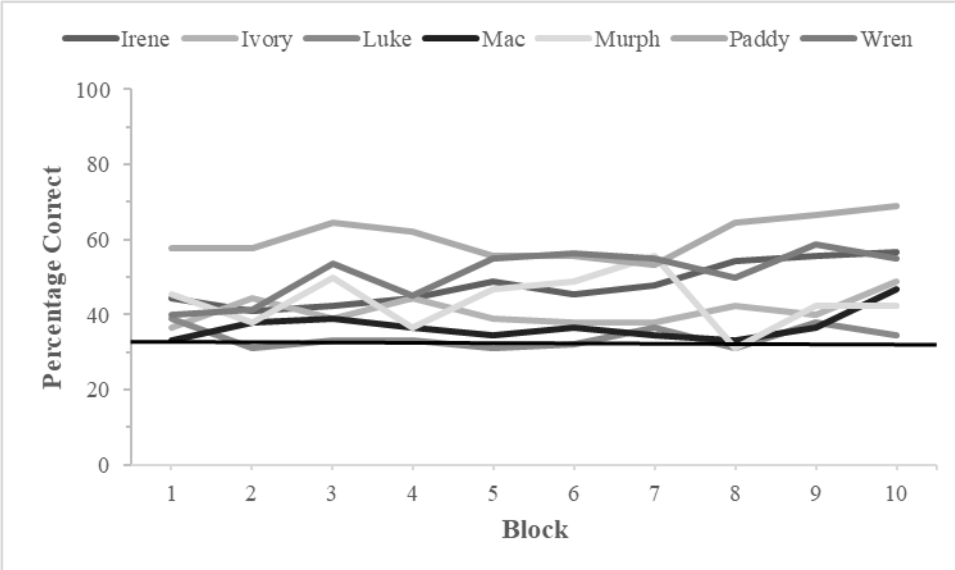

Figure 6-5 Individual Performance During the Correction Phase as a Function of Block

Note. Performance reflects the percentage of correct responses, defined as choosing the optimal cursor relative to chance. The straight black horizontal line represents the 33% chance level. Data were collapsed across task set and task type. Luke, Mac, and Murph are adult male rhesus macaques. The remaining four monkeys are adult female capuchin monkeys.

6.4.3.3 Cursor Bias

As shown in Table 16, seven participants completed the Correction phase. Six of these participants did not meet the criteria to move on to Generalization Testing. To explore whether the monkeys were displaying the same cursor bias that was observed during Tool Testing (i.e., choosing the same cursor that was freely chosen on the previous trial), a one-sample t-test was conducted to compare the percentage of trials for which the chosen cursor matched the cursor chosen on trial $n - 1$ to chance levels of responding (33%). This analysis was conducted with the subset of animals that did not pass the Correction phase ($n = 6$). The percentage of time the monkeys chose the previous trial's cursor ($M = 49.55$, $SD = 23.16$) did not significantly differ from chance, $t(5) = 1.72$, $p = .15$. Alternatively, since cursor choices were corrected after incorrect responses, monkeys may have been choosing the same cursor that was used to complete the previous trial. To test this, a one-sample t-test was conducted to compare the percentage of trials for which the chosen cursor matched the correct cursor on trial *n*-1 to chance levels of responding. Once again, this analysis was conducted with the subset of animals that did not pass the Correction phase ($n = 6$). The percentage of time the monkeys chose the previous trial's correct cursor ($M = 32.48$, $SD = 2.32$) did not significantly differ from chance, $t(5) = -0.90$, $p = .41$. The monkeys were not relying on whichever cursor they had most recently used (the correct cursor for the previous trial's task) to direct their cursor choices. However, when choice behavior was examined for the five monkeys that completed both the Tool Testing and Correction phases, their choice behavior during the Tool Testing phase ($M = 53.61$, $SD = 25.27$) did not show the original bias that was observed at the group level during that phase, $t(4) = 1.80$, $p = .15$. Therefore, the absence of a trial $n - 1$ cursor bias during the present phase is because these specific monkeys never showed the original bias, as opposed to an effect of correction experience.

On the group level, the frequency of participants' cursor choices across tasks (Figure 6-6) was also calculated to identify if their choices were changing over the 900 trials, specifically for biases toward or away from certain cursors. Across all ten blocks, the monkeys chose the fast cursor approximately 33% of the time. There was a slight difference in the frequency of cursor choices between the big and transparent cursors during blocks one through five, with the big cursor often being chosen above chance levels and the transparent cursor below chance levels. However, across blocks six through ten all of the cursors were being used equally. So, while there was an initial bias to use the big and fast cursors over the transparent cursor, by the second half of the Correction phase there was no overall bias to choose the same cursor across blocks. Instead, monkeys were using all of the cursors but not in a manner that aligned with the correct task.

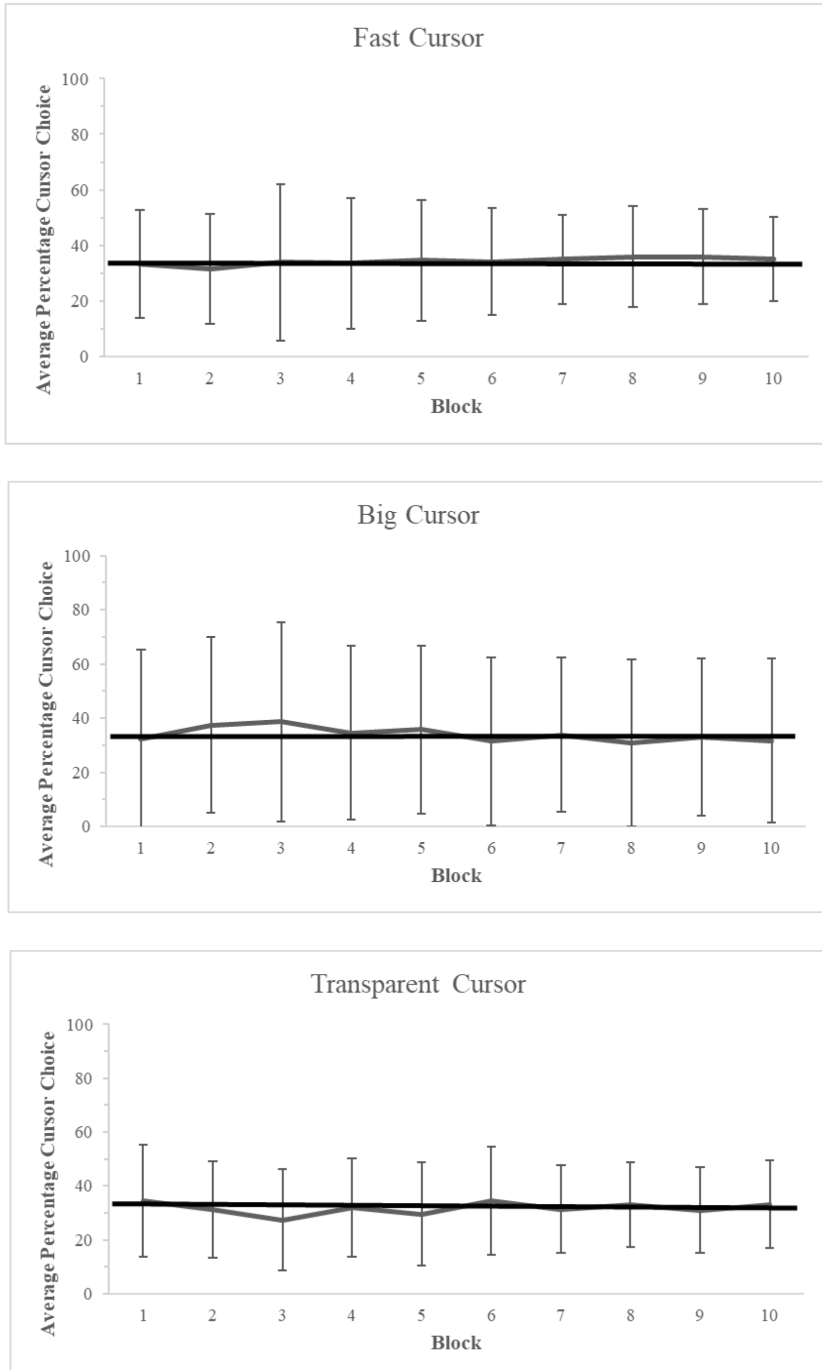

Figure 6-6 Frequency of Cursor Choice During the Correction Phase as a Function of Block

Note. Errors bars represent ± 1 standard deviation. The black horizontal line represents the 33% chance level. Data were collapsed across task set and task type.

6.4.4 Generalization Testing

6.4.4.1 Overall Performance

Only one monkey (Paddy, assigned to task set 1) made it to the Generalization Testing phase (after passing the Correction phase in 450 out of 900 trials). Binomial tests indicated that she performed significantly above chance levels of responding (33%) on the BARRIER (58.4%) and SIDE (52.3%) tasks, both $p < .001$. However, she scored significantly below chance on the CATCH task (13.1%), $p < .001$. Consistent with the previous phases and experiments, the CATCH task was the hardest. However, she was able to identify the optimal cursor for two of the transfer tasks.

6.4.4.2 Generalization

To determine whether Paddy's performance on BARRIER and SIDE was due to generalization, her choices during only the first three trials of each task were assessed. Performance on each task during these early trials was 33.3%, which is not significantly different from chance, $p = .44$. This suggests that she learned through experience (associatively) which cursor was optimal for these two tasks during the Generalization phase. As opposed to generalizing her knowledge of the cursors' functionalities, associative learning throughout testing likely played a large role in her choice behavior.

6.4.4.3 Cursor Bias

A qualitative analysis of Paddy's cursor choices revealed that she had a bias for the transparent and fast cursors. Table 20 shows her cursor selections during Generalization Testing across tasks. During Generalization Testing, Paddy chose the fast cursor (44.8%) and transparent cursor (47.2%) at similar rates. She used the big cursor for only 8.0% of the trials. In comparison, Paddy showed a bias for the big cursor during Tool Testing (Table 21), using it on

60.9% of trials compared to the fast cursor (16.7%) and transparent cursor (22.4%). In other words, during Generalization Testing, Paddy did not return to her original bias once differential feedback from the Correction phase was removed. Instead, experience during the Correction phase resulted in underutilization of the big cursor during Generalization Testing, resulting in her performance being significantly above chance (33%) for BARRIER and SIDE, but not CATCH.

Table 20 Cursor Choice Behavior During Generalization Testing

	Big Cursor		Fast Cursor		Transparent Cursor		
Task	*n*	%	*n*	%	*n*	%	Total
BARRIER	1	0.3	123	41.3	**174***	**58.4**	298
CATCH	**39**$^{+}$	**13.1**	121	40.6	138	46.3	298
SIDE	32	10.5	**159***	**52.3**	113	37.2	304
	72	8.0	403	44.8	425	47.2	900

Note. The optimal cursor for each task is bolded. Chance = 33%. * denotes significantly above chance ($p < .001$) $^{+}$ denotes significantly below chance ($p < .001$)

Table 21 Cursor Choice Behavior During Tool Testing for Capuchin Monkey Paddy

	Big Cursor		Fast Cursor		Transparent Cursor		
Task	*n*	%	*n*	%	*n*	%	Total
FENCE	51	20.4	81	32.4	**118***	**47.2**	250
ERASE	**275***	**84.9**	24	7.4	25	7.7	324
CHASE	222	68.1	**45**$^{+}$	**13.8**	59	18.1	326
	548	60.9	150	16.7	202	22.4	900

Note. The optimal cursor for each task is bolded. Chance = 33%. * denotes significantly above chance ($p < .001$) $^{+}$ denotes significantly below chance ($p < .001$)

7 GENERAL DISCUSSION

The present study investigated whether adult humans, children, and monkeys exhibited cognitive control in the specific context of a computerized conditional discrimination task or relied more heavily on associative learning though trial by trial feedback to succeed. It expanded upon existing research and theory about the interplay of cognitive and associative factors in determining behavior by utilizing a novel task across species and age ranges.

Associative cues from differences in cursor efficiency were predicted to lead to learning of the optimal cursor for three primary tasks for adults, children, and monkeys. Approximately 62% of the adults failed to learn which stimuli were optimal across all three tasks during the Tool Testing phase. Nearly all the children (93%) and all of the monkeys also required correction experience. In the case of the children and monkeys, the analyses determined that they often were choosing the cursor that was chosen on the previous trial instead of the optimal cursor for the present task. Once correction procedures were incorporated, however, optimal behavior was observed to various extents across groups. While adult participants were largely successful (only 7% of participants did not pass the Correction phase), almost half of the children failed to learn the task-cursor contingencies even after correction experience and nearly all of the monkeys failed to learn the optimal cursor for each task even in the presence of 20-sec time-outs.

This lack of learning may have been because the simple difference in speed of reinforcement used in the present study was not a salient enough associative cue to support learning. That is, the differences in optimality, defined as shorter response times when using the optimal cursor compared to when not using it, may not have been a strong enough consequence to facilitate discrimination. This may be especially true considering that there was not a statistically significant difference in response times between the correct and incorrect cursors for the speed tasks (CHASE and SIDE) completed by the children. Additionally, the continuous reinforcement schedule used during Tool Testing may have established habitual behavior that was resistant to change.

Alternatively, children and monkeys may have had difficulty learning the present task because it was more difficult than a typical conditional discrimination problem. Andrews et al. (2012) reported that requiring too many relational representations or stimulus configurations within a problem may result in lower performance. Conditional discrimination tasks traditionally only involve three components (two stimuli that must be discriminated from each other and the conditional discrimination cue). The present study, however, utilized six components (three stimuli and three distinguishing stimuli functions) that all had to be learned simultaneously. Given that Andrews et al. found fluid intelligence to be positively correlated with relational processing in preschool children, the pattern of findings in the present study suggest that executive functions (such as working memory, inhibition, etc.) may have played an important role in participants success (or lack thereof). The fact that adults, who have more developed working memory compared to children, for example, had much more success after receiving correction experience may also support this. Anecdotally, the children often verbally expressed frustration when they struggled to complete the tasks. They also seemed to understand that certain cursors had something special about them and, in some cases, they even understood that they should be used only for certain tasks. For example, children would say things such as “the purple one chases” and “when it’s the erases one I need the big one” while working. However, they would sometimes express this knowledge about the cursors in relation to the wrong tasks or the wrong cursor feature (e.g., “this one is slow, this one is fast, and this one moves through walls” even though the identifying features were size, speed, and transparency). Together, this may be evidence that the children were trying to learn the response contingences, but the relational complexity of the conditional discriminations were more than they could keep track of (Golin & Liss, 1962; Zelazo et al., 2003). Instead, the children that did not pass the Correction phase continued to choose the cursor that was freely chosen on the previous trial, which was an ultimately unsuccessful strategy.

Another important distinction between the novel testing paradigm used in this study compared to a standard conditional discrimination task is the use of multiple stimuli that led to reinforcement (i.e., there was no singular S+). Conditional discrimination tasks typically incorporate positive or negative feedback after the correct or incorrect discrimination was made, respectively. In comparison, the present study distinguished the cursors from each other based on optimality. So, even though there was an optimal cursor, any of the cursors could have been used to successfully complete each task. As a result, there were multiple "correct" options to complete each trial. This point is especially important to consider given that some participants across all three experiments may have been making satisficing cursor choices by primarily using the single cursor that was the fastest across all tasks, instead of making conditional discriminations.

This possible explanation raises an interesting question about how adults, children, and monkeys handle potential "redundancy" among options. For example, chimpanzees have a large repertoire of different sticks that they use as tools (puncturing and brush-tipped sticks for termite fishing, heavy and thick sticks to open beehives, and thin, long sticks for honey dipping, Boesch, 2013; Byrne et al., 2013). In comparison, bearded capuchin monkeys are skilled at using stone hammers and anvils to crack open nuts and other encased foods (e.g., Visalberghi & Fragaszy, 2012, 2013). However, de A. Moura and Lee (2004) and Mannu and Ottoni (2009) reported observations of bearded capuchin monkeys using stones to also dig for tubers, roots, or insects, open hollow branches, and break apart tubers and cacti. The use of stones for a wide array of functions, as opposed to different specialized tools for each task, may highlight an important difference between these species.

Successful generalization of optimal choice behavior to the transfer tasks, presumably driven by cognitive control, was predicted for adults, children, and monkeys. While the adults showed generalization of their knowledge of the cursors' functionality (as evidenced by optimal performance significantly above chance levels within the first three trials of each task), the children and the one capuchin monkey that progressed to the Generalization Testing phase did not. Alternatively, the children and monkey seemingly re-learned through associative experience with the transfer tasks which cursors were optimal for some of the different tasks. While the adult humans' performance was suggestive of cognitive control and a conceptual understanding of cursor functionality.

The lack of evidence for cognitive control in monkeys and children was the most surprising. Previous research with nonhuman primates, for example, suggested that they should have had much more success with the present task than they demonstrated. Monkeys have shown capabilities for flexible behavior on numerous occasions, such as self-control (e.g., Beran, Perdue, Rossettie, et al., 2016; Beran, Rossettie, et al. 2016), executive attention (e.g., French et al., 2018; Hassett & Hampton, 2022), and navigation of mazes (e.g., Beran et al., 2015; Washburn & Astur, 2003). While further research would be needed to parse this explanation out further, children and monkeys may have struggled more than adults to generalize the cursors' functionalities because the problem space (i.e., identifying functionality) was abstract, as opposed to stimulus based. Conceptual behavior, as defined by Thomas and Kerr (1976), are "selective responses to stimuli which are consistently correct in terms of predetermined and discoverable reinforcement contingencies but which do not depend upon prior experience with the specific stimuli presented on a given trial" (p. 335). Additionally, the response rate of generalization to a new stimulus changes as a factor of physical similarity to the originally reinforced stimulus. Conditional discrimination tasks often use a specific stimulus feature as the contextual cue (e.g., background color, shape, etc.). However, the present task used the general task environment as the cue. Tasks were designed to have limited physical similarity (e.g., color, orientation, etc.) and instead had similar task demands (e.g., need to move fast). Previous studies that have used conceptual information as conditional cues (e.g., Burdyn and Thomas, 1984; Flemming, 2011; Schrier et al., 1984; Thomas & Kerr, 1976) used exemplar stimuli as the discriminative stimulus, which is still a visual cue. However, forming a cognitive representation of "slowness" to identify the relevant aspects of each task would presumably be more difficult. In a review of categorical perception and conceptual judgments, Thompson and Oden (2000) reported that monkeys had a harder time forming complex abstract concepts (such as analogical conceptual judgements of relations-between-relations) than apes and humans.

Additionally, rhesus monkeys and capuchin monkeys are primarily local processers (De Lillo et al., 2005; Hopkins & Washburn, 2002; Spinozzi et al., 2006; Spinozzi et al., 2003), unlike humans who show more frequent global-to-local precedence when assessing compound stimuli (e.g., Navon, 1977, 1981; Thompson & Oden, 2000). This difference may have limited the monkeys' representation of the broader problem space in a way necessary for generalization in this task. Rezvani et al. (2020) has also shown that global precedence changes as a factor of variables in the perceptual environment (e.g., stimulus size, shape type, sparsity, solidness, etc.). While these are perceptual factors, as opposed to abstract concepts, they would still affect the cognitive representation participants are able to form. Therefore, while much more research would be needed to investigate this further, it is possible that the conceptual or perceptual nature of the task, in the absence of exemplars, was potentially not sufficient information for the children and monkeys to form cognitive sets that directed behavior in a conditional manner.

It was also predicted that there would be individual differences in the extent of success between the individuals that advanced to the Generalization Testing phase. Adult participants performed above chance levels for an average of two tasks. Closer inspection of these results revealed that an equal number of participants (approximately 45%) performed above chance levels for two tasks, as did those who performed above chance for three tasks. In comparison, majority of the children (42%) performed above chance on only one task. While an equal number of children (17%) performed above chance for two and three tasks. Only one monkey advanced to the Generalization Testing phase. Therefore, as predicted, adults showed the highest level of success, followed by children, and then monkeys. Individual differences in working memory capacity potentially explain these individual differences (Andrews et al., 2012; Shipstead et al., 2016), since success in the present study relies on keeping track of and holding in memory multiple relational or stimulus configurations. However, other possible explanations outside of executive functions (maturation, psychomotor skills, tendency to generalize information to novel contexts, etc.) must also be considered.

Finally, instructions about the cursors' functionality were predicated to facilitate learning and generalization in adults. Receiving explicit instructions did improve adults' overall performance on the transparency tasks during Tool Testing. However, instructions did not predict whether the participants required correction experience after completing the Tool Testing phase nor their overall performance during the Generalization Testing phase. These limited effects of instruction are interesting. Given that the transparent cursor is the only cursor that you must use while interacting with another object (i.e., move through a barrier) in order to identify its salient feature, it makes sense that performance was facilitated by instructions for the transparency tasks only. However, since the other cursors have more obvious features (size and speed) that likely did not need instructions in order to notice, a larger overall effect was not observed past initial use of the cursors during the Tool Testing phase.

Overall, within the specific paradigm of the present study, humans demonstrated learning and generalization to a limited extent, while monkeys did not. Some adult humans identified the optimal cursor across five (of six) tasks during the Generalization Phase and did so immediately, suggesting generalization facilitated by cognitive control. Children and monkeys struggled to a much greater extent, identifying the optimal cursor for only one (of six) and two (of three) tasks, respectively, in a manner that suggested associative learning mechanisms drove their performance. While it is possible that the present paradigm is beyond the capabilities of the participants in the latter two groups, the methodological limitations of the present study must be considered before claims of such qualitative differences can be made with any level of certainty.

7.1 Limitations

Given the complex cognitive abilities that have been demonstrated by children and nonhuman primates, as well as previous research that has shown their ability to solve conditional discrimination problems, the results of the present study were unexpected. Several methodological constraints of the present study have been identified that may have influenced results. As previously discussed, differences in the speed of reinforcement as designed in the present study may not have been a salient enough consequence to facilitate learning of the cursors' differences in optimality. Additionally, since the same physical exertion was required to choose and use each cursor (i.e., joystick manipulation), it is possible that the present study did not evoke taxing enough consequences for a psychomotor task. Modifying the current conditional discrimination task into a manual task would potentially address this. By incorporating a tactile element to the task, there would be additional sensory feedback available when discriminating stimuli and navigating the problem space. Physically moving through space as opposed to using computerized cursors would likely make the differences in effort and response time more salient to participants' choices, since there would now be a bigger cost to not performing optimally. A manual task would also likely match the natural behaviors and ecology of the monkeys consistent with their innate systems of behavior, as suggested by Timberlake (1993).

Additionally, the continuous reinforcement schedule used during Tool Testing may have inadvertently established habitual behavior. This is likely especially true for the monkeys, who experienced high trial counts of reinforced behavior during the Tool Testing phase (900 trials) before moving to the Correction phase. That is, continuous reinforcement may have established a bias in responding that overshadowed the correction feedback for majority of the monkeys. The main effect of block in monkeys' overall performance during the Correction phase suggested that this may be the case, since performance only started to improve during the final two blocks of correction experience (i.e., after 720 trials for the monkeys that ultimately did not pass correction). It would, therefore, be imperative to see if choice behavior in the animals would have continued to improve with extended correction experience. However, the extensive correction experience needed by the monkeys further suggests the key role that associative experience played in these animals' performance in the present task. Alternatively, reinforcement during the Tool Testing phase could be adjusted by decreasing the number of trials required and/or using different reinforcement contingencies. The latter could be established by incorporating negative feedback into this initial phase of testing (like in a typical conditional discrimination task) instead of only during the Correction phase or by using different reinforcers. For example, differential outcomes have been shown to have a positive effect on conditional discrimination learning (e.g., Maki et al., 1995; Martínez et al., 2009). Pairing unique rewards with each individual cursor may bring attention to the fact that the cursors also differ in other ways (i.e., function).

It was also clear that there were particular tasks that were not as easy to complete as others. Across experiments, participants struggled to make optimal choices when completing the CATCH task. Monkeys working during the Tool Testing phase, for example, had an overall performance of only 8.2% correct. Even adults during Generalization Testing performed at chance levels for this task. Given the same pattern of behavior across age and species, it is clear that the CATCH task was likely not the best assessment of behavior, and that a different task could have been used. The cursors also differed in that only the transparent cursor had to be experienced coming into contact with another object in order to see that it could move through barriers. The salient features of the other cursors (size and speed) can be discerned by simply looking at or moving it around the screen, respectively. It should also be noted that the particular monkeys that participated in this study have extensive experience with computerized cognitive testing. These studies are usually designed with the cursor on the screen being a small red circle (see the cursor depicted in Figure 4-1). Therefore, the square cursor that was optimal for the size tasks was more novel for these animals than the circular cursors for the speed and transparency tasks.

Additional information from a post-task assessment, such as participants' individual rating of the cursors' value, as well as their rank-ordered cursor preferences, would have provided insight into how the tasks and cursors were perceived in comparison to how they were designed. Incorporating post-task measures of the children's memory and task awareness would have also been very informative. For instance, if children were unable to remember what each cursors' important feature was, despite being told so at the beginning of the test session, then it is unlikely that they would have been able to make controlled decisions based on this information. Asking the adult and children participants if they were using any specific strategies to select cursors (e.g., "what do you choose when…?"), as well as if they were aware of the outcome contingencies (e.g., "did it take you longer when you used the…") would also be important information in discerning if the present results were due to true differences in performance or a larger design problem of the present study.

Finally, the sample size was small for both the children and monkeys. Since so few children required correction experience after Tool Testing, there were certain analyses that could not be calculated. Potential age effects in the children could also not be explored. Considering that research with both conditional discrimination (e.g., Andrews et al., 2012; Rudy et al., 1993) and cognitive control (Best et al., 2009; Best & Miller, 2010) tasks have shown age effects, this was an important element that was missing from the present study. Along with traditional assessments of age differences, it would have also been interesting, for instance, to see if there would have been any differences in performance between children in preschool (3 to 4 years old) and pre-kindergarten (4 to 5 years old) given the differences in classroom structure and instruction. Additionally, due to the low number of monkeys that advanced to (and beyond) Tool Testing, rhesus macaques and capuchin monkeys were assessed together for all data analyses. This did not allow for any important species comparisons to be assessed. Given that capuchin monkeys are known to be prolific tool-users (e.g., Visalberghi & Fragaszy, 2012, 2013) in ways not observed in rhesus macaques, it is logical to assume that capuchin monkeys may have shown better proficiency using specialized cursors with different tasks.

7.2 Future Directions

There are several interesting lines of research that would be ideal follow-up studies to the present one. For example, an eye-tracking study with the present paradigm would show how the problem space is being explored during decision-making. For example, whether participants look more at the presented task or the various cursor options more before making their cursor selection could be investigated.

If the nontraditional and more complex design of the present study's conditional discrimination problem was the reason children and monkeys struggled to learn, segmentation of the problem space, or having participants first complete reversal problems, may help improve performance in the conditional discrimination task. Previous research (e.g., Andrews et al., 2012) has shown that breaking down complex tasks into less components reduces complexity and processing load for participants. For example, they could be trained first on the three individual discriminations before exposure to the conditional discrimination task. Learning binary discriminations in succession, instead of simultaneous contingencies, may help scaffold learning and improve performance in the final task. Reversal learning would expose participants to changing contingencies, which may alert them that the correct answer can change across trials.

Another approach that could be taken to investigate task complexity would be to reverse the problem space. That is, instead of participants choosing which cursor they would use to complete a task, they would select which of the three tasks to complete based on a cursor that had been chosen for them. Differences in performance between these two task designs may provide interesting insights into the role that inhibition played in the present design. Presently, the tasks were fully shown on the computer screen but inaccessible until a cursor choice was made. By utilizing task symbols to make task choices and/or having the specific cursor chosen for you, participants may be able to inhibit prepotent responses that they were unable to with the reverse design.

Alternatively, incorporating multiple cues in addition to cursor functionality may reduce the abstract nature of the present discrimination. The use of multiple cues during conditional discrimination tasks has been shown to have an additive effect on problem-solving in monkeys (Warren, 1964). This additional layer of information may help participants form a more concrete representation of the differences between cursors and improve their conditional discriminations.

Given the individual differences observed in performance by the adults and previous correlations between relational processing and fluid intelligence (Andrews et al., 2012), an assessment of executive functioning (e.g., attention control, response inhibition, working memory capacity, etc.) for the participants who were more likely engaged in cognitive control would provide potential correlates for performance on this conditional discrimination task. An established executive functioning battery, or a combination of tasks such as the Flanker (Eriksen & Eriksen, 1974), Go/No-Go (e.g., Koek et al., 2015), and operation span tasks (e.g., Conway et al., 2005) could be used. Importantly, this would also elucidate the relation of this novel task to already established tasks to assess the validity of the present paradigm as an assessment of cognitive control. Similarly, incorporating a cognitive load and/or memory demand to assess potential capacity limitations would also be interesting. The presence or absence of performance interference would be indicative of cognitive or stimulus control, respectively.

Testing older children and great apes (e.g., chimpanzees) would also be a useful comparison to the data presented here. Pérez-González and Serna (2003) found that 10 to 17-

year-old children were able to generalize conceptual cues during a conditional discrimination task. Additionally, increased encephalization among nonhuman primates affected learning behavior during transfer index tasks (Rumbaugh & Pate, 1984) and executive functioning improves significantly with age and development of the frontal lobe (e.g., Bunge et al., 2002; Brydges et al., 2013; Dempster, 1992). Toates (1998) also proposed that the locus of control over the various processes driving behavior changes with age. Children of an older age and chimpanzees, for example, would allow for a closer comparison to the more developed prefrontal cortex of adults, who were the most successful in the present study.

7.3 Conclusion

In summary, the conditional discrimination task used in the present study assessed the extent to which cognitive control and associative processing seemingly influenced behavior. Adult humans, but not children, seemed to demonstrate cognitive control over behavior in the form of conceptual generalization. Similarly to the children, monkeys also relied on associative processing. Cognitive control and stimulus control are clearly both capable control mechanisms over behavior that can lead to success in the present task. Defining flexible, goal-driven, and adaptive responding as *only* cognitive seems to be a misnomer. Stimulus generalization and discrimination are efficient methods of adaptive behavior in many situations. Cognitive control should not be considered “superior” to stimulus control, and reliance on stimulus control should not be considered a handicap to performance. Instead, both of these control mechanisms are modes of processing that can lead to appropriate responses. Studying the interaction of these processes, including a continuum along which the locus of control lies, should be prioritized in research. Participants that may have been making satisficing cursor choices is a good example of this. This decision-making heuristic required discrimination of the fastest cursor and then generalization of this behavior across all tasks. As a result, flexible and adaptive behavior grounded in stimulus control produced optimal cursor choices when optimality is defined in another way.

The novel task design used in the present study assessed flexible and goal-directed behavior by incorporating task-switching and transfer tasks with equivalent task demands, creating a unique need for generalization. While the present study was not designed to parse out the different executive functions of cognitive control, conceptual conditional discrimination tasks such as this one are a useful method to study problem-solving.

REFERENCES

Abrahamse, E., Braem, S., Notebaert, W., & Verguts, T. (2016). Grounding cognitive control in associative learning. *Psychological Bulletin*, *142*(7), 693-728. http://dx.doi.org/10.1037/bul0000047

Addessi, E., Paglieri, F., & Focaroli, V. (2011). The ecological rationality of delay tolerance: Insights from capuchin monkeys. *Cognition*, *119*(1), 142-147. https://doi.org/10.1016/j.cognition.2010.10.021

Anderson, M. C., & Levy, B. J. (2009). Suppressing unwanted memories. *Current Directions in Psychological Science*, *18*(4), 189-194. https://doi.org/10.1111/j.1467-8721.2009.01634.x

Andrews, G., & Halford, G. S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology*, *45*(2), 153-219. https://doi.org/10.1016/S0010-0285(02)00002-6

Andrews, G., Halford, G. S., & Boyce, J. (2012). Conditional discrimination in young children: The roles of associative and relational processing. *Journal of Experimental Child Psychology*, *112*(1), 84-101. https://doi.org/10.1016/j.jecp.2011.12.004

Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, *126*(6), 925–945. https://doi.org/10.1037/0033-2909.126.6.925

Barge Jr, E. M., & Thomas, R. K. (1969). Conditional discrimination in the squirrel monkey. *Psychonomic Science*, *15*(3), 161-162. https://doi.org/10.3758/BF03336258

Basile, B. M., & Hampton, R. R. (2013a). Dissociation of active working memory and passive recognition in rhesus monkeys. *Cognition*, *126*(3), 391-396. http://dx.doi.org/10.1016/j.cognition.2012.10.012

Basile, B. M., & Hampton, R. R. (2013b). Recognition errors suggest fast familiarity and slow recollection in rhesus monkeys. *Learning & Memory*, *20*(8), 431-437. https://doi.org/10.1101/lm.029223.112

Basile, B. M., & Hampton, R. R. (2017). Dissociation of item and source memory in rhesus monkeys. *Cognition*, *166*, 398-406. http://dx.doi.org/10.1101/lm.029223.112

Basile, B. M., Hampton, R. R., Suomi, S. J., & Murray, E. A. (2009). An assessment of memory awareness in tufted capuchin monkeys (*Cebus apella*). *Animal Cognition*, *12*(1), 169-180. http://dx.doi.org/10.1007/s10071-008-0180-1

Beran, M. J. (2015). Chimpanzee cognitive control. *Current Directions in Psychological Science*, *24*(5), 352-357. http://dx.doi.org/10.1177/0963721415593897

Beran, M. J., & Evans, T. A. (2006). Maintenance of delay of gratification by four chimpanzees (*Pan troglodytes*): The effects of delayed reward visibility, experimenter presence, and extended delay intervals. *Behavioural Processes*, *73*(3), 315-324. http://dx.doi.org/10.1016/j.beproc.2006.07.005

Beran, M. J., Menzel, C. R., Parrish, A. E., Perdue, B. M., Sayers, K., Smith, J. D., & Washburn, D. A. (2016). Primate cognition: Attention, episodic memory, prospective memory, self-control, and metacognition as examples of cognitive control in nonhuman primates. *Wiley Interdisciplinary Reviews: Cognitive Science*, *7*(5), 294-316. http://dx.doi.org/10.1002/wcs.1397

Beran, M. J., Parrish, A. E., Futch, S. E., Evans, T. A., & Perdue, B. M. (2015). Looking ahead? Computerized maze task performance by chimpanzees (*Pan troglodytes*), rhesus monkeys (*Macaca mulatta*), capuchin monkeys (*Cebus apella*), and human children

(*Homo sapiens*). *Journal of Comparative Psychology*, *129*(2), 160-173. https://doi.org/10.1037/a0038936

Beran, M. J., Perdue, B. M., Bramlett, J. L., Menzel, C. R., & Evans, T. A. (2012). Prospective memory in a language-trained chimpanzee (*Pan troglodytes*). *Learning and Motivation*, *43*(4), 192-199. http://dx.doi.org/10.1016/j.lmot.2012.05.002

Beran, M. J., Perdue, B. M., Church, B. A., & Smith, J. D. (2016). Capuchin monkeys (*Cebus apella*) modulate their use of an uncertainty response depending on risk. *Journal of Experimental Psychology: Animal Learning and Cognition*, *42*(1), 32–43. https://doi.org/10.1037/xan0000080

Beran, M. J., Perdue, B. M., Rossettie, M. S., James, B. T., Whitham, W., Walker, B., Futch, S. E., & Parrish, A. E. (2016). Self-control assessments of capuchin monkeys with the rotating tray task and the accumulation task. *Behavioural Processes*, *129*, 68-79. https://doi.org/10.1016/j.beproc.2016.06.007

Beran, M. J., Rossettie, M. S., & Parrish, A. E. (2016). Trading up: Chimpanzees (*Pan troglodytes*) show self-control through their exchange behavior. *Animal cognition*, *19*(1), 109-121. http://dx.doi.org/10.1016/j.beproc.2016.06.007

Beran, M. J., & Smith, J. D. (2011). Information seeking by rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*). *Cognition*, *120*(1), 90-105. http://dx.doi.org/10.1016/j.cognition.2011.02.016

Beran, M. J., Smith, J. D., & Perdue, B. M. (2013). Language-trained chimpanzees (*Pan troglodytes*) name what they have seen but look first at what they have not seen. *Psychological Science*, *24*(5), 660-666. https://doi.org/10.1177/0956797612458936

Beran, M. J., Washburn, D. A., & Rumbaugh, D. M. (2007). A Stroop-like effect in color-naming of color-word lexigrams by a chimpanzee (Pan troglodyte). *The Journal of General Psychology*, *134*(2), 217-228. http://dx.doi.org/10.1177/0956797612458936

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641-1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review*, *29*(3), 180-200. https://doi.org/10.1016/j.dr.2009.05.002

Bodrova, E., & Leong, D. J. (2008). Developing self-regulation in kindergarten. *Young Children*, *63*(2), 56-58.

Boelens, H., Smeets, P., Hoogeveen, F., & Lancioni, G. (1989). Effect of conditional discrimination learning on subsequent multiple-cue discrimination in children. *Behavioural processes*, *19*(1-3), 143-153. https://doi.org/10.1016/0376-6357(89)90037-5

Boesch, C. (2013). Ecology and cognition of tool use in chimpanzees. In C. M. Sanz, J. Call, & C. Boesch (Eds.), *Tool use in animals: Cognition and ecology* (pp. 21-47). Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511894800.003

Brady, R. J., & Hampton, R. R. (2018a). Nonverbal working memory for novel images in rhesus monkeys. *Current Biology*, *28*(24), 3903-3910. http://dx.doi.org/10.1016/j.cub.2018.10.025

Brady, R. J., & Hampton, R. R. (2018b). Post-encoding control of working memory enhances processing of relevant information in rhesus monkeys (*Macaca mulatta*). C*ognition*, *175*, 26-35. http://dx.doi.org/10.1016/j.cognition.2018.02.012

Brady, R. J., & Hampton, R. R. (2021). Rhesus monkeys (*Macaca mulatta*) monitor evolving decisions to control adaptive information seeking. *Animal Cognition*, *24*(4), 777-785. http://dx.doi.org/10.1007/s10071-021-01477-5

Bramlett, J. L., Perdue, B. M., Evans, T. A., & Beran, M. J. (2012). Capuchin monkeys (*Cebus apella*) let lesser rewards pass them by to get better rewards. *Animal Cognition*, *15*(5), 963-969. http://dx.doi.org/10.1007/s10071-012-0522-x

Brown, E. K., Basile, B. M., Templer, V. L., & Hampton, R. R. (2019). Dissociation of memory signals for metamemory in rhesus monkeys (*Macaca mulatta*). *Animal cognition*, *22*(3), 331-341. http://dx.doi.org/10.1007/s10071-019-01246-5

Brown, E. K., & Hampton, R. R. (2020). Cognitive control of working memory but not familiarity in rhesus monkeys (*Macaca mulatta*). *Learning and Behavior*, *48*(4), 444-452. http://dx.doi.org/10.3758/s13420-020-00432-7

Brydges, C. R., Anderson, M., Reid, C. L., & Fox, A. M. (2013). Maturation of cognitive control: Delineating response inhibition and interference suppression. *PLOS ONE*, *8*(7), e69826. https://doi.org/10.1371/journal.pone.0069826

Bub, D. N., Masson, M. E., & Lalonde, C. E. (2006). Cognitive control in children: Stroop interference and suppression of word reading. *Psychological Science*, *17*(4), 351-357. https://doi.org/10.1111/j.1467-9280.2006.01710.x

Bugg, J. M., & Crump, M. J. C. (2012). In support of a distinction between voluntary and stimulus-driven control: A review of the literature on proportion congruent effects. *Frontiers in Psychology*, *3*, 367. http://dx.doi.org/10.3389/fpsyg.2012.00367

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron*, *33*(2), 301-311. https://doi.org/10.1016/S0896-6273(01)00583-9

Burdyn, L. E., & Thomas, R. K. (1984). Conditional discrimination with conceptual simultaneous and successive cues in the squirrel monkey (*Saimiri sciureus*). *Journal of Comparative Psychology*, *98*(4), 405–413. https://doi.org/10.1037/0735-7036.98.4.405

Byrne, R. W., Sanz, C. M., Morgan, D. B. (2013). Chimpanzees plan their tool use. In C. M. Sanz, J. Call, & C. Boesch (Eds.), *Tool use in animals: Cognition and ecology* (pp. 48-63). Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511894800.004

Cirino, P. T., Ahmed, Y., Miciak, J., Taylor, W. P., Gerst, E. H., & Barnes, M. A. (2018). A framework for executive function in the late elementary years. *Neuropsychology*, *32*(2), 176-189. https://doi.org/10.1037/neu0000427

Cohen, J. D. (2017). Cognitive control: Core constructs and current considerations. In T. Egner (Ed.), *The Wiley handbook of cognitive control* (pp. 3-28). John Wiley & Sons Ltd. https://doi.org/10.1002/9781118920497.ch1

Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, *123*(4), 354–373. https://doi.org/10.1037/0096-3445.123.4.354

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769-786. https://doi.org/10.3758/BF03196772

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547-552. http://dx.doi.org/10.1016/j.tics.2003.10.005

D'Amato, M. R., Salmon, D. P., & Colombo, M. (1985). Extent and limits of the matching concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, *11*(1), 35–51. https://doi.org/10.1037/0097-7403.11.1.35

D'Amato, M. R., Salmon, D. P., Loukas, E., & Tomie, A. (1985). Symmetry and transitivity of conditional relations in monkeys (*Cebus apella*) and pigeons (*Columba livia*). *Journal of the Experimental Analysis of Behavior*, *44*(1), 35-47. https://doi.org/10.1901/jeab.1985.44-35

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037-2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

De Lillo, C., Spinozzi, G., Truppa, V., & Naylor, D. M. (2005). A comparative analysis of global and local processing of hierarchical visual stimuli in young children (*Homo sapiens*) and monkeys (*Cebus apella*). *Journal of Comparative Psychology*, *119*(2), 155-165. https://doi.org/10.1037/0735-7036.119.2.155

de A. Moura, A. C., & Lee, P. C. (2004). Capuchin stone tool use in Caatinga dry forest. *Science*, *306*(5703), 1909-1909. https://doi.org/10.1126/science.1102558

De Lillo, C., & Visalberghi, E. (1994). Transfer index and mediational learning in tufted capuchins (*Cebus apella*). *International Journal of Primatology*, *15*, 275-287. https://doi.org/10.1007/BF02735277

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review*, *12*(1), 45-75. https://doi.org/10.1016/0273-2297(92)90003-K

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood:

Cognitive functions, anatomy, and biochemistry. In D. Stuss & R. Knight (Eds.), *Principles of frontal lobe function* (pp. 466-503). Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195134971.003.0029

Diamond, A. (2011). Biological and social influences on cognitive control processes dependent on prefrontal cortex. *Progress in Brain Research*, *189*, 319-339. https://doi.org/10.1016/B978-0-444-53884-0.00032-4

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Duncan, J., Parr, A., Woolgar, A., Thompson, R., Bright, P., Cox, S., Bishop, S., & Nimmo-Smith, I. (2008). Goal neglect and Spearman's g: Competing parts of a complex task. *Journal of Experimental Psychology: General*, *137*(1), 131-148. https://doi.org/10.1037/0096-3445.137.1.131

Duverne, S., & Koechlin, E. (2017). Hierarchical control of behaviour in human prefrontal cortex. In T. Egner (Ed.), *The Wiley handbook of cognitive control* (pp. 207-220). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118920497.ch12

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19-23. http://dx.doi.org/10.1111/1467-8721.00160

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory (*Vol. 44, pp. 145–199). Elsevier Science. http://dx.doi.org/10.1016/S0079-7421(03)44005-X

Engle, R. W., Kane, M. J., Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working* *memory: Mechanisms of active maintenance and executive control* (pp. 102-134). Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139174909.007

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309–331. https://doi.org/10.1037/0096-3445.128.3.309

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143-149. https://doi.org/10.3758/BF03203267

Evans, T. A., & Beran, M. J. (2012). Monkeys exhibit prospective memory in a computerized task. *Cognition*, *125*(2), 131-140. http://dx.doi.org/10.1016/j.cognition.2012.07.012

Evans, T. A., Beran, M. J., Paglieri, F., & Addessi, E. (2012). Delaying gratification for food and tokens in capuchin monkeys (*Cebus apella*) and chimpanzees (*Pan troglodytes*): when quantity is salient, symbolic stimuli do not improve performance. *Animal Cognition*, *15*(4), 539-548. http://dx.doi.org/10.1007/s10071-012-0482-1

Evans, T. A., Perdue, B., & Beran, M. J. (2014). The relationship between event-based prospective memory and ongoing task performance in chimpanzees (*Pan troglodytes*). *PLoS One*, *9*(11), e112015. http://dx.doi.org/10.1371/journal.pone.0112015

Evans, T. A., Perdue, B. M., Parrish, A. E., & Beran, M. J. (2014). Working and waiting for better rewards: Self-control in two monkey species (*Cebus apella* and *Macaca mulatta*). *Behavioural Processes*, *103*, 236-242. http://dx.doi.org/10.1016/j.beproc.2014.01.001

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition:

Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223-241. https://doi.org/10.1177/1745691612460685

Evans, T. A., & Westergaard, G. C. (2006). Self-control and tool use in tufted capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, *120*(2), 163–166. https://doi.org/10.1037/0735-7036.120.2.163

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906-911. https://doi.org/10.1037/0003-066X.34.10.906

Flemming, T. M. (2011). Conceptual thresholds for same and different in old (*Macaca mulatta*) and new-world (*Cebus apella*) monkeys. *Behavioural Processes*, *86*(3), 316-322. https://doi.org/10.1016/j.beproc.2011.01.001

Fobes J. L., & King J. E. (1982). Measuring primate learning abilities. In J. L. Fobes & J. E. King (Eds.), *Primate behavior* (pp. 289–326). Academic Press.

French, K., Beran, M. J., Espy, K. A., & Washburn, D. A. (2018). Simians in the Shape School: A comparative study of executive attention. *Learning & Behavior*, *46*(3), 281-293. http://dx.doi.org/10.3758/s13420-017-0310-1

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186-204. http://dx.doi.org/10.1016/j.cortex.2016.04.023

Gardner, R. W. (1961). Cognitive controls of attention development as determinants of visual illusions. *The Journal of Abnormal and Social Psychology*, *62*(1), 120-127. https://doi.org/10.1037/h0040838

Gardner, R. W., Lohrenz, L. J., & Schoen, R. A. (1968). Cognitive control of differentiation in the perception of persons and objects. *Perceptual and Motor Skills*, *26*(1), 311-330. http://dx.doi.org/10.2466/pms.1968.26.1.311

Gardner, R. W., & Long, R. I. (1960). Errors of the standard and illusion effects with L-shaped figures. *Perceptual and Motor Skills*, *10*(2), 107-109. http://dx.doi.org/10.2466/pms.1968.26.1.311

Gardner, R. W., & Long, R. I. (1961). Selective attention and the Mueller-Lyer illusion. *The Psychological Record, 11(4)*, 317–320. http://dx.doi.org/10.1007/BF03393417

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*(1), 31-60. https://doi.org/10.1037/0033-2909.134.1.31

Gollin, E. S. (1965). Factors affecting conditional discrimination in children. *Journal of Comparative and Physiological Psychology*, *60*(3), 422–427. https://doi.org/10.1037/h0022583

Gollin, E. S., & Liss, P. (1962). Conditional discrimination in children. *Journal of Comparative and Physiological Psychology*, *55*(5), 850–855. https://doi.org/10.1037/h0045951

Guttman, N. (1963). Laws of behavior and facts of perception. In S. Koch (Ed.), *Psychology: A study of a science* (pp 114-178). McGraw-Hill.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, *21*(6), 803–831. https://doi.org/10.1017/S0140525X98001769

Hammond, K. R., & Summers, D. A. (1972). Cognitive control. *Psychological Review*, *79*(1), 58–67. https://doi.org/10.1037/h0031851

Hampton, R. R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences*, *98*(9), 5359-5362. http://dx.doi.org/10.1073/pnas.071600998

Hampton, R. R., Engelberg, J. W., & Brady, R. J. (2020). Explicit memory and cognition in monkeys. *Neuropsychologia*, *138*, 107326. http://dx.doi.org/10.1016/j.neuropsychologia.2019.107326

Hampton, R. R., Hampstead, B. M., & Murray, E. A. (2005). Rhesus monkeys (*Macaca mulatta*) demonstrate robust memory for what and where, but not when, in an open-field test of memory. *Learning and Motivation*, *36*(2), 245-259. http://dx.doi.org/10.1016/j.lmot.2005.02.004

Hampton, R. R., Zivin, A., & Murray, E. A. (2004). Rhesus monkeys (*Macaca mulatta*) discriminate between knowing and not knowing and collect information as needed before acting. *Animal cognition*, *7*(4), 239-246. http://dx.doi.org/10.1007/s10071-004-0215-1

Hanson, H. M. (1959). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, *58*(5), 321–334. https://doi.org/10.1037/h0042606

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, *56*(1), 51–65. https://doi.org/10.1037/h0062474

Harlow, H. F. (1951). Learning theories. In W. Dennis, R. Leeper, H. F. Harlow, J. J. Gibson, D. Krech, D. M. Rioch, W. S. McCulloch, & H. Feigl (Eds.), *Current trends in psychological theory* (pp. 57–84). University of Pittsburgh Press. https://doi.org/10.1037/11318-003

Hassett, T. C., & Hampton, R. R. (2022). Control of attention in rhesus monkeys measured using a flanker task. *Attention, Perception, and Psychophysics*, *84*(7), 2155-2166. http://dx.doi.org/10.3758/s13414-022-02452-z

Hedden, T., & Park, D. C. (2003). Contributions of source and inhibitory mechanisms to age-related retroactive interference in verbal working memory. *Journal of Experimental Psychology: General*, *132*(1), 93–112. https://doi.org/10.1037/0096-3445.132.1.93

Herrmann, E., & Tomasello, M. (2015). Focusing and shifting attention in human children (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *129*(3), 268-274. https://doi.org/10.1037/a0039384

Hill, S. D. (1962). Chronological age levels at which children solve three problems varying in complexity. *Perceptual and Motor Skills*, *14*(2), 254-254. https://doi.org/10.2466/pms.1962.14.2.254

Hopkins, W. D., & Washburn, D. A. (2002). Matching visual stimuli on the basis of global and local features by chimpanzees (*Pan troglodytes*) and rhesus monkeys (*Macaca mulatta*). *Animal Cognition*, *5*, 27-31. https://doi.org/10.1007/s10071-001-0121-8

Hull, C. L. (1930). Knowledge and purpose as habit mechanisms. *Psychological Review*, *37*(6), 511–525. https://doi.org/10.1037/h0072212

Hull, C. L. (1939). The problem of stimulus equivalence in behavior theory. *Psychological Review*, *46*(1), 9–30. https://doi.org/10.1037/h0054032

Hull, C. L. (1943a). *Principles of behavior: An introduction to behavior theory.* Appleton-Century.

Hull, C. L. (1943). The problem of intervening variables in molar behavior theory. *Psychological Review*, *50*(3), 273–291. https://doi.org/10.1037/h0057518

Hull, C. L. (1947). The problem of primary stimulus generalization. *Psychological Review, 54*(3), 120–134. https://doi.org/10.1037/h0061159

Jordan, C. R., Pilgrim, C., & Galizio, M. (2001). Conditional discrimination and stimulus equivalence in young children following three different baseline training procedures. *Experimental Analysis of Human Behavior Bulletin*, *19*, 3-7.

Judge, P. G., & Essler, J. L. (2013). Capuchin monkeys exercise self-control by choosing token exchange over an immediate reward. *International Journal of Comparative Psychology*, *26*(4), 256. http://dx.doi.org/10.46867/ijcp.2013.26.04.05

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*(3), 213-233. http://dx.doi.org/10.1007/s11065-007-9040-z

Kane, M. J., Conway, A. R., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. *Variation in Working Memory*, *1*, 21-48. http://dx.doi.org/10.1093/acprof:oso/9780195168648.003.0002

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 336–358. https://doi.org/10.1037/0278-7393.26.2.336

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin and Review*, *9*(4), 637-671. https://doi.org/10.3758/BF03196323

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A.

(2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, *144*(11), 1147. http://dx.doi.org/10.1037/bul0000160

Karr, J. E., Rodriguez, J. E., Goh, P. K., Martel, M. M., & Rast, P. (2022). The unity and diversity of executive functions: A network approach to life span development. *Developmental Psychology*, *58*(4), 751-767. http://dx.doi.org/10.1037/dev0001313

Kaufman, A. B., Call, J., & Kaufman, J. C. (Eds.). (2021). *The Cambridge handbook of animal cognition*. Cambridge University Press. http://dx.doi.org/10.1017/9781108564113

Kinoshita, M., Ohta, H., & Matano, S. (1997). Age change in learning ability of Japanese monkeys (*Macaca fuscata*) measured by the transfer index. *Psychological Reports*, *80*(2), 467-473. https://doi.org/10.2466/pr0.1997.80.2.467

Koek, W., Gerak, L. R., & France, C. P. (2015). Effects of amphetamine, morphine, and CP 55,940 on go/no-go task performance in rhesus monkeys. *Behavioural Pharmacology*, *26*(5), 481-484. https://doi.org/10.1097/FBP.0000000000000146

Kornell, N., Son, L. K., & Terrace, H. S. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychological Science*, *18*(1), 64-71. http://dx.doi.org/10.1111/j.1467-9280.2007.01850.x

Lucas, G. A. (2019). Adaptive systems influence both learning and conscious attention. *Behavioural Processes*, *168*, 103871. https://doi.org/10.1016/j.beproc.2019.05.018

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59-80. https://doi.org/10.1348/026151003321164627

Maestripieri, D. (Ed.). (2003). *Primate psychology*. Harvard University Press.

http://dx.doi.org/10.4159/9780674040427

Maki, P., Overmier, J. B., Delos, S., & Gutmann, A. J. (1995). Expectancies as factors influencing conditional discrimination performance of children. *The Psychological Record*, *45*(1), 45-71.

Mannu, M., & Ottoni, E. B. (2009). The enhanced tool-kit of two groups of wild bearded capuchin monkeys in the Caatinga: Tool making, associative use, and secondary tools. *American Journal of Primatology*, *71*(3), 242-251. https://doi.org/10.1002/ajp.20642

Marsh, R. L., & Hicks, J. L. (1998). Event-based prospective memory and executive control of working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 336–349. https://doi.org/10.1037/0278-7393.24.2.336

Martinez, L., Estévez, A. F., Fuentes, L. J., & Overmier, J. B. (2009). Improving conditional discrimination learning and memory in five-year-old children: Differential outcomes effect using different types of reinforcement. *Quarterly Journal of Experimental Psychology*, *62*(8), 1617-1630. https://doi.org/10.1080/17470210802557827

Meador, D., Rumbaugh, D. M., Pate, J. L., & Bard, K. (1988). Learning, problem solving, cognition, and intelligence. In G. Mitchell, & J. Erwin (Eds.), *Comparative primate biology* (Vol. 2 Part B, pp. 17-84). Alan R. Liss.

Meier, M. E., & Kane, M. J. (2017). Attentional control and working memory capacity. In T. Egner (Ed.), *The Wiley handbook of cognitive control* (pp. 50-63). John Wiley & Sons Ltd. http://dx.doi.org/10.1002/9781118920497.ch3

Menzel, C. R. (1999). Unprompted recall and reporting of hidden objects by a chimpanzee (*Pan* *troglodytes*) after extended delays. *Journal of Comparative Psychology*, *113*(4), 426–434. https://doi.org/10.1037/0735-7036.113.4.426

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, *244*(4907), 933-938. http://dx.doi.org/10.1126/science.2658056

Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience*, *1*(1), 59-65. https://doi.org/10.1038/35036228

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167-202. https://doi.org/10.1146/annurev.neuro.24.1.167

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8-14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100. http://dx.doi.org/10.1006/cogp.1999.0734

Moore, J. (1999). The basic principles of behaviorism. In B. A. Thyer (Ed.), *The philosophical legacy of behaviorism* (pp. 41-68). Springer Netherlands.

Moore, J. (2017). John B. Watson's classical S–R behaviorism. *The Journal of Mind and Behavior*, *38*(1), 1-34. https://www.jstor.org/stable/44631526

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, *11*(1), 56-60. https://doi.org/10.1080/17470215908416289

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353-383. https://doi.org/10.1016/0010-0285(77)90012-3

Navon, D. (1981). The forest revisited: More on global precedence. *Psychological Research*, *43*(1), 1-32. https://doi.org/10.1007/BF00309635

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & Shimamura, A. P. (Eds.), *Metacognition: Knowing about knowing* (pp. 1-26). The MIT Press. http://dx.doi.org/10.1016/S1364-6613(03)00028-7

Nigg, J. T. (2017). Annual research review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *58*(4), 361-383. https://doi.org/10.1111/jcpp.12675

Nissen, H. W. (1951). Analysis of a complex conditional reaction in chimpanzee. *Journal of Comparative and Physiological Psychology*, *44*(1), 9–16. https://doi.org/10.1037/h0058252

Parrish, A. E., James, B. T., Rossettie, M. S., Smith, T., Otalora-Garcia, A., & Beran, M. J. (2018). Investigating the depletion effect: Self-control does not waiver in capuchin monkeys. *Animal Behavior and Cognition*, *5*(1), 118-138. https://doi.org/10.26451/abc.05.01.09.2018

Parrish, A. E., Perdue, B. M., Evans, T. A., & Beran, M. J. (2013). Chimpanzees (*Pan troglodytes*) transfer tokens repeatedly with a partner to accumulate rewards in a self-control task. *Animal cognition*, *16*(4), 627-636. http://dx.doi.org/10.1007/s10071-013-0599-x

Perdue, B. M., Bramlett, J. L., Evans, T. A., & Beran, M. J. (2015). Waiting for what comes later: Capuchin monkeys show self-control even for nonvisible delayed rewards. *Animal cognition*, *18*(5), 1105-1112. http://dx.doi.org/10.1007/s10071-015-0878-9

Perdue, B. M., Evans, T. A., Williamson, R. A., Gonsiorowski, A., & Beran, M. J. (2014). Prospective memory in children and chimpanzees. *Animal Cognition*, *17*(2), 287-295. http://dx.doi.org/10.1007/s10071-013-0661-8

Pérez-González, L. A. (1994). Transfer of relational stimulus control in conditional discriminations. *Journal of the Experimental Analysis of Behavior*, *61*(3), 487-503. https://doi.org/10.1901/jeab.1994.61-487

Pérez-González, L. A., & Serna, R. W. (2003). Transfer of specific contextual functions to novel conditional discriminations. *Journal of the Experimental Analysis of Behavior*, *79*(3), 395-408. https://doi.org/10.1901/jeab.2003.79-395

Posner, M. I., & DiGirolamo, G. J. (1998). Executive attention: Conflict, target detection, and cognitive control. In R. Parasuraman (Ed.), *The attentive brain* (pp. 401-423). MIT Press.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55-85). Erlbaum.

Postle, B. R., Brush, L. N., & Nick, A. M. (2004). Prefrontal cortex and the mediation of proactive interference in working memory. *Cognitive, Affective, and Behavioral Neuroscience*, *4*(4), 600-608. http://dx.doi.org/10.3758/CABN.4.4.600

Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, *55*(2), 233-244. http://dx.doi.org/10.1901/jeab.1991.55-233

Rezvani, Z., Katanforoush, A., & Pouretemad, H. (2020). Global precedence changes by environment: A systematic review and meta-analysis on effect of perceptual field variables on global-local visual processing. *Attention, Perception, and Psychophysics*, *82*(5), 2348-2359. https://doi.org/10.3758/s13414-020-01997-1

Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., Manes, F., & Duncan, J. (2010). Executive function and fluid intelligence after frontal lobe lesions. *Brain*, *133*(1), 234-247. https://doi.org/10.1093/brain/awp269

Roitblat, H. L. (1984). Representations in pigeon working memory. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 79–97). Erlbaum.

Rudy, J. W., Keith, J. R., & Georgen, K. (1993). The effect of age on children's learning of problems that require a configural association solution. *Developmental Psychobiology*, *26*(3), 171-184. https://doi.org/10.1002/dev.420260304

Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2004). Attentional control and self-regulation. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory, and applications* (2nd ed., pp. 284-299). The Guilford Press.

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences*, *102*(41), 14931-14936. http://dx.doi.org/10.1073/pnas.0506897102

Rumbaugh, D. M. (1969). The transfer index: An alternative measure of learning set. In C. R. Carpenter (Ed.), *Proceedings of the second international congress of primatology: Behavior* (Vol. 1, pp. 267-273). Karger-Basel.

Rumbaugh, D. M. (1970). Learning skills of anthropoids. In L. A. Rosenblum (Ed.), *Primate* *behavior: Developments in field and laboratory research* (Vol. 1, pp. l-70). Academic Press.

Rumbaugh, D. M., Bell, C. L., & Gill, T. V. (1972). Two discrimination test apparatuses for primates. *Behavior Research Methods & Instrumentation*, *4*, 6-10. https://doi.org/10.3758/BF03209962

Rumbaugh, D. M., & Gill, T. V. (1973). The learning skills of great apes. *Journal of Human Evolution*, *2*(3), 171-179. https://doi.org/10.1016/0047-2484(73)90073-0

Rumbaugh, D. M., & Pate, J. L. (1984). The evolution of cognition in primates: A comparative perspective. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 569-585). Lawrence Erlbaum.

Sayers, K., & Menzel, C. R. (2012). Memory and foraging theory: Chimpanzee utilization of optimality heuristics in the rank-order recovery of hidden foods. *Animal Behaviour*, *84*(4), 795-803. http://dx.doi.org/10.1016/j.anbehav.2012.06.034

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1–66. https://doi.org/10.1037/0033-295X.84.1.1

Schrier, A. M. (1970). Effect of location of the conditional cue on conditional discrimination learning by monkeys (*Macaca mulatta*). *Learning and Motivation*, *1*(2), 207-217. https://doi.org/10.1016/0023-9690(70)90087-1

Schrier, A. M., Angarella, R., & Povar, M. L. (1984). Studies of concept formation by stumptailed monkeys: Concepts humans, monkeys, and letter. *A Journal of Experimental Psychology: Animal Behavior Processes, 10*(4), 564–584. https://doi.org/10.1037/0097-7403.10.4.564

Schwartz, B. L. (2005). Do nonhuman primates have episodic memory? In H. S. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 225-241). Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195161564.003.0009

Schwartz, B. L., Colon, M. R., Sanchez, I. C., Rodriguez, I., & Evans, S. (2002). Single-trial learning of" what" and" who" information in a gorilla (*Gorilla gorilla gorilla*): Implications for episodic memory. *Animal Cognition*, *5*(2), 85-90. https://doi.org/10.1007/s10071-002-0132-0

Schwartz, B. L., & Evans, S. (2001). Episodic memory in primates. *American Journal of Primatology*, *55*(2), 71-85. http://dx.doi.org/10.1007/s10071-002-0132-0

Shepard, R. N. (1958). Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychological Review, 65*(4), 242–256. https://doi.org/10.1037/h0043083

Shettleworth, S. J. (2009). The evolution of comparative cognition: Is the snark still a boojum? *Behavioural Processes*, *80*(3), 210-217. http://dx.doi.org/10.1016/j.beproc.2008.09.001

Shields, W. E., Smith, J. D., Guttmannova, K., & Washburn, D. A. (2005). Confidence judgments by humans and rhesus monkeys. *The Journal of General Psychology*, *132*(2), 165.

Shields, W. E., Smith, J. D., & Washburn, D. A. (1997). Uncertain responses by humans and Rhesus monkeys (*Macaca mulatta*) in a psychophysical same–different task. *Journal of Experimental Psychology: General*, *126*(2), 147–164. https://doi.org/10.1037/0096-3445.126.2.147

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127–190. https://doi.org/10.1037/0033-295X.84.2.127

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2015). Working memory capacity and the scope and control of attention. *Attention, Perception, and Psychophysics*, *77*(6), 1863-1880. http://dx.doi.org/10.3758/s13414-015-0899-0

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2016). Working memory capacity and fluid intelligence: Maintenance and disengagement. *Perspectives on Psychological Science, 11*(6), 771-799. https://doi.org/10.1177/1745691616650647

Shipstead, Z., Lindsey, D. R., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language, 72*, 116-141. http://dx.doi.org/10.1016/j.jml.2014.01.004

Skinner, B. F. (1953). Some contributions of an experimental analysis of behavior to psychology as a whole. *American Psychologist, 8*(2), 69–78. https://doi.org/10.1037/h0054118

Skinner, B. F. (1963). Operant behavior. *American Psychologist, 18*(8), 503–515. https://doi.org/10.1037/h0045185

Skinner, B. F. (1974). *About behaviorism*. Knopf Doubleday Publishing Group.

Skinner, B. F. (1981). Selection by consequences. *Science, 213*(4507), 501–504. https://doi.org/10.1126/science.7244649

Skinner, B. F. (1984). An operant analysis of problem solving. *Behavioral and Brain Sciences*, *7*(4), 583–591. https://doi.org/10.1017/S0140525X00027412

Smith, L. D. (1990). Models, mechanisms, and explanation in behavior theory: The case of Hull versus Spence. *Behavior and Philosophy*, *18*(1), 1-18. https://www.jstor.org/stable/27759206

Smith, J. D., Beran, M. J., Redford, J. S., & Washburn, D. A. (2006). Dissociating uncertainty responses and reinforcement signals in the comparative study of uncertainty monitoring. *Journal of Experimental Psychology: General, 135*(2), 282-297. https://doi.org/10.1037/0096-3445.135.2.282

Smith, J. D., Coutinho, M. V. C., Church, B. A., & Beran, M. J. (2013). Executive-attentional uncertainty responses by rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: General*, *142*(2), 458–475. https://doi.org/10.1037/a0029601

Smith, J. D., Shields, W. E., Allendoerfer, K. R., & Washburn, D. A. (1998). Memory monitoring by animals and humans. *Journal of Experimental Psychology: General, 127*(3), 227–250. https://doi.org/10.1037/0096-3445.127.3.227

Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences, 26*(3), 317-339. http://dx.doi.org/10.1017/S0140525X03000086

Smith, J. D., Shields, W. E., Schull, J., & Washburn, D. A. (1997). The uncertain response in humans and animals. *Cognition*, *62*(1), 75-97. http://dx.doi.org/10.1016/S0010-0277(96)00726-3

Spaet, T., & Harlow, H. F. (1943). Solution by rhesus monkeys of multiple sign problems utilizing the oddity technique. *Journal of Comparative Psychology, 35*(2), 119–132. https://doi.org/10.1037/h0059354

Spinozzi, G., De Lillo, C., & Salvi, V. (2006). Local advantage in the visual processing of hierarchical stimuli following manipulations of stimulus size and element numerosity in monkeys (*Cebus apella*). *Behavioural Brain Research*, *166*(1), 45-54. https://doi.org/10.1016/j.bbr.2005.06.043

Spinozzi, G., De Lillo, C., & Truppa, V. (2003). Global and local processing of hierarchical visual stimuli in tufted capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, *117*(1), 15-23. https://doi.org/10.1037/0735-7036.117.1.15

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Templer, V. L., Brown, E. K., & Hampton, R. R. (2018). Rhesus monkeys metacognitively monitor memories of the order of events. *Scientific Reports*, *8*(1), 11541. http://dx.doi.org/10.1038/s41598-018-30001-y

Templer, V. L., Gazes, R. P., & Hampton, R. R. (2019). Co-operation of long-term and working memory representations in simultaneous chaining by rhesus monkeys (*Macaca mulatta*). *Quarterly Journal of Experimental Psychology*, *72*(9), 2208-2224. http://dx.doi.org/10.1177/1747021819838432

Templer, V. L., & Hampton, R. R. (2012). Rhesus monkeys (*Macaca mulatta*) show robust evidence for memory awareness across multiple generalization tests. *Animal Cognition*, *15*(3), 409-419. http://dx.doi.org/10.1007/s10071-011-0468-4

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(4), 799–806. https://doi.org/10.1037/0096-1523.20.4.799

Thomas, R. K., & Kerr, R. S. (1976). Conceptual conditional discrimination in *Saimiri sciureus*. *Animal Learning & Behavior*, *4*, 333-336. https://doi.org/10.3758/BF03214060

Thompson, R. K., & Oden, D. L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive Science*, *24*(3), 363-396. https://doi.org/10.1207/s15516709cog2403_2

Thorndike, E. L. (1933). A proof of the law of effect. *Science*, *77*(1989), 173-175. https://doi.org/10.1126/science.77.1989.173.b

Timberlake, W. (1993). Behavior systems and reinforcement: An integrative approach. *Journal of the Experimental Analysis of Behavior*, *60*(1), 105-128. https://doi.org/10.1901/jeab.1993.60-105

Timberlake, W. (1994). Behavior systems, associationism, and Pavlovian conditioning. *Psychonomic Bulletin & Review*, *1*, 405–420. https://doi.org/10.3758/BF03210945

Timberlake, W., & Lucas, G. A. (1989). Behavior systems and learning: From misbehavior to general principles. In S. B. Klein, & R. R. Mowrer (Eds.), *Contemporary learning theories: Instrumental conditional theory and the impact of biological constraints on learning* (pp. 237-276). Lawrence Erlbaum Associates.

Toates, F. (1998). The interaction of cognitive and stimulus–response processes in the control of behaviour. *Neuroscience & Biobehavioral Reviews*, *22*(1), 59-83. https://doi.org/10.1016/S0149-7634(97)00022-5

Tobin, H., Logue, A. W., Chelonis, J. J., Ackerman, K. T., & May, J. G. (1996). Self-control in the monkey *Macaca fascicularis*. *Animal Learning and Behavior*, *24*(2), 168-174. https://doi.org/10.3758/BF03198964

Tolman, E. C. (1932). *Purposive behavior in animals and men*. University of California Press.

Tolman, E. C. (1933). Sign-Gestalt or conditioned reflex. *Psychological Review*, *40*(3), 246–

255. https://doi.org/10.1037/h0071472

Tolman, E. C. (1934). Theories of learning. In F. A. Moss (Ed.), *Comparative psychology* (pp. 367–408). Prentice-Hall, Inc. https://doi.org/10.1037/11453-012

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55*(4), 189–208. https://doi.org/10.1037/h0061626

Tolman, E. C., & Honzik, C. H. (1930). "Insight" in rats. *University of California Publications in Psychology, 4,* 215–232.

Tomasello, M., & Call, J. (Eds.). (1997). *Primate cognition*. Oxford University Press.

Tu, H.-W., & Hampton, R. R. (2014). Control of working memory in rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Learning and Cognition*, *40*(4), 467–476. https://doi.org/10.1037/xan0000030

Tu, H.-W., Pani, A. A., & Hampton, R. R. (2015). Rhesus monkeys (*Macaca mulatta*) adaptively adjust information seeking in response to information accumulated. *Journal of Comparative Psychology*, *129*(4), 347–355. https://doi.org/10.1037/a0039595

Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science*, *2*(3), 67-70. http://dx.doi.org/10.1111/1467-8721.ep10770899

Visalberghi, E., & Fragaszy, D. (2012). What is challenging about tool use? The capuchin's perspective. In T. R. Zentall & E. A. Wasserman (Eds.), *The Oxford handbook of comparative cognition* (pp. 777–799). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195392661.013.0039

Visalberghi, E., & Fragaszy, D. (2013). The etho-*cebus* project: Stone-tool use by wild capuchin monkeys. In C. M. Sanz, J. Call, & C. Boesch (Eds.), *Tool use in animals: Cognition and ecology* (pp. 203-222). Cambridge University Press. https://doi.org/10.1017/CBO9780511894800.013

Wais, P. E., & Gazzaley, A. (2011). The impact of auditory distraction on retrieval of visual memories. *Psychonomic Bulletin and Review*, *18*(6), 1090-1097. https://doi.org/10.3758/s13423-011-0169-7

Wang, M. Z., & Hayden, B. Y. (2021). Latent learning, cognitive maps, and curiosity. *Current Opinion in Behavioral Sciences, 38*, 1-7. https://doi.org/10.1016/j.cobeha.2020.06.003

Warren, J. M. (1960). Solution of sign-differentiated object and positional discriminations by rhesus monkeys. *The Journal of Genetic Psychology*, *96*(2), 365-369. https://doi.org/10.1080/00221325.1960.10534304

Warren, J. M. (1964). Additivity of cues in conditional discrimination learning by rhesus monkeys. *Journal of Comparative and Physiological Psychology, 58*(1), 124–126. https://doi.org/10.1037/h0047425

Washburn, D. A. (1994). Stroop-like effects for monkeys and humans: Processing speed or strength of association? *Psychological Science*, *5*(6), 375-379. http://dx.doi.org/10.1111/j.1467-9280.1994.tb00288.x

Washburn, D. A. (2016). The Stroop effect at 80: The competition between stimulus control and cognitive control. *Journal of the Experimental Analysis of Behavior*, *105*(1), 3-13. http://dx.doi.org/10.1002/jeab.194

Washburn, D. A., & Astur, R. S. (2003). Exploration of virtual mazes by rhesus monkeys (*Macaca mulatta*). *Animal Cognition*, *6*, 161-168. https://doi.org/10.1007/s10071-003-0173-z

Washburn, D. A., Hopkins, W. D., & Rumbaugh, D. M. (1989a). Automation of learning-set testing: The video-task paradigm. *Behavior Research Methods, Instruments, & Computers, 21*, 281-284. https://doi.org/10.3758/BF03205596

Washburn, D. A., Hopkins, W. D., & Rumbaugh, D. M. (1989b). Video-task assessment of learning and memory in macaques (*Macaca mulatta*): Effects of stimulus movement on performance. *Journal of Experimental Psychology: Animal Behavior Processes*, *15*(4), 393–400. https://doi.org/10.1037/0097-7403.15.4.393

Washburn, D. A., & Rumbaugh, D. M. (1991). Rhesus monkey (*Macaca mulatta*) complex learning skills reassessed. *International Journal of Primatology*, *12*, 377-388. https://doi.org/10.1007/BF02547618

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, *20*(2), 158–177. https://doi.org/10.1037/h0074428

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 601–621. https://doi.org/10.1037/0096-1523.10.5.601

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 121–134. https://doi.org/10.1037/0096-1523.16.1.121

Zakrzewski, A. C., Perdue, B. M., Beran, M. J., Church, B. A., & Smith, J. D. (2014). Cashing out: The decisional flexibility of uncertainty responses in rhesus macaques (*Macaca mulatta*) and humans (*Homo sapiens*). *Journal of Experimental Psychology: Animal Learning and Cognition*, *40*(4), 490–501. https://doi.org/10.1037/xan0000041

Zelazo, P. D., & Frye, D. (1998). Cognitive complexity and control: II. The development of executive function in childhood. *Current Direction in Psychological Science*, *7*(4), 121-126. https://doi.org/10.1111/1467-8721.ep10774761

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., Chiang, J. K., Hongwanishkul, D., Schuster, B. V., Sutherland, A., & Carlson, S. M. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, *68*(3), i-151. https://www.jstor.org/stable/1166202

Zentall, T. R., Smith, A. P., & Beckmann, J. (2019). Differences in rats and pigeons suboptimal choice may depend on where those stimuli are in their behavior system. *Behavioural Processes*, *159*, 37-41. https://doi.org/10.1016/j.beproc.2018.11.012

Zentall, T. R., Wasserman, E. A. (Eds.). (2012). *The Oxford handbook of comparative cognition*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195392661.001.0001

Ziafar, M., & Namaziandost, E. (2019). From behaviorism to new behaviorism: A review study. *Loquen: English Studies Journal*, *12*(2), 109-116. https://doi.org/10.32678/loquen.v12i2.2378